9-2018

The Philosophical Foundations of PLEN: A Protocol-theoretic Logic of Epistemic Norms

Ralph E. Jenkins
The Graduate Center, City University of New York

Recommended Citation

Jenkins, Ralph E., "The Philosophical Foundations of PLEN: A Protocol-theoretic Logic of Epistemic Norms" (2018). *CUNY Academic Works.*
https://academicworks.cuny.edu/gc_etds/2825

The Philosophical Foundations of PLEN: A Protocol-theoretic Logic of Epistemic Norms

by

Ralph Jenkins

A dissertation submitted to the Graduate Faculty in Philosophy in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York

The Philosophical Foundations of PLEN:
A Protocol-theoretic Logic of Epistemic Norms

by

Ralph Jenkins

This manuscript has been read and accepted for the Graduate Faculty in Philosophy in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

Date	[Sergei Artemov, Ph.D.] Chair of Examining Committee
Date	[Nickolas Pappas, Ph.D.] Executive Officer

Supervisory Committee:

Advisor: Graham Priest

Sergei Artemov

David Papineau

THE CITY UNIVERSITY OF NEW YORK

Abstract

The Philosophical Foundations of PLEN: A Protocol-theoretic Logic of Epistemic Norms

by

Ralph Jenkins

Advisor: Graham Priest

In this dissertation, I defend the protocol-theoretic account of epistemic norms. The protocol-theoretic account amounts to three theses: (i) There are norms of epistemic rationality that are procedural; epistemic rationality is at least partially defined by rules that restrict the possible ways in which epistemic actions and processes can be sequenced, combined, or chosen among under varying conditions. (ii) Epistemic rationality is ineliminably defined by procedural norms; procedural restrictions provide an irreducible unifying structure for even apparently non-procedural prescriptions and normative expressions, and they are practically indispensable in our cognitive lives. (iii) These procedural epistemic norms are best analyzed in terms of the protocol (or program) constructions of dynamic logic.

I defend (i) and (ii) at length and in multi-faceted ways, and I argue that they entail a set of criteria of adequacy for models of epistemic dynamics and abstract accounts of epistemic norms. I then define PLEN, the protocol-theoretic logic of epistemic norms. PLEN is a dynamic logic that analyzes epistemic rationality norms with protocol constructions interpreted over multi-graph based models of epistemic dynamics. The kernel of the overall argument of the dissertation is showing that PLEN uniquely satisfies the criteria defended; none of the familiar, rival frameworks for modeling epistemic dynamics or normative concepts are capable of satisfying these criteria to the same degree as PLEN. The overarching argument of the dissertation is thus a theory-preference argument for PLEN.

Procedural epistemic norms are essentially rules that conditionally instruct the execution of epistemic procedures formed out of epistemic actions. That these kinds of norms characterize the requirements of epistemic rationality and do so ineliminably can be seen by examining, first, a number of problems for existing models of epistemic dynamics and, second, a set of important desiderata on formal frameworks for modeling epistemic dynamics and epistemic norms. Chapter 1 considers the problems while outlining the rest of the dissertation, and Chapter 2 establishes the desiderata, which take the form of principles about epistemic rationality and its norms that must be adequately formalized, as each desideratum is grounded in best solutions to a number of important epistemological disputes. Together, Chapters 1 and 2 argue that this set of problems and desiderata forms a sort of core or minimal theory of epistemic rationality to which any formal account must be adequate.

Formally, PLEN is a version of propositional dynamic logic, a logic of programs that encodes procedural operations on relatively basic computational steps with syntactic constructions and that interprets them over labeled transition systems or directed graphs in which transition relations or edges represent the traversal of states of computation according to programs. Protocols are, in PLEN, syntactic constructions that encode procedural operations on sets of basic epistemic actions. Epistemic actions are represented in the semantics of PLEN by combining tools from dynamic logic, epistemic logic, and belief revision. Rather than transition relations, sets of pairs of states, epistemic actions - and processes built by sequencing or other procedural operations on actions - are represented in PLEN by paths (sets of n-tuples) that build in information about distinct actions or processes that may link the same states in a transition. This results in the models of PLEN resembling multi-graphs rather than the directed graphs that underly traditional dynamic logics. This is a modest innovation with deep philosophical benefits with respect to solving the problems raised in Chapter 1 and capturing the desiderata of Chapter 2. The semantic rules of PLEN map protocols to (structures that represent) processes that result in change of epistemic states and states of the world, thereby mapping protocols to unique paths as well as preconditions and postconditions of execution. Protocols thus perspicuously represent certain features of epistemic rationality norms such as their procedural complexity, their verdicts with respect to various epistemic factors (e.g., whether a process, transition, or state is permissible or accords with a norm), their conditions of application, and the results of carrying out processes that accord with them. This machinery provides the basis for powerful and flexible logics with which to encode valid reasoning about these features of epistemic rationality and its norms. Chapter 3 defines a simplified version of PLEN and makes note of several important results. A separate technical document establishes these results as well as those deployed in Chapter 4 as part of the theory preference argument.

The defense of PLEN's theoretical preferability consists primarily of establishing (a) that PLEN solves the problems delineated in Chapter 1 and satisfies the desiderata articulated in Chapter 2, and (b) that none of the rival formalisms from belief revision, dynamic logic, dynamic epistemic logic, or deontic logic solve all of these problems or satisfy all of these desiderata equally well. Chapter 4 establishes the former of these claims by exploiting some of PLEN's technical features. It's nearly mechanical to show that PLEN suffers from none of the problems and satisfies all of the desiderata. Chapter 5 establishes the latter by examining the various families of rival formalisms one by one and showing the various ways in which they fail to either address the problems or formalize the desiderata. A final chapter summarizes the results of the dissertation and scouts future extensions of the research programme.

Front Matter

Preface

This is not a thesis about how to change your mind. No, wait... maybe it is. This is a thesis about devising and exploring a logical framework for thinking about the norms of epistemic rationality and the requirements they impose on agents that would be rational. Given that epistemic rationality norms define requirements not just on the synchronic relations among the items of content in our minds but on the dynamic processes by which we change them, this must be a thesis (at least partially) about how to change your mind.

The norms of epistemic rationality are best represented and reasoned about in terms of *protocols*, as they are understood in dynamic logics. That is the main thought of this dissertation. To put it a bit more precisely, protocols are formal devices that analyze procedural operations on spaces of basic actions and processes. Interpreted over structures that represent epistemic action, protocols are formal devices that analyze procedural operations on spaces of basic *epistemic* actions and *epistemic* processes. The models of dynamic logic - with some simple upgrades - provide just the structures to represent epistemic action. The requirements of epistemic rationality are defined by norms - rules, policies, or principles that dictate, prohibit, permit, or recommend epistemic factors (over others) under various conditions. Procedural epistemic norms - rules, policies, or principles that apply to epistemic actions and processes - are epistemologically important and unify many other kinds of epistemic norms as well as the epistemic normative propositions and prescriptions that articulate them. These procedural norms can essentially be captured by complex instructions formed out of procedural operations on epistemic actions and processes. Hence, the main thought.

To develop and defend the main thought, I devise a formal framework based on dynamic logics that deploys protocols to analyze epistemic rationality norms. The core of the framework is a "protocol-theoretic" logic of epistemic norms: PLEN. It's worth noting that there are several frameworks in dynamic and epistemic logic that take a general approach to norms that is similar to PLEN in some respects or others. Some systems have nearly identical formal languages to PLEN and/or domains of structures that are isomorphic to a subset or superset of the models of PLEN's languages.[1] Some systems even provide technical devices that are explicitly incorporated into PLEN's final form [114, 184]. Really, almost all of PLEN's innovations are parts of some alternative system or another. So, what's the point of developing PLEN rather than deploying an existing system?

The answer, in short, is that no existing system puts all of the pieces together the way that PLEN does. First, PLEN is a relatively simple logical framework that can:

- (i) provide devices for *explicitly* representing and reasoning about (protocols that encode procedural) norms of epistemic rationality,[2]
- (ii) capture the distinction between processes of epistemic change and the changes of epistemic state they bring about when completed, and
- (iii) formalize features of epistemic rationality that are inadequately recognized in both the formal and traditional epistemology literatures.

Properties (i) through (iii) capture, in an attenuated form, criteria of theoretical adequacy on formal - or informal - accounts of epistemic rationality. This claim is argued for over the

---

[1] See Chapters 5, 9, and 10 of [113] for something of an overview. See [174, 175] for some uses of a complementation operator similar to the parallel execution operator used in PLEN and for the analysis of normative - specifically deontic - expressions. See [30, 28] for discussion of the general form of the structures that model the languages. See [114, 184] for systems with models that bear important similarities to those of PLEN.

[2] PLEN is an *exogenous* logic [113] (p. 157). The structures that determine the features of the dynamic evolution of a system - programs, for instance - are explicit in the logical language.

course of Chapters 1 and 2. That PLEN's semantic structures and logic suffice for (i)-(iii) is argued for at length through Chapters 3 and 4.

Second, PLEN *uniquely* contains logical resources for (i)-(iii); no rival formalism can compete on all of these tasks. Frameworks for analyzing belief revision [5, 92, 67, 69, 84], credence update [182, 153], and dynamic evolutions of knowledge (as in dynamic epistemic logics and other frameworks [65, 86, 34, 233, 148, 131, 97, 74, 50, 158, 162, 245]) give models of how reasoning unfolds within the constraints imposed by specific sets of epistemic norms (deductive or probabilistic coherence, entrenchment, etc.) or normative assumptions (e.g., time-slice epistemology, probabilism, etc.) underlain by specific assumptions about the nature of epistemic dynamics. However, none of these frameworks provide general analyses of epistemic norms, and they don't provide explicit resources for reasoning about the logical properties of epistemic norms. More, they nearly universally fail with respect to (ii) and (iii). Deontic logics and normative systems do take on the project of analyzing norms and their logical properties.[3] However, even when they are tailored to the challenges of the epistemic case, they fail to account for (ii) and (iii) despite having resources that can be applied toward (i). The details of the foregoing claims and arguments for them are housed primarily in Chapter 5, with some preliminary remarks in Chapters 2, 3, and 4.

Defending a formal framework for philosophical purposes is a daunting task. This dissertation can really only be an opening move and one of limited scope, at that. Given that this is primarily a philosophical thesis, the technical exploration of the formal framework is neither comprehensive nor all that sophisticated. I've settled for merely developing the properties of PLEN related to the soundness of its deductive system and its effectiveness at treating the considerations in (i)-(iii). Beyond very basic observations, e.g., soundness and some theorems relating the constructions in PLEN to other novel constructions, I neglect rigorous exploration of the mathematical properties of PLEN that are standardly of interest in studying logical systems (e.g., completeness, complexity, compactness, decidabil-

---

[3] For a handbook approach to the entire field, see [89]. See [4] for the seminal introduction and [3] for a crossover with belief revision.

ity, expressive power).[4] This has been exchanged for extended philosophical examination and defense of the basics of the system and application of the system to the main task of analyzing epistemic rationality norms.

On the other hand, the amount of technical exploration required by the project entails that not all of the philosophical considerations that are germane to the project can be discussed, let alone engaged with at appropriate levels of detail. Defenses of substantive principles of epistemic rationality, robust investigations of connections with established projects in formal epistemology (e.g., learning theory) and the ethics of belief, and examination of connections to the broader formal and informal study of normativity and practical rationality, (e.g., the usefulness of PLEN as a general model of norms) will, along with further technical examination, have to be relegated to future work.

Where this project lies in the grand scheme of things is in the intersection of epistemology, applied logic, and philosophy of logic. The main interests and upshots of the project are epistemological, much of the core work is in applied logic, and the project embodies certain perspectives in the philosophy of logic. What binds these threads is a guiding interest in finite, flawed epistemic agents, their goals, and the tools they use to achieve them. Some epistemic agents, for instance, aim to achieve the goals of inquiry - things like truth, knowledge, and rationally justified belief. But every observable instance of such an agent is prone to performance error, subject to bias, and limited in cognitive resources. In order to meet their epistemic goals, such reasoners must find ways to recover from errors, to detect bias and course correct, and to overcome their limitations. In other words, the question is: how can beings like us overcome our frailties to meet the goals of inquiry? The answer, I think, is *epistemic methodology*.

Epistemic methodology consists of (1) systems of norms that we endorse and deploy to evaluate and regulate our epistemic behavior and (2) our reflection thereon and deliberation thereover. We don't (always) blindly endorse and deploy procedures and standards for

---

[4] The technical exploration of PLEN is carried out primarily in a separate technical document. [135]

carrying out our epistemic projects; we reflect on their properties (e.g., the effects of following them, the restrictions they impose, the commitments they entail), and deliberate over which of them to keep endorsing and deploying. We even devise and adopt new and better epistemic norms from time to time. Imagine trying to think clearly about the effects of policy decisions without statistics or working through a complex argument without using some of the formal techniques of logic. Extensive, deliberate use of these things is hardly native endowment.[5]

This project is part of the analysis of epistemic methodology. As such, the approach is that of procedural epistemology.[6] The focus of the project is not the analysis of fundamental elements of epistemic value. Rather, it focuses on the kinds of normative structures that are needed in epistemic regulation or epistemic advice. These, it turns out, must focus on the distinctive features of procedures and processes involved in changes of mind. Ultimately, this thesis is designed to contribute something to the reflective and deliberative activities involved in epistemic methodology; it's an attempt to contribute to the general project of helping us think carefully and to think of ways that we might do better.

Acknowledgements

I owe much to the dynamic logic paradigm and its diverse elaborations in philosophy, logic, and computer science. Among these, see [30, 175, 184, 187, 229, 227, 231, 34, 131]. These are merely some of the many sources of technical invention and philosophical insight that have informed this project, and to which I owe a deep debt of gratitude. Oversights, errors in argument, technical failings, expository foibles, strange locutions, misguided attempts at humor, and other forms of misstep are my own. This is reflected in a stylistic choice. Despite the focus on logic in this thesis, I opt for the direct, first-person “I” rather than the “royal ’we”’ that is common in formal literature. One can implicitly distribute credit to those I've

---

[5] See the rather large empirical literature on what can, with some controversy [100, 101, 102], be called “irrationality”, starting from the seminal [136] and continuing through all of [2, 213, 242, 238, 73] and beyond.

[6] Arguably, this puts the project in line with the true mainstream of epistemology, in the long run. For influential modern and recent examples, see: [64, 173, 16, 17, 194, 196, 197].

learned from when things go right in this thesis. Else, the civilian “I” directs responsibility where it properly belongs.

This project couldn’t have ever come about if not for the patience and guidance of the faculty at the Graduate Center, especially the supervisor of this dissertation: Graham Priest. Aside from directly providing virtually endless technical guidance and philosophical insight, Graham’s continued defense and development of dialetheism over the years in ever more diverse and surprising ways is an exemplar of the “hard core” research strategy. It’s an example that has fundamentally re-shaped my philosophical goals and professional aspirations.

There are very many ways in which this project could have failed, whether from technical or philosophical error or from the numerous ways in which extra-academic factors get in the way. Somehow, I haven’t quite run into them face-first, yet. The explanation lies partly in the combination of guidance and freedom from constraint provided by Graham and the Graduate Center. Another part of the explanation is Rutgers-Newark, whose small, excellent, and deeply encouraging philosophy department stoked a flame that, via Jeff Buechner, spread to the Grad Center.

A final note about notation: I frequently employ nonstandard and ad hoc notation. My conventions in this work are explicitly set wherever it isn’t an obvious matter of trying to select a letter, symbol, or subscript that is easy to remember. Further, the indices of enumerated claims like “(i) ..., (ii)”, “(a) ..., (b) ...,” and so on are restricted to the sections in which they first appear. Thus, references to “claim (i)” or the like refer only within explicitly titled sections or subsections. If a claim requires reference across sections, it is given a distinctive label or it is cross-referenced with its home section, as in “claim (i) of section x” or “clause (ii) of the Preface”.

Contents

List of Figures

List of Tables

Part I

The Protocol-theoretic Account of Epistemic Rationality Norms

Chapter 1

Epistemic Dynamics, Protocols, and Epistemic Rationality

1.1 Introduction

Epistemic rationality imposes various kinds of requirements on the epistemic factors of reasoners. That is, where "epistemic factors" encompasses the belief or acceptance states and ways of changing belief/acceptance states (inter alia) of reasoners, the thought is that reasoners must ensure that their epistemic factors meet certain conditions, fall within certain parameters, or have certain properties; if a factor fails to meet those conditions, the agent is not as she rationally should be, and the factor itself may be irrational.[1]

In general, there are distinct kinds of epistemic requirement that focus on the distinct kinds of epistemic factors. One kind of epistemic requirement focuses on belief states. There are, of course, irrational states of belief, as the ascendancy and weaponization of internet

---

[1] In all of what follows, I focus on the rationality - broadly construed - of epistemic factors. It might be thought that it is reasoners themselves that are the unique appropriate objects of epistemic appraisal - particularly where rationality is concerned. This can't be right. It's a commonplace feature of the concept of rationality that we distinguish the rationality of a view that agent i has or a move i has made in reasoning from the rationality of i herself. While it is plausible that appraisals of rationality apply to agents themselves in addition to their epistemic factors, I ignore this complication in all of what follows. There should be no special difficulty extending the work here to agents.

conspiracy theorizing has made dramatically and depressingly clear. More formally, there are probabilistically and logically incoherent states of belief as well as states of belief featuring individual beliefs or sets of beliefs for which there is inadequate evidence or overwhelming counterevidence. These are forms of irrational belief state, and epistemic rationality requires that one avoid being in such belief states or that one find one's way out of them. A second kind of epistemic requirement focuses on ways of changing belief states; there are irrational processes and moves in changing states of belief, such as wishful thinking and fallacious or biased reasoning.[2] Epistemic rationality *requires* that one avoid or halt these processes of belief change and execute other - normative - processes of belief change instead.[3]

Call the rules, policies, or principles that define these requirements on reasoners the *norms of epistemic rationality* or, interchangeably, *epistemic rationality norms*.[4] This dissertation is an extended defense of the *protocol-theoretic account* of epistemic rationality norms. The core thought of the protocol-theoretic account is that epistemic rationality norms are best represented, analyzed, and reasoned about in terms of *protocols*, as they are defined in dynamic logics.[5] There are two main reasons why I think that this account should be

---

[2] At this point, I'm assuming that there are what Kolodny [147] calls "process requirements" of epistemic rationality. This is sometimes thought to be controversial on the basis of commitment to what I will call *epistemic staticism*, which is the view that there are really only requirements on the synchronic features of epistemic states. Some threads of staticism include time-slice epistemology [179, 127] and some forms of probabilism [53] and epistemic consequentialism [57, 83]. It has been impressed on me from workshopping the idea in dissertation seminars that the idea that there are dynamic requirements is pre-theoretically clear and plausible, so, for those inclined toward epistemic staticism, please bear with me. Epistemic staticism receives due attention in Chapter 2.

[3] Rationality may even require reversing or otherwise correcting the effects of prior processes or foreclosing the possibility of alternative processes. The infamous recovery principle of AGM [214] is a prima facie plausible example of the reversing or correction, but more complex belief change algorithms may require such reversal, e.g., in cases where inferences were drawn from conclusions that themselves were drawn in error or "learning by erasing" [99, 161].

[4] Perhaps it would be better to define epistemic rationality norms as the regularities that characterize the requirements in question (perhaps indirectly by characterizing an ideally rational model of epistemic performance), so as not to create the impression that the rules are written and then deliberately implemented. As I'll argue in Chapter 2 and in the technical document [135], it is the case that epistemic rules are at least sometimes explicitly written, adopted, and implemented, but it's plausible that at least some epistemic norms are more like evolved patterns of personal or sub-personal cognitive behavior than rules that are understood and adopted. The idea is that, in order to be rational, sometimes we are required to instantiate some of these patterns, which in turn imposes further requirements. In any case, the question of what I take norms to be is briefly taken up in Chapter 3.

[5] To be precise, the term "protocols" is prevalent in dynamic epistemic logic [74, 65]. Literature on propositional and first-order dynamic logic, or dynamic logic in general, seems to favor the term "programs",

accepted. *First*, the protocol-theoretic account is developed with a formal framework - PLEN - that models epistemic rationality norms with protocols, and PLEN can be shown to satisfy theoretical desiderata relevant to understanding epistemic rationality. *Second*, PLEN does so uniquely; no rival, predecessor, or otherwise alternative formalism captures the features of epistemic dynamics and epistemic rationality norms that PLEN does. More precisely, the PLEN framework models epistemic dynamics with a level of accuracy that no other framework does, and it formalizes an epistemologically important set of features of epistemic rationality and its norms that rival and predecessor frameworks fail to recognize. Hence, rational theory preference dictates (ceteris paribus) adoption of the PLEN framework over its rivals and predecessors, and this implies acceptance of the protocol-theoretic account.[6]

This opening chapter serves as both an introduction and as a piece of substantive argumentation aimed at laying the groundwork for the overarching arguments. The thrust of the chapter is that there are a number of problems that models of epistemic dynamics must solve in order to serve as an underlying model for theories of epistemic norms. Section 1.2 lays out some important features of the conceptual framework underlying the project. Much of this section can be skipped by readers familiar with the basic concepts and models of epistemic dynamics, the basic formal idea of protocols or programs, and the perspective of procedural epistemology. The essential parts are sections 1.2.1 (including 1.2.1.1), 1.2.2 (but not the subsections), and 1.2.3. The rest can be skipped with relative safety if the reader is confident that they see what I'm up to.

Section 1.3 is the core of the chapter, as it develops several problems for the modeling of epistemic dynamics that are basically damning to formalisms that are to be deployed for representing and reasoning about epistemic norms. The upshot of 1.3 is that there is a

---

which is more consonant with the origins PDL as a logic of programs. I am primarily following [184] in deploying the term "protocols", though, formally, protocols and programs are equivalent constructions.

[6] I mean *acceptance*, as understood by Cohen [55] (also Stalnaker [235] and perhaps Restall [210]). "[T]o accept that p is to have or adopt a policy of deeming, positing, or postulating that p — i.e. of including that proposition or rule among one's premisses for deciding what to do or think in a particular context, whether or not one feels it to be true that p." (p. 4) Equivalently, I think the two main arguments justify acceptance of the account as a working hypothesis and deploying PLEN as a framework for developing the hypothesis.

theoretically important distinction to be drawn between *processes* of epistemic state change and the resulting changes to the contents of epistemic states. It is argued in 1.3 that models of epistemic dynamics that do not draw this distinction (a) fail to accurately model epistemic dynamics, and (b) this failure results in errors in thinking about the requirements of epistemic normativity. Conversely, it is argued that failure to account for the requirements of epistemic normativity also lead to inaccurate models of epistemic dynamics. This bleeds over into a first stab at articulating the desiderata or criteria of adequacy that PLEN will be shown to uniquely satisfy. Section 1.4 integrates the foregoing problems and the putative criteria of adequacy into an articulation of the fundamental problem that the protocol-theoretic account and PLEN are devised to address. Section 1.5 details the chapter plan for the rest of the dissertation.

1.2 Preliminaries

There are three big ideas underneath the project, and these ideas will help to clarify the protocol-theoretic account and lay the ground for the main arguments of the thesis. The first big idea is the concept of epistemic dynamics [86, 91]: that there are numerous possible epistemic states that a reasoning epistemic agent could be in at any time and numerous possible paths or trajectories through the space of epistemic states that such a reasoner can follow in transitioning from one state to another by execution of actions or processes. That is, reasoning is a dynamic process; our epistemic states (belief states, acceptance states, etc.) evolve over time along routes that are determined by the ways that we reason. The ways that we reason can be restricted or enabled in various ways. Some of these restrictions are due to cognitive or environmental limitations. Others are normative or methodological restrictions. An important distinction in thinking about dynamics is that between transitions among states or even paths of states and the actions, processes, or procedures that are used to bring about those transitions and paths of states. That is, a process p might take an agent from one state, s, to another, s', that is distinct because it has distinct content, but so might a process q. These processes are distinct despite their identical start points and end points.

The second idea is that dynamic logics are interpreted over plausible and useful models of epistemic dynamics. Formally, the models are equivalently described as labeled directed graphs [93], labeled transition systems [35], or multi-modal Kripke structures [33]. These are abstract models of dynamic systems; systems comprising sets of states that change over time by various sorts of events, actions, or processes. The nodes or worlds in these structures abstractly represent states (without internal structure) of the system, the edges represent changes of state, and the labels or modalities represent processes or actions that result in change of state. These systems naturally represent such dynamic systems as computers traversing various computational states as they execute programs, evolutions of knowledge bases by learning processes, or the unfolding of information in the minds of groups of agents in communicative situations [30, 34, 65, 113]. Given the first idea, these are transparently useful models of epistemic dynamics: the states are epistemic states, the edges connecting states are changes of epistemic state, and the labels are epistemic actions or epistemic processes.

The third idea is that there are procedural epistemic norms, and these are central to the analysis of epistemic rationality. This claim goes beyond the idea that there are diachronic or, more generally, dynamic norms that specify that some transitions or movements among states are (e.g.) rational. As noted above, epistemic actions and processes are not interchangeable with the changes of the content of epistemic states. Procedural norms carry information that is distinct from and additional to that carried by "merely" dynamic norms.[7] Epistemic actions and processes can be broken down into sequences, combinations, choices, and other procedural operations on more basic actions and processes. Some sequences, combinations, choices, and other operations are rational, others not.[8] There are epistemic rationality norms

---

[7] Though, as I'll argue in Chapter 4, the important normative information carried by dynamic norms - and even static norms that assign normative statuses to epistemic states - can be "read off" from procedural norms, universally.

[8] This oversimplifies in two salient ways. First, the list of operations isn't exhaustive. As we will see, there are several other basic combinatorial instructions that are useful in specifying the procedural content of norms. It's important to note that these procedural operations are recursive, and so very complex procedural

that define the rational procedural operations; they carry information or instructions about what procedural operations on actions and processes to execute, modulo one's epistemic circumstances, in order to be rational. This is, perhaps, not the exclusive kind of epistemic norm, but it is undeniably a kind, and it's a kind that is indispensable for the function of epistemic norms in cognitive life. More, every epistemic norm is logically associated with this kind of procedural norm, so the procedural norms unify epistemic norms and normative expressions.[9]

I discuss epistemic dynamics and the basic formal insights for thinking about epistemic dynamics in the next two subsections, 1.2.1 and 1.2.2. These subsections introduce some important ideas about the basic model of dynamics underlying this project as well as the notion of protocols and its use for thinking about constraints on dynamics of various kinds, especially constraints of capacity and feasibility. Section 1.2.3 addresses a philosophical objection that has been with the project since its earliest days and clarifies the philosophical perspective from which epistemic rationality norms are approached. The objection is targeted to the emphasis on procedural epistemic norms, and I argue that it can be easily set aside by taking the perspective of procedural epistemology as well as by noting an important technical point raised later in the dissertation.

1.2.1 Epistemic Dynamics

Call any system that deliberately reasons, learns, or acquires and processes information an *epistemic agent*, *epistemic system*, or, interchangeably, a *reasoner* or a *cognizer*.[10] Epistemic

---

instructions can be encoded with protocols. Second, some operations on actions are more rational than others; that is, rationality sometimes appears to be scalar and other times comparative rather than categorical. There are also rational deontics: some operations are rationally obligatory, others are forbidden. The remarks in the main text apply equally well to these more nuanced verdicts of epistemic rationality.

[9] The third big idea is defended by means of both (a) the arguments in Chapter 2 concerning the features of epistemic norms necessary to the cognitive function of epistemic norms and (b) the reduction thesis defended in Chapter 4.

[10] The term "deliberately" is not accidental, here. There are worries - e.g., Alston's classic argument in [7] - that discussions of epistemic normativity - at least deontic kinds of normativity - are moot because epistemic actions are not voluntary. At this stage, I'm simply declaring commitment to the claim that this worry is mistaken. In Chapter 2, I defend the claim that much epistemic action - and most epistemic action

agents are usefully analyzed in terms of two factors: *epistemic states* and *epistemic dynamics.* Epistemic states are reasoners' representational states, the models of the world that the agent constructs, manipulates, and transforms in order - among other things - to guide its actions. These will be discussed a bit more in Chapters 2 and 3, but, at a minimum, epistemic states are composed of *epistemic judgments* and their *contents.* Contents are usefully encoded with sentences of a language, natural or formal. Standard taxonomies of epistemic judgments include full and partial belief (credence), acceptance, and other cognates or alternatives. So, for instance, one's epistemic states can be partially described by a set of believed sentences or by a credence function on a set of sentences. See Table 1.1 for a sample of several kinds of partial epistemic state description with some common (and uncommon) bits of notation for them.[11]

Epistemic states change over time as a result of various processes. *Doxastic* or belief-related states, for instance, can be changed by processes such as reasoning or observation. Full beliefs are expanded, contracted, revised, or sustained over reasoning processes, credences are updated, and so on. Take a toy example:[12]

The Detective A detective is investigating the head, K, of a murderous criminal organization.

At time t_1, she believes, on the basis of the available forensic evidence, that K uses

---

of interest in this project - is voluntary. For the impatient, my response to the classic involuntarist arguments is rather similar to that of [171].

[11] The notation $b(P)$ is mine, and, though I use it here to denote belief, I use it in later sections in a noncommittal fashion - belief or acceptance. In doxastic and epistemic logic, belief is often denoted with a modal operator. In any case, I omit a full taxonomy of notations. However, do note that taking on different categories of epistemic judgment as sort of a founding conceptual assumption is common in various fields. Full or all-or-nothing belief is prominent in, e.g., Epistemic Logic [129], Doxastic Logic [162, 158, 224], Belief Revision [118, 91, 92], and in epistemology, generally [121, 1, 7, 23, 54, 57, 83, 104, 164, 167, 173, 194, 195, 196, 217]. See, especially, Staffel's [234] critical discussion of its prominence and Leitgeb's work relating full belief to credences [155]. Credences are prominent, of course, in Bayesian epistemology [182, 53, 155, 153] and also in certain branches of belief revision [84, 149] and doxastic logic [20]. Acceptance and rejection are discussed independently by Cohen [55] and Restall [210], with roots in philosophy of science, re: theoretical commitments [204]. Supposition is a medieval logical notion [207] given an interesting analysis in an inductive context by Levi [160]. Perhaps the list of epistemic judgments is even more expansive. Some potential add-ons: predicting P, suspecting P, guessing P, having the hunch that P, and so on. This is one of many points that suggests future work in analyzing epistemic dynamics.

[12] Any noticeable lack of realism is probably due in part to the distorting effects of simplification and to my having augmented a scenario from a comic book: *Deathnote* by Tsugumi Ohba and Takeshi Obata. Nonetheless, it serves as an example of complex epistemic dynamics with interactions among information generating and inferential actions.

Table 1.1: Epistemic Judgments

Judgment	Content	Notation	References
(Full) Belief	P, Q, R,...	$b(P)$, $b(\neg P)$, $P \in K$ Informally: P (resp. $\neg P$) is in one's belief set, P is fully believed	[91, 121]
(Full) Disbelief	P, Q, R,...	$d(P)$, $d(\neg P)$, $P \notin K$, $\neg P \in K$ Informally: P (resp. $\neg P$) is not in one's belief set, P is fully disbelieved	[91, 121]
Credence (Partial (Dis)Belief)	P, Q, R,...	$Pr(P)$, $Pr(\neg P)$, $Pr(P \mid Q)$ Informally: P (resp. $\neg Q$) has subjective probability Pr (given Q)	[86, 182, 53, 155]
Acceptance	P, Q, R,...	$P \in X$, $\neg P \in X$, $P \notin X$ Informally: P is accepted	[55, 210]
Rejection	P, Q, R,...	$P \in Y$, $\neg P \in Y$, $P \notin Y$ Informally: P is rejected	[55, 210]
Supposition	P, Q, R,...	$P \in T(K, P)$ - with $T(K, P)$ the suppositional corpus supported by belief set K Informally: P is a supposition supported by one's belief set, K $P > Q$ - conditional belief on the basis of supposition Informally: Supposing P, Q rationally follows (were one to fully believe P, one would be committed to fully believing Q)	[160, 161]

(exclusively and exhaustively) either City A or City B as a kind of home-base. This is information she needs in order to improve her chances at apprehending K. At t_1 she also believes, on the basis of the best available psychological profile, that if she leaks information that is insulting or embarrassing to K in a way that can only be accessed in A, and K is based in A, then the frequency of crimes matching the modus operandi of K will increase (as a form of response to the insult, let's say); else, no increase in A will be seen.[13] Between t_1 and t_2, she acquires an independent stock, E, of inductive or defeasible evidence that K is based in A, and she also knowingly leaks insulting information in a way that renders it accessible only in A. Having not jettisoned any beliefs between t_1 and t_2, she ends up at time t_2 with a new doxastic state containing all of the foregoing beliefs. After t_2, she consults crime statistics, from sources S1 and S2, concerning the period since her leak. S1 reports no increase in the frequency of the relevant crimes, and S2 reports a dramatic increase. Not trusting S2, she comes to believe that K-like crimes did not increase in frequency. This process, as all do, takes time, so at t_3, some time after t_2, the detective has a new set of beliefs that includes all of her prior beliefs and the belief that K-like crimes have not increased in frequency. Suppose that she has several options going forward from t_3: she might deduce that K is in B, she might infer inductively that K is in A due to her evidence E (perhaps revising away the apparent inconsistency), or she might conclude that K is in B on the basis of consulting a psychic.[14] Doing any of these would take her to a new state after some time. She deduces that K is in B, ending up with a final belief state at t_4 containing the prior beliefs and her final conclusion about K's base of operations.

---

[13] Suppose that such increases provide no additional evidence about K's home-base location.

[14] To be exact about her deductive options: Assume classical logic. Since there was no increase, it follows by modus tollens that either there was no leak of insulting information or that K is not based in A. Since she knows that she leaked the insulting info, disjunctive syllogism entails that K is not based in A. From the exclusive disjunction of "K is based in A" and "K is based in B", she deduces that K is based in B.

The detective's epistemic state at t_1 is partially described by her beliefs (or at least the propositions she accepts), her *doxastic state.* The content differences between the detective's doxastic states at some pair, t_i and t_{i+1}, entail that the detective has distinct epistemic states at t_i and t_{i+1} in virtue of the nonidentical doxastic states. The additions of belief at t_2, t_3, and t_4 are processes that thus result in changes from one epistemic state to another over time. Call changes from one epistemic state to another *epistemic transition*s.

The actual and possible epistemic transitions, actual and possible processes of epistemic transition resulting from sequences and choices among actions, and the properties of and relations among the foregoing items constitute a system's *epistemic dynamics.*[15] A final terminological note: *epistemic factors* will be my catch-all term for epistemic states and the various parts that compose epistemic dynamics, e.g., changes of epistemic states, paths of epistemic state evolution, epistemic actions and processes of epistemic state change, and methods or procedures for structuring actions and building processes.

The Detective case illustrates a number of important features of epistemic dynamics. First, it illustrates the way that complex processes can be constructed out of more basic actions. The transition from her first belief state to her final belief state is effected by sequencing multiple, smaller processes, often making choices among alternatives. Second, it illustrates why it's important for epistemological analyses to focus on the alternative, unactualized trajectories as well as their relations with each other and with the actual one.

Presumably, the detective's beliefs evolved from earlier states to those at t_1 and will unfold past t_4, but even if she was born with innate knowledge at t_1 and she dies at t_4, the actual sequence of doxastic changes is merely a part of her epistemic dynamics that matters for epistemological purposes. Given the stipulation that she could have based her belief on S2's data or carried out those other processes at t_3, there are other possible trajectories for

---

[15] The phrase "epistemic dynamics" can also denote the study or theory of these epistemic changes. I'll let context disambiguate. I won't distinguish between *epistemic kinematics* (the "laws of motion" taking states to other states given inputs or forces of change) and epistemic dynamics as some kind of more comprehensive analysis of the epistemological features of inputs as well as the kinematics. Both [86, 91] discuss the distinction usefully.

her through the space of possible doxastic states. There are properties of and relations among these other possible processes and trajectories worth noting. For instance, some processes *enable* other processes. The deduction at t_3 required the acquisition of the belief that the crimes in A did not increase in frequency; the latter action enabled the former because no comparable deduction could have been performed from any of the detective's other beliefs. This obviously matters for practical or cognitive control purposes, and it also matters if, for instance, epistemic ought implies can; given an ought implies can principle, it can't be the case that the detective ought to deduce that K is in B if her prior actions did not enable the deduction.

Further, there are some general relations among possible epistemic processes that stand out. Some processes are more complex or more difficult to carry out than others. Some distinct pairs of processes effect identical changes of state. Some processes are normatively correct, and others are not. These properties might interact; the complexity or difficulty of processes might, for instance, impact their normative statuses. Putting these observations together, note that distinct processes can result in the same doxastic changes and yet still be normatively distinct. The detective's deduction gives the same conclusion as accepting the outputs of the psychic consultation, despite their manifest epistemological differences.[16] This observation carries some weight, as we will see below and in more detail in Chapter 2.

1.2.1.1 Labeled Transition Systems

It's useful to think of epistemic dynamics in terms of *labeled transition systems* (LTSs) or (polymodal) *Kripke structures*, though slightly more complicated structures will be introduced to model epistemic dynamics in future chapters due to problems for standard

---

[16] I take it for granted that the deduction is rationally acceptable in ways that the acceptance of psychic testimony isn't - assuming a world like ours. There is much more to say about this toy example. One question to raise: what are we to say about her evidence E? Is it uncontroversially more rational to conclude that K is based in B on the basis of the deduction described than it is to conclude that K is based in A given E? The stipulations of the case alone don't answer this question, of course, but it does enough to raise the issue of reasoning from bodies of beliefs that contain numerous and prima facie conflicting logical and evidential relationships. These are important questions, and I won't be answering them in this project.

LTS-based models of epistemic dynamics.[17] A labeled transition system, $\langle S, \rightarrow_{a \in A} \rangle$, is a non-empty set, S, with a set of relations, $\rightarrow_{a \in A}$, on S. Each element of $\rightarrow_{a \in A}$ is a binary relation on S, $\rightarrow_a$, that is indexed to a member of the set A. These are the *transition relations*. Each ordered pair of elements of S, $\langle s, s' \rangle$, is a *transition*, which represents a change from the left element, s, to the right element, s'. These bare transitions are distinct from *labeled transitions*, which are triples, $\langle s, a, s' \rangle$. The distinction between labeled and "bare" transitions is important to note. The labeled transitions $\langle s, a, s' \rangle$ and $\langle s, b, s' \rangle$ are clearly distinct labeled transitions, but they connect a single pair of elements, and so are both related to the unique transition $\langle s, s' \rangle$. Labeled transition systems can be defined in terms of both binary transition relations, $\rightarrow_{a \in A} \subseteq S \times S$, or labeled (ternary) transition relations, $\rightarrow_{a \in A} \subseteq S \times A \times S$. This choice is not of any importance at this stage of the exposition, but will become a matter of significance in later sections and in Chapters 3 - 5.

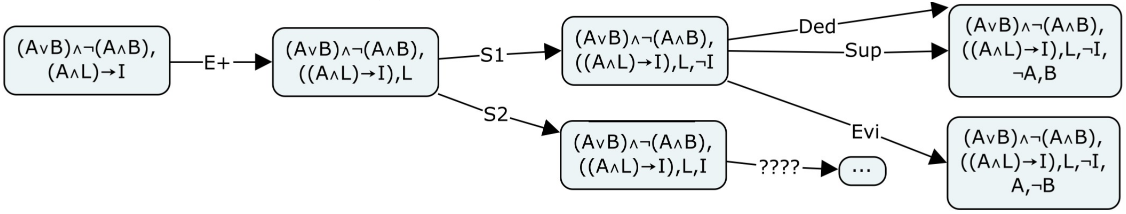
Interpret S as a set of possible epistemic states, A as a set of basic actions that can effect changes from one epistemic state to another, and the transition relations as showing us what transitions among states result from what actions or, equivalently, what states can be reached by what actions from any given state. Each transition can be thought of as a change of state and each labeled transition as a token process or act that takes a system from one state to the next. An LTS then displays a space of possible trajectories through the space of epistemic states and the actions (and courses of action) that could take a system through each one. Letting states be associated with sets of formulae of propositional logic (say, via an interpretation function), the Detective's epistemic dynamics can be represented with the LTS in Figure 1.1.

Formal frameworks for analyzing epistemic dynamics tend to broadly resemble the basic structure of transition systems. LTSs are a form of *branching time model* of dynamics [131] which stand in prima facie contrast to *input assimilating model*s [118]. The former represent dynamics with directed graphs or trees where the nodes are states and the edges are ways of

---

[17] See section 1.3.

Figure 1.1: The Detective's LTS



The detective starts at the leftmost state and moves right over time. Let each column of states from left to right represent the possible states for the detective at $t_1 - t_4$, respectively. First, she acquires evidence E (by action E+) that K is based in City A and recognizes that she made the leak, transforming her state at t_1 to the state at t_2. From there, she can deploy S1's data or S2's data to update her beliefs concerning the crime rates matching K's MO to take her to either of two distinct states at t_3. From either of the resulting states, she can deploy several processes to draw a conclusion about where K is based at t_4. She might reason from the psychic consultation (action Sup), she might deduce (Ded) that K is in B from her beliefs at t_3, or she might revise in accordance with evidence E (action Evi) to end up believing that K is based in A. Note that she hasn't yet figured out that the last move results in inconsistency. This particular LTS might encode only a fragment of her epistemic life, and so be a small subtree out of a vastly larger LTS. What actions are available at t_4 or t_3 after trusting S1, for instance, are left unspecified.

moving from state to state. In these models, we can represent histories or sequences of states, actions, or events through which a system unfolds or that an agent experiences or carries out. We can talk about what histories converge to a certain point and what possible histories unfold from that point. Epistemic Temporal Logic (ETL) [74, 65, 131] is a standard-bearer in thinking of epistemic dynamics in these terms.

Input-assimilating models take the form of functions from a pair comprising an input and a state, to a new state. Typically, these models have more complex and useful structures for representing epistemic states than mere nodes in a graph or LTS. That is, they typically have formal objects that can represent distinct contents or parts of epistemic states. Dynamic Epistemic Logic (DEL) [65] and AGM Theory (AGM) [4, 91] take this form.

Branching time models and input assimilating models can be merged, as van Benthem et al [37] and Hoshi [131] show for DEL and Epistemic Temporal Logic (ETL); every DEL model generates an ETL model by means of their construction. Independent of the details of this result, it's not difficult to think about input assimilating models in terms of branching time models. As Gardenfors notes, it's elementary that any function from input-state pairs to states can be treated as a function from states to states indexed to inputs or, perhaps, functions from inputs to pairs of states [91]. Such functions will be equivalent to transition relations and so related to subsets of paths in a branching time model - those paths that start at one element of the transition and end at the other. This point will be more carefully articulated in Chapter 4.

There are numerous formal frameworks for analyzing epistemic dynamics in general and for analyzing specific kinds of epistemic dynamics. Many of the most well-known and well-developed take the form of input assimilating models. To get a feel for models of epistemic dynamics, I briefly explicate the most basic and important parts of a few of the most well-developed input assimilating frameworks for epistemic dynamics to acquaint the reader with the general idea of thinking of epistemic dynamics formally and to pave the way in the next section for highlighting where they fall short.[18]

1.2.1.2 Dynamic Epistemic Logic (DEL)

One of the best and most well-known general formal frameworks for epistemic dynamics is Dynamic Epistemic Logic.[19] DEL is based on Epistemic Logic (EL) which applies modal logic to the analysis of knowledge and related concepts.[20] *Epistemic models* are triples, $\langle S, R, V\rangle$, of a non-empty set, S, a set, R, of binary transition relations $R_i \subseteq S \times S$ such that each transition relation is indexed to elements of a set, I, and a valuation function, V.

---

[18] It's useful in the next sections to try to visualize the models of dynamics as directed graphs. This is a bit misleading, though. First, it's hard to visualize anything but very simple graphs for the dynamics in DEL, because product updates connect epistemic models, which are just graphs themselves. We get graphs with nodes that are graphs! Second, the following systems are input assimilating models, and so don't have very complex dynamics - just functions from state to state. However, as Hoshi and van Benthem showed, these kinds of systems can be connected to branching time models, and there, visualizing the states in these systems developing as though in a graph is helpful.

[19] DEL is one of the best in the sense that numerous other models of epistemic dynamics can be embedded within DEL, and so subsume them. AGM, for instance, has been shown to be equivalent to a fragment of DEL; one just goes through the possible worlds models of AGM [108] to arrive at epistemic models. This is shown in Chapter 3 of deeply useful textbook [65]. If DEL is equivalent to a fragment of PLEN, then PLEN will be, in this sense, be even better than DEL. This is a plausible conjecture, but one I leave unexamined until future work.

[20] See Hintikka [129], of course.

The elements of S are interpreted as possible ways the world could be, and the relations, R_i, are interpreted as indistinguishability for an agent $i \in I$. The valuation functions, V, map sets of expressions of the language of epistemic logic L_{EL} to worlds in S. This gives an interesting definition of knowledge: *i knows A* iff A is true in every world linked by i's indistinguishability relation. In other words, i's knowledge is supposed to allow her to distinguish possible ways the world could be. If A is false at some world, w', that i can't tell from w, then she doesn't know A at w. This kind of model enables the definition of knowledge for all of the agents in I, as well as notions of common knowledge, belief and other doxastic states, and a basic logic of those concepts.

DEL captures the dynamic aspects of epistemic state change by encoding the epistemic states of multiple agents with epistemic models and by modeling events and their informational effects on epistemic models. An *event model* is a triple $\langle E, \rightarrow_i, pre\rangle$ where E is a set of events e, $\rightarrow_i$ is an indistinguishability relation among events for each agent, i, and *pre* is a mapping of sets of expressions of the language of DEL, L_{DEL}, to events in E. Event models encode sets of possible events and whether or not agents can tell events apart. $e_1 \rightarrow_i e_2$ tells us that i can't tell whether e_1 or e_2 happened; when e_1 happens, i thinks it's possible that e_2 happened. The *precondition* function, *pre*, tells us what the preconditions are of each event. If $pre(e) = A$, then e can occur only if A is true.

DEL provides an input-assimilating model of epistemic dynamics with the addition of product update functions. *Product updates* are functions from pairs of epistemic model and event model, $M = \langle S, R, V\rangle$ and $E = \langle E, \rightarrow_i, pre\rangle$, to an epistemic model, $M' = \langle S', R', V'\rangle$ such that:

- $S' = \{(s, e) \mid s \in S, e \in E, pre(e) \subseteq V(s)\}$,
- $(s, e)R'(s', e')$ iff sR_is' in M and $e \rightarrow_i e'$ in E, and
- $V'(s, e) = V(s)$ for all atomic formulae of L_{DEL}

The input of event e to a model, M, results in M'.

Epistemic models represent epistemic states by encoding (e.g.) knowledge and belief in the coupling of their indistinguishability relations and their valuations. What an agent knows depends on where (at what world) in the epistemic model the agent is and what is true at the indistinguishable worlds. Belief can be defined in similar ways. Event models represent information-changing events. Product updates encode the effects of various kinds of information-changing events on agents' indistinguishability relations, and thus on what agents can know (believe, etc.). That is, product updates represent the transitions among epistemic models that result from informational events.[21] DEL thus provides a model of epistemic dynamics. All of the possible changes of epistemic model due to any possible informational event are mapped out by the semantic structures of DEL.

Next, I consider two well-known frameworks for analyzing subtypes of epistemic dynamics: rational belief revision and rational partial belief update.

1.2.1.3 AGM Theory (AGM)

One subtype of epistemic dynamics is rational belief revision. An important core model of this kind of epistemic dynamics is given AGM theory [5]. Like DEL, AGM is an input-assimilating model. The relevant epistemic states are belief sets, K, that are sets of sentences of a propositional language, L, closed under entailment. That is, $K \subseteq L$ such that $K = Cn(K)$ where $Cn(K)$ is the set of logical consequences of K. AGM identifies a typology of three kinds of epistemic change: expansion K^+_A, contraction K^-_A, and revision K^*_A, and defines *epistemic commitment functions* of the form $B_L \times L \longrightarrow B_L$, such that B_L is the subset of the powerset of L that contains only belief sets. Each commitment function is defined so that it satisfies philosophically defended *rationality postulates* based in epistemic entrenchment.[22]

---

[21] This is an elementary point, but I refer the reader to any of the useful diagrams in the relevant chapters of any of [65, 34, 131] for help in visualizing this.

[22] See the first half of [91] for an exploration of AGM, the commitment functions, and the rationality postulates. In particular, Gardenfors explores entrenchment conditions and proves a couple of representation theorems between the rationality postulates and a set of entrenchment postulates.

The basic thought is that each belief set describes an agent's doxastic state in equilibrium under the forces of logical consistency and closure. These equilibrium states transition to new equilibrium states according to the epistemic commitment functions when disturbed by the external force of epistemic inputs from (e.g.) perception or reasoning. That is, starting at K, observing that A when $A \notin K$ forces a disturbance of equilibrium; one must add the belief that A to restore equilibrium. But this isn't enough, as now we must also be committed, on AGM, to adding the logical consequences of $K \cup \{A\}$. So K_A^+ is a new belief set equal to $Cn(K \cup \{A\})$. If we end up finding out that our beliefs are contradicted by (e.g.) empirical reality, then we must revise our beliefs and restore equilibrium (consistency and closure). The AGM commitment functions define a mechanism for transitioning from any belief set by means of an epistemic input to a new belief set that is restored to equilibrium that satisfies conditions of rationality. Epistemic states are belief sets and the actual and the possible rational changes of epistemic state are those transitions that are encoded by the epistemic commitment functions.

1.2.1.4 Bayesian Rationality Theories (BRT)

The rational update of partial beliefs according to total evidence is another subtype of epistemic dynamics. Subjective Bayesianism, providing a simple core model of this kind of dynamics, represents epistemic states with probability functions over a set of sentences.[23] Let L be a language and let $Pr : L \longrightarrow [0, 1]$ be a probability assignment to each sentence of L. The probability $Pr(A)$ represents a reasoner's degree of belief in A or the degree of confidence that she has in A. These states of partial belief are assumed to obey systems of axioms, one standard example of which is:

(A1) For all $A \in L$, $0 \leq Pr(A) \leq 1$.

(A2) If $\vdash A$, then $Pr(A) = 1$, and if $\vdash \neg A$, then $Pr(A) = 0$.

---

[23] I'm not targeting any particular system, here, but rather to the basic thrust of Bayesian epistemology. For greater detail, see, inter alia, [182, 53, 153, 155].

(A3) For all $A, B \in L$, if A and B are mutually exclusive, $Pr(A \vee B) = Pr(A) + Pr(B)$.

Dynamics involves transitioning from one probability function on L to another probability function. Rules of conditionalization, of varying levels of simplicity, give *update* functions from rational partial belief states to rational partial beliefs states:

(Simple Conditionalization) Where probability functions Pr_i and Pr_j represent a root and terminal state of an epistemic transition, respectively, rational changes follow the rule: If $Pr_i(E) > 0$ and $Pr_j(E) = 1$, then for all A, $Pr_j(A) = Pr_i(A \mid E)$.

The basic thought is that each probability function or partial belief state describes an agent's doxastic state in equilibrium under the forces of probabilistic coherence given by the axioms A1-A3. These equilibrium states transition to new equilibrium states according to the epistemic commitment function supplied by the conditionalization rule when disturbed by the external force of epistemic inputs from (e.g.) perception or reasoning (which raise the partial belief in E to certainty). That is, starting at Pr_i, observing that E (a conjunction of evidence statements) when $Pr_i(E) > 0$ forces a disturbance of equilibrium; one adds the certain belief that E. But this isn't enough, as now we must also be committed, on BRT, to restoring probabilistic coherence. Conditionalization defines a mechanism for transitioning from any partial belief state by means of acquired evidence to a new partial belief set that is restored to equilibrium.

1.2.2 Protocols and Some Frameworks for Protocols

The frameworks for epistemic dynamics above are all important and informative, but what they leave out or leave entirely implicit is essential to analyzing epistemic dynamics. Epistemic states change by the execution of some kinds of actions, procedures, or processes, and these executions are subject to various constraints. The analysis of epistemic dynamics must account for these constraints, or else be misleading or incomplete. This point will be familiar from the development of dynamic epistemic logics, and will be expanded upon in 1.3 [65, 66].

To model constraints, protocols are introduced. The basic, informal idea of a protocol is a rule that determines what actions can be performed at what states of the world. There are, of course, a number of formal tools that can be used to represent and think about protocols. Protocols are deployed to analyze constraints on epistemic dynamics in formal frameworks based on dynamic logics or, generally, in branching time models. There are several formal frameworks for analyzing dynamics governed by protocols.[24] Some are frameworks for general dynamics into which protocols are integrated. Others merge frameworks for epistemic dynamics with frameworks for protocols. In any case, many of the most well-known and well-developed take the form of branching time models. I briefly describe a few of the frameworks most directly related to the project in this dissertation, omitting most of the technical detail. I focus on briefly explicating the framework's conception of protocols and analyses of constraints of possibility and feasibility. The purpose of these digressions is to indicate, in a preliminary fashion, the location of PLEN in the space of formalisms, and to clarify the basic concept of protocols by showing some ways in which the concept has been formalized.[25] Remarks about the limitations of these formalisms are expanded in Chapter 5.

1.2.2.1 Epistemic Temporal Logic (ETL)

Epistemic Temporal Logic [74, 65] is, in some ways, the basic model of epistemic dynamics - or dynamics in general - under constraints. The structures of ETL are branching time models; ETL represents epistemic dynamics in terms of sequences of events. Let E be a set of events. A *history*, h, is a finite sequence of events in E. A history that leads up to the event e is denoted he. The notion of a prefix can be defined: h is a *prefix* of h' iff $h' = he$ for some e. The set of all histories constructed from E is $E*$. An ETL model is a triple, $\langle E, H, \rightsquigarrow_i, V\rangle$, where E is a set of events, $H \subseteq E*$, $\rightsquigarrow_i \subseteq H \times H$ is an indistinguishability

---

[24] See [65, 34, 30, 255, 231, 227, 175, 184, 74] for a smattering.

[25] The four systems chosen for discussion here are idiosyncratic. A fuller repertoire of systems is discussed in Chapter 5. The systems here are not meant to represent the best, most well-developed, or most influential systems in their respective fields or even for this application. They were chosen to represent theories that intuitively and simply convey the concept of protocols employed in this project.

relation among histories for each agent i, and V is a valuation function mapping expressions in the language of L_{ETL} to histories.

A set of histories, H, is a *protocol* in ETL. Thinking of events as possible actions, the histories of each ETL, H, represent a constraint on the set, $E*$, of possible histories (up to the null constraint where $H = E*$). According to H, histories can intersect and diverge. Any given event, e, can be followed in H only by a subset of possible sequences of events that follow it in $E*$. The protocol, H, shows us what the constraints are on the overall space of dynamics defined by $E*$.

ETL models provide simple and intuitive models of epistemic dynamics under constraint. Each node in an ETL is a temporal moment in the history of a reasoner's epistemic state development. At each node, there are multiple possible histories extending from it. That is, there are multiple possible paths of epistemic change that can be carried out from any given point in a history. But only a subset of possible sequences of events can begin at any given point. For example, as the Detective case and Muddy Children [65] show us, some events can only occur after some others and cannot occur after some others; there several possible histories that extend from t_1 in The Detective. In general, there are possible courses of action or sequences of events that cannot occur given that others have occurred. This kind of constraint isn't present in $E*$ itself, which is simply set of all possible finite sequences of events. It takes protocols to introduce constraints.

ETL doesn't have the machinery necessary to say much of interest about changes of epistemic states, and also lacks machinery for distinguishing procedures from transitions, and this is a serious limitation, as will be shown below. It does, however, shed light on capacity and feasibility, and overcomes the limitations of DEL in modeling the development of a particular state over time and the consequent enabling and disabling of processes. Additionally, the representation of states of a system with nodes of a graph is an abstract way of representing any sort of epistemic dynamics, and some additional structure is desirable for representing changes in the executability of actions from state to state.

1.2.2.2 Propositional Dynamic Logic (PDL)

PDL [113] is the central formal framework of the dynamic logic paradigm. Originally devised as a logic of programs, it provides an abstract model of general dynamics just as well. Essentially a modal logic with an action algebra on top, PDL has two simultaneously defined languages, L and P. L is a relatively standard propositional language plus a couple of operators defined with expressions from P: $[\pi]$ and $\langle\pi\rangle$. P is a program language. It begins with a set of atomic actions *Act*, that are composed by the operations of sequential composition ;, choice $\cup$, iteration $*$, and test $(A)?$ which is formed from expressions of L. The idea is that, if atomic actions are relations (linking states of a system), then the *program control operators* are the regular operations on these relations in relational algebra. The semantics of PDL are labeled transition systems or polymodal Kripke structures, $\langle S, \{R_a \mid a \in Act\}, v\rangle$ such that S is a non-empty set, each $R_a \subseteq S \times S$ is a binary transition relation indexed to members of *Act*, and v is a valuation function mapping subsets of L to each S. Each program is mapped to complex graphs formed from S with edges in R_a by rules for successful execution of the programs.

The original thought of PDL is that each structure displays a space of possible computations out of which a subgraph shows the possible computational processes that accord with each (possibly indeterministic) program. The programs are instructions formed from basic actions by various operations. The sequential composition of $a; b$ is the instruction to do a and then do b. Choice $a \cup b$ is the indeterministic choice operation; either a or b can be completed in order to execute this program construct. Iteration $a*$ is the instruction to iterate a some infinite number of times. And test is an atomic action that, if successful, always turns up a true A. From these, more complex instructions can be composed. For instance, the conditional program instruction of conditional choice "If A, do π_1; else, do π_2" is formalized in PDL as $((A)?\pi; \pi_1) \cup ((\neg \langle (A)?\pi\rangle \top)?; \pi_2)$. The instruction, guarded iteration "While A, do π" becomes $((A)?\pi_1; \pi)*; (\neg A)?\pi_1$. Each protocol acts as a constraint on the possible space of computations of the system. On this basis, a sound, complete, and decidable logic is defined. The well-known Segerberg axioms form the basis of the logic.

The models of PDL turn out to capture the exact thought that my earlier identification of dynamics with transition systems captures; any dynamic system can be usefully thought of in terms of states, transitions, and complex procedures for combining transitions. Programs capture what protocols are meant to capture; constraints on the possible ways that a dynamic system can unfold. The execution conditions tell us what programs can be executed at what states, and the program syntax tells us what the dynamics of the system must look like if the program is to be executed. Programs define what states their constituent actions can be executed at, how complex procedures are formed from basic actions (the *computation sequences* of the programs), and which processes (of computation, of dynamics) are instantiations of the program. Consequently, PDL has the important feature that the execution conditions of programs can encode the changes of possibility and feasibility by state or by time. Think of the executions of a program as those sequences of actions of state change that unfold under a constraint. From any state, programs tell us (via the execution conditions in PDL) what actions can be executed at that state according to the program, hence, under the constraint. These can be directly compared to the transition relations of each PDL structure.

Nevertheless, PDL fails to provide adequate models of epistemic dynamics. The representation of epistemic states in PDL, as states with valuations, is limited in some fairly important ways. It doesn't, for instance, allow the interpretation of epistemic or doxastic operators in any principled or interesting way. The basic language of PDL also lacks some interesting and intuitive control operators of PDL that are useful to analyzing epistemic dynamics. It may, for instance, be highly important to the modeling of epistemic dynamics that something like intersection of programs be added. More sophisticated variants of PDL are desirable.

Worse, PDL cannot recognize the distinction between procedures and transitions because it identifies actions with binary relations. However, by replacing binary relations with ternary relations, $S \times Act \times S$, the distinction between labeled and unlabeled transitions can be used to distinguish procedures from transitions. The failure to do this raises some problems in the analysis of reasoning about protocols, however, as we will see in Chapters 3, 4, and 5. The importance of this limitation will be argued below. Doing so, however, introduces some complications worth considering. I hold this for the technical [135].

1.2.2.3 Epistemic Protocol Logic (EPL)

EPL [184] is a nearby variant on PDL explicitly devoted to the analysis of protocols and reasoning about protocol information. EPL begins with *basic protocols*, which are finite trees or finite, pointed, directed, acyclic graphs without converging edges where each edge is labeled with an atomic action in the set *Act*. More precisely, each protocol is a *finite tree*, $T = \langle S, \{R_a \mid a \in Act\}, s_0\rangle$ where S and each R_a are as in PDL, but each protocol is a pointed model such that the transition relations are irreflexive, antisymmetric, and each state has a unique predecessor. The state s_0 is the *root* of the tree, and each sequence of edges starting at the root is finite. Thus, there is a *frontier* or a set of leaves, the set of end states of each path starting at s_0.

For each protocol, EPL defines a *protocol expression* $\langle x, a_i, t_{ai}\rangle + ... + \langle x, a_n, t_{an}\rangle$ where x is a node variable (a variable naming some node in some basic protocol), a_i is an action (labeled edge) taking x to the root of t_{ai}, which is a basic protocol. The atomic case is simply a node variable. The next simplest case is a single action tree, $\langle s_0, a_i, s_1\rangle$ where the frontier is just s_1. More complex trees can be formed by connecting multiple states by action labeled edges to single states. Protocol expressions can name complex trees whose root state is x formed out of the subtrees $t_{ai}, ..., t_{an}$ and linked to x by actions $a_i, ..., a_n$. These expressions are then paired with a syntax similar to PDL's. Protocol expressions are combined with sequential composition ;, choice $\cup$, and iteration $*$. These are then interpreted by mapping them to complex protocols that are generated by combining the protocols of the constituent protocol expressions.

The protocols and protocol language are then paired with *finite arenas with imperfect* *information*, $\langle S, \rightarrow, \rightsquigarrow \rangle$, where S is a set of states, $\rightarrow$ is a set of sets of binary transitions indexed to actions in Act, and $\rightsquigarrow$ is an uncertainty relation linking states that are interpreted as being indistinguishable from an epistemic point of view. The idea is that these structures represent all of a system's possible trajectories and an epistemic agent's knowledge about what trajectories are possible from any state. A protocol, complex or basic, can be *embedded* in an arena, which is to say that it's mapped injectively to a subtree of the arena. Such a protocol is *enabled* at the state that is the root of this subtree. On this basis, a sound and weakly complete epistemic logic encoding what agents can know about protocols as well as facts about the enabling of protocols is defined and axiomatized.

EPL was explicitly designed to analyze dynamics that are constrained by such factors as feasibility and possibility. The arenas of EPL represent a system's dynamics. The protocols identify subtrees of arenas as the dynamics of the system constrained by possibility and feasibility. That is, the protocols identify what actions an agent can or cannot execute, and the arenas with imperfect information identify what they know about what they can accomplish given what actions they can or cannot execute.

EPL bears all of the strengths and some of the weaknesses of PDL in modeling constraints on epistemic dynamics. EPL does well to distinguish procedure from epistemic change as well as differences in possibility and feasibility over state, agent, and time. More, EPL usefully recaptures important parts of PDL because there are protocol expressions in EPL that correspond to every finitary program in PDL formed from sequential composition and choice. However, the use of structureless nodes to represent epistemic states is troubled by their simplicity, while the system is overall somewhat too indebted to PDL for it's own good, as Chapter 5 argues; even EPL fails to do everything that is required of a fully adequate model of epistemic dynamics.

1.2.2.4 Epistemic Logic for Rule-based Agents (ELRA)

Some formalisms focus on the ways in which epistemic states change. Jago's Epistemic Logic for Rule-based Agents (ELRA) [134] focuses on the execution of epistemic transitions in accordance with formal rules. We begin with a denumerable language L. Let it be standard in all other ways for now. Let A_i and B be metavariables ranging over atomic propositions and their negations. Let *rules* be structures of the form $A_1, ..., A_n \rhd B$. Let $A_1, ..., A_n$ be the condition of the rule. Consider transition graphs, $\langle S, R, Lab\rangle$, formed from a non-empty set, S, a serial binary relation, R, relations among states, $R \subseteq S \times S$, and a labeling function, $Lab : S \longrightarrow 2^L$. That is, the graphs display a space of transition relations among states each of which is associated with a set of sentences of L. For a rule, r, say that a state is $r - matching$ iff for all A_i in the condition of the rule are in $Lab(s)$. Then Jago directs our attention to the class of models such that:

(M1) For all $s \in S$ and all rules $r = A_1, ..., A_n \rhd B$, if s is $r - matching$, then there is an s' such that sRs' and $Lab(s') = Lab(s) \cup \{B\}$.

(M2) For any state s that matches no rules, there is a modally equivalent state s' such that sRs'.

(M3) For any s, s', sRs' only if either there is some r such that s is $r - matching$ and (M1) is satisfied or there are no rules that s matches and (M2) holds of it.

(M4) All states match the same rules.

These are the models of ELRA over which epistemic, doxastic, and modal language can be interpreted. ELRA is, as might be easily noticed, a branching time model of dynamics.

ELRA has numerous interesting features. First, it represents epistemic states purely sententially, thereby bypassing many of the unfortunate aspects of propositional or possible worlds based models of epistemic states.[26] Second, it builds in information about the natures

---

[26] See [148] and [134]. More on this in Chapters 3 and 5.

of various kinds of transitions; this information is about the preconditions and postconditions of the transitions that can be carried out in a model. This allows the interpretation of claims concerning what holds after the indeterministic “firing” of a rule. That is, we can formulate statements about the informational effects of the execution of various rules. Seeing a set of rules as a nondeterministic program (in which, at any state, one must execute some matching rule out of the set, in no particular order), ELRA enables the formalization and analysis of statements about (e.g.) the effects of carrying out the program from a given state.

The basic thought is that each state, coupled with a set of sentences, is an epistemic state. The sentences that label each state can be interpreted, for instance, as those believed by an agent. By firing various rules, an agent carries out changes of epistemic state, modifying the set of beliefs at the state at which the rule is fired in a manner according with the rule. ELRA explicitly attempts to model the dynamics of systems that unfold according to rule sets. These provide models of constraints on dynamics insofar as the models of ELRA can be compared to the developments of epistemic state that take place in a “wild” transition graph that does not satisfy M1-M4, and so doesn’t “follow” the rules that its states match. That is, rule sets define programs or protocols that determine constraints on the unfolding of epistemic states. So, ELRA enables the distinction between rule-based procedures for state change and the total set of transitions among states and it enables comparison among rule sets. It also nicely analyzes preconditions of the application of (deductive) rules, and so gives a very nice partial account of epistemic dynamics and, perhaps, certain kinds of norms.

ELRA isn’t, however, sufficient to model constraints on epistemic dynamics in general, as it simply isn’t general; it’s an analysis of reasoners that change epistemic state by the application of rules of a certain form. As such, it offers no analysis of procedures that do not execute rules of that form.

1.2.3 Procedural Epistemology

The third big idea introduced above is that the requirements of epistemic rationality are defined by procedural norms that rule in and rule out processes formed out of more basic processes in certain ways. As I'll argue in Chapters 2 and 4, these procedural norms are integral to epistemic rationality. The attention paid to these procedural norms might seem alien to some perspectives in epistemology. I frequently defend putative procedural requirements by appeal to means-ends considerations rather than what might be thought of as "fundamental" epistemological considerations. In the rest of Chapter 1, I appeal to plausibility arguments (intuitions) concerning the hypothesis that epistemic rationality distinguishes between epistemic processes even where they have identical start and end states. In Chapter 2, I argue that there are procedural norms of epistemic rationality - norms that require or prefer some epistemic processes over others because the required/preferred process is built out of more basic epistemic actions and processes by means of procedural operations of sequencing and choice in some specified way. Some of these arguments are based on the observation that procedural features (like the sequencing of subprocesses) of our epistemic processes impact how well we can attain epistemic goals.

To this general approach, it might be objected that the "epistemic" norms that I focus on at such length are merely practical rules for achieving properly epistemic ends or complying with properly epistemic norms; they are not fundamental nor properly epistemological. Consider, for instance, the monistic view that truth is the sole, fundamental epistemic norm; it's the only standard against which epistemic factors can be evaluated. On this view, beliefs can be indirectly evaluated with respect to evidential support and coherence in virtue of the relations between these features of belief systems and truth, and reasoning processes can be indirectly evaluated with respect to their reliability for producing true beliefs. On this view, there might be rules that pick out some processes from others in virtue of having telltale features associated with reliability. But these rules will not be properly epistemological; they'll be strictly practical rules the normative force of which reduces to conditional imperatives of the form: "If you want to have true beliefs, then you ought to embody such-and-such mental processes, since they are reliable for producing truths."[27]

My response to this objection is twofold. First, it's not really an objection. My project is compatible with the view above; I just think that the derivative and "not fundamental and properly epistemological" procedural or pragmatic norms are extremely important for any theorist that wants to understand epistemic dynamics and methodology, and I use the phrase "epistemic norms" to talk about them. The basic epistemological perspective in this thesis is that of procedural epistemology. [196] As Papineau [185] puts it, I think of epistemology "...as a practical enterprise, concerned to ensure that the processes we embody are reliable for truth, as opposed to a purely theoretical enterprise within the realm of belief." (p.xviii) In other words, I'm concerned with epistemic methodology - the kinds of rules, principles, and concepts that limited epistemic agents can used to achieve the aims of inquiry. To that end, I pay special attention to considerations about deliberate epistemic self-regulation, which requires norms governing procedural construction - or so I argue in Chapters 2 and 4. If these concerns are not properly epistemological, epistemology is poorer for ignoring them.

Second, as I argue in Chapter 4, if I'm right about the abstract form of epistemic norms concerning belief states as well as the logical form of epistemic procedures, there is a clear sense in which the "not properly epistemological norms" are indispensable to epistemology for reasons that go beyond the analysis of self-regulation. I show that there is a sense in which the important normative information that is carried by epistemic belief state norms can be read off of procedural norms and that the converse fails. If that's all correct, whether we call procedural norms "properly epistemic" or not is of little relevance, as a model of epistemic rationality that rigorously analyzes procedural norms will get a good analysis of the "fundamental and properly epistemic" norms for free.

---

[27] See [185] (p.xi) for this imperative. Papineau isn't a proponent of the view described. On Papineau's view, one fundamental epistemic norm of truth is one too many: true belief is desirable because it's a prerequisite for successful action; there are no fundamental epistemic norms [185, 186].

1.2.4 The Protocol-theoretic Account

With the three underlying ideas in hand, the protocol-theoretic account of epistemic rationality norms can be stated in a bit more detail.[28] Epistemic rationality norms are, in general, best analyzed by explicitly representing procedural epistemic norms with protocols while representing epistemic dynamics with (upgraded versions of) the structures over which dynamic logics are interpreted.[29] In dynamic logics, protocols are objects that syntactically encode complex instructions formed out of procedural operations on basic epistemic actions and processes. This alone makes them natural representatives of procedural norms, as, per the third idea, procedural norms function as instructions that dictate, permit, or forbid such procedural operations. But there's more!

The semantics of protocols associates the syntax of protocols to operations on dynamic objects like transitions among states, extended processes of state change, and complex procedures built out of basic actions. These features of the structures of dynamic logics perspicuously represent corresponding features of epistemic dynamics - operations on transitions of epistemic state, extended epistemic processes, and complex epistemic procedures. These are exactly the features of epistemic dynamics that are associated with the requirements specified by procedural norms.

On the basis of the foregoing structural similarities, important features of epistemic rationality norms can be formalized and reasoned about in protocol-theoretic logics of epistemic norms. In particular, a formal framework that builds in the protocol-theoretic account of epistemic norms can support logics that codify valid reasoning about (i) the preconditions and postconditions of complying with them, (ii) equivalence relations among them, and (iii) their normative verdicts with respect to states, transitions, and processes.

---

[28] Given that procedural epistemic norms are the key to unifying epistemic norms generally, the thesis scopes over epistemic norms in general.

[29] The upgraded models are labeled directed multi-graphs [93], subsets of which are isomorphic to labeled transition systems, or multi-modal Kripke structures (plus or minus sundry modifications) [33].

1.3 Problems and Solutions

The closing observations of some of the prior sections reveal serious limitations on all of the foregoing frameworks for epistemic dynamics and numerous others besides. Models of epistemic dynamics must account for constraints of feasibility and capacity. As we will see, there is a more serious problem: if a formal framework conflates actions, procedures, or processes of epistemic change for the changes of epistemic state that might result from them, the model cannot be complete and accurate. This point can be seen by considering an aspect of the Detective case.

1.3.1 Representing Constraints on Dynamics

Epistemic states change by the execution of some kinds of actions, procedures, or processes, and these executions are subject to various constraints. The analysis of epistemic dynamics must account for these constraints, or else be misleading or incomplete.

Constraints on epistemic dynamics specify what an agent *can* do given any particular state.[30] There are multiple, distinct notions of what agents can do: call two of them *capacity* and *feasibility*. Some actions or procedures are not executable by an agent under any condition.[31] These actions are not within the agent's *capacity*. Say alternatively that executing those actions is *not possible for the agent*. If there are conditions under which an agent can carry out an action, it is possible for her or within her capacity. Other actions are not executable by an agent under some particular conditions, even if they are executable under others. When an action can't be executed under a specified condition, it's not *feasible* for the agent with respect to that condition. It's feasible when it can be executed.

---

[30] In Hoshi's [131] terms, constraints on epistemic dynamics are encompassed by "protocol information". While this usage gels with the conventions of the DEL and dynamic logics terminology, it wouldn't work quite so well with some of my terminological choices. I will prefer "restrictions" when talking about protocol information associated with norms, reserving "constraints" to talk about capacity and feasibility.

[31] This can be multiply specified, of course. Among others, we might specify that some actions are not executable by an agent under any logically possible condition, any nomologically possible condition, under any psychologically possible conditions, etc..

There are four important things to note at this stage about constraints of capacity and feasibility:

- First, capacity and feasibility are distinct. An action a may be possible for an agent or within her capacities, but not feasible under a specified set of conditions.
- Second, capacity and feasibility may differ for distinct agents. Some agents can do under condition C what others cannot under C or at all.
- Third, the conditions that contribute to the feasibility of an action can be external facts in the world or they can be facts about what epistemic states the agents are in. That is, external and internal conditions can determine whether an action is feasible for an agent at a time.
- Finally, there are changes in what actions can be carried out by what agents over time. For example, given that epistemic states are a factor in determining what actions are feasible, the basic idea of epistemic dynamics implies that the set of feasible actions for an agent need not be constant from state to state. Learning and environmental changes can enable new actions or disable actions normally within one's repertoire.

Standard illustrations of these claims include, inter alia, the familiar Muddy Children puzzle:[32]

Muddy Children Several children are playing outside. After playing they come inside and their father says, "At least one of you has mud on your forehead." Each child can see the other children's forehead but not his/her own. Their father repeats the following question, "Do you know whether or not you have mud on your forehead?" The children are very intelligent and honest, so they answer father's question at the same time. Can the children come to know, over rounds of the father's question, whether they have mud on their foreheads? If they can, how many question rounds are needed? (p. 6)

---

[32] See [165] for the classical formulation. I explicitly quote Hoshi [131] for this formulation.

Each of the children is capable of figuring out who has muddy foreheads, but cannot feasibly do so without the father's declaration, and one of them can't figure it out without the information provided by the other child's answer. For a Muddy Children case where, of two children, A and B, A is the muddy one, the first round of questioning produces "Yes" and "No" answers from A and B, respectively. The second round produces uniform "Yes" answers. It is possible for each child to figure out whether she is muddy, but at the first round figuring it out is feasible for A and not for B. Thus, capacity and feasibility come apart by condition and by agent. A's ability to answer correctly was dependent on what A could see (that B wasn't muddy) and on the information gleaned from father's announcement. Not being able to observe the same facts, B couldn't perform the same deduction. What deductions were feasible depends on what information each child has *as well as the actions the child can perform to get information from the environment.* That is, even idealizing away differences in deductive ability, the children's epistemic states and repertoires of action impose constraints on feasibility. Finally, the announcements and inferences that the muddy children can perform at early stages are not those that can be performed at later stages. A could figure it out at round one while B couldn't, but at round 2, everyone is caught up. There are changes in what actions are feasible or possible for agents over time. Both learning and changes of environment enable or disable actions.

Reconsider the detective example. Had the detective not had access to S1's data, she would not have formed the key belief that was the basis for the deductive inference at t_3. This inference is possible for her but not feasible under the condition that she can't get S1's data.[33] The leak of insulting information was an action that *enabled* the deduction by transitioning the detective to a new epistemic state where she possessed the requisite premises. If an effective leak wasn't in the detective's repertoire, the deduction wouldn't have been available at t_3. The leak was, in essence, an information generating action that had significant epistemological effects, including enabling her final deduction. Also, she couldn't

---

[33] Assuming there is no other method of forming the appropriate belief for the deduction.

knowingly have generated whatever information was generated if she didn't believe the info she leaked to be insulting to K's leadership. So, epistemic states make a difference not only to feasibility, but to additional properties of action such as the possibility of deliberate or knowing execution. That the deduction becomes executable only after acquiring the belief that K's crimes haven't increased shows that the feasible actions are not constant from state to state.

Given that these features are ubiquitous or even merely typical, representation of them is clearly required of the analysis of epistemic dynamics. However, many frameworks for epistemic dynamics don't even attempt to represent the features of epistemic dynamics illustrated above. Of the frameworks described above, only DEL was developed with an eye to modeling the aspects of epistemic dynamics under consideration, and this it only does in a limited fashion.[34]

AGM and BRT give little thought to dynamics under constraints.[35] Neither AGM nor BRT contain any rigorous concept of epistemic action as distinct from the epistemic commitment functions or update functions that define dynamics in these systems. Thus, any possible permutation of the contents of epistemic state s that meets the formal conditions in these theories is a perfectly possible change from any possible belief state or probability function. There is no sense in which a permutation of the contents is ruled out by any considerations other than the rationality postulates or the laws of probability, and these are not, in any sense, constraints of feasibility or capacity. Thus, due to the conflation of processes of epistemic change with formally possible epistemic changes, no serious consideration of constraint on action is an option in these theories.[36]

---

[34] This point is more fully articulated and defended in Chapter 5. But a broad argument is drawn in the main text to this effect.

[35] This should be qualified. The models of dynamics in these systems *can* be interpreted as representing constraints in the sense that the revision and update functions identify a proper subset of the set of all possible transitions (ordered pairs) of epistemic states defined in the frameworks. This is what makes them interesting and useful as theories of regulative ideals or norms for knowledge or rational belief. However, this way of modeling epistemic norms, while an insight that motivates the project, is completely inadequate given the principles that I argue must be captured by any theory of epistemic rationality and its norms. This is a point to be expanded upon in Chapters 2 and 5.

[36] Of course, modified versions of AGM and BRT that incorporate dynamic epistemic languages and de-

At best, we can compare the class of transitions that satisfy the definitions of the commitment and update functions to the subsuming class of all possible pairs of epistemic states in each system.[37] This will define a broad notion of constraint insofar as the class of transitions that satisfies the relevant axioms, postulates, or rules are thought of as those satisfying the requirements of rationality, but no other kinds of constraint on available actions (even restricting attention to inferential actions) will be intelligible in these systems, and this notion of constraint will be a far cry from an adequate analysis of capacity and feasibility. For instance, there is no representation of child A's public answer to father's question, nor of the limitations on B's ability or inability to acquire beliefs from which to deduce the correct answer.

To be fair, the foregoing frameworks aim only to give a basic model of epistemic dynamics in order to get at the logical conditions on knowledge, rational full belief, or partial belief and changes thereon. All the same, they do give basic models of epistemic states and mechanisms for epistemic state change that are formally precise and not without great insight. Nonetheless, the systems generally fail to account for constraints on dynamics, and they fail especially with respect to dynamics in relation to epistemic rationality.[38]

DEL can go further on this note than either AGM or BRT; the Muddy Children puzzle is a standard example in DEL textbooks, after all. In DEL, actions may be considered a subset of events, and these are related to transitions of epistemic state by product updates. This has two benefits. First, since the preconditions of events can come apart, there is a modeling of constraints of capacity and feasibility. An agent, i, might be able to carry out

---

ductive closure under a DEL validity relation will provide models that perfectly well mirror the inferences made in (e.g.) the muddy children case. The point, however, is not that the languages of standard AGM and BRT systems are inadequate to mirroring some kinds of inference, but that they have very little, without serious modifications (e.g., [6, 20, 99, 68]) to say about the cognitive abilities, opportunities, and limitations of the agents in these kinds of cases. See the next paragraph as well as Chapter 5 for more detailed spelling out of this argument.

[37] That is, if we consider the full sets of transitions, $B_L \times B_L$ (where B_L is the set of belief sets defined over artificial language L), and $Pr(L) \times Pr(L)$ (where $Pr(L)$ is the set of probability functions defined over language L), including transitions from and to incoherent or irrational states. The revision functions and the conditionalization rule clearly define subsets of these transition relations. Thus, they define a restriction of a sort on this more abstract transition relation.

[38] Again, a promissory note: this will be argued for in much more detail in Chapter 5.

some event, e, at a given state, while some other agent, j, can't, separating the feasibilities of agents. Some agents might be able to carry out some actions at some states but not at others (those at which $pre(e)$ hold), separating feasibility and capacity. Second, in DEL, there is the possibility that two distinct event models, E_1 and E_2, result in identical transitions given an epistemic model M. This suggests that DEL does not commit the conflation that AGM and BRT commit of identifying actions or processes with the changes they cause. However, this is all too fast. DEL is still limited for analyses of epistemic dynamics and insufficient to provide a framework for epistemic rationality norms, as we will see in detail in Chapter 5.

1.3.1.1 Protocols

Protocols are useful tools for modeling constraints. Cognate concepts include plans, policies, and especially programs, but these carry connotations that are misleading when talking about constraints of capacity and feasibility.

The basic, informal idea of a protocol is a rule that determines what actions can be performed at what states of the world. Somewhat more formally, thinking in terms of dynamic systems, protocols are rules that tell us what the options are - at any given state of a dynamic system - for transitioning to a new state. To state it a little more formally, say that *epistemic processes* are chains or sequences of *epistemic actions*.[39] I assume the usual type-token distinction holds for processes and actions [257]. Say that *procedures* are process types; they are sets of processes defined by relations between process tokens like procedural operations or construction out of a shared set of actions. For instance, a procedure might be the set of processes such that actions of specified types are always sequenced or chosen freely among under certain conditions.

Then, with that terminology in hand, *protocols* define procedures; they define what states actions can be executed at (or, equivalently, what actions can be executed at each state), what procedural operations can be used to combine basic actions or processes into more

---

[39] Epistemic action is understood in a rather liberal fashion, as will be commented on in Chapters 2, 3, and 4.

complex actions or processes (which determine the relations among processes mentioned above), what processes instantiate or adhere to a procedure, and what sequences of states are traversed by executing or complying with a protocol.

1.3.2 The Processes and Constraints Problems

Recall the Detective Case from section 1.2.1. The process of deducing that K is in City B from the available evidence and the process of simply believing the psychic's testimony result in identical changes of belief. For any theory that conflates (token) changes in epistemic states with actions or processes of change, the differences between these two (token) processes are invisible. They are, after all, merely expansions of the detective's belief set with the belief that "K is in B".[40] This is a barrier to modeling the epistemic dynamics of an individual agent with a level of fidelity that is relevant to representing and reasoning about epistemic rationality. The point will be expanded upon in the next section and in Chapter 2, but for now, simply note that if believing psychic testimony and deduction from the best available evidence are epistemically distinct (one way of reasoning being more rational than the other, for instance), then, in order to see how an agent's token reasoning processes are rational or not, our model must be able to distinguish between them even if they result in the same changes of epistemic state. From another angle, all LTSs provide are distinct labels, and we need systematic recipes for encoding interesting things about the distinctness of those labels.

The prior point (in combination with the observations about constraints on epistemic

---

[40] Strictly speaking, according to AGM, this final belief should have been added to the detective's belief set when she added "If I leak embarrassing information about K to media in A only, then the rate of K-like crimes in A will increase", "I leaked embarrassing information about K to media in A only", and "There is no increase of K-like crimes in A". This is due to the fact that expansions are closed under deduction. Nonetheless, AGM faces a dilemma: either we identify processes of epistemic change with commitment functions (hence, there are only three kinds of process of epistemic change and no analysis of the process of deductive reasoning at all), or we identify processes of epistemic change at a finer grain. The first option renders the main devices of AGM implausible as a model of epistemic dynamics for precisely the reason that the two processes of change in the main text are clearly psychologically, phenomenologically, and epistemologically distinct. AGM can't tell us anything about how they're different. The inclusion of "K is in B" in a belief set must be either a trivial consequence of the expansion above or the trivial consequence of expanding with "K is in B". AGM doesn't tell us anything interesting about how the same belief could be added by the process of deduction as distinct from the process of belief by psychic testimony. The second option requires machinery that is not present in AGM. This point gets spelled out and expanded upon in a bit more detail in Chapter 5.

action) implies that the following properties of formal systems are problems that any model of epistemic dynamics must overcome or avoid:

(Processes) There are no structures for representing actions, procedures, or processes as distinct from transitions between (models of) states.

(Constraints) There are no structures for representing constraints on actions, procedures, or processes as opposed to constraints on transitions.

The Processes problem was pointed out in this section; a model of epistemic dynamics should distinguish actions, procedures, or processes from transitions among states for the reason that two (token) processes may, for a particular agent, time, or set of conditions, generate identical transitions and yet be epistemically distinct. This point is expanded upon below and defended at some length in Chapter 2. The Constraints problem notes the interaction of the prior point with the observation of the constraints pointed out in subsection 1.3.1. A model without the means of drawing the line between processes and transitions can't draw the line between constraints on processes and constraints on transitions. A model of epistemic dynamics must do both if it's to give a good analysis of epistemic dynamics. This is also expanded upon below.

As we will see, there are further demands of a theory of epistemic dynamics and/or epistemic rationality that, in concert with these problems, suggest the following related problems:

(Constraint Unfolding) There are no structures representing how epistemic dynamics unfolds under constraints as contrasted to how it would unfold without the constraints.

(Process Reasoning) There are no languages for encoding reasoning about the distinctions in (Processes).

(Constraint Reasoning) There are no languages for encoding reasoning about the constraints in (Constraints).

(Constraint Unfolding Reasoning) There are no languages for encoding reasoning about how epistemic dynamics unfolds under constraints as contrasted to how it would unfold without the constraints.

The Constraint Unfolding problem is a corollary of the Constraint problem. If a model cannot represent constraints at all, it certainly won't be able to distinguish how a system develops under constraints from how it might have developed in the absence of those constraints. Conversely, if the latter can be represented, then the former can be represented indirectly in virtue of the representation of the latter.[41]

The Reasoning problems are corollaries of the other problems: models that don't provide structures for representing these distinct objects can't provide explicit methods for reasoning about those distinctions in their languages. At least, they can't provide explicit formal expressions that represent statements that would draw the distinctions while also giving those expressions semantic values in the models. As I argue below, PDL (for instance) doesn't seem able to interpret statements that draw the distinction between accepting psychic testimony and deducing the same information from one's own premises. Suppose that we give these token acts distinct labels in PDL. Then they must be embedded in the program expressions or the propositional expressions in which program expressions are embedded. But now, which expressions in PDL can plausibly represent the claim that accepting the testimony is irrational and deducing the same information is rational? Nothing in the language of PDL suffices. Nothing in the dynamic deontic expansions of PDL [174] works, either, since they function by identifying subsets of states (i.e., "red states" or states that verify "violation constants" will clearly be unable to draw the distinction). The general point here is expanded upon below and in detail in Chapters 2, 4, and 5.

These problems result in a number of other serious theoretical failings; if a formal framework has the Processes, Constraints, Unfolding, or Reasoning properties, then it fails to be a

---

[41] Suppose that a model of dynamics can tell us how a system would look if it executed action a at state s and also that a cannot be executed at state s. Then whatever constraint determines that a can't be done at s is at least partially characterized by this fact.

complete and accurate model of epistemic dynamics. Generally, the analysis of actions and processes exclusively in terms of transitions among epistemic states results in the failure of a model of epistemic dynamics to accurately characterize important features of dynamics, including some that are relevant to the analysis of epistemic rationality norms. To illustrate this, consider PDL as a test case.[42] If the binary transition relations mapped to a and b in all PDL structures are identical, then a and b are equivalent in all contexts in the language of PDL; they are perfectly intersubstitutable. Formally, if the binary transition relation, R_a, mapped to a is identical to that mapped to b (R_b) for every structure of PDL, then $\models_{PDL} A(a) \equiv A(b)$ where $A(x)$ is any formula of PDL containing x. But this is wrong. There are classes of assertions that may be true of action a that are false of b even if a and b are mapped to equivalent transition relations.[43]

The main thought is that it's possible for epistemic actions to be associated with identical binary transition relations despite being distinct in many other ways. Consider the following case.

(Honest Twin) Suppose that I know methods for discerning the validity of arguments in first-order logic that are infallible if carried out correctly. Suppose that I have a twin that knows all of the same methods I know but is infallibly honest and diligent when asked about validity; so much so that he walks me through every proof construction

---

[42] The problems articulated here generalize to many forms of DEL, belief revision theories, and, generally, any theory that conflates epistemic actions and processes with epistemic transition relations or operations on epistemic states. In any system in which $A(\pi)$ and $A(\pi')$ are interpreted exclusively over transition relations or operations on states associated with π and π', the problems in the main text will result. This is shown - or at least sketched - in more detail in Chapter 5.

[43] Note that some researchers define PDL in terms of ternary transition relations; see Ch. 4 of [30]. This is not equivalent to the usual definition of PDL [113]. Define two program equivalence relations: Rp and Rr such that: (1) $\pi Rp \pi'$ is true at s iff all labeled transitions (sets of triples or ternary transition relations) or π arrows starting at s are π' arrows and vice versa, and (2) $\pi Rr \pi$' iff all binary transitions (sets of pairs, binary transition relations) starting at s are π transitions and vice versa. Then $\pi Rp \pi'$ can fail at states at which $\pi Rr \pi'$ holds. Consider a ternary or labeled transition system with only states s and s' and labeled transitions $\langle s \pi s' \rangle$ and $\langle s \pi' s' \rangle$. Note that $\langle s \pi s' \rangle$ and $\langle s \pi' s' \rangle$ are distinct triples; $\langle s \pi s' \rangle$ may be in the set of π transitions but not in the set of π' transitions despite the fact that $\langle s \pi' s' \rangle$ is in the set of π' transitions. Thus, in that case, $\pi Rp \pi'$ fails at s. However, if the set of binary transitions mapped to π contains $\langle s, s' \rangle$ and the set of binary transitions contains $\langle s, s' \rangle$, then $\pi Rr \pi'$ holds at s. So, shifting from binary transition relations to ternary transition relations isn't an insignificant move; though the language of PDL can't distinguish these structural nuances, the language of PLEN can. It's not clear how best to extend execution conditions defined in terms of ternary relations to complex protocols, as we will see in [135].

> if I but ask.[44] In any situation, for any argument, if I were to successfully carry out a method for discerning validity, the execution of the methods I know would tell me exactly what my twin would tell me, were I to ask, and vice versa. In cases where there are multiple proofs, my twin always carries out the one that I would have, had I carried out the process. Let a represent the process of testing arguments using my deductive and model-checking methods. Let b represent the process of asking my twin about the validity of first order arguments and coming to believe what he says. These two processes are certainly equivalent in a sense; given an argument, they return identical verdicts if carried out correctly; they even give the same proofs. Thus, they terminate with identical epistemic states from any initial state. Let R_a be the transition relations associated with a; the set of all pairs of initial states and output states of carrying out a. Let R_b be defined in a parallel fashion. Then $R_a = R_b$.

The Honest Twin case - and any like it - generates a number of kinds of error for any system representing epistemic dynamics that analyzes actions and processes in terms of transitions.

First, conflating actions and/or processes with transitions will get claims about methodological inclusion wrong. Let $a!$ and $b!$ be instructions for carrying out a and b. Let $M_i(a!)$ say that the instruction encoded by $a!$ is part of an explicit methodology that I've learned, M. Now, in a model of epistemic dynamics in which actions and processes are analyzed in terms of transitions only, the statements about actions and processes will be interpreted with the same truth rule schema as any other formula in PDL: $s \Vdash M_i(a!)$ iff (... R_a ...) where the righthand side of the definition specifies some condition that R_a satisfies.[45] Thus,

---

[44] I, of course, I am the dishonest twin. An alternative case: Suppose that I have a power of clairvoyance that enables me to directly query an omniscient database. For any A, I can add it truly to my belief set - if it's true - by consulting the database, and I can eject it from my belief set by consulting the database and finding out it's false. Still, I go about revising my beliefs in a piecemeal fashion. Suppose also that I've got a hotline to God, and he always answers for me. I can revise my beliefs the same way, querying God one belief at a time. Querying God and querying the database will always turn out the same answers, under any conditions. The arguments in the main text concerning methodological inclusion and procedural features apply as well to this case.

[45] The alternative is to assign semantics to such formulae without reference to the parts of the models that interpret action or program expressions. But any such route will fail to solve the embedding problem raised in [135] with respect to epistemic normative terms.

whatever the rule says, $R_a = R_b$ implies that $s \Vdash M_i(a!)$ iff $s \Vdash M_i(b!)$ for any s in any model. Thus, for instance, $\models_{PDL} M_i(a!) \equiv M_i(b!)$. Of course, this equivalence fails in the Honest Twin Case; $b!$ may not be in my explicit methodology even if $a!$ is. Additionally, let $\langle M_i \rangle \top$ be a statement to the effect that some procedure in my methodology M has been executed. This will be true if $a! \in M$ and a has been executed. But if a and b are equivalent just if they lead to the same transitions, then executing b suffices for executing a. So, even if b is not in M (let $\{a\} = M$), $\langle M_i \rangle \top$ will be true whenever b is executed. But the deductive methodology I've learned may not include (and could perhaps explicitly prohibit or instruct the avoidance of executing) the instruction to ask my twin, and so asking will not carry out anything in my methodology (and may actually constitute a breach of the methods). As shown, conflating processes with transitions will get this wrong. Thus, the analysis of actions with transitions also gets facts about the feasibility of following our methodologies wrong as well as the foregoing facts about methodological inclusion.

Second, procedural information about a and b is incorrectly equated by the identity of binary transition relations. If any of the propositions of PDL encode procedural information, e.g., information about the number of steps a procedure requires, how those steps are combined via procedural operations, what facts hold over the course of carrying a procedure out, or information about particular substeps of carrying out a procedure, then the equivalence of binary relations implies false equivalences among these assertions.[46] The action a may require numerous substeps to complete while b is relatively basic. If $Com(a)$ says, in PDL, that a is complex, then the above fact that $\models_{PDL} A(a) \equiv A(b)$ has a false instance: $Com(a) \equiv Com(b)$. However, $Com(b)$ may be false while $Com(a)$ is true for all that identity of transitions establishes. Asking my twin does not, of course, need to be as complex a process as figuring out a proof myself. More, there are senses in which a might require execution of an action c while b does not. If a is actually a complex action, $c; d$, while b is atomic, then

---

[46] That this is not visible in the usual form of PDL is an expressive limitation; its languages don't have any formulae that can distinguish transitions from procedures. The arguments here essentially examine what happens when you add new expressions that do so.

it's true that a requires execution of c but b does not.[47] More, the process of asking my twin certainly contains sub-actions that doing the proof doesn't, namely, asking my twin. If there are expressions in PDL for encoding this, like $Com(a = c; d)$ and $Com(b = e; f)$, the semantics of PDL will simply get them wrong because, as above, it must interpret them in terms of R_a and R_b. Such an interpretation is already intuitively absurd; of course information about transitions doesn't reveal information about procedures. More directly, though, asking my counterpart and constructing a proof do not involve the same procedural operations on the same sets of basic actions; the equivalence is simply false. Thus, the conflation of processes with transitions fails to accurately capture procedural information.

Third, and most important for the purposes of this project, the analysis of actions and processes with transitions only will mischaracterize the verdicts of epistemic norms. The point can be seen most readily by considering a case where a token action a and a token action b would, if either is carried out, result in the same transition $t = \langle s, s' \rangle$. Let $N_{i,C}(a)$ say that norm N specifies that this particular token of a is rational for i to carry out under condition C. These statements must, on the "transitions only" view that I've been criticizing, be interpreted over the transitions associated with a and b. Thus, at s, $N_{i,C}(a) \equiv N_{i,C}(b)$. But this is is false. According to N, a may be rational and b may not be despite the fact that these actions result in the same transition. There are plenty of examples to be had. Consider inference by Modus Ponens from the set of statements $\{A, A \supset B, B \supset C, C\}$ as opposed to inference by Affirming the Consequent on the same set. It's clear that one is a rational inference, and the other not despite the fact that any statement about rationality must be interpreted only over the transitions involved, which, for this initial state, will be the same for both forms of inference.

If this argument just applied to token actions that result in the same token transition, it would suffice to establish that transitions-only models mischaracterize normative verdicts. However, as the Honest Twin Case shows, it's possible that two actions (action types) are

---

[47] I'm abusing notation a bit. Here, a and b are being used as variables for the perhaps complex procedures deducing and asking my twin, respectively. They are not meant to be atomic actions.

associated with identical transition relations, not just a shared transition. Explicitly, let $N_{i,C}(a)$ say that norm N specifies that a is a rational action (type) for i to carry out under condition C. Then, as above, $R_a = R_b$ implies that $\models_{PDL} N_{i,C}(a) \equiv N_{i,C}(b)$. This is a false equivalence. In situations like that described in the Honest Twin Case, $N_{i,C}(b)$ may be false while $N_{i,C}(a)$ is true for all that identity of transitions establishes. The transition-equivalent actions, a and b - as opposed to the merely transition-convergent actions in the modus ponens vs affirming the consequent case - are rationally distinct.

Under some circumstances, it may be rational to carry out a proof myself and not rational to ask my counterpart. For instance, if I am unaware of my counterpart's inexhaustible honesty or even have strong but defeasible - and ultimately misleading - reason to distrust him but I have extremely powerful reasons for trusting my own reasoning processes, it is more rational to deploy my own proof methods. The converse case works just as well. If I have good evidence my twin is better at proofs than I am and strong but defeasible - and ultimately misleading - evidence for distrusting my own reasoning, I have a powerful reason to ask my twin rather than rely on my own thinking. In the selection of epistemic actions, there is something epistemically irresponsible from the agent-relative perspective with selecting the action that is an unknown quantity or apparently the worse option [77, 78].

There are, additionally, differences between procedures in Honest Twin-like cases that are of epistemic relevance. Among these are features of procedures like (e.g.) efficiency, problem-solving power, speed, and likelihood of error. Consider likelihood of error. In a noisy environment, asking my twin may be more likely to be erroneously executed than deploying my own proof processes, however faster or more efficient asking might be. As we try to guard against this likelihood of error in a noisy environment, we will encounter a tradeoff with efficiency. More, different actions may have cognitive effects that are not part of the representation of epistemic states. Suppose that total procedural knowledge is not represented in a model's representation of epistemic states. Then, it's possible that a and b have different effects on my procedural knowledge; being walked through a proof by my honest twin may not provide the same sort of procedural knowledge that working through the proof on my own provides (or vice versa). This is epistemically relevant, as it affects what future epistemic actions are available to me, as well as the likelihood of performance errors, efficiency, problem solving power, and speed.

The foregoing arguments will find further articulation and defense in the next three subsubsections and in Chapter 2; explicitly, a number of examples of plausible epistemic norms that declare some processes rational and others irrational despite defining identical transition relations are given. Thus, there are numerous ways in which conflating transitions for processes gets epistemic dynamics and epistemic rationality wrong. More important, these are all features of epistemic dynamics that we want to think about in analyzing epistemic rationality and its norms for the transparent reason that these assertions all characterize various aspects of epistemic norms and the relations of factors of epistemic dynamics to norms. This should become clearer in future chapters.

Each of these failings is a result of the Processes problem. The central point is that, if a formal framework contains assertions about the foregoing properties of processes but interprets them in terms of sets of transition relations, it erroneously validates various classes of invalid inferences and verifies classes of false equivalences. Of course, the defense of a model of dynamics may simply rely on the fact that it is focused on some particular kind of epistemic process or that it analyzes dynamics at a level of abstraction at which processes drop out of the picture. In that case, the inferences and equivalences won't be within the scope of things the model aims to get right. That's fine; theoretical interests can vary from theorist to theorist, but it indisputably leaves the resulting model incomplete. Such a model, constricted in its interests, leaves elements of epistemic dynamics unanalyzed. Thus, any one of these problems is a black mark on the adequacy of a system as a general model of epistemic dynamics; a model with these problems is simply incomplete or misleading. However, if a model of epistemic dynamics is to be put to analyzing epistemic rationality, these problems may be decisive.

1.3.3 From Bad Models of Dynamics to Bad Models of Norms

I want to elaborate on the third problem in the foregoing section. This problem presupposed that there are epistemic norms that distinguish process tokens (and types) despite those tokens (resp. types) corresponding to the same transition tokens (resp. types) among epistemic states. I argue in Chapter 2 that this is indeed the case, but for now, let's simply take the following putative examples for granted:

(Bad Reasoning) IF: (i) a belief set at t is $s = \{B, A \rightarrow B, C \rightarrow A, C\}$ where $\rightarrow$ is a validly detachable conditional, and A is consistent with B and C, (ii) the belief set at $t+1$ is $s' = \{B, A \rightarrow B, C \rightarrow A, C, A\}$, and (iii) transitioning from s to s' can be carried out by two process types:

(Process 1) Detach A from the conditional $C \rightarrow A$ by affirming the antecedent, C, and

(Process 2) Detach A from the conditional $A \rightarrow B$ by affirming the consequent, B, and *not* performing any sort of explanatory or confirmatory inference to the same conclusion,

THEN: carrying out Process 2 in lieu of Process 1 is irrational, and carrying out Process 1 is rational.[48]

(Wishful Thinking) IF: (i) a belief set at t is $s = \{B, A \rightarrow B, C \rightarrow A, C\}$ where $\rightarrow$ is a validly detachable conditional, and A is consistent with B and C, (ii) the belief set at $t+1$ is $s' = \{B, A \rightarrow B, C \rightarrow A, C, A\}$, (iii) transitioning from s to s' can be carried out by two process types:

(Process 1) Detach A from the conditional $C \rightarrow A$ by affirming the antecedent C, and

---

[48] Perhaps more plausible, suppose that Process 1 involves explicitly, formally representing my belief set, plugging my beliefs into the valid argument schema of conditional detachment, and thereby recognizing the validity of such an argument, then updating my beliefs according to a rule like Defeasible Closure 1, thereby expanding my beliefs with A. Suppose that Process 2 follows suit, but with affirming the consequent.

(Process 2) Infer A without any regard to prior belief because it feels good to believe A,

THEN: carrying out Process 2 in lieu of Process 1 is irrational, and carrying out Process 1 is rational.

(Non Reasoning) IF: (i) a belief set at t is $s = \{B, A \rightarrow B, C \rightarrow A, C\}$ where $\rightarrow$ is a validly detachable conditional, and A is consistent with B and C, (ii) the belief set at $t+1$ is $s' = \{B, A \rightarrow B, C \rightarrow A, C, A\}$, and (iii) transitioning from s to s' can be carried out by two process types:

(Process 1) Detach A from the conditional $C \rightarrow A$ by affirming the antecedent C, and

(Process 2) Arrange medically unnecessary brain surgery that is known to cause belief that A (assume the surgery succeeds),

THEN: carrying out Process 2 in lieu of Process 1 is irrational, and carrying Process 1 is rational.

Note that (a) each of these norms conditions a verdict of irrationality on the availability of alternative processes with certain features, and, more important for the point at hand, (b) each example delivers the verdict that a process is irrational despite its delivering the same epistemic transition as a rational process. Assume these are genuine or plausible epistemic norms. This means that there are epistemic rationality requirements on processes of change of epistemic state as well as on states and transitions. Explicitly, the foregoing problems imply the following theses:

(D) Epistemic rationality is dynamic; there are epistemic requirements on transitions among epistemic states, not just states or contents of states.

(P) Epistemic rationality is procedural; there are epistemic requirements on epistemic processes and procedures, not just states or transitions among states.

If I'm right, a formal model of epistemic rationality norms must be adequate to (D) and (P). How can a formal framework for epistemic rationality norms make sense of these features of epistemic rationality?

To be adequate to (D), one's model of epistemic rationality norms must have an underlying model of epistemic dynamics in addition to an underlying model of statics. To be adequate to (P), the underlying model of epistemic dynamics must be able to carve transitions apart from processes. More precisely, if a model fails to distinguish transitions from processes, it will have no machinery for drawing the distinction between transitional and procedural requirements of epistemic rationality, and so no capacity to account for (P).

Generally, if a framework conflates the transitions associated with a and b with the actions a and b themselves, then it must interpret assertions in its language about the actions a and b on the basis of only transitions and features of transitions that omit procedural information (i.e., relations of the transitions to edges or actions other than being in R_x). But if that's the case, then the identity of the transitions associated with a and b will wrongly validate or verify schemata like $A(a) \equiv A(b)$. If the language of the framework contains expressions, $A(a)$ and $A(b)$, that can be interpreted as statements about the rational permissibility of a and b (with respect to norm N), then the framework will validate falsehoods.[49] Explicitly, in Wishful Thinking, Non-Reasoning, and Bad Reasoning, there is only the transition from s to s' associated with each process. Thus, if $A(a)$ and $A(b)$ are interpreted in terms of transitions, $A(a) \equiv A(b)$ will hold at s. However, if $A(a)$ says, intuitively, that a is rational (according to, e.g., Wishful Thinking, Non-Reasoning, or Bad Reasoning), then $A(a) \equiv A(b)$ is false; $A(b)$ may be false while $A(a)$ is true for all that identity of transitions establishes. Thus, conflating processes and actions with transitions will either get these kinds of procedural norms wrong by misconstruing their verdicts or it will have to just not contain expressions encoding verdicts like "a is rational (according to procedural norm N)". Further, in Chapters 2-5, it is

---

[49] This argument is phrased in general terms, but it applies more acutely to cases in which the a and b transitions unfolding from a root state are identical. There, $A(a) \equiv A(b)$ will be true in the framework while the equivalence it represents will be false.

shown that this argument extends to central features of epistemic rationality and its norms including logical relations among epistemic norms, e.g., the failure to distinguish transitions from processes results in failure to adequately analyze (e.g.) the equivalence conditions of procedural norms. So, it's generally worse for an account of epistemic rationality if its models fail to distinguish transitions and procedures.[50]

The strongest form of this argument depends on the class of "transition-equivalent" pairs of actions being non-empty. It might be objected that the basic premise of these arguments, that there are any actions a and b such that $R_a = R_b$, is false.[51] Specifically, it might be argued that, in any case like the Honest Twin, a and b can't be identified with identical transitions because at the states that a reaches, one believes that one reached the state by a, and at the states reachable by b, one believes that one reached it by b. This won't work. First, note that if a and b are analyzed strictly in terms of their transition relations then the beliefs in question are identical or logically equivalent with each other. In the first case, $R_a = R_b$. In the second case, if epistemic states are closed under equivalence, $R_a = R_b$. If epistemic states aren't so closed, then R_a and R_b may be nonidentical. But, second, in that case or the case in which a and b aren't strictly analyzed in terms of their transition relations, one can simply construct a modification of the Honest Twin case in which a' and b' either give mixed up beliefs about what process was carried out or they simply fail to give beliefs about what processes were carried out.

---

[50] Of course, this all depends on there really being plausible procedural norms. Chapter 2 attempts to establish this.

[51] In the main text, I only consider the objection from direct differences in the contents of the terminal states of a and b. I won't consider differences in, for instance, basing relations between the terminal states for three reasons. First, I consider basing relations in Ch.2, and what I say there, if it works, undermines the objection from basing relations here. Second, the Honest Twin case explicitly builds in the stipulation that a and b provide exactly the same proofs, hence, exactly the same non-causal basing relations. Third, to insist on a difference of causal basing relations will reinstate the very problem of normative verdicts above: differences of process (defined in terms of causal structure) are operative in normative judgments about beliefs. Such insistence is also either misguided or provides a direct reason to favor models that distinguish processes from transitions. If the causal basing relation is defined in terms of the prior states, in which case the causal basing relations just are identical between a and b because those processes apply in all of the same situations, then $R_a = R_b$. If causal basing relations are defined in terms of something more fine-grained, then R_a is not R_b but only because differences in the processes are operative in normative judgments about belief, which returns us to the normative verdicts problem.

There is no reason to think that all epistemic processes give such beliefs in addition to the beliefs they directly add. Many beliefs about the processes that generate belief either fade over time or are not scrutable in the first place.[52] To see this, simply consider a modification of the Honest Twin Case that extends the processes over long periods of time. Let a' be the process of carrying out the deduction or model-check and then going on with one's life as normal for ten years, and let b' be the counterpart composed of asking my twin and then going on with life for ten years. Assume that, as normal, I forget the details of how I came to the conclusion about validity that I came to by doing a', but I remember the outcome. Or assume that I form the correct belief about validity on the basis of misremembering which process I used, but correctly recalling the outcome. In that case, $R_{a'} = R_{b'}$ because in the long term, the two processes converge on exactly the same states. Schematically let $a' = a; c$ and let $b' = b; c$ where c is a process of forgetting what processes led to what beliefs, and the original argument goes through.

Now, the argument formulated above needn't be as generally formulated as it is - it's a matter of how serious a problem a model of epistemic dynamics faces. The Honest Twin case is a limit case; $R_a = R_b$ for a and b. However, the argument above can be modified to apply to the more general class of transition-convergent cases - cases where two actions merely overlap in associated transition relations despite being distinct in numerous, epistemologically relevant ways. Define $S = \{s \mid \langle s, s' \rangle \in R_a \cap R_b\}$ and then consider an equivalence, $A(a) \equiv_S A(b)$, such that $s' \Vdash A(a) \equiv_S A(b)$ iff $(s' \Vdash A(a)$ iff $s' \Vdash A(b))$ for all $s' \in S$. Then the S-variants of all of the equivalences above will be true at all states in S, and these equivalences will still be false for the reasons articulated above. Thus, even if there are no transition-equivalent pairs of actions, the analysis of actions with transitions will fail to get the foregoing features of epistemic dynamics and epistemic normativity correct.

---

[52] See, for instance, the general inscrutability of many heuristic and subpersonal processes [136, 238, 242, 105, 73], and Lam's considerations [154].

1.3.4 From Bad Models of Norms to Bad Models of Dynamics

The norms of epistemic rationality come in at least three types corresponding to the three distinct epistemic factors so far implicated. There are norms defining requirements on epistemic state, norms defining requirements on epistemic transitions, and norms defining requirements on epistemic action or process. Importantly, the norms described above define constraints distinct from those of capacity and feasibility. This is all argued for explicitly in Chapter 2. If this is right, then the evolution of an agent's epistemic states must unfold along a subset of the possible trajectories, and normative evolution is distinct from evolution by arbitrary selection of feasible moves. Thus, any model of epistemic dynamics that aims at a reasonable semblance of completeness and accuracy must analyze the constraints defined by epistemic norms in addition to and in contrast with the constraints of capacity and feasibility. Call the absence of resources for this the problem of First-Class Citizenship:[53]

(First-Class Citizenship) There are no structures that represent epistemic norms, no structures that represent compliance with epistemic norms, and no languages encoding reasoning about epistemic norms.

Two of the subclauses of the problem bear spelling out:

(Normative Unfolding) There are no structures representing how epistemic dynamics unfolds under normative restriction as contrasted to how it would unfold (under constraints of capacity and feasibility) without the normative restriction.

(Normative Reasoning) There are no languages for encoding reasoning about the properties of epistemic rationality norms, e.g., how epistemic dynamics unfolds under normative restriction as contrasted to how it would unfold without the normative restriction,

---

[53] As with the Processes and Constraints problems above, the reasoning problem is a simple corollary: if a framework is to give a logic of dynamics, it needs to provide a system for reasoning about these things. The terminology, here, is inspired by van Benthem [30, 34, 28], whose logical dynamics programme aims to, among other things, make actions first class citizens in logic. I think that formal analyses of epistemic rationality have to do the same for norms.

> what norms determine about specific courses of action, what norms can be derived from others, what sets of norms are equivalent to each other, etc..

The Normative Unfolding problem is simply a matter of spelling out the clause in the FCC about representing compliance with epistemic norms. Modeling how a system unfolds under normative restriction is just a way of representing compliance with norms. The Normative Reasoning problem is a simple corollary to FCC in that it's explicitly claimed in FCC. The absence of any languages for encoding reasoning about epistemic norms implies the absence of languages for encoding reasoning about the specific properties of epistemic norms detailed.

The First-Class Citizenship problem undermines models of epistemic dynamics. Normative courses of action - those in compliance with a set of norms - are typically distinct from those that an agent is capable of at any given state or time. Thus, if a formal framework for modeling epistemic dynamics fails to solve FCC, then it is misleading or incomplete with respect to epistemic dynamics. The thought is that normative constraints are genuine constraints that are distinct from constraints of capacity and feasibility. Normative courses of action - those in compliance with a set of norms - are typically distinct from those that an agent is capable of at any given state or time. Let a and b both be feasible processes or epistemic actions that a reasoner could carry out from a given state. The explanation of why a rational agent's actual history contains execution of procedure a rather than b may be that a complies with some normative canon and b complies with none. In order to adequately capture this feature of epistemic dynamics, a model needs some devices to represent the distinction between normative constraints and constraints of feasibility and capacity.

From another angle, the thought is that models of epistemic dynamics must distinguish between normative constraints and the constraints of feasibility and capacity because epistemic agents are capable of violating the normative constraints, and normative constraints typically identify proper subsets of feasible options. Thus, normative constraints are distinct from constraints of feasibility and capacity and require distinct representations.

The Processes and Constraints problems exacerbate the FCC. If a model fails to analyze epistemic norms by failing to draw the distinctions between static and dynamic and between transitional and procedural norms, it fails to adequately analyze epistemic dynamics because it fails to represent the requirements of (e.g.) procedural norms. Consider any of the norms in the last subsection. The explanation of why an epistemic agent's history contains an execution of procedure a rather than b may be that a complies with a deductive or inductive canon and b is bad reasoning, wishful thinking, or non-reasoning. A model needs to be able to represent the distinction between norms that constrain transitions and norms that constrain processes or procedures in order to represent the difference between a and b. The inability to do this is an inability to fully or accurately represent the epistemic dynamics of rational agents.

So, any general or complete models of dynamics require some features for representing and reasoning about epistemic norms and their requirements. In order to do this, epistemic norms ought to be "first-class citizens" in a framework in the sense that there is something in the structures that represents them - at least indirectly through representation of compliance with them - and parts of the languages of the logic for reasoning about them.

1.4 The Framework Problem and PLEN

1.4.1 Introducing The Desiderata

Recall that the protocol-theoretic account is supported by a theory preference argument: the formalism by which the protocol-theoretic account is developed (PLEN) satisfies theoretical desiderata that no other formalism does. The argument begins with the thought that formalization in philosophy is theoretically valuable; all else being equal, a partially formalized theory is to be preferred to a strictly informal theory. Formalization in philosophy forces theorists to explicitly articulate hidden assumptions and empowers them to rigorously draw difficult or obscure conclusions.[54] Both of these activities are useful for the evaluation and improvement of theories. Formalization is consequently a sort of crucible through which good philosophical theories would do well to pass, and so a desideratum on theoretical adequacy.[55]

Now, despite the virtues of formalization, formalization doesn't suffice to transform bad assumptions about a subject matter into a good theory. If an account of epistemic rationality is fundamentally incomplete or mischaracterizes its target, formalization may only carry these properties over despite its other virtues. This is the Garbage-in, Garbage-out principle.[56] Thus, anchoring the formalization to the most defensible philosophical characterization of the target domain available is paramount. Accordingly, the defense of PLEN and, consequently, the defense of the protocol-theoretic account depend on the philosophical characterization of epistemic rationality norms.

The foregoing problems partially characterize epistemic rationality norms by introducing important features of epistemic dynamics. I'll outline the rest of the core account of epistemic rationality norms below and defend it in Chapter 2. Call the constraints on epistemic dynamics defined by the requirements of epistemic rationality *restrictions*, for the sake of setting them apart from capacity and feasibility. The main idea of the account is that epistemic rationality restricts the space of possible epistemic factors that an ideally rational reasoner can navigate.[57] Requirements on epistemic states restrict the space of possible states for an epistemically rational cognizer to be in. Requirements on epistemic transitions restrict the space of possible transitions or trajectories of state change that an epistemically rational agent can evolve along. Requirements on epistemic processes restrict the space of possible epistemic actions and processes that a rational reasoner can deploy to evolve her states.

Formal frameworks that get epistemic dynamics wrong by ignoring the difference between

---

[54] See Hansson's useful discussion [112] of the procedure and virtues of formalization, as well as its vices. For additional insights about formalization, see [18].

[55] To be clear, it's a desideratum of non-zero weight. That, of course, doesn't imply that formalization is a minimal requirement of theoretical adequacy.

[56] See [195] wherein Pollock devises very much the same kind of motivation for his project.

[57] The notion of restriction will be more precisely characterized in Chapter 2.

transitions and procedures get epistemic norms wrong, and frameworks that get epistemic norms wrong get epistemic dynamics wrong. I have, hopefully, established that much. However, the examples in prior sections suggest specific conditions on modeling epistemic normativity. As argued above, given that the distinction between state, transition, and process requirements carves epistemic rationality at real joints, formal frameworks that ignore the difference between static and dynamic requirements - and that between transitional and procedural requirements - imposed by norms of rationality are either incomplete or else simply misleading about epistemic rationality *and* epistemic dynamics (insofar as rationality requirements define ideally rational epistemic state change). The following observations thus lie among the principles of epistemic rationality that any framework for analyzing epistemic rationality must capture:

D1 Epistemic rationality is dynamic; there are epistemic restrictions on transitions among epistemic states, not just states or contents of states, and there are characteristic relations among restrictions on states and restrictions on transitions.

D2 Epistemic rationality is procedural; there are epistemic restrictions on epistemic actions, processes, and procedures, not just states or transitions among states, and there are characteristic relations among restrictions on transitions and restrictions on processes.

D1 and D2 primarily rehearse D and P above but specify some further requirements on avoiding the First-Class Citizenship problem; there are relations among the different kinds of restrictions. More on this in Chapter 2. But D1 and D2 are merely some of the principles that the examination of epistemic rationality norms uncovers. The observations above concerning capacity and feasibility partially suggest that:

D3 Epistemic rationality norms define restrictions on epistemic dynamics that model (a) ideally rational evolution of epistemic state, (b) ideally rational changes of epistemic state, (c) ideally rational procedures and processes of epistemic state change, and that (d) can be distinguished from constraints of capacity and feasibility.

As D1 and D2 imply, there must be epistemic rationality norms that carry some information that defines rational dynamics, and this content distinguishes rational changes of state from rational processes of change. All of this work can be done by explicitly defined restrictions on epistemic dynamics, or so Chapters 2 and 4 will argue. These restrictions must be distinct from the constraints of capacity and feasibility for the simple fact that neither the capacity nor the feasibility of an action, process, or epistemic change imply its rationality.

It will be argued in chapter 2 that restrictions on epistemic procedure frequently, even ubiquitously, involve restriction on the construction of procedures. Given the core array of procedural operations, this implies the following corollary of D3:

Corollary Epistemic rationality norms restrict the construction of procedures; they determine how basic epistemic actions are to be sequenced, iterated, combined, chosen among, conditioned on tests, avoided, reversed, and cautiously executed.

D3 and its corollary put further flesh on the account of epistemic rationality norms that must anchor formal frameworks for epistemic rationality. I'll forego developing the individual parts of that account further, here.

In Chapter 2, each of D1-D3 is articulated in detail and defended as a condition on any putative logic of epistemic norms. As will be seen there, these principles arise in connection with central epistemological debates concerning the existence of dynamic epistemic norms and the function of epistemic norms in cognition. If the arguments of Chapter 2 are right, then D1-D3 are integral to the characterization of epistemic rationality and its norms. D1-D3 thus specify conditions of adequacy on addressing the First-Class Citizenship problem because they carve the nature of epistemic rationality and its norms at real joints. The structures and logical resources of a framework for representing and reasoning about epistemic rationality norms are thus incomplete or plainly incorrect if there aren't properties of the structures or formal assertions that adequately represent D1-D3.

1.4.2 The Framework Problem

If the foregoing arguments work, the analysis of epistemic rationality demands formalization and the formalization must address the Processes, Constraints, and First-Class Citizenship problems in a way that adequately captures D1-D3. Hence, the *Framework Problem*:

(Framework Problem) Any formal framework for reasoning about epistemic dynamics or the norms of epistemic rationality, F, that either fails to adequately represent epistemic dynamics or fails to adequately represent epistemic norms is either seriously incomplete or inaccurate with respect to its target domain. If a framework for epistemic dynamics fails to adequately formalize epistemic norms, it fails to be adequate with respect to epistemic dynamics. If a framework for norms fails to adequately represent epistemic dynamics, it fails to be adequate with respect to epistemic norms. Now, it is established in Chapter 1 that any framework that suffers the Processes, Constraints, or First-Class Citizenship problems (or related subproblems) fails to adequately formalize epistemic dynamics. It is established in Chapter 2 that any framework that fails to formalize desiderata/principles D1-D3 fails to adequately formalize epistemic norms. Thus, the fundamental problem is to devise a formal framework for representing and reasoning about epistemic dynamics, epistemic rationality, and epistemic rationality norms that (I) solves the Processes, Constraints, and First-Class Citizenship problems and related sub-problems, and (II) adequately formalizes principles D1-D3.

Now, the primary argument can be restated: we should understand epistemic rationality norms in terms of protocols because the *protocol-theoretic logic of epistemic norms*, PLEN, solves the Framework Problem.[58] Showing that PLEN does so and is the only formalism in

---

[58] The most useful method for defending a philosophical account of something is showing that it does better than its rivals relative to a set of desiderata (ex. D1-D3) articulated within a theory preference procedure (ex. Priest's multi-factor model [199, 202]). Now, this way of doing things sometimes allows a great degree of pluralism; the desiderata I want to satisfy may be anchored to goals like modeling a particular aspect of a complex object of study. A model aimed at understanding a different aspect of the same object need not answer to my desiderata. Rather than establishing a theory to be correct or objectively best in any general sense, this way of arguing carves out a niche in the ecology of theories and shows how a theory fills it. The argument of this thesis really is about niche-carving.

its neighborhood that does so is the primary project of this dissertation.

1.5 The Shape of Things to Come

This dissertation has three parts. Part I articulates the protocol-theoretic account, outlines the whole project, and provides a piecemeal defense of an informal, minimal theory of epistemic rationality that grounds the account. This minimal theory provides the philosophical criteria of adequacy against which formal frameworks for epistemic norms will be evaluated. In the course of spelling out the account and main lines of argument for it, Chapter 1 has articulated a number of problems that mar formal models of epistemic dynamics. These problems implicitly characterize epistemic rationality as a matter of dynamic processes constrained in various ways. Of the ways that epistemic processes are constrained, those related to norms of epistemic rationality are of obvious epistemological importance. Chapter 2 examines the nature of epistemic rationality and its norms more closely, giving a detailed defense of a minimal theory in the form of three principles: D1-D3. In Chapter 2, D1-D3 are articulated and defended on the basis of arguments arising from disputes about the static or dynamic nature of epistemic rationality and about the function of epistemic norms in epistemic life.

Part II is the technical core of the project; it defines PLEN and shows how PLEN solves the Framework Problem. Chapter 3 provides a detailed formulation of PLEN and discusses some of its important properties. The syntax and semantics of PLEN are discussed in some detail both formally and intuitively. The discussion focuses on getting a feel for the system and for interpreting the formalism as an abstract theory of the nature, structure, and content of epistemic rationality norms. To those ends, a sound axiomatization of PLEN is given. Chapter 4 shows how PLEN solves the Framework Problem, and so completes the first half of the theory preference argument for the protocol-theoretic account. Each of the Processes, Constraints, and First-Class Citizenship problems are solved in turn, and D1-D3 are formalized by relying on the technical properties of PLEN. The argument of Chapter 4 relies on a technical result that is more philosophically interesting if an analysis of deontic operators can be devised in PLEN. To that end, there is a brief appendix that gives a plausible - if preliminary - analysis of deontic operators in PLEN.

Part III completes the defense of PLEN and considers the road ahead. Chapter 5 locates PLEN relative to the formalisms that PLEN draws on, showing that rival formalisms fail to satisfy at least some of the foregoing desiderata. It is shown that PLEN occupies an interesting place; it solves the problems articulated in Chapter 1, it formalizes D1-D3 with great fidelity, and it's a relatively basic framework that is easy to work with and modify. Rival and predecessor formalisms, while fruitful for thinking about epistemic dynamics and about norms and protocols, tend to fall short of PLEN on one or the other of these dimensions. Chapter 6 recapitulates and suggests a number of ways in which PLEN can be complicated or modified to improve its usefulness or to take on other philosophical projects.

Chapter 2

Desiderata: The Shape of Epistemic Dynamics and the Structure of Epistemic Norms

2.1 Introduction

The Framework Problem articulated in Chapter 1 was partially the problem of developing a formal framework for theories of epistemic rationality that adequately formalizes the following set of principles:

D1 Epistemic rationality is dynamic; there are epistemic restrictions on transitions among epistemic states, not just states or contents of states, and there are characteristic relations among restrictions on states and restrictions on transitions.

D2 Epistemic rationality is procedural; there are epistemic restrictions on epistemic actions, processes, and procedures, not just states or transitions among states, and there are characteristic relations among restrictions on transitions and restrictions on processes.

D3 Epistemic rationality norms define restrictions on epistemic dynamics that model (a) ideally rational evolution of epistemic state, (b) ideally rational changes of epistemic state, (c) ideally rational procedures and processes of epistemic state change, and that (d) can be distinguished from constraints of capacity and feasibility.

Corollary Epistemic rationality norms restrict the construction of procedures; they determine how basic epistemic actions are to be sequenced, iterated, combined, chosen among, conditioned on tests, avoided, reversed, and cautiously executed.

These principles comprise a sort of core, pre-theoretic account of epistemic rationality. The account doesn't commit to any substantive norms or to any particular account of the structure of epistemic states (i.e., quantitative beliefs or full beliefs). Importantly, the account of epistemic rationality comprising D1-D3 is entirely independent of the protocol-theoretic account. It's intended merely to enumerate features of epistemic rationality that PLEN - or any other formal theory - must account for. In other words, D1-D3 articulate an intuitive theory of the structure or form of epistemic norms that formal theories are to be measured against.

Formalizing these features is a condition on the adequacy of formal frameworks for epistemic rationality. This was partially motivated in Chapter 1 regarding D1 and D2: if these principles carve epistemic rationality at real joints, then formal frameworks that fail to capture them in some way will be either incomplete or misleading about epistemic dynamics, epistemic rationality, or the nature and structure of epistemic norms. This chapter aims to fully substantiate that conditional with regard to each of D1, D2, and D3 and its corollary.

The expository plan of the chapter is to articulate and defend these principles in turn, thereby articulating the core account of epistemic rationality and its norms that formal frameworks must capture. Doing so takes us through examination of a number of important questions about epistemic rationality:

1. Is epistemic rationality strictly or fundamentally static or dynamic? What do norms of epistemic rationality govern? What do their conditions of application concern?

2. What kind of information is carried by epistemic norms? Propositions? Imperatives? Programs?
3. What is the primary role of epistemic norms in cognition? Factual judgments? Post hoc evaluations? Epistemic advice?

The sections of the chapter are organized straightforwardly around principles D1-D3, but they overlap with respect to the questions above. First, a preliminary section outlines my assumptions and terminological choices. From there, each section is composed of three subsections: a subsection in which the principle is explicated, a subsection in which the principle is argued for, and a subsection in which the principle is discussed. The articulation and argumentation subsections are relatively self-contained. The discussion subsections do the explicit work of explaining how the principles answer questions 1-3, and how the principles hang together. This results in there being a bit of a complicated weave of interconnections between these subsections.

2.1.1 Preliminaries

2.1.1.1 Epistemic Rationality and its Norms

The requirements of epistemic rationality are the conditions that one's epistemic factors must meet in order for them to be rational. The *norms* of epistemic rationality specify these requirements; they are the cognitive or conceptual items that function in our cognitive economies in order to allow us to think about and coordinate our actions with the requirements of epistemic rationality. Where the requirements of rationality are the things that we need to do or the conditions we need to meet in order to be rational, epistemic rationality norms are the rules, plans, or policies that spell out those requirements.

The fundamental target for analysis in this thesis is epistemic rationality norms. Epistemic rationality norms are broadly construed, here. The norms of rational belief revision that distinguish acceptable forms of belief formation, retention, or change fall under the rubric of epistemic rationality norms, as do norms that distinguish rational from irrational belief sets. Now, there are a great variety of proposed epistemic rationality norms across the fields of epistemology and philosophy of logic. The debate about these often centers on whether the norms are themselves rational - or, if one prefers, plausible - as norms for controlling belief change. See, for instance, the debates about (e.g.) deductive closure, conditionalization, or even logical consistency [2, 53, 121, 199]. I will not be directly weighing in on these debates. Rather, I'll be pulling examples of epistemic norms liberally from the literature as well as the aether in order to illustrate concepts or make arguments. While it will sometimes matter whether these examples are cases of rational epistemic rationality norms, it will be assumed that they are all norms of rationality as opposed to norms of some other kind.[1]

Epistemic rationality norms are often encoded via "norm-kernels":

(Norm Kernel) If C, then $R(x)$.

These rule-like expressions are a focal point of analysis [261, 164, 167] because they plausibly articulate the essential structure and properties of norms. It is this form of rule that enables norm-kernels to define the requirements of rationality.

The condition C is commonly called the *condition of application* of the norm; it gives us

---

[1] As for what distinguishes epistemic norms from non-epistemic norms, I wish to remain non-committal. At a first glance, it seems that epistemic norms, as opposed to non-epistemic norms, condition their valence (positive or negative, rational or irrational) on "epistemic" concepts including (inter alia) logical and probabilistic relations among the contents of epistemic states. Very roughly, say that, if the valence of an epistemic factor x according to norm N is conditioned on logical or evidential properties of x or its relations to epistemic procedures, then N is an epistemic norm. If the valence that N assigns is conditioned on practical or moral properties of epistemic factors, then, roughly speaking, it's a non-epistemic norm. This rough and ready rule will pick out the norms I focus on in this thesis. However, it may be possible to reduce such norms to other moral, practical, or aesthetic norms. I won't pursue this line of thought. Nothing hangs on the impossibility or even contingent failure of reduction of epistemic to non-epistemic norms because my targets are clear enough. Even if, as Papineau argues in [186], epistemic norms are really - at bottom - grounded in practical, moral, or aesthetic considerations, there is a subset of norms concerning cognitive states and activities that directly involve no practical, moral, or aesthetic concepts. The putative requirement that one's beliefs be probabilistically coherent, for instance, doesn't directly appeal to any means, ends, morality, or aesthetic values, though we might justify following a norm imposing such a requirement by appeal to the desire not to get Dutch Booked or to some other practical considerations. Additionally, see the arguments below distinguishing the project of analyzing strictly the non-reducible norms from the project of analyzing the norms that are important for epistemic life.

a condition under which the normative claim, $R(x)$, holds. The variable x stands in for a term denoting something in the domain of the norm, something to which the norm applies. In other words, x is the *object* of the norm. The objects of norms are what they govern. The consequent, $R(x)$, is a normative claim of some kind, often a deontic proposition. $R(x)$ states the *valence* or *verdict* of the epistemic appraisal. The object of the norm may be declared (by the norm), under condition C, to be rational or irrational, rationally obligatory or permissible, more or less rational relative to a standard or an alternative object, and so on.[2] In other words, epistemic rationality delivers verdicts about what states, transitions, and processes are rational (inter alia), and the norms of epistemic rationality are the rules that structure the verdicts by giving conditions under which their objects possess some valence or another.

There is room for norm-kernels of a converse form:

(Converse Norm Kernel) If $R(x)$, then C.

These norm-kernels specify necessary conditions for a valence rather than sufficient conditions. The components of the anatomy are the same, though. Conditions of application and objects are easy enough to identify, though the logical connection between them in converse norm-kernels is different. Epistemic norms are often not merely statements of sufficient conditions but statements of exclusive conditions. Probabilistic coherence is sometimes thought to be both necessary and sufficient for rational belief. These norms can be expressed in pairs of converse norms or pairs giving conditions for dual valences: “If state s contains credences that violate the axioms of probability, then s is irrational”, and “if s contains credences that satisfy the axioms of probability, then s is rational”. Assuming that valences obey usual duality rules like those that (e.g.) deontics and modals do, the standard norm-kernel form indirectly specifies requirements (i.e., necessary conditions).

Norms involving deontic valences are of special importance because the requirements of rationality frequently seem to require or permit actions or states. For example, “If the

---

[2] The relations among these valences are discussed briefly in Chapter 4.

conditional probability of H given E is above threshold t, then it is rationally permissible to believe H given that you believe E." For another, "If the inference (s, s') leads to an inconsistent belief state, then inference (s, s') is forbidden." And a third, "If you learn (only) B between t_1 and t_2, then it is rationally obligatory to update your credence in A at t_2, $Pr(A)_2$, so that $Pr(A)_2 = Pr(A \mid B)_1$. In general, norm-kernels like "if C, then $O(x)$" tell us that rationality requires us to carry out the action x or be in state x. Permissive norm-kernels, "If C, then $P(x)$" work in tandem to define requirements. Consider a set of norm-kernels, $N_1, ..., N_n$, all of which condition on C and permit $a_1, ..., a_n$. If this set exhausts the permission norms, executing some a_i is required (or obligated) by $N_1, ..., N_n$; one is required to act permissibly, and $a_1, ..., a_n$ exhausts the set of permissible actions. Prohibitions negatively define the set of permissible actions, and thereby shape requirements indirectly as well directly requiring that one avoid some actions.

Three remarks are in order. First, deontic and other valences are interdefinable. It's plausible that, if a norm N is a norm of epistemic rationality, then if N declares x impermissible under condition C, this is equivalent to declaring x irrational under C. It isn't possible to do x, according to N, while being rational. That is, if N declares x irrational under C, in order for an agent to be rational under C, she can't do x (should x be an action) or be in x (should x be a state). But this is just to say that x impermissible according to the norms of rationality. So, it seems to be possible to reduce the customary catalog of valences - rational, irrational, more/less rational, and "there is a reason to" - to the deontic valences - obligation, permission, and prohibition.[3]

Second, some valences of an object may preclude other valences of that object or they may preclude other valence-object pairs. The obligatoriness of x may preclude the permissibility of y where x and y are incompatible in certain ways, say, if x and y require use of the same scarce resources, are time-sensitive, or cause the enabling conditions of one another to

---

[3] This idea will be taken on at length in Chapter 4. One obvious hitch is that "more/less rational" and "there is a reason to" seem difficult to reduce to deontic valences. Prima facie, it requires consideration of degrees of obligation and permissibility or robustness of permissibility or consideration of how deontic operators rank actions and how this works in contexts of choice.

fail. The converse also holds. A norm obligating x under C functionally prohibits y under C, because the execution of y implies the failure of the duty to x. From another angle, if omissions can be counted as kinds of action, if x is an action and y is the omission of x, x's obligatoriness entails that y is impermissible.

Third, norm kernels may come in comparative form, as well:

Norms Kernel If C, then $R(x, y)$.

Normses Kernel If C, then $R(x)$ and $R'(y)$.

Norms of both forms can give us comparative information about the objects x and y. The condition of application, C, may state comparative information like "the epistemic state reached by x is more coherent than that reached by y", or "x is a more coherent state than y". The valence $R(x, y)$ tells us that there is a comparative valence between x and y. For example, "If belief that-A is more probable given evidence E than belief that-B, then, given belief in E, it is *more rational* to believe A than B." The conjunction $R(x)$ and $R'(y)$ is straightforward. Such a norm just tells us that if these conditions hold of x and y, then x and y have such and such differential normative properties. This is often the form of comparative norms. Consider the following norms as instances:

(Evidence Over Consistency 1) If inference (s, s') leads to a contradictory but properly evidence-apportioned belief state, and inference (s, s'') leads to a consistent but not properly evidence-apportioned belief state, then it's more rational to carry out inference (s, s') than to carry out (s, s'').

(Evidence Over Consistency 2) If inference (s, s') leads to a contradictory but properly evidence-apportioned belief state, and inference (s, s'') leads to a consistent but not properly evidence-apportioned belief state, then it's (rationally) permissible to carry out inference (s, s') and (rationally) impermissible to carry out (s, s'').

The plausibility of these norms depends on how proper evidence apportionment is spelled out, of course, but they illustrate different forms of comparative assignment of epistemic verdicts. In all of what follows, I make no use of alternative forms of norm-kernel, but it shouldn't be difficult to see how what is considered applies to these different forms of normative expression.

Epistemic rationality norms can be more or less informative. For instance, it's commonly thought that logical consistency or probabilistic coherence are requirements of rationality. A consistency norm may be a rule or plan dictating the detection and revision of logically or probabilistically inconsistent belief sets; a rule that tells us to identify inconsistency and repair it. Such a rule needn't be as informative as, for instance, an algorithm for detecting and revising away inconsistency. That is, a consistency norm might be a rule that not only tells us to identify and repair inconsistency but that tells us exactly how. A norm that carries no information about exactly how to detect those conditions or how to revise belief upon detection will be less precise and consequently less useful in guidance of epistemic action than one that carries that information. Compare the rules:

R1 If A and $\neg A$ are both in your belief set, revise by rejecting A or $\neg A$ and everything that entailed the belief you rejected.

R2 If your belief set is inconsistent, remove the inconsistency.

R2 is less precise in two ways. First, it carries no information about how to tell when one's belief set is inconsistent, in contrast to R1. R2 may, for all of the information it carries, require revision away from some idiosyncratic notion of inconsistency (e.g., inconsistency with the revealed truth of the scriptures) rather than logical inconsistency. Second, R2's recommendations are less precise. R1 dictates one of three possible revisions: revising A, $\neg A$, or both. For all that R2 tells us, there may be countlessly many ways to comply with R2, including emptying one's belief set entirely, deploying one of the above revisions alongside a number of arbitrary other revisions (reject A and also B, C, and all beliefs about lizards), or simply rejecting A, ignoring the possibility of reclosing one's belief set. More precise, informative norms are necessary for deliberate application of norms, or so I'll argue below.

I will say that rationality or its norms *distinguish* objects of epistemic appraisal x and y when x and y differ in the valences assigned to them. I will also say, equivalently, that x and y are *rationally distinct* or *rationally distinguishable.* I will say that rationality *distinguishes x and y on the basis of C* when x and y are rationally distinct and C is what makes the difference in valence. I will say that, thinking in terms of transitions among states in an LTS, epistemic changes are transitions from one state to another. In a transition from state s to state s', s is the *root state* and s' is the *terminal state.* If two different transitions $\langle s, s' \rangle$ and $\langle s, s'' \rangle$ are discussed, it means that both transitions can be *executed* from state s or both transitions are *available* at s. Take this to mean that some process resulting in these transitions is feasible for a reasoner with state s.

For convenience, say that a normative property is *dynamic* if the objects that bear the property are changes of states or processes of state change (or sets thereof). Say that a normative property is *static* if the objects that bear the property are epistemic states (or sets thereof). The *synchronic properties* of epistemic states are the properties that a state, s, has at any particular time, t, or that relate s to other things (e.g., the parts of s, the contents of s, other epistemic states, s') only at t.

Epistemic norms frequently specify requirements on *epistemic action.* Epistemic action is deployed as a capacious term. The set of epistemic actions is the set of actions that generate, transform, reveal, or obscure information in any ways relevant to epistemic cognition. This includes actions as diverse as those overt actions involved in the construction and manipulation of external representations (e.g., carrying out calculations or derivations on a blackboard, crafting a physical model and manipulating it), methodical evidence acquisition or generation (i.e., experimentation or physical manipulations of the environment that otherwise generate information), or in more internally oriented but still voluntary cognitive behavior like consciously attempting to think of alternative explanations, directing attention to particular bodies of evidence, or the actions involved in deliberate self-stimulation of inference or supposition [144, 181, 171, 77].[4]

As a result of this conception of epistemic actions, I'll be setting aside any worries about doxastic involuntarism and related issues. One might argue, as Alston [7] does, that epistemic norms cannot require or impose normative requirements on things like inference or belief formation because normative requirements can only apply to things that agents have control over, and inference and belief formation are not among these things. However, even if we accept the claim that inference and belief formation are not actions over which agents wield direct control, an indirect form of control over inference and belief change can be exerted by performing actions that trigger these involuntary cognitive processes [171]. For instance, the deliberate direction of attention to logical relations among believed propositions can trigger expansion or contraction of beliefs. More generally, the deliberate direction of attention to specific features of argument structure, evidence, environments, or our models can function as self-stimulating actions that trigger inference and belief change. The deliberate construction, modification, and manipulation of models can do the same - if only because deliberately building, fiddling with, and using models can direct attention to logical relations, evidence, etc..

By these means one can *see to it that* one infers or forms a belief, as long as there are regular connections between actions that one can directly control and those cognitive processes. That is, one can wield control over otherwise involuntary cognitive processes by carrying out epistemic actions in the broad sense above. So, in place of a full engagement with doxastic voluntarism and involuntarism, I'll just stipulate that epistemic actions subsume

---

[4] This usage is especially consilient with lines of empirical research on epistemic action, e.g., [144] and [169], that show that manipulation of models can improve performance on cognitive tasks. For instance, the manipulation of physical models of arrangements of blocks improves performance on later abstract block arrangement tasks, specifically as exhibited in playing the game Tetris. There is also much philosophical research on the construction and manipulation of models in philosophy of science [178, 87, 212]. This research tends to focus on what epistemic gains are made by manipulating models despite their admitted distance from the phenomena they model. However, that such gains are made coupled with the fact that model manipulation is a volitional activity implies the point being made: that the sorts of non-inferential or non-cognitive actions involved in manipulation of models are epistemically relevant and useful despite not being (e.g.) inferences, perhaps because they trigger inferences or other cognitive processes. For instance, turning over cards, looking under cups, reading letters, or any of the other standard actions in the DEL literature, re: Ch. 5 of [65].

both the kinds of actions described above as well as the kinds of action that reliably stimulate inference or belief. So, where I assert that norms require inference or formation of belief under some conditions, the doxastic involuntarist can re-phrase this as saying that norms really require that one see to it that one infers or forms beliefs under those conditions by carrying out the requisite self-stimulation actions.

Epistemic actions can be strung together to form epistemic processes. The epistemic actions out of which a process is formed are basic, with respect to the processes they comprise. Complex processes can be strung together to form more complex processes. These combinatorial assumptions - perhaps controversially - run together processes of epistemic cognition and processes formed out of sequenced epistemic actions, as defined above. Such epistemic processes might involve the interleaving of cognitive processes and actions on the environment. One might wish to draw some sharp line between processes that affect the outside world and those that don't. However, I doubt that this boundary is much of one for the reasons that, as argued above, (i) actions that affect the world generate, reveal, or conceal information for cognitive processes to use and (ii) they also stimulate or trigger such processes. A defense of this line of thought is in [171]. The boundary is thoroughly questioned independent of my worries: [144, 181].

Finally, a reminder of the key background conceptual assumptions. First, much of the argumentation in this chapter is dependent on thinking in terms of epistemic dynamics and LTSs. Implicitly, epistemic rationality is thought of as telling theorists and reasoners about which states, transitions, actions, processes, or procedures of change of state are (epistemically) good or bad. The arguments are thus formulated in terms of classes of possible epistemic states, transitions among these, and actions, processes, and procedures that result in such transitions. Second, the requirements of epistemic rationality are the conditions that one's epistemic factors must meet in order for them to be rational. Finally, the norms of epistemic rationality specify these requirements; they are the cognitive or conceptual items that function in our cognitive economies in order to allow us to think about and coordinate our actions with the requirements of epistemic rationality. Where the requirements of rationality are the things that we need to do or the conditions we need to meet in order to be rational, norms are the rules, plans, or policies that spell out those requirements.

2.2 D1: Epistemic Dynamicism I

2.2.1 Principle

D1 Epistemic rationality is dynamic; there are epistemic restrictions on transitions among epistemic states, not just states or contents of states, and there are characteristic relations among restrictions on states and restrictions on transitions.

This principle is straightforward. The thought is that epistemic rationality imposes requirements on what transitions between states reasoners carry out or, in other words, epistemic rationality distinguishes some transitions from other transitions. Given an array of potential transitions among epistemic states, epistemic rationality norms deliver (conditional) verdicts about them, carving them into distinct classes corresponding to normative valences - rational, irrational, rationally obligatory, rationally permissible, rationally forbidden. This partitioning into classes is what it means to restrict transitions. Suppose that, at some particular state, s, there is a set of possible transitions, $T = \{t_1, ..., t_n\}$ from s to various other states, $\{s_1, ..., s_n\}$, that a reasoner can realize with her repertoire of epistemic actions, abilities, skills, etc.. D1 tells us that, in at least some cases like this, there are at least some epistemic rationality norms that restrict this set of transitions by dictating that some of these transitions are (e.g.) rational and others are not. The norms restrict the set of transitions in two senses: (i) the norm maps the state or the conditions of application to a subset, T', of T, and (ii) assuming that the reasoner wants to carry out rational transitions, she's got to stay within the "restricted area" of T' by carrying out only the transitions in T'.

Now, if there are norms that restrict transitions by delivering verdicts about them, then, given the conception of norm-kernels, it follows that:

(1) There are norms of epistemic rationality that take epistemic transitions as their objects.

This is trivial from the form of norm-kernels. The normative claim in the consequent of a norm-kernel that delivers a verdict about transitions does so by attributing some valence to transitions. More, there are norms that condition their verdicts with respect to transitions and with respect to states on the properties of epistemic transitions. This entails more than merely the synchronic properties of states, for instance, being the determiners of which transitions are rational or irrational. Explicitly:

(2) There are norms of epistemic rationality that take properties of epistemic transitions as their conditions.

This will be argued for below.

Now, as to the characteristic relations among restrictions, only some can be easily seen with the material in hand. For instance, here are two characteristic relations between restrictions on states and restrictions on transitions:

(StoT) Norms that restrict epistemic states can generate norms that restrict epistemic transitions.

(TtoS) Norms that restrict epistemic transitions can generate norms that restrict epistemic states.

Now, "generation" is applied in a pretty murky way. What I mean is that, a norm that restricts states generates a norm that restricts transitions iff there exists some norm-kernel that takes a verdict on s with respect to a state-restricting norm as its condition of application and, in the consequent, assigns the same verdict to a transition t. Replace the relevant terms for the other generation claim. Take an arbitrary consistency norm, N_C. One can "generate" a restriction on transitions from N_C by means of a norm-kernel rule of this form:

(GenStoT) If s is rational according to N_C and $t = \langle x, s \rangle$, then t is rational.

This kind of generation will be of note below with respect to “reduction strategies”. The idea is intuitive; if one has rules telling rational from irrational states, the general policy of being in rational states and avoiding irrational states carves transitions into rational and irrational classes. One can generate a restriction on states as easily with the rule:

(GenTtoS) If t is rational according to N_T and $t = \langle x, s \rangle$, then s is rational.

Now, StoT and TtoS were phrased imprecisely with “can” as the key modal. It is possible, in general, to carve the class of transitions on the basis of restrictions on states and vice versa. However, this is very different from the much stronger and more interesting claim that every restriction on transitions can be derived from some Gen-like rule applying to states. This fails in one direction:

TtoS Irreducibility: Some norms that restrict epistemic transitions cannot be generated by norms that restrict epistemic states.

Counterexamples are given in the next section. It’s essentially a trivial consequence of the fact that some rationally distinct transitions from a state s converge in their terminal states. The other direction doesn’t fail, with some qualifications:

Qualified StoT Reducibility: Every norm that restricts epistemic states that excludes states that cannot be reached by any transitions can be generated by a norm that restricts epistemic transitions.

Almost every partitioning of states into rational and irrational classes can be generated by GenStoT, assuming that every state can be reached by some transition. This is easy to see. Assume there is some N that dictates that s is rational. If s is the terminal state of any transition, t, then there is some arbitrary partitioning, N_A, of transitions into rational and irrational sets such that t is rational according to N_A; just plug N_A into GenTtoS. However, if there is no t with s as a terminal state, the reduction fails. The idea is that, if reasoners start off in some initial state, s_I, and some norm N tells us that s_I is rational (or irrational), will there be no t with s_I as the terminal state to plug into GenTtoS.

This principle and the claims spelling it out are strongly at odds with a cluster of commitments that I will net together under the title of Epistemic Staticism.

2.2.2 Arguments for D1

Epistemic Staticism is a general orientation to epistemic rationality that holds that the requirements of epistemic rationality are strictly static.[5] Defenders of epistemic dynamics or, if you will, Epistemic Dynamicism, deny this, maintaining that the requirements of epistemic rationality are dynamic[6]. There are two independent dimensions of epistemic statics to distinguish:

Objectual Epistemic Statics (OES) There are no epistemic rationality norms that take anything other than epistemic states as objects.

Conditional Epistemic Statics (CES) There are no epistemic rationality norms in which the conditions of application concern anything other than the properties of epistemic states.

OES states that the objects of epistemic rationality are strictly *static objects* like doxastic states or credences, as opposed to *dynamic objects* like transitions from state to state or processes of state change. One can think of OES in terms of properties. OES states that epistemic rationality is a strictly static property. CES states that it may be the case that epistemic rationality is a dynamic property, but the conditions under which it holds of a dynamic object factor in exclusively the properties of states. The strongest form of Epistemic

---

[5] This is a widely held view. See, for a small set of examples, [179, 127], [83], [259], and [236].

[6] See Lam's Dissertation [154] as well as [146, 147], [217], Chapters 4-7 of [105], and [194, 196] for various forms of epistemic dynamicism. My case against staticism owes much to these sources, and Lam's in particular. Framing the argument in terms of reduction strategies is an insight I owe directly to Lam's discussion.

Staticism is the conjunction of both OES and CES. OES and CES individually amount to weaker forms of staticism.

Note that CES conditions valences on the properties of epistemic states in general, rather than on merely synchronic properties. This means that CES is prima facie more permissive than forms of Epistemic Staticism that condition valences only on synchronic and intrinsic properties of epistemic states[7]. By CES, the rationality of a state could be conditioned on (e.g.) relations between the state and another state or on its diachronic properties. I will not remark on these less permissive forms of staticism. The primary argument against the permissive form of CES works, a fortiori, against any more restrictive form of CES that focuses attention on subsets of static properties.

Epistemic Staticism is to be rejected along both axes. Let s, s', and s'' be epistemic states, and let pairs of epistemic states, (s, s') and (s, s''), be transitions from the state s to s' and from s to s'', respectively. Then OES and CES are to be rejected on the basis of the following schematic arguments:

(Against OES1) There are classes of pairs of transitions, $\langle s, s'\rangle$ and $\langle s, s''\rangle$, that are rationally distinct. If rationality distinguishes x and y, then there is a norm that takes x and y as objects. Thus, transitions are the objects of some epistemic rationality norms.

(Against CES1) Some of the foregoing pairs of transitions, $\langle s, s'\rangle$ and $\langle s, s''\rangle$, differ only in the properties of the transitions themselves. Thus, $\langle s, s'\rangle$ and $\langle s, s''\rangle$ are distinguished on the basis of the properties of transitions. If rationality distinguishes x and y on the basis of C, then there is a norm in which C is the condition of application. Thus, the conditions of application of some epistemic rationality norms concern properties of transitions.

The key premise in both arguments is the one that asserts that the rational distinguishability of two objects implies the existence of certain kinds of norms. This is trivial if there are

---

[7] Re: Time-slice Epistemology [179, 127]

any regularities governing the rational distinguishability of the objects of epistemic norms. Any such regularity can be stated as a rule taking the form of a norm-kernel. There is no other form that a regularity of rational distinguishability could take than that of a rule that connects some conditions to rational valences with some logical force. These rules are norms of epistemic rationality as defined in the introduction; they specify the regularities in rational distinguishability (and its requirements) or, in other words, they are rules that specify requirements of epistemic rationality. Thus, if the coming examples capture genuine regularities in rational distinguishability, then there must be some rule detailing the conditions of application, the valence, and the object of such a regularity. No more robust existential claim is being made than that.

The first premise is where controversy lies. To establish the first premise, I provide a set of straightforward examples of dynamic norms. These norms rationally distinguish transitions rather than states and that condition their valences, irrespective of object, on the properties of transitions rather than merely the properties of states. The counterexamples thus show that attributions of rationality distinguish dynamic objects sometimes on the basis of dynamic properties. The assumption of OES or CES precludes the explanation of these observations, and no means of explaining them away by reduction strategy works. As such, there is no plausible account of epistemic rationality that explains them or explains them away. So, dynamic norms of both kinds must be countenanced as parts of an account of epistemic rationality.

There are classes of pairs of transitions that are rationally distinct. This is an undeniable pattern in the attribution of rationality that supplies the first premise of Against OES. Consider these schematic norms:[8]

Defeasible Closure 1 (DC1) Let a belief set, s, contain $B_1, ..., B_n$, the true belief that $B_1, ..., B_n \vdash A$, and let it contain no counterevidence for A. Let s' be identical to s but that

---

[8] All of these examples are presented in a simplified form; rationality and irrationality are presented as mutually exclusive, binary properties of states, transitions, etc.. Each of these cases may be re-articulated with "more rational" and "less rational" in place of "rational" and "irrational" respectively, to recognize a perhaps more plausible comparative conception of the requirements of epistemic rationality.

it contains A. Let s'' be identical to s except that it contains $\neg A$. If these conditions hold, it is rational for an agent to transition from s to s', and it is not rational to transition from s to s''.

Demand for Evidence 1 (DE1) Let a belief set, s, contain only $B_1, ..., B_n$, none of which individually or collectively bears logical, probabilistic, explanatory, or any other evidential relations to A. Let s' be identical to s except that it contains A. Let s'' be identical to s but that it does not contain A (or, if your account of epistemic states can handle it, it contains explicit suspension of judgment on A). If these conditions hold, it is not rational for an agent to transition from s to s', and it is rational to transition to s''.

Minimal Mutilation (MM) Let a belief set, s, contain $B_1, ..., B_n$ and be logically consistent. Let s' be some belief set with an empty intersection with s. Let s'' be any belief set that satisfies every condition of minimal mutilation, conservatism, or preservation of entrenchment relative to s defended in the epistemological literature. If these conditions hold, it is not rational for an agent to transition from s to s', and it is rational to transition to s''.

These cases seem clearly to state norms that distinguish epistemic transitions. Given that these rules are actually rules of epistemic rationality, it's clear that OES is false; the requirements of epistemic rationality take more than epistemic states as their objects; they apply equally well to changes of epistemic state. More, some of them condition the rational distinction of transitions on the properties of the transitions such as whether they embody valid inferences, minimally mutilate, or constitute complete non sequiturs. Thus, CES is false, as well.

2.2.2.1 The Failure of Reduction Strategies

Against OES could fail if there are successful strategies for reducing each example to some more fundamental static norm. The thought would be that, if we accept the claim that the only norms that are real are the fundamental norms, then a successful reduction strategy straightforwardly shows that the dynamic norms above aren't real, and so can't plausibly impose any conditions on the adequacy of formal reconstructions of epistemic rationality.[9]

Reduction strategies will fail to undermine Against OES and Against CES. First, there is no reason at all to accept the claim above about the reality of derivative norms. This is a point I will return to at the end of the next section. For now, just note that, after all, brains are derivative entities; they reduce to their atoms. However, they seem perfectly real, and it's absurd to think that any worthwhile theory of the things in the world could exclude them without loss. Just the same, dynamic epistemic norms seem perfectly real even if reducible to static norms, and, as the next few sections will argue, it is their properties that a theory of epistemic rationality must account for independent of whether they are fundamental.

Second, there are no plausible reduction strategies. Several philosophers have defended the principle that:

Reduction 1 (i) Epistemic states are rational iff they satisfy the requirements of whatever canon of synchronic rationality one prefers (e.g., probabilistic coherence). (ii) A change of epistemic state from s to s' is rational iff s' is rational.[10]

The idea is that the rationality of the change from s to s' depends only on the rationality of s'. Supposing that this is the sole requirement of epistemic rationality on changes of epistemic state, this would offer a formula for determining whether any particular change of states is rational. Despite the fact that Reduction 1 proposes that there is at least one

---

[9] "Real" is used in a non-committal way to capture the idea that dynamic norms, if reduction strategies are successful, are not worth modeling in a formal theory of epistemic rationality. Perhaps "legitimate" or "acceptable" would be appropriate cognates.

[10] Moss [179] takes this approach explicitly, as does Foley [83]. This is also essentially the strategy behind the rationality postulates of the belief revision paradigm [92, 91] and probabilism [53].

norm that takes transitions as objects, the formula could, prima facie, be defended even by the staunchest epistemic staticist. The rationality of any transition ultimately or really only depends on the rationality of its terminal state, and this is a strictly static notion. No dynamic norm (other than Reduction 1 itself) is fundamental enough, then, for a staticist to enter it into her ontology.

Unfortunately for such defenders of Epistemic Staticism, Reduction 1 is clearly mistaken. Consider:

Non-Arbitrariness (NA) Let a belief set, s, contain $B_1, ..., B_n$ and an inconsistency, $A \wedge \neg A$. Let some subset, C, of s be counterevidence for A (not including its negation). Let s' be identical to s but that it does not contain A and it satisfies the canons of synchronic rationality as fully as possible. Let s'' be a non-empty belief set satisfying the synchronic canon of rationality as fully as possible that also has an empty intersection with s. If these conditions hold, it is not rational for an agent to transition from s to s'', and it is rational to transition to s'.

Even where s'' satisfies the conditions of synchronic rationality, it's still clearly irrational to move, willy nilly, to just any random statically rational state or another with no consideration of the evidence one has already acquired. This is not a rational change of belief. Thus, because these two transitions are indistinguishable in terms of the rational distinguishability of the relevant states, it must be the properties of the transitions that makes the difference.

The strongest response to counterexamples like this is to deny that there are any cases where the relevant states are indistinguishable. This takes the discussion down one of two paths: one takes us into the territory of minimal mutilation or epistemic conservatism of epistemic states and the other takes us into the territory of *epistemic basing relations* [248, 151, 150]. Let's take these paths in turn.

A tempting response to the examples above is the replacement of Reduction 1 with considerations of conservatism or minimal mutilation:

Reduction 2 (i) An epistemic state, s', is rational iff it satisfies the canons of synchronic

> rationality and it represents a minimal mutilation or a maximal conservation of the state that precedes it in the course of the evolution of a reasoner's epistemic states. (ii) A change of epistemic state from s to s' is rational iff s' is rational.

Reduction 2 is schematic, and we can let minimality be cashed out in terms of some preferred kind of comparison between the content of s and s' with respect to, e.g., entrenchment, conservatism, etc.. There are competing accounts of minimality in the epistemological literature, and nothing hangs on the details of these. The idea is that, as long as minimal mutilation or conservatism rules out all states with an empty intersection with its predecessor, then the distinctions that Non-Arbitrariness appears to make really just come down to Reduction 2 and the specification of mutilation.

This reduction strategy doesn't work. First, if correct, it falsifies CES - it doesn't reduce MM above to static notions only. More on this in the next section. Second, note that Reduction 2 fails to handle the following case:

Defeasible Consistency 1 (DC1) Let a belief set, s, contain $B_1, ..., B_n$, the true belief that $B_1, ..., B_n \vdash A$, and counterevidence for A, $C_1, ..., C_n$. Let s' be identical to s but that it contains A and $\neg A$. Let a distinct belief set, s'' be identical to s but that doesn't contain A, it contains $\neg A$, does not contain some B_i, and contains defeaters for B_i. If both of these transitions are feasible for a reasoner, then it is rational to transition from s to s'', and it is not rational to transition from s to s'. If the transition $\langle s, s'' \rangle$ is not feasible for a reasoner, it is rational to transition to s'.

This is a schema for non-paradoxical Preface-like cases. Let A be the conjunction of all B_i. Let the counterevidence for A, $C_1, ..., C_n$, amount to all available evidence for fallibilism about A - that any suitably large class of beliefs I have is overwhelmingly likely to contain a false belief and that $B_1, ..., B_n$ satisfies the size requirement. Standard Preface cases show that, arguably, it is more rational to maintain the inconsistency that results from adding A

despite the counterevidence for it (the evidence for fallibilism about A) than it is to randomly reject some of the particular conjuncts for which I have strong evidence. It is crucial to these cases that rejection of any B_i would be random. The reasoner in the Preface case cannot find a way to rationally select any of her B_i for revision; else, there isn't much of a problem at all. If acquiring counterevidence to some B_i is feasible for an agent, then there is no serious conflict. In any Preface case, if we could undermine the evidence for one or other of the belief sets in tension, there would be no reason to accept the inconsistency. Now, despite the fact that the terminal states differ across these transitions, it's clear that the static properties are not doing the work. Which transition is rational depends on which transitions are available for execution, not merely the static states involved. Thus, CES fails: it is the properties of transitions that matter for at least some norms.

Next, I consider basing relations. The usual thought [151] is that epistemic states inherently carry information about the evidential bases of beliefs. The epistemic basing relation is the relation that holds between a set of beliefs $B_1, ..., B_n$ and some other belief A just when the belief A is based on $B_1, ..., B_n$. According to the basing theorist, it is only relative to the beliefs on which a belief is based that it is rational or justified. Only states with appropriate basing relations are rational, and basing relations track (inter alia) inferential relationships. Consider a schematic reduction strategy in terms of basing:

Reduction 3 (i) An epistemic state is rational only if the *epistemic basing relation* of that state satisfies (for my purposes, arbitrary) condition C. (ii) A change of epistemic state from s to s' is rational iff s' is rational.

Deploying the basing concept, Reduction 3 appears to effectively reduce Non-Arbitrariness 1. Willy nilly changes to synchronically rational belief states are prohibited by Reduction 3 - the willy nilliness of the transition would be recorded in inappropriate basing relations. The resulting states would be irrational, so Reduction 3 would do the distinguishing work that Non-Arbitrariness seemed to do despite taking only the properties of states as fundamental. No need to appeal to any non-derivative dynamic notions.

The thought of Reduction 3 depends on the idea that, in these cases, the terminal states are not actually identical or even rationally indistinguishable. The terminal states that result from the processes in these cases fail to have identical basing relations, and this results in rational distinguishability between them. On the basis of this difference, the transitions are rationally distinguishable.

This strategy fails. First, it defends neither OES nor CES. The clause (ii) in Reduction 3 states a dynamic restriction. More, the condition of application of the norm ineliminably appeals to a property of transitions, namely the identity of the terminal states of transitions. That a transition has an irrational terminal state is a property of the transition. Thus, the reduction strategy fails to do the work it was needed for. This will be a problem as long as the epistemic staticist even acknowledges restrictions on transitions; any strategy for reducing the apparent restriction on transitions to properties of some class of states will depend on such states playing a specific role in the transitions. Thus, the strategy will generate a norm that conditions its verdicts on the properties of transitions. This objection applies not just to Reduction 3, but to any reduction strategy that attempts to explain away apparent restrictions on transitions in terms of properties of states. Second, the only form of basing relation that will even appear not to appeal to properties of transitions or processes will be doxastic basing relations. And these will fail to produce a viable reduction strategy.

There are three broad families of accounts of the epistemic basing relation, causal accounts, counterfactual causal accounts, and doxastic accounts. The details of these views are irrelevant to the thrust of argument, here.[11] The doxastic basing relation is the only version of the epistemic basing relation that works for Reduction 3. Suppose epistemic basing is a matter of the causal relations between states, actual or counterfactual, whatever preferred account one gives along these lines. Reduction 3 would then tell us that a transition is rational iff the terminal state is such that it is linked to the root state by an appropriate causal process. It is transparent that this could not support CES; the rationality of epis-

---

[11] For these irrelevant details, see: [150, 151, 248].

temic states would be conditioned on whatever properties of the causal processes that result in them constitute appropriateness. Thus, only the doxastic account of basing could possibly be deployed in a successful reduction strategy for defending CES.

Unfortunately for CES, Reduction 3 can't work. It fails to handle DC1. There, the norm distinguished transitions on the basis of the availability of other transitions, not on anything reducible to doxastic basing. The next section on D2 introduces more cases that undermine Reduction 3. So, neither path of Reduction 3 handles the counterexamples to CES and OES that I've provided.

2.2.3 Discussion

Epistemic staticism has an interesting and laudable pedigree.[12] It is directly defended by Williamson [259], Foley [83], Hedden [127], and Moss [179]. It is perhaps indirectly defended - by means of reduction strategies - in any thread of epistemology that grounds epistemic normativity strictly in the features of doxastic states. A proper taxonomy of such views is beyond the scope of this discussion, but it's important to note the prevalence of staticism. Epistemic Staticism subsumes arguably any form of epistemic consequentialism [236], including forms of - oddly enough - process and method reliabilism [105], and Bayesian epistemology [53]. These might be surprising to treat as staticist views, but the thought is that they support forms of Reduction 1 or something in the neighborhood of Reduction 1 that analyzes the rationality of a process in terms of the output states of a process satisfying static evaluative criteria.

Epistemic consequentialism is the clearest candidate for a reduction strategy, as it identifies the rationality of methods and processes straightforwardly in terms of their consequences - which is to say, whether the states they produce tend to be true, coherent, etc.. Reliabilism, for instance, typically grounds the rationality (read: justification) of a belief in terms of the reliability of the method or process the belief was produced by, and the reliability

---

[12] Which is to say, I don't think I'm swinging at ghosts.

of a method/process is typically grounded in the properties of the states it outputs. That is, even process reliabilists can endorse weakened versions of OES and CES that assert that the only fundamental or irreducible norms are those that take states and properties thereof as objects. After all, the reliability of a process reduces to the properties of the states it produces in the limit (or across nearby possible worlds, etc..).

The rationality of a Bayesian update - by simple conditionalization, for instance - is dependent on the rationality of its output state (i.e., probabilistic coherence and proper relation to prior states). As Moss [179] and Hedden [127] argue, the apparently diachronic conditionalization rules are actually just matters of proper basing. Failure of conditionalization is not irrational because the diachronic transition fails to have the right features; it's that the output state doesn't contain the proper relations between credences.

Desideratum D1 is grounded in what I think is the best resolution of the dispute between epistemic staticism and epistemic dynamicism: qualified - if polemical - rejection of staticism. Exactly what the qualifications are will be apparent after examining D2. At this stage, it suffices to note that establishing D1 effectively falsifies staticism as articulated above in CES and OES. There are more subtle forms of staticism, however, and the relations among D1 and D2 and more subtle static views is not as clear cut a clash.[13]

2.3 D2: Epistemic Dynamicism II

2.3.1 Principle

D2 Epistemic rationality is procedural; there are epistemic restrictions on epistemic actions, processes, and procedures, not just states or transitions among states, and there are characteristic relations among restrictions on transitions and restrictions on processes.

This principle is as straightforward as the first. It claims the class of objects of epistemic norms contains not only states - as per the Epistemic Staticist - and transitions - as per

---

[13] See, for instance, the form of staticism used as a foil in the next section.

the Epistemic Dynamicist - but also actions, processes, and even complex procedures. The idea is that, given an array of potential actions, $\{a_1, ..., a_n\}$, (resp. processes$\{p_1, ..., p_{n'}\}$, or procedures $\{P_{\pi 1}, ..., P_{\pi n''}\}$) there are epistemic norms that deliver verdicts about them, carving them into rational and irrational subsets, and thereby defining a restriction on actions (resp. processes and procedures).[14] An important part of that claim is that epistemic norms deliver verdicts about procedures. Explicitly, some epistemic norms deliver verdicts about procedural operations on basic actions. That is, if one wants to be rational, epistemic norms show us that there is a restricted space of actions and potential procedural constructions over those actions within the total space of possible actions and procedural constructions that we must keep to. Explicitly:

(1) There are norms of epistemic rationality that take epistemic actions, processes, and procedures as their objects.

That this is the case will be argued below. It will also be argued that:

(2) The norms of epistemic rationality take properties of epistemic actions, processes, and procedures as their conditions of application.

Now, as to the characteristic relations among restrictions, only some can be easily seen with the material in hand. For instance, here are three characteristic relations between restrictions on actions, processes, and procedures and restrictions on transitions:

(TtoA) Norms that restrict epistemic transitions can generate norms that restrict epistemic actions.

(TtoP) Norms that restrict epistemic transitions can generate norms that restrict epistemic processes.

---

[14] This general idea is shown below to apply as well to more subtle partitions, for instance, division of epistemic factors by deontic status: obligatory, permissible, impermissible, and prima facie reason to carry out or not to carry out.

(TtoPr) Norms that restrict epistemic transitions can generate norms that restrict epistemic procedures.

Letting p be a process and P_π be a procedure, the schemes for showing this are familiar:

(GenTtoA) If t is rational according to N_T and performing a at s results in a transition from s to s', $t = \langle s, s' \rangle$, then a is rational.

(GenTtoP) If t is rational according to N_T and completing process p at s results in a transition from s to s', then p is rational.

(GenTtoPr) If t is rational according to N_T and implementing procedure P_π at s results in a transition from s to s', then P_π is rational.

However a norm carves up the transition space, there is a rule that carves up the space of actions, processes, or procedures in a way that matches in the sense that if t has valence R, then any a, p, or P_π meeting the conditions in the Gen rules has valence R.

Finally, the converses of all of the foregoing claims hold:

(APPrtoT) Norms that restrict epistemic actions, processes, or procedures can generate norms that restrict epistemic transitions.

The scheme is, again, familiar:

(GenAPPRtoT) If a (resp. p, P_π) is rational according to N and performing a (resp. completing p, implementing P_π) at s results in a transition from s to s', then $t = \langle s, s' \rangle$ is rational.

This section predictably mirrors the prior principle articulation section. More, the generation facts of the prior section show that there's a kind of transitivity; for norms that restrict actions, processes, and procedures can generate norms that restrict states.

The prior claims are about what kinds of rules "can" generate what other kinds of rules. However, this is very different from the much stronger claim that every restriction on actions, processes, or procedures is equivalent to some Gen-like rule for states or transitions:

(APPrtoS Irreducibility) Not all norms restricting epistemic actions, processes, or procedures can be generated from norms restricting epistemic states.

(APPrtoT Irreducibility) Not all norms restricting epistemic actions, processes, or procedures can be generated from norms restricting epistemic transitions.

Counterexamples are given in the next section.

As with D1, D2 suffices for the failure of epistemic staticism.

2.3.2 Arguments for D2

The basic argument schemata for D1 can be modified and applied to get D2:

(Against OES2) There are classes of pairs of processes, $\langle s, a, s' \rangle$ and $\langle s, b, s'' \rangle$, that are rationally distinct. If rationality distinguishes x and y, then there is a norm that takes x and y as objects. Thus, processes are the objects of some epistemic rationality norms.

(Against CES2) Some of the foregoing pairs of processes, $\langle s, a, s' \rangle$ and $\langle s, b, s'' \rangle$, differ only in the properties of the transitions they result in or the processes themselves. Thus, (s, a, s') and (s, b, s'') are distinguished on the basis of the properties of transitions or processes. If rationality distinguishes x and y on the basis of C, then there is a norm in which C is the condition of application. Thus, the conditions of application of some epistemic rationality norms concern properties of transitions or processes.

The key premise is given by examples:

Bad Reasoning (BR) Let some reasoner's belief set at t be $s = \{B, A \rightarrow B, C \rightarrow A, C\}$ where $\rightarrow$ is a validly detachable conditional, and A is consistent with B and C. Let the belief set at $t+1$ be $s' = \{B, A \rightarrow B, C \rightarrow A, C, A\}$. Consider two possible process types for transitioning from s to s':

(Process 1) Detach A from the conditional $C \rightarrow A$ by affirming the antecedent C.

(Process 2) Detach A from the conditional $A \rightarrow B$ by deliberately affirming the consequent, B, and *not* performing any sort of explanatory or confirmatory inference to the same conclusion.

If these conditions hold, Process 2 is bad reasoning; an agent carrying out Process 2 in lieu of Process 1 is not executing a rational epistemic process.

Non Reasoning (NR) Let some reasoner's belief set at t be $s = \{B, A \rightarrow B, C \rightarrow A, C\}$ where $\rightarrow$ is a validly detachable conditional, and A is consistent with B and C. Let the belief set at $t+1$ be $s' = \{B, A \rightarrow B, C \rightarrow A, C, A\}$. Consider two possible process types for transitioning from s to s':

(Process 1) Detach A from the conditional $C \rightarrow A$ by affirming the antecedent C.

(Process 2) Hold an umbrella out during a lightning storm in the hope that it causes belief that A and, by chance, get struck by lightning and end up with the belief that A.

If these conditions hold, Process 2 is not even reasoning; an agent carrying out Process 2 in lieu of Process 1 is not executing a rational epistemic process.

Non-Arbitrary Reasoning 1 (NAR1) Let a belief set, s, contain $B_1, ..., B_n$ and be logically consistent. Consider two possible processes:

(Process 1) Subtract every B_i such that i is a multiple of 3, and restore consistency by removing any element of s that implies any such B_i. Let the resulting state be s'.

(Process 2) Transition to s'', which is an arbitrary, consistent expansion of s.

Process 1 is an utterly arbitrary change of belief. If these conditions hold, then an agent carrying out Process 1 in lieu of Process 2 is not executing a rational epistemic process.

Non-Arbitrary Reasoning 2 (NAR2) Let a belief set, s, contain $B_1, ..., B_n$ and be logically consistent. Let s' be a belief set with an empty intersection with s that satisfies any static conditions of rationality one wishes. Let s'' be any belief set that satisfies every condition of minimal mutilation, conservatism, or preservation of entrenchment relative to s defended in the epistemological literature. Consider two possible processes:

(Process 1) Subtract every B_i such that i is a multiple of 3, and restore consistency by removing any element of s that implies any such B_i until arriving at s''.

(Process 2) Expand s with an observation, C_i that contradicts some B_i, then remove the B_i and anything implying it to restore consistency. Iterate this process until reaching s'.

Process 1 is a process of chaining utterly arbitrary changes of belief while every step of Process 2 is a rational belief change. If these conditions hold, then an agent carrying out Process 1 in lieu of Process 2 is not executing a rational epistemic process.

Non-Arbitrary Reasoning 3 (NAR3) Let a belief set, s, contain $B_1, ..., B_n$ and be logically consistent. Let s' be a belief set that satisfies any static conditions of rationality one wishes, as well as every condition of minimal mutilation, conservatism, or preservation of entrenchment relative to s defended in the epistemological literature. Consider two possible processes for reaching s':

(Process 1) Subtract every B_i such that i is a multiple of 3, and restore consistency by removing any element of s that implies any such B_i until arriving at s'.

(Process 2) Expand s with an observation, C_i that contradicts some B_i, then remove the B_i and anything implying it to restore consistency. Iterate this process until reaching s'.

Process 1 is a process of chaining utterly arbitrary changes of belief while every step of Process 2 is a rational belief change. If these conditions hold, then an agent carrying out Process 1 in lieu of Process 2 is not executing a rational epistemic process.

2.3.2.1 Continued Failure of Reduction Strategies

It's straightforward to translate the reduction strategies discussed above in terms of processes rather than transitions. Note that each of the examples given directly undermines Reduction 1. There is no sense in which any of the terminal states above is irrational, unless one adverts to talk of minimal mutilation or basing relations. We can consider both of these options together. BR, NR, and NAR3 describe rationally distinct processes that terminate with the same synchronically rational state. NAR3 explicitly stipulates that the terminal state is reached with minimal mutilation, and, in each case, the identity of states can simply be stipulated to preserve basing relations. Thus, neither Reduction 2, in terms of minimal mutilation, nor Reduction 3, in terms of basing relations can possibly distinguish these processes. But they are clearly rationally distinct. Hence, none of the reduction strategies considered above will help the epistemic staticist.

There are still further reasons not to side with the staticist. I now want to return to the dismissal of non-fundamental norms that the epistemic staticist must commit to, and strengthen the case against it. The argument, Against OES, delivers the first parts of D1 and D2:

D1.1 There are norms of epistemic rationality that take transitions as objects.

D2.1 There are norms of epistemic rationality that take processes as objects.

That is, epistemic rationality is dynamic. There are epistemic rules that declare some transitions and processes rational and others not. Note the distinction between the conjunction of D1.1 and D2.1 and something a bit like them:

D+ Epistemic rationality is *fundamentally* dynamic; there are irreducible norms of epistemic rationality that take epistemic transitions or processes as objects or that condition valences of states, transitions, or processes on the properties transitions or processes.

Against OES (1 and 2) establishes D1.1 and D2.1 but does not deliver D+. It is, for all D1.1 and D2.1 say, possible that all norms that distinguish transitions and processes from one another can be derived from static norms by some Reduction, and so they are compatible with the failure of D+. Now, even if D+ utterly failed, it might still be the case that D1.1 and D2.1 would be features of epistemic rationality that formal accounts must capture. After all, if there are dynamic and procedural norms, they might play a role in epistemic cognition that is distinct from that played by static norms or they might play a role that is important to capture irrespective of their reducibility to static norms.

Even assuming a workable reduction of procedural norms to static norms, the derivative, procedural norms would still be norms to be accounted for - if only by giving an account of the fundamentals in which they are grounded. As the following few sections argue, dynamic rules (including procedural rules) are indispensable for studying normative constraints on epistemic dynamics, which, in turn, are indispensable for analyzing epistemic rationality in any way that captures its primary role in epistemic life. As reasoners, we are frequently more interested in dynamic norms than in static norms as components of epistemic advice. These kinds of norms are simply more useful to reasoners. Telling me that I ought not be in a contradictory belief state is not helpful. I need advice outlining strategies for resolving or coping with the tension. This advice may be grounded by static considerations of logical relationships, but the actual, working advice that I can receive and deploy will have a dynamic character; it will dictate or recommend procedures, actions, or processes for me to carry out (or at least dictate reasons to select some procedures, etc., over others): thinking of alternative explanations, confounders, countermodels, arguments and counterarguments, seeking out new evidence, seeking out colleagues to ask about the matter, etc.. I need direction for what I ought to do in the face of inconsistency, after all. So, insofar as logical frameworks are useful for generating or validating epistemic advice, they will have to say something, at least indirectly, about dynamic norms. However, they'll be far more useful for this cause if they can explicitly represent dynamic norms and their properties and provide valid principles governing them.

The thought is that normative constraints on epistemic dynamics are worthy of study as a matter of providing theoretically grounded epistemic advice or thinking about how we use epistemic norms in cognition (more on this in the next full section, 2.4). That is, even if staticists are successful in their reductive aims and the dynamic rules are entirely derivative on some reduction strategy, D1.1 and D2.1 would stand as features of epistemic rationality to be captured by a logical framework. Thus, a formal account of epistemic rationality norms would still need to provide resources for representing and reasoning about dynamic and procedural norms.

The prospects of procedural epistemic norms need not be so qualified, though, for the staticist cannot be successful in their reductive aims. First, note that each reduction strategy requires a dynamic norm in order to work. The second clause of each strategy is a norm-kernel that conditions the rationality of transitions, actions, processes, or procedures on the rationality of their terminal states; it's straightforwardly a rule that conditionally assigns a verdict to a dynamic or procedural object. Thus, each reduction strategy deploys a dynamic or procedural norm-kernel. A reduction of all other apparent rules for distinguishing (e.g.) transitions to (e.g.) Reduction 1 coupled with an account of the rationality of states would not eliminate all fundamental dynamic norms. It would leave exactly one fundamental dynamic norm: clause (ii) of Reduction 1. So, even an otherwise perfectly effective reduction strategy using (e.g.) Reduction 1 would falsify OES as well as generate a fundamental dynamic norm in Reduction 1. This would entail D+, and thereby entail D1.1 and D2.1. This point generalizes; any rule that generates constraints on dynamics from static constraints will just be equivalent to some dynamic constraint. All that is effected is a reduction to a fundamental dynamic norm-kernel. The section on D3 will strengthen this point. Second, the staticist cannot even go as far as reducing all dynamic and procedural norms to some static norm by means of any reduction strategy. It is shown in Chapter 4 that there can be no reduction of restrictions on actions, processes, and procedures to restrictions on transitions or states. This is actually a relatively simple consequence of the distinction between transitions and processes. Thus, D+ is correct in addition to D1 and D2, and there can be no escape from the requirement of accounting for at least D1 and D2.

2.3.3 Discussion

This completes the qualified polemic against epistemic staticism. The main thought is that, even if there are no fundamental dynamic norms in the sense that - with some ingenuity - one can get all of the dynamic norms from a specification of what makes epistemic states rational or not (which is itself extremely dubious), there are still strong reasons for any formal framework for norms to provide an account of dynamic and procedural norms. There manifestly are such norms, so, if they play an interestingly distinct or important role in epistemic cognition, an account of epistemic rationality should provide means of capturing their main features.

More, the arguments above connect to the problems for models of epistemic dynamics noted in Ch.1. Explicitly, Bad Reasoning, Non-Reasoning, and Non-Arbitrary Reasoning 3 imply the Processes and Constraints problems. Each of those putative epistemic rules rationally distinguishes token processes, p_1 and p_2, despite those processes being bookended by the same transition $t = \langle s, s' \rangle$. Assuming those norms really carve epistemic rationality at its joints (the plausibility of the examples suggests they do), a formal model of epistemic dynamics without the means of distinguishing processes from transitions will be literally incapable of representing the different verdicts that these norms give to p_1 and p_2. They will thus be incapable of representing the distinct constraints on epistemic dynamics that those norms specify, and so any system that suffers Processes and Constraints will be less than complete with respect to its representation of normative constraints.

Finally, this section is the first to emphasize procedural epistemic norms, norms that explicitly specify requirements on what processes to carry out. The putative norms given above to further bury OES show us that epistemic rationality distinguishes processes. Now, it may not be clear in what sense these norms are "procedural" rather than merely conditions on unanalyzed processes. The term perhaps suggests that epistemic norms specify something more structured, like particular ways of ordering or composing actions. Merely delivering verdicts on processes doesn't transparently imply this. But it does imply this. I'll argue below (and in Ch.4) that restrictions on processes do imply restrictions on what epistemic actions to perform and how to string them together with procedural operations. This should be clearer after consideration of D3 (especially the D3 Corollary).

2.4 D3: Restrictions on Epistemic Dynamics

2.4.1 Principle

D3 Epistemic rationality norms define restrictions on epistemic dynamics that model (a) ideally rational evolution of epistemic state, (b) ideally rational changes of epistemic state, (c) ideally rational procedures and processes of epistemic state change, and that (d) can be distinguished from constraints of capacity and feasibility.

Corollary Epistemic rationality norms restrict the construction of procedures; they

> determine how basic epistemic actions are to be sequenced, iterated, combined, chosen among, conditioned on tests, avoided, reversed, and cautiously executed.

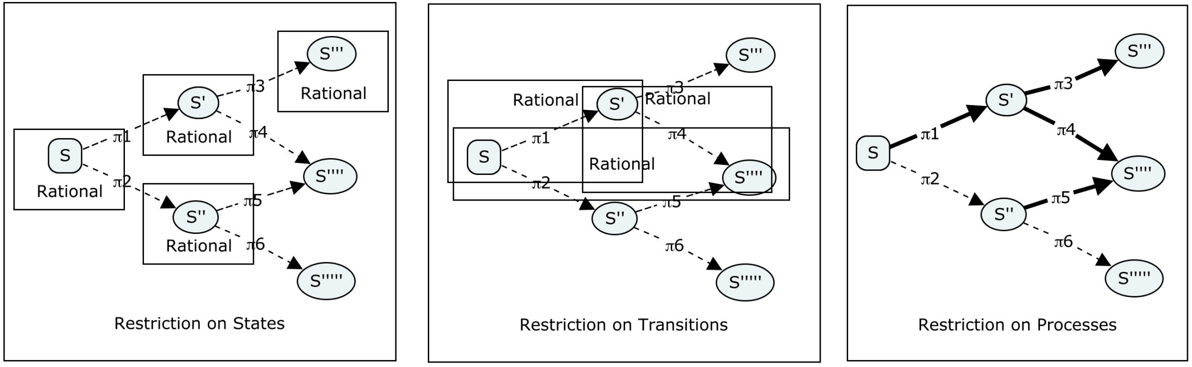
This principle identifies the functional content of epistemic rationality norms, the information associated with such norms that enables them to play the roles in cognition, decision-making, and evaluation that they play. In order to be epistemically rational, a cognizer's epistemic factors must satisfy certain conditions; else, the epistemic factors fall short of epistemic rationality. That is, epistemic rationality rules in some states, transitions, processes, and procedures and rules out others. Epistemic rationality norms define the conditions under which epistemic factors fall short or don't. They are the rules that do the ruling in and ruling out. But this is just to say that epistemic rationality norms define *restrictions on epistemic dynamics*, restrictions on which states, transitions, and processes can be a part of an ideally rational epistemic history. This includes imposing restrictions on how basic actions are conjoined by procedural operations.

What does it mean to define a restriction on epistemic dynamics? Here is a formal mechanism for precisifying restrictions on epistemic dynamics.

(Restrictions on Epistemic Dynamics) Consider some set S of components of epistemic dynamics; these could be states, transitions, processes, or some specified subset thereof. Let $P(S)$ be the powerset of S. Consider some partial order $\prec$ on the elements, P_i, of $P(S)$. Call some designated element of $P(S)$, T_I, the standard of irrationality. Now, let R_S and I_S be subsets of S such that: $\forall P_i \in P(S) \forall s \in P_i (P_i \preceq T_I \text{ iff } s \in I_S)$, and $\forall P_i \in P(S) \forall s \in P_i (P_i \succeq T_I \text{ iff } s \in R_S)$ Then: every $R_{\prec} = (P(S), \prec, T_I, R_S, I_S)$ is a *restriction* on epistemic dynamics.

The basic idea of restriction is a generalization of the notion of rational distinction or rational valence. Technically, each restriction comprises a (not necessarily exhaustive or exclusive) sorting of components of epistemic dynamics into R_S and I_S and an ordering over the powerset of a set of components of epistemic dynamics. The thought is that the subsets, R_S and

Figure 2.1: Restrictions on Epistemic Dynamics



If the restriction is on epistemic states, the boundary encircles the rational states, placing the irrational states out of bounds, not to be trespassed upon. If the restriction is on transitions, the boundary encircles the rational transitions, placing the irrational states out of bounds, and cutting across any paths to their terminal states, blocking access. If the restriction is on processes, the boundary encircles the rational processes; these are the processes that are to be executed, even if a transition to a process' terminal state is rational, the boundary may place some ways of carrying out the transition out of bounds. With the subsets R_S and I_S, restrictions codify this kind of structure.

I_S represent the way that categorical norms distinguish the objects to which they apply. R_S is the set of rational elements of S according to $R_{\prec}$, and I_S is the irrational set of elements of S. Ignore the other parts of the structures for now; they'll receive comment in due course. The core part of restrictions on epistemic dynamics for the purposes of this project is the division of components of dynamics to R_S and I_S.

Think of restrictions as boundaries separating different classes of epistemic factors. See Figure 2.1 for diagrammatic help.

Norms are sometimes simple in that they categorically assign components of epistemic dynamics exclusive statuses like "rational" and "irrational", “permissible” and “impermissible”. Restrictions codify the content of these kinds of norms. In the simplest case, there is a partition of the set of elements into two disjoint subsets, one of which is the rational or legitimate subset, the other not. For example, the set of epistemic states can be partitioned into those with inconsistent belief sets and those without. Formally, let $P(S)$ partition the set of epistemic states, S, into just P_C and P_{IC}, where P_C is the set of consistent belief sets and P_{IC} is the set of inconsistent belief sets. Let $P_{IC} \prec P_C$, $P_{IC} = T_I = I_S$, and $R_S = P_C$. Then the following structure is a restriction on epistemic states: $R_{C\prec} = (P(S), \prec, P_{IC}, P_C, P_{IC})$. A restriction like the foregoing completely captures the content of a full-stop consistency norms such as:

(Full-Stop Consistency) For any belief state, s, s is rational iff there is no sentence A such that both A and $\neg A$ are believed in s.[15]

It's easy to see that this norm defines a class of restrictions exactly like $R_{C\prec}$, though perhaps varying in the properties of their ordering relations. More on this in a moment. FSC quantifies over all belief states and partitions them into rational and irrational subsets on the basis of inconsistency. The ordering, $\prec$, of a restriction need only satisfy the condition that $P_{IC} \prec P_C$. Full-Stop Consistency (FSC) tells us that the set of all and only the consistent states is R_S, and the set of inconsistent states is to be identified with both I_S and T_I. Every other set of belief states may be incomparable in the ordering; the norm says nothing about states on the basis of their other properties.

The ordering relation on the powerset provides a representation of more nuanced structures. Continuing with the boundaries metaphor, the ordering over the elements of the powerset of components works like a concentric series of boundaries; this enables the representation of of comparative normative statuses (x is outside fewer boundaries than y) and thresholds (some of the boundaries are fine to cross as long as one doesn't go too far or pass a cutoff point). That restrictions order powersets of elements of epistemic dynamics also enables the construction to resemble norms in the way that they single out elements of epistemic dynamics by their properties. Explicitly, the thought is that each $P_i \in P(S)$ represents the extension of a property of epistemic factors. The ranking may be so defined that it equates the possession of some property with a normative verdict, e.g., consistency or degree of entrenchment with greater rationality. After all, epistemic norms, expressed via norm-kernels, typically tell us that entire classes of objects (e.g., states, transitions, or pro-

---

[15] This norm is, for familiar reasons [167, 121], both implausibly permissive and implausibly strict. Still, it serves as a suitable formal example.

cesses) possess some normative valence if they satisfy C, the condition of application. Given that the ordering is partial, there may be nothing else to say about the other elements of the power set with respect to that restriction - the ordering may leave them all incomparable. This is a virtue in terms of the representation of norms; a consistency norm like FSC may tell us nothing about a mixed set of consistent and inconsistent belief states other than that the consistent subset is a set of rational states.

More complex restrictions can be defined by more subtle norms. A strictly negative consistency norm, for instance, might declare inconsistent belief states irrational but not declare consistent belief states automatically rational. A more complex division of S and ordering thereon might then be defined by the norm. Whatever the level of nuance, each epistemic rationality norm defines some restriction on epistemic dynamics of some level of complexity. For a first example, note that the threshold element, T_I, can be fiddled with in a number of ways. The definition of restrictions builds in the notion of a threshold element, T_I, that does intuitive work: epistemic rationality norms are frequently categorical, as FSC is; the threshold element just matches the ordering to the categorical verdict that results from dividing states into R_S and I_S. However, for more nuanced norms that issue no all-or-nothing verdicts with respect to their objects, the threshold T_I may be safely ignored while the ordering is employed; for any restriction, there is always a restriction with an identical rational subset and ranking but that $T_I = I_S = \emptyset$. These ranking relations then can represent preference relations on epistemic factors. Thus, subtler epistemic norms than FSC that merely recommend some epistemic factors over others can be captured by restrictions. Of course, T_I might also represent a threshold with respect to the ranking.

Norms that deliver deontic verdicts can also be modeled with restrictions on epistemic dynamics. Deontics can be encoded with restrictions by taking I_S to be the prohibited components, R_S to be the set of all permissible components, and some $\prec$ maximal subset of R_S to be the obligatory components. An alternative and simpler encoding of obligation is this: any component, a, that is a part of every component in R_S or otherwise accompanies every permissible component is obligatory. The prohibited actions are thus all of those actions that preclude the execution of a. To illustrate: if one's norms permit all and only the courses of action p_1, p_2, and p_3, and there is a fourth course of action, q, that must be carried out in order to carry out any of the foregoing p_i, then q is obligated by those norms. This will be the case if q is a subprocess of each p_i, if q must be carried out prior to carrying out any p_i, or if one must do q concurrently with any permissible p_i. In other words, the obligatory actions are simply those actions such that either (i) they are the uniquely permissible actions under some condition or (ii) they must be executed in addition to any other permissible action. All one needs to analyze obligation in terms of restrictions is machinery for thinking about these kinds of features of action. With obligation translated into restrictions, permission and prohibition come along for free by the usual intertranslations of deontics follows easily. More on this at the end of Chapter 4.

Now, D3 tells us that epistemic rationality restricts states, transitions, and processes, and that these restrictions are distinct from the constraints of capacity and feasibility. In other words, rationality requires not being in some of the states agents can end up in and not doing some of the things agents can do. The thought is simply that rational agents are not infallibly rational; their dynamics can diverge from the requirements of rationality. Agents are typically capable of being in states that are not rational, making transitions that are not rational, and carrying out actions and processes that aren't rational. The requirements of rationality typically pick out a proper subset of the states an agent can be in, a proper subset of the transitions an agent can undergo from any particular state, and a proper subset of the actions and processes that an agent can carry out under a particular circumstance. At any given moment, there are plenty of things that are perfectly feasible for me that I rationally ought not do. I am generally capable of thinking wishfully when making predictions about the likely outcomes of my decisions, and, at a given time, such thinking may be feasible for me. A model of my capacities and feasibilities will contain these factors, a model of my ideally rational epistemic dynamics will not.

The D3 Corollary amounts to the claim that, for a set of possible ways to construct procedures out of a set of more basic actions, there is a restriction on that set - there is a carving of that set of possible procedures into rational and irrational (et al) subsets. It breaks down into separate subclaims for all of the possible ways to put actions together into procedures. For instance, it entails at least these three claims:

(1) Epistemic rationality requires the execution of sequential and iterative procedures under some conditions.

(2) Epistemic rationality requires parallel execution of procedures under some conditions.

(3) Epistemic rationality requires nondeterministic choice among procedures under some conditions.

Sequential procedures are procedures formed from more basic actions such that one performs one basic action and then another in sequence. Iteration is performing the same more basic action some number of times. Parallel execution is exactly what it sounds like: it's carrying out two distinct actions at once. Nondeterministic choice is the most transparent of all; the thought being that epistemic rationality constricts our choices but permits or delivers equal verdicts to some sets of actions, among which one can choose freely. By "more basic", I mean simply that the basic actions out of which procedures are built are relatively simpler than the procedures themselves. It may be the case that a procedure is itself formed out of complex procedures. Consider laboratory protocols, data analysis techniques, or the proof construction and model construction methods that compose the techniques that one aims to inculcate in much logic instruction. The list of claims about specific procedural operations entailed by D3 can be extended to exhaust the set of procedural operations.

The thought underneath the corollary and the subclaims fleshing it out is simply that it's epistemically rational or epistemically better to put together epistemic actions in some ways and not others under certain conditions. With respect to sequencing and iteration, it's helpful to briefly consider some examples. For example, one can partially guard against p-hacking in data analysis - applying techniques of data gathering and analysis that make spurious correlations appear statistically significant - by discretely separating the data acquisition and data analysis steps of a study [126]. The analysis of ad hockery in formal learning theory exposes other examples of rational vs irrational sequencing [85, 140, 142]. Systems of multiple norms may require parallel execution if one is to abide by all of them at once. If one is simultaneously to abide by norms of consistency and deductive closure, one might have to carry out consistency detection and restoration tasks while ensuring that one's belief set is closed. For an example of epistemic rationality requiring choice, consistency norms are useful examples. There are typically at least two different ways to restore consistency to any inconsistent belief set; if A and B are inconsistent, one can contract A (and what implies it) or B (and what implies it). Some closure norms do the same [22, 23]. More detailed and just plain more examples are given in the next section.

Finally, note that the following "generation" facts hold. That is, as with D1 and D2, one can easily find a rule to take a restriction on procedural construction to a restriction on, in these cases, paths through epistemic state space:

(4) Sequential and iterative procedures determine sequentially composed paths through epistemic state space.

(5) Parallel execution procedures determine converging and corresponding paths through epistemic state space.

(6) Procedures instructing nondeterministic choice determine branching paths through epistemic state space.

These facts follow easily from plausible assumptions. First, assume that each action corresponds to a path through epistemic state space. Then, if a procedure is formed from sequentially executing one action, a, and then another action, b, the two paths through the space of epistemic states corresponding to them, call them q and r, form a concatenated path qr. If a and b must be executed in parallel, then a and b are to be started at the same state and concluded at the same state.[16] Finally, if a and b are distinct procedures corresponding to distinct paths, both of which are equally rational or the choice of which is rational, the rational paths branch. But all of this is just to say that if epistemic rationality requires procedures of these forms, then it requires that one follow the corresponding paths. More precisely, we can give generation schemata as with the prior two desiderata:

(PrCtoP) If a procedure formed by sequentially composing two actions, a and b, is rational, and $p = qr$ where q corresponds to a and r corresponds to b, then p is rational.

(PrCtoP) If a procedure formed by parallel execution of two actions, a and b, is rational, and $p = q$ where q corresponds to a or $p = r$ where r corresponds to b, and q and r share root and terminal states, then p is rational.

(PrCtoP) If a procedure formed by choosing between two actions, a and b, is rational, and p corresponds to a or p corresponds to b, then p is rational.

Each of these rules defines a restriction on paths through the space of epistemic states - the paths that result from the rational procedures are distinguished from those that don't.

Now, why should we take it that epistemic rationality norms define restrictions on epistemic dynamics at all, let alone with the properties described?

2.4.2 Arguments for D3

Three independent arguments converge to support D3. The first is a direct argument from the robust conception of epistemic norms that emerges from many, independent lines of epistemological work. The main thought is that many lines of thought in epistemology converge on the idea that norms specify or define restrictions on epistemic dynamics. The second is an argument from the cognitive function of epistemic rationality norms. Epistemic norms play a role in deliberation about and guidance of epistemic action, and this can be done

---

[16] It's also quite plausible that doing a and b in parallel must take an agent through a unique sequence of states, since it's far from clear that one can be in more than one epistemic state at once.

only if they provide information in the form of restrictions on epistemic factors. The third argument appeals to the arguably standard (or at least highly influential) Dual Processes Model (DPM) or Default Interventionist Model (DIM) [73, 238, 2] of reasoning to suggest that if norms do define or specify restrictions on dynamics, they can play an important explanatory role for the DPM (resp. DIM). Finally, I argue for the D3 corollary directly from D3 and from several lines of thought that parallel the cognitive function arguments for D3.

2.4.2.1 Norms Specify Restrictions on Epistemic Dynamics

First, norms define the requirements of epistemic rationality. I take it that this is simply built into the concept of norms. Just as the norms of morality specify the moral requirements on action, epistemic rationality norms specify the requirements of epistemic rationality. They are, transparently, the rules or policies that specify what conditions an epistemic factor is to meet in order to be rational, what conditions suffice for irrationality, and so on.

Second, the requirements of rationality divide states, changes of state, actions, and processes that effect changes of state into rational and irrational classes or perhaps into more complex divisions involving rankings or preferences among these epistemic factors. I take it that this is simply built into the concept of normative requirements; for any given requirement, the set of relevant objects is divided into those that meet it and those that don't, those that are assigned some particular verdict and those that aren't. Sometimes requirements are looser or more nuanced; they might divide objects into a ranking according to how many requirements are met or how heavily weighted the met requirements are. Nonetheless, such requirements divide the objects of norms into various categories, though perhaps not exclusively or exhaustively. Thus, the norms that specify requirements of rationality divide the epistemic factors that they take as objects into rational and irrational subsets or they effect some other more complex division.

Finally, any division of epistemic factors into rational and irrational classes or epistemic rationality ranking of epistemic factors (with or without a threshold that assigns the locations in the ranking lower than it to the irrational class) defines restrictions on epistemic dynamics. For any norm N, define an order on the powerset of N's domain, $\preceq$, such that for all $x, y \in Dom(N)$, and all P_i and P_j in $P(Dom(N))$, $x \in P_i$, $y \in P_j$, and $P_i \succ P_j$ iff either (i) x is rational according to N, and y is irrational according to N, or (ii) x is more rational according to N than y. Then define a restriction on N's domain, $R_{\preceq} = (P(S), \preceq, T_I, R_S, I_S)$, where S is $Dom(N_i)$, and (iii) for all such P_i and P_j , $P_i \subseteq R_S$ and $P_j \subseteq I_S$, or (iv) the ordering over the powerset of the domain with respect to the norm is identical to the ordering in the restriction. Thus, given the construction of restrictions in the prior section, epistemic norms define restrictions on epistemic dynamics.

2.4.2.2 Robustness Analysis

That epistemic rationality divides epistemic factors into sets or rankings isn't merely a matter of pretheoretical judgment; it's the result of a "robustness analysis" argument. The notion of robustness analysis arises from consideration of what makes the sciences so distinctly reliable in generating knowledge. One way of explaining the distinct reliability of science is that the sciences frequently deploy multiple, independent methods of investigation, measurement, or discovery to triangulate their way onto verdicts about the existence or character of a phenomenon. As Wimsatt [260] puts it:

> The family of criteria and procedures which I seek to describe in their various uses might be called robustness analysis. They all involve the following procedures:
>
> 1. To analyze a variety of independent derivation, identification, or measurement processes.
> 2. To look for and analyze things which are invariant over or identical in the conclusions or results of these processes.

3. To determine the scope of the processes across which they are invariant and the conditions on which their invariance depends.
4. To analyze and explain any relevant failures of invariance. (p.44)

It is the deployment of these procedures of robustness analysis that underly the reliability and epistemic value of the sciences. There are varying accounts as to why this is the case, but for the purposes of this argument, the key upshots are these:

- If multiple, independent lines of investigation, measurement, or discovery converge on the same model, then the model is robust.
- Robust models are the best guides to the reality of the domain of inquiry.

One way to think of this is that by deploying multiple, independent methods or modes of investigation and measurement to generate a model of something, one is maximizing chances of discovering that certain features of a model are mere artifacts of particular methods. Convergence across methods rules out this possibility, given the independence of those methods - at least for the set of methods employed. Thus, a robust model is the best guide to the reality of the thing because it has the lowest chance of characterizing the thing inaccurately by way of modeling artifacts.

Multiple, independent lines of investigation in formal and traditional epistemology converge on the idea that epistemic rationality sorts epistemic factors into rational and irrational subsets. Thus, a model of epistemic rationality as a force that sorts epistemic factors emerges as robust. In showing this, I'll settle for only a small handful of examples.

Classical belief revision theories in the AGM mould explicitly articulate restrictions on epistemic factors. The consistency and closure requirements on belief sets articulate a restriction on belief sets. Now, in the standard formulations of AGM, the deductive closure condition defines belief sets; there are, technically, no belief sets in AGM that are not deductively closed. However, belief sets are explicitly deployed as models of ideally rational doxastic state, and these are characterized in the machinery of AGM exclusively by means of sets of sentences of a propositional language, L. Thus, the powerset of L implicitly characterizes the space of all possible doxastic states; every K such that $K \subseteq L$ is a doxastic state. Deductive closure rules out a great swathe of these potential doxastic states as irrational. To the extent that the trivial belief set is also to be avoided, closure imposes a kind of consistency requirement on doxastic states; no belief sets (rational doxastic states) that aren't trivial are inconsistent.

More, the epistemic commitment functions of expansion, contraction, and revision articulate restrictions on the set of possible transitions among belief sets. These functions explicitly represent ideally rational change of doxastic state. Those transitions in the extension of the commitment functions are a small, proper subset of the possible transitions among belief sets, which themselves are a small, proper subset of possible transitions among doxastic states.

These comments apply, mutatis mutandis to every other belief revision theory based even remotely on the same basic model of AGM. Spohn's ordinal conditional functions [233], Doyle's truth maintenance systems [67], Ellis' rational belief systems [69], Forrest's belief dynamics [84], and more [11, 18, 63, 74, 77, 99, 118, 134, 140, 149, 155, 160, 195, 211, 216, 20, 245] all essentially support these same claims. These models are not entirely independent of the AGM framework. However, the following models are.

Bayesian credence update theories [53, 182] also do the same work of dividing epistemic factors into rational and irrational sets. Rational epistemic states are modeled by probabilistically coherent credence functions. Rational epistemic transitions are a proper subset of possible transitions from credence function to credence function. The rational transitions are those that maintain probabilistic coherence according to some favored diachronic update function.

Dynamic epistemic logic [65] characterizes the evolution of knowledge under logical closure. In responding to the standard arguments against the closure of knowledge, many logicians and philosophers involved with DEL and epistemic logic more generally have adverted to a view of DEL as giving a model of ideal evolutions of knowledge states. That is, actual knowledge may not be closed under deduction, but it ought to be! Thus, in DEL, product updates give a model of ideal evolution of the knowledge states of epistemic agents. That is, where the true propositions involving an individual knowledge operator K_i in one epistemic model, M, represent an agent i's epistemic state, the product update for some event e identifies a transition to a new epistemic model, M', and thus a new epistemic state for i. These transitions from state to state are a proper subset of the possible pairs of such states, and so a proper subset of the epistemic transitions possible for i. They are the correct transitions with respect to the ideal evolution of knowledge.

Traditional epistemology follows suit. Basic norms of evidence apportionment do the same sorting work. Strict, anti-dogmatic norms such as Clifford's Rule divide belief states into those that are rational (those for which all beliefs have sufficient evidence) and those that aren't on the basis of a demanding standard [54]. Belief states such that there is sufficient evidence for each constituent belief are a proper subset of belief states. These kinds of static rules can be lifted to dynamic contexts easily: adding a belief without sufficient evidence is irrational. Such transitions are, of course, a proper subset of doxastic transitions. More permissive or conservative norms provide different sortings but nonetheless divide states into rational and irrational sets [160, 16, 173, 167]. Contractions of belief, according to conservative theories, without discovery of a sufficiently serious problem with the prior belief state, are irrational [121, 122, 160].

An independent line of inquiry that converges on this model is that analyzing the normative role of logic. Disputes about logic's normative role often assume a model of epistemic rationality as a force for dividing states and transitions into rational and irrational sets despite entirely different starting points. Consider such norms for rational inference as those put forward by Beall, Field, and MacFarlane:

(Beall) In the context of multiple-conclusion logic, where X and Y are theories (set of propositions): If $X \vdash Y$, then it's irrational to accept (all of) X and reject (all of) Y.[17]

(Field) Letting $Cr(A)$ be the degree of credence in A: If $A_1, ..., A_n \vdash B$, then $Cr(B) \geq \sum_i Cr(A_i) - n + 1$.[18]

(MacFarlane) **wo-** If $A, B \vdash C$, then you ought to see to it that if you believe A and you believe B, you do not disbelieve C.

wr+ If $A, B \vdash C$, then you have reason to see to it that if you believe A and you believe B, you believe C.[19]

These norms straightforwardly divide rational from irrational inferences. In general, inferences from one's present information that accord with the relevant logical laws comprise a proper subset of the possible transitions one can make among epistemic states. Interestingly, each is indeterministic: according to these rules, there are multiple correct or rational ways of inferring conclusions from one's present theory or information that comply with logic. Still, these represent only a few out of many possible moves.

In Beall's rule, the only verboten path is accepting all of X and rejecting all of Y. Where $Y = \{B_1, ..., B_n\}$, transitioning to an epistemic state where one accepts $X \cup B_i$ is, for each B_i, a rational transition. So is rejecting all of Y and rejecting some part of X. That is, if $X = A_1, ..., A_m$, rejecting all B_i and some A_j is a rational transition. By Field's rule, any transition from a credence state is rational as long as one's credence in the consequences of $A_1, ..., A_n$ is not lower than the sum of one's credences in each A_i. This allows a number of

---

[17] This is Beall's "normative constraint role" of logic [23]. This rule is formulated in terms of Beall's preferred multiple-conclusion LP [22]. The rule generalizes the single conclusion variant: If $X \vdash A$, then it's irrational to accept all of X and reject A.

[18] Field [77, 78, 79, 80] generalizes this rule in a number of ways. First, to the multiple conclusion case involving both degrees of credence $Cr(A)$ and disbelief $Dis(A)$: If $A_1, ..., A_n \vdash B_1, ..., B_m$, then $\sum_{i \leq n} Dis(A_i) + \sum_{j \leq m} Cr(B_j) \geq 1$. Second, to the case of conditional belief: Letting $A \mid C/D$ denote the conditional probability of A given full acceptance of C and full rejection of D, then: If $A_1 ..., A_n \vdash B_1, ..., B_m$ then for all C and C, $\sum_{i \leq n} Dis(A_i|C/D) + \sum_{j \leq m} Cr(B_j|C/D) \geq 1$.

[19] MacFarlane, in one of the more useful and interesting unpublished papers [167] I've ever come across, examines a wide array of inferential constraints constructed according to the schema: If $A, B \vdash C$, then (Normative Claim). By varying a number of parameters like scope and valence in normative claim, he generates an interesting array of possible bridge principles between logic and rationality. He ultimately defends the two in the main text.

options for compliance. One can, of course, raise one's credences to meet the constraint or, if one's credence in B is quite low, one can lower one's credence in some A_is. MacFarlane's rules both provide multiple options for compliance. For wo-, if you believe A and B, you can expand your beliefs to encompass C or come to disbelieve A or B. For wr+, if you believe A and B, you have a defeasible or pro tanto reason to believe C. Exactly what options comply with a pro tanto reason are unclear, but they include at least the following options: if one has reason to believe C and no reason not to, then one ought to believe C, and if one has reason not to believe C, then one is perfectly well permitted to continue not believing C.

I'll let the enumeration of implicit and explicit models of epistemic rationality's requirements that converge on divisions of epistemic factors rest with this small handful of samples.[20] I take the foregoing to at least have made it prima facie plausible that restrictions on epistemic dynamics provide a robust model of the requirements of epistemic rationality. These multiple investigatory lines converge on the division of epistemic factors by requirements of epistemic rationality. Thus, given the construction of restrictions in the prior section and the conceptual claim that epistemic norms define the requirements of epistemic rationality, epistemic norms define or specify restrictions on epistemic dynamics. Thus, the model of epistemic rationality norms as specifiers of restrictions on epistemic dynamics is a robust model, and so a reliable guide to the nature of epistemic norms. This, I take it, establishes the core part of D3: that epistemic rationality norms define restrictions on epistemic

---

[20] Really, this is a shallow sampling. My conjecture is that this is a conceptual feature of epistemic norms generally, so it's to be expected that every putative norm in the epistemology literature exhibits this feature. Here, I want to gesture at even more work at analyzing epistemic norms that supports this conjecture. These include the well-known rationality and entrenchment postulates of AGM theory [5, 91], Bayesian norms of credence update [53, 153], the closure and updating constraints on dynamic epistemic logics [65], the upshots for rationality of Tennant's finite dependency networks [245], and more [23, 45, 67, 69, 74, 84, 108, 184, 233]. One might question their independence, given that they all attempt formal analyses of rational belief revision, but (a) they are developed under sometimes very different conceptions and assumptions, and (b) that's not all! For accounts of epistemic rationality or norms thereof that converge on restriction outside of formal rational belief revision theories, see: Field's triple of articles, [77, 78, 79], Beall [22, 23], Restall [210], any first-order account of norms of rationality that employs the deontic notion of permission (e.g., [164]), most forms of evidentialism [57], specific accounts of rationality norms like the knowledge norm [1], Pollock's defense of internalism [194], Gibbard's meta-ethics [98], Schafer's doxastic planning account of at least some rationality norms [217], or Kolodny's explicit defense of process-requirements [146, 147]. Finally, Pettigrew's entire epistemic utility theory research programme construes various methodological, logical, and credence norms as constraints on (e.g.) the formation and update of credences [189, 191, 190, 192].

dynamics.

2.4.2.3 Epistemic Dynamicism

Now, the content of the restrictions on dynamics that epistemic norms define is a matter of some controversy, as noted in defending D1 and D2. Epistemic norms perform sorting for at least epistemic states. This almost needn't be argued insofar as one thinks that there are any epistemic rationality norms. Observations of norms that rule in epistemic states with some features and rule out others are ubiquitous. One need only find a consistency or closure norm of some kind or remind oneself that a "wise man apportions his belief to the evidence" [132].

However, it is not the case that all of the models enumerated above converge on restrictions on all of these factors of epistemic dynamics. This may simply be a case of limited scope. The models under consideration may never have been applied to the general analysis of epistemic rationality. Rather, they have been guided by specific observations or problems about dynamics and rationality. Nevertheless, D1 and D2 expand the epistemic factors over which epistemic norms define restrictions to transitions and processes. In arguing for D1 and D2, I really only argued for the claim that epistemic rationality divides the sets of possible transitions and possible actions and processes into rational and irrational sets or imposes a ranking on them. I called these "restrictions" without supplying a technical definition for the term. Given the construction of restrictions on epistemic dynamics in this section, D1 and D2, formulated in terms of divisions of transitions and processes into distinct sets, directly imply that there are restrictions on transitions and processes in the more precise sense of this section. This point is strengthened by the fact that most of the models above converge on the restrictions on dynamics rather than merely on restrictions of states.

In sum, multiple, independent lines of epistemological and philosophical investigation converge on restrictions on epistemic dynamics as models of the requirements of epistemic rationality. These requirements are specified by epistemic norms. From the definition of restrictions on dynamics and the observations defended as principles D1 and D2, it follows that epistemic norms specify restrictions on epistemic states, transitions, actions, and processes. Finally, that these restrictions are distinct from the constraints of capacity and feasibility is trivial from the observation - made in the discussion section - that reasoners are capable of doing irrational things.

2.4.2.4 The Cognitive Function of Epistemic Norms

Shifting gears to the second argument, a powerful rationale for restrictions as a formalization of the content of epistemic norms is that epistemic normativity is to play a role in the dynamic control of epistemic state change. This is the cognitive role or function of epistemic norms, and for playing this role, it's necessary for epistemic norms to be associated with some kind of restriction (whether in the form of categorical verdicts or rankings) with regard to possible states to be reached or transitions and processes of change to be applied. Else, norms don't really provide information with which to guide us in figuring out what to do next, epistemically. The "what to do next" problem is, fundamentally, the challenge of finding a function from arbitrary epistemic states with multiple successors reachable by multiple feasible processes to some subset of those successors or to some subset of the feasible processes. This argument is consilient with the fact that, it is, after all, not rare or controversial to conceive of epistemic norms doing exactly this kind of ranking or preference defining work [117, 6, 133, 34].

Now, as to the guidance aspect of norms, if one sees two potential states ahead of one's current state one of which is consistent while the other is not or sees that one is in an inconsistent state, both the ordering built into restrictions based on (e.g.) consistency norms and the assignment of objects to rational and irrational sets can provide the requisite ranking. The ordering and categorical assignments determine (or represent a regularity with respect to) what transition one must make from one's current state in those circumstances. The thought is that such a norm ranks the set of consistent states more highly than the set of inconsistent states (or places it directly among the rational states as opposed to the irrational states). If one is to be in more highly ranked or rational states, this settles choices (at least along one axis of decision). A general function, "From any state, select the highest items in the ordering" where the rational set is uniformly higher than the irrational set then settles the "what to do next" problem, perhaps indeterministically (depending on the nature of the ordering). Apply this to transitions, actions, and processes mutatis mutandis. Thus, restrictions provide information capable of providing regulative guidance of epistemic actions; they provide answers to the "what to do next" problem. I will argue below that epistemic norms must define or otherwise be associated with restrictions if they are to provide regulative guidance.

There are two obvious roles for epistemic norms to play in cognition:

(Descriptive) Epistemic norms could be regularities that show up in the best (idealized) models of an agent's cognition. They "govern" ideal epistemic dynamics the way the laws of physics govern the evolution of physical systems.

(Normative) Epistemic norms could be objects that play a role in evaluating and guiding cognition, not just in describing regularities. They govern evaluations of epistemic dynamics the way that laws govern legal evaluations of behavior, and they guide epistemic dynamics the way that professional methodologies guide professional behavior.

The description of the normative role of epistemic norms suggests two kinds of normative role for norms to play:

(Post Hoc Evaluation) Assessment epistemology: evaluating an epistemic state, change of epistemic state, or process of epistemic state change after it's occurred. Epistemic norms can be used as standards or normative ideals that a system can realize, fail to realize, or approximate to some greater or lesser degree. That is, norms can be used to determine, post hoc, that a system is in the wrong state or that it got to its current state by the wrong means.

(Ante Hoc Regulation) Directive epistemology: deployment of norms to select epistemic actions, procedures, or processes. Epistemic norms can be used as standards or normative ideals that a system's current or future states and potential courses of action can realize, fail to realize, or approximate to some greater or lesser degree. The relations that current and future states and courses of action bear to these standards can be used in selecting future courses of action.

To illustrate post hoc evaluation, we can consider the normal uses of various kinds of standards. For instance, we can use the dictionary to look up "comprise" and "compose" in order to determine whether a sentence with one of these terms that has already been spoken was grammatically correct. Additionally, we often want to tell whether someone has followed an experimental protocol or proper methods of data analysis in presenting us with a new finding or whether someone's already constructed proof actually works. Post hoc evaluation of reasoning processes is also a standard practice. Knowing that I'm prone to wishful thinking and recognizing that wishful thinking is a reasoning process that frequently guarantees error, I can ask whether I got to my conclusion by merely being wishful.

Ante hoc regulation is about projecting what norms require of future actions, and trying to ensure that our actions accord with those requirements. For instance, we can look up "comprise" and "compose" in order to determine how to construct a sentence with one of these terms (rather than the other). With some conception of how the sentence ought to come out, use the dictionary to guide the process of constructing the sentence in a way that matches grammatical usage. For further parallel examples, empirical scientists employ experimental protocols to guide their information generating actions in the lab, statisticians use established methods of data analysis to guide the way they draw conclusions from data sets, and mathematicians use methods of structural and transfinite induction to guide their construction of proofs. Again, ante hoc guidance of reasoning processes is also standard practice. Knowing that I'm prone to wishful thinking and recognizing that wishful thinking is a reasoning process that frequently guarantees error, I can suppress this process and substitute better ways of reasoning.

A useful framework for thinking about ante hoc regulation is epistemic planning in artificial intelligence research [115, 95]. Essentially, epistemic planning is just a special case of planning. Planning is conceived of in terms of planning problems and solutions to planning problems. The basic background is the by now familiar conception of dynamic systems. There is a set of possible states of the relevant system - whether the system is a robot, set of interacting agents, or single agent in an environment - and a set of actions that the system can perform to transition from state to state. An action library is a subset of the set of all possible actions. A goal state is a state of the system that has some specified set of properties that encode a goal for the system. Planning problems are triples of initial state, goal state, and action library.

The intuitive thought is that the planning problem is the problem of figuring out how to get from the initial state, by some sequence of actions in the action library, to the goal state. Solutions to planning problems are thus plans or instructions built out of procedural operations on the action library that suffice to transition the system from the initial state to the goal state. An epistemic planning problem is simply a planning problem that builds epistemic normativity into the goal state by specifying that the state have certain normative features. For instance, the goal state may be any state that satisfies some condition like consistency, deductive closure, or proper basing relations. As D2 implies, a planning problem might also impose restrictions on the kinds of plans that can be used to reach the goal state. That is, the goal state might specify that only a certain kind of procedure was used to reach it.

With that in mind, the following argument directly establishes D3:

1. If epistemic rationality norms can be deployed or applied in deliberate, ante hoc regulation or epistemic planning, then they must define restrictions on epistemic dynamics that model (a) ideally rational evolution of epistemic state, (b) ideally rational changes of epistemic state, (c) ideally rational procedures and processes of epistemic state

change, and that (d) can be distinguished from constraints of capacity and feasibility.

2. Epistemic rationality norms can - and, if they are to have any role in cognition other than description of regularities in practice, must - be used in deliberate, ante hoc regulation/epistemic planning.

3. $\therefore$ Epistemic rationality norms define restrictions on epistemic dynamics that model (a) ideally rational evolution of epistemic state, (b) ideally rational changes of epistemic state, (c) ideally rational procedures and processes of epistemic state change, and that (d) can be distinguished from constraints of capacity and feasibility.

The basic model of deliberate ante hoc regulation can be thought of in terms of planning problems. Suppose that an agent, α, endorses a set of norms, $\{N_1, ..., N_n\}$, and sets out to follow them. At any initial state, s, that α can be in, there is a set of possible epistemic states, $\{s_1, ..., s_m\}$, that α can reach by applying some set of epistemic processes, $\{p_1, ..., p_k\}$, formed out of the actions in α's action library. According to $\{N_1, ..., N_n\}$, some of these epistemic states are rational, others irrational, some transitions from s to s_i are rational, others not, and some processes are rational, others not. The goal state can involve arriving at a normative state, carrying out a normative state change, or carrying out a normative action or process. Ante hoc regulation is deliberating about which of the p_i will meet the conditions imposed by $\{N_1, ..., N_n\}$ on the goal state, and then selecting and carrying out the right p_i. To do this, α must answer certain arrays of questions about $\{N_1, ..., N_n\}$. These questions are displayed in Table 2.1.

The necessity of these questions in deliberate ante hoc regulation is somewhat self-explanatory. If one is trying to deliberately guide one's belief formation toward a rational state, one must figure out which rational states one can reach by chaining actions in the action library. In order to deliberately select a normative action or process, one must be able to figure out what feasible actions and processes are ruled in by the norms one endorses. If one cannot distinguish the normative courses of action from those that are merely feasible,

Table 2.1: Solutions to Epistemic Planning Problems

States	**Transitions**	**Processes**
Whether a possible epistemic state, s, is rational or not with respect to N_i. • What the static indicators of rationality and irrationality are with respect to N_i. • Whether the static indicators of rationality and irrationality with respect to N_i hold for s	Whether a possible epistemic transition, t, is rational or not with respect to N_i. • Whether the current state is rational or not with respect to N_i. • What the dynamic indicators of rationality and irrationality are with respect to N_i. • Whether the dynamic indicators of rationality and irrationality with respect to N_i hold for t.	Whether a possible epistemic process, p, is rational or irrational with respect to N_i. • What the procedural indicators of rationality and irrationality are with respect to N_i. • Whether the procedural indicators of rationality and irrationality with respect to N_i hold for p.
Whether the current state is rational or not with respect to N_i.	Whether a transition that is rational with respect to N_i is feasible (available or executable).	Whether a process that is rational with respect to N_i is feasible (available or executable).
Which feasible (available or executable) transitions and processes lead to rational states, and which lead to irrational states.	What the consequences of executing a rational transition with respect to N_i will be (e.g., what state one ends up at).	What the consequences of executing a rational process with respect to N_i will be (e.g., what state one ends up at, what transition is effected).

the use of norms in action guidance will be frustrated. One would figure out what one's norms required, confuse those with the merely feasible actions, and be no better equipped to act normatively. Supposing that epistemic norms are useful in ante hoc regulation, the only plausible use is in providing answers to questions like these. Thus, in addition to dividing epistemic factors into sets by valence, norms useful for ante hoc regulation must provide information that distinguishes the merely feasible actions from the rational ones.

From the definition of restrictions on epistemic dynamics, it follows that any norm that provides answers to these questions specifies restrictions on epistemic dynamics. More precisely, if these questions, with respect to norm N_i, even have answers, then N_i divides epistemic factors into distinct subsets and identifies them as rational or irrational, more rational or less, etc.. Thus, the definition of restrictions on dynamics trivially implies that there is a restriction on epistemic dynamics that agrees with N_i. For any norm N_i, define an order on the powerset of N_i's domain, $\preceq$, such that for all $x, y \in Dom(N_i)$, and all P_i and P_j in $P(Dom(N_i))$, $x \in P_i$, $y \in P_j$, and $P_i \succ P_j$ iff either (i) x is rational according to N_i, and y is irrational according to N_i, or (ii) x is more rational according to N_i than y. Then define a restriction on N's domain, $R_{\preceq} = (P(S), \preceq, T_I, R_S, I_S)$, where S is $Dom(N_i)$, and (iii) for all such P_i and P_j , $P_i \subseteq R_S$ and $P_j \subseteq I_S$, or (iv) the ordering over the powerset of the domain is identical to the ordering in the restriction. This restriction will be exactly the one that answers the questions in the deliberation table above. Finally, as epistemic norms deliver verdicts about all of these epistemic factors (as argued above), an ideally rational agent will thus find itself in the rational states, traversing the rational transitions, and employing the rational processes; else, there is no sense in which it's ideally rational.

In sum, to deploy a norm N_i in ante hoc regulation, one must answer the questions above. Answering the questions above is straightforwardly limning a restriction on epistemic dynamics that meets conditions (a)-(c) of D3. The restriction outlined by answering the deliberation questions must be distinct from the constraints of capacity and feasibility if it's to be useful, so the restriction meets (a)-(d) of D3.

The second premise is easy to see. Epistemic norms would be entirely inert in cognitive life if they had no role in ante hoc regulation. Suppose that they didn't. Then epistemic norms might still be useful to post hoc evaluation; they could be the standards by which we make post hoc evaluative judgments. But what could possibly be the point of post hoc evaluations in the absence of ante hoc regulation? If post hoc evaluations could be used to deliberately guide future epistemic behavior, it's unclear how this wouldn't just be deliberate ante hoc regulation. Suppose that post hoc evaluation shows that state s_{i-1} was irrational according to norm N_i. An agent might bypass directly reasoning about the verdicts of N_i by reasoning about future states by analogy with s_{i-1}. Future state s_{i+1} might have exactly the features of s_{i-1} that made it irrational according to N_i, and so a reasoner might seek to avoid s_{i+1}. But this is just a circuitous way of answering the questions in the table above, and so just a circuitous form of ante hoc regulation. If post hoc evaluations couldn't be used in at least this way, epistemic norms can be nothing other than tools for pointless score keeping.[21]

The pointlessness of such scorekeeping is worth contemplating. If there were only post hoc evaluation according to norms, errors in reasoning would be entrenched and incorrigible. To suppose that there are no means of ante hoc regulation is to suppose that there is no way to drive the negative post hoc evaluations into future corrections. Any such correction requires some way of guiding future action according to the verdicts of our epistemic norms. Suppose that post hoc assessment of my prior reasoning reveals that I drew conclusions on insufficient evidence. To assume that there is no way to employ norms in ante hoc regulation is just to say that there is no way for me to then deliberately ensure that my future belief states accord with norms of sufficient evidence. Deliberately doing so would require answering questions in the table above with respect to such evidence norms and then, somehow, driving the selection of future courses of inference in accordance with those answers. Thus, if there

---

[21] More, even post hoc evaluation is a dynamic process that can be done correctly or incorrectly. I may, after all, find out that I misapplied an evaluative norm. The norms of epistemic rationality must be used to guide this process, or else all we can do is keep score about our score keeping behavior.

were nothing that counts as ante hoc regulation, then there would simply be no way to guide my future epistemic actions. In the absence of some way of suppressing my erroneous dispositions to act and supplying a normative course of action, I would, at best, only be able to recognize that I have made errors while being utterly powerless to do anything about them. In particular, without the ability to guide my future evaluations, which would require ante hoc regulation of evaluation processes, I would be hopelessly sentenced to a future of commission of errors despite being capable of recognizing all of them. This is a little too fast; there might be ways of normatively driving epistemic action without deliberate ante hoc regulation. I consider this point below. For now, it suffices to note that this kind of epistemic fatalism isn't merely troubling, it's demonstrably not our epistemic situation.

To see that epistemic norms can be deliberately applied, all that's necessary is (a) noting that empirically supported models of cognition and cognitive behavior presuppose deliberate application of epistemic norms, and (b) direct observation of epistemic behavior reveals deliberate application of epistemic norms. First, note the vast empirical literature on cognitive biases and de-biasing strategies [2, 136, 238]. The thought that runs through this research is that our innate cognitive endowments are "fast and frugal" but lead us into error in complicated or novel situations. In order to override these error-prone strategies, cognizers must halt their cognitive processes and enact a normative alternative. These normative alternatives can be supplied by explicit knowledge of logical and statistical methods. That many epistemically damaging biases can be counteracted by training in various formal methods establishes the prevalence and usefulness of effortful, reflective, and otherwise deliberate ante hoc regulation [71, 72, 180].

Second, the sketch of deliberate application of epistemic norms above is, in many respects, notably similar to the EMIL-A model of the cognitive dynamics of norm compliance [59].[22] The EMIL-A model directly provides a place for restrictions on dynamics to fit; basically,

---

[22] It may be interesting to note that the sketch of deliberate application was outlined independently of EMIL-A, and that there are a number of other independent sketches of norm adoption that basically capture the contours of the model. See, for instance, the synoptic power of EMIL-A in [59]. This might suggest that there is something robust to the main ideas of the sketch.

the array of features that need to be discovered in the procedure above correspond to beliefs that must be developed in the EMIL-A model.

The EMIL-A model is broken into an epistemic and a pragmatic component. The epistemic component - comprising the norm recognition module - transforms information inputs into normative beliefs. Normative beliefs come in three varieties: main normative beliefs, norm pertinence beliefs, and norm enforcement beliefs. Main normative beliefs essentially comprise the information contained in norm-kernels: beliefs about the verdicts of norms under their conditions of application, beliefs about whether the conditions of application hold, about what specific actions in the agent's repertoire are assigned what deontic operators under the present conditions, and other beliefs, including those about the authority or sources of the norm. Norm pertinence beliefs are about whether the norm binds the agent. Norm enforcement beliefs are about the sanctions underwriting the norm.

Normative beliefs are transformed by the norm-adoption module in EMIL-A into normative goals by the pragmatic component. Normative goals are then sent to a planning system that solves planning problems in which the goal states are shaped by normative goals, which generates normative intentions and, finally, actions. I overlook alternative models of norm compliance here, for the reason that EMIL-A synthesizes the insights of available alternative accounts of norm compliance.[23]

The parallels between the ante hoc regulation procedure and the EMIL-A model are clear: both models require figuring out the features of norms outlined in the table above as part of synthesizing a normative plan of action. The main normative beliefs generated by the norm recognition module are explicit answers to the questions articulated in that table. Restrictions on epistemic dynamics are exactly the kind of construction from which those answers can be derived. The EMIL-A model thus requires the use of representations that carry the information that can be derived from restrictions on epistemic dynamics.

Third, recall the examples above about ante hoc regulation. Empirical scientists employ

---

[23] This is argued at some length in [59].

experimental protocols to guide their information generating actions in the lab. Statisticians use established methods of data analysis to guide the way they draw conclusions from data sets. Mathematicians use methods of structural and transfinite induction to guide their construction of proofs. Philosophers deploy all of the above from time to time as well as established argumentative methods to approach theses proposed by colleagues and forerunners. More homely examples abound. Knowing that I'm prone to wishful thinking and recognizing that wishful thinking is a reasoning process that frequently guarantees error, I can suppress this process and substitute better ways of reasoning. All of these examples involve both pre-emptive commitment to certain courses of action as well as self-correction or the halting of natural or intuitive processes and supplying normative courses of action. These kinds of occurrence of ante hoc regulation have not only been empirically catalogued but can be readily witnessed. The second premise of the argument is secure.

2.4.2.5 Cognition and Restrictions on Epistemic Dynamics

Finally, I want to consider a third way of arguing for the core of D3. The argument is this:

1. D3 can solve the problem of supplying alternative courses of action in Default Interventionist (DI) or Dual Processes (DPM) models of cognition.
2. Default Interventionism/DPM is a theoretically well-supported theory of rational cognition.
3. If a principle solves problems in theoretically well-supported theory, T, without worsening T, then there is a powerful reason to accept the principle.
4. $\therefore$ There is a powerful reason to accept D3.

Default Interventionism is the main idea of the Dual Process Models of rational cognition.[24]

---

[24] This argument is framed in terms of the Dual Processes Model and/or the Default Interventionist Model. These, in many ways, come to exactly the same thing - heuristic processes can, upon reflection and with effort, be halted and substituted for normative processes. See, for overview and defense: [73, 72] and the relevant chapters of [2]. The DPM can be said to underly the DI model of rational cognition. A largely

There are numerous particular theories of rational cognition that are grouped together as Dual Process Models, but they share enough and are, collectively, prevalent enough that it's safe to call the standard model of rational cognition the Dual Processes Model (DPM). The DPM divides epistemic processes into two types: Type 1 processes (also called The Automated Set of Systems (TASS)) and Type 2 processes.

The defining feature of Type 1 processes is that they are autonomous under the stimulus conditions that trigger them; they proceed in a mandatory fashion without input from higher level cognition. These processes tend to be automatic, executable in parallel, and fast and cheap, from a processing power standpoint. These processes or systems also tend to be unreliable for solving abstract, complex, or novel cognitive problems, but we nonetheless resort to them by default.

By contrast, the defining feature of Type 2 processes is that they are non automatic, serial, and require a greater degree of processing power than those of Type 1. Type 2 processes are also frequently called rule-based or rule-guided processes because (i) they adhere more closely to normative rules (e.g., probability and logic) than Type 1 processes (though this is contested [100, 101, 102, 152]), and (ii) are typically impacted by an agent's explicit knowledge of such rules [168, 239]. Type 2 processes tend to be far more reliable for solving abstract, complex, and novel cognitive problems [136, 242, 2].

The DPM requires something like ante hoc regulation. The defining interaction between the two types of processes is the suppression and substitution of Type 1 by Type 2 processes. Type 2 processes are, in other words, the processes of reflective thought that override and correct our instinctive or intuitive ways of thinking. This is Default Interventionism; the default Type 1 processes are intervened upon by higher, normative processes.

Rationality in cognitive science is standardly thought of in terms of divergence from ideal epistemic dynamics. In particular, cognitive scientists break rationality into two kinds, epistemic and instrumental. Epistemic rationality is standardly modeled with probabilistically

---

parallel argument could be framed in terms of theories of the cognitive dynamics of norm compliance like EMIL-A.

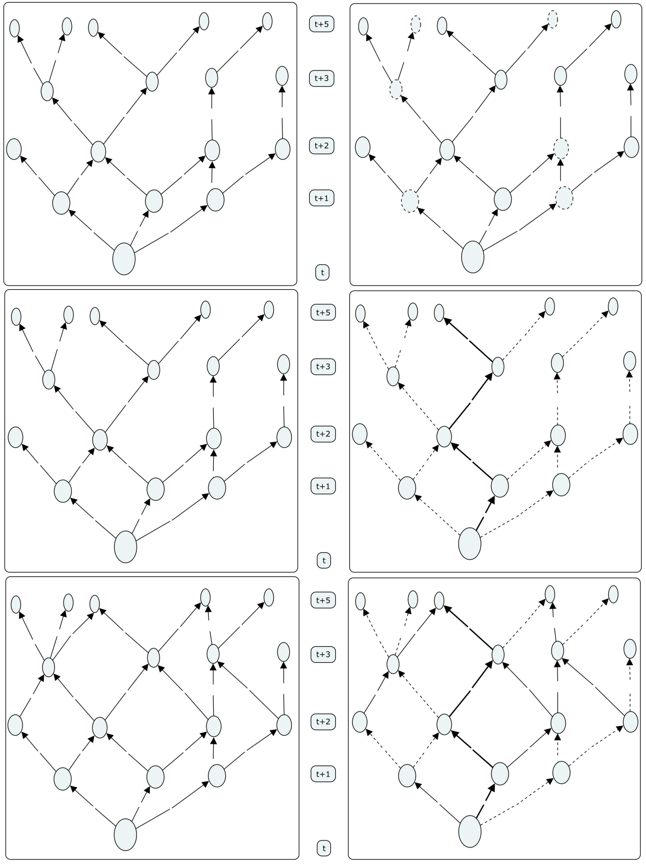
coherent degrees of belief, and instrumental rationality with expected utility optimization. The rationality of particular cognizers is measured as divergence from these two ideal models. The Dual Processes Model holds that Type 1 processes must be overridden and replaced by Type 2 processes because the heuristic features of Type 1 processes cause irrationality (or divergence from the ideals just noted). However, in order for it to be cognitively useful to override Type 1 processes, there must be normatively superior processes to substitute for them. This is where restrictions on dynamics come in.

An efficient solution to the problem of supplying superior alternative processes is for agents to be able to use norms to model normatively correct processes and then drive them into behavior. Let me unpack that. The proposal consists of two ideas. First is the idea that a kind of Type 2 reasoning about epistemic norms is part of any efficient solution: explicit representation or modeling of the norms and (logical) derivation of the elements of the set of normatively correct processes given an initial state. The second idea is merely schematic: there must be some way of driving the alternative processes derived from representation of norms into epistemic behavior. I won't say much about the second idea because (i) the problem can be thought of in terms of well-developed theories of plan-synthesis and execution (e.g., solving a planning problem and then acting on plans/complex intentions), and (ii) the first idea is what does the work for this argument for D3.

In order for epistemic norms to provide solutions to the "supply problem" in the DPM, epistemic norms must carry information that the agent can use about what epistemic factors are normatively correct and under what conditions they are correct. But this just to say that epistemic norms divide epistemic factors into rational and irrational subsets or otherwise supply rankings of them, and this implies that they specify restrictions by the arguments above. Thus, epistemic norms must specify restrictions on epistemic dynamics in order to solve the supply problem. To illustrate the idea of supplying normative courses of action, consider Figure 2.2.

To sum the key thought: epistemic norms map the normative actions, processes, and

Figure 2.2: Supplying Normative Solutions



These graphs illustrate the relevance of restrictions on dynamics to ante hoc regulation. The shaded ovals are epistemic states. Imagine the states on a plane, with the bottom closest to us in time (our actual epistemic state at t), and the rest receding away from us (hence, appearing smaller) in time. In the first and second pairs of illustrations, the arrows between the states represent transitions among states. In the third, arrows represent distinct actions or processes of state change that may lead to the same states. Dotted lines indicate irrationality with respect to norms, N_1, N_2, and N_3. If a state has a dotted outline, it's irrational according to N_1. If an arrow is dotted in the second pair of graphs, then it's irrational according to N_2. In the third pair of graphs, dotted arrows $\preceq$ thin arrows $\preceq$ thick arrows according to (the restriction defined by) N_3, where the latter is a norm restricting processes rather than transitions. The idea is that, where a (relatively simple) norm N restricts epistemic states (by attributing flat out rationality or irrationality), an agent, looking forward from her state at t can use these restrictions to see what future states, transitions, and processes are verboten or not with respect that norm (or to see what the ranking is). This is (a visual representation of) the kind of information that enables the substitution of normative courses of epistemic action in the Dual Processes/Default Interventionism Model. From a given state, N_1 tells us what future potential states are ruled in and which are ruled out. N_2 tells us which potential future transitions are ruled in and which are ruled out. N_3 tells us which processes are to be preferred over the others at future states. Were an agent to halt her heuristic processes at the bottom most state and search for a normative course of action, each of these norms gives her an an array of potential solutions to the planning problem of complying with N_i.

successor states to each state, but they can only do this if they define restrictions. If there were no restriction on (e.g.) states or actions correlated to each norm, they couldn't do this work. Suppose that there were no restrictions associated with any N_i. Then, from the perspective of the bottom-most state, there is no structure that distinguishes the rational or rationally preferable states from those that aren't. The same holds for the transitions, actions, and processes. If there is such a structure, it is equivalent in verdicts to some restriction on dynamics. The availability of this information about an agent's library of available actions enables the substitution of normative courses of epistemic action.

Now, the argument could fail if there were some way to blindly guide epistemic practice in response to post hoc evaluation. Consider a trial and error model of how post hoc evaluation could result in ante hoc regulation. Let $a_1, ..., a_n$ be some set of epistemic changes - expansions, contractions, and revisions of full belief, let's say. Let $A_1, ..., A_m$ be a set of assertions. Suppose that a reasoner has a multi-track disposition in which conditions $A_1, ..., A_m$ comprise the conditions of characteristic manifestation, $a_1, ..., a_n$ comprise the characteristic manifestations, and there is some way of mapping the A_j to the a_i that captures the tracks of the disposition.[25] Suppose that there is an epistemic norm, N, that says if A_5 holds, then a_4 is a rationally permissible belief change, but that if A_5 and A_9, a_4 is rationally impermissible. Suppose further that, upon carrying out an action, the reasoner can perform post hoc evaluations, and these post hoc evaluations modify the disposition, weakening or strengthening connections between conditions of manifestation or even inhibiting or disinhibiting the disposition. In particular, suppose that when the reasoner carries out a_4 when A_5 and A_9 hold, and she reflects on this performance, she gets a negative post hoc evaluation that, over time, inhibits the disposition to execute a_4 when she recognizes that A_5 and A_9 hold. Alternately, suppose that verdicts of rational permissibility strengthen dispositions to carry out some A_i under the appropriate A_j.

Given this model, it might seem as though post hoc evaluations are all we need in order

---

[25] See [249, 51] for the gist of the terminology deployed, here.

to correct reasoning processes. After all, if one's disposition initially dictates doing a_4 under A_5, but post hoc evaluation reveals that this is an error and inhibits the future disposition to do a_4 when A_5 and A_9 hold, one will come to do (or avoid) exactly what the norm requires.

While this would obviate the need for deliberate ante hoc regulation, it wouldn't tell against ante hoc regulation in general nor would it provide a role for epistemic norms such that they would fail to specify restrictions on epistemic dynamics. First, this method just is one of ante hoc regulation. Call the initial disposition d_1; it is a disposition to act in a way that violates norm N. After post hoc evaluation of reasoning that involves doing a_4 under condition A_5 and A_9, the reasoner comes to deploy a new disposition, d_2. The disposition d_2 is characterized by the rule: If A_5 and A_9, then don't execute a_4! The reasoner has developed an internalized plan of epistemic action that adheres to norm N. The trial and error mechanism for revision of dispositions just is a case of using epistemic norms to control one's future epistemic actions in a way that adheres with the norm in question. It is of course, not a mechanism that enables knowing adherence with norms. Second, post hoc evaluation itself requires that epistemic norms divide epistemic factors into rational and irrational subsets or otherwise impose an ordering on them. From the definition of restrictions on epistemic dynamics, it follows that even norms useful only for post hoc evaluation specify restrictions on epistemic dynamics. Finally, the direct observation of deliberate ante hoc regulation noted above suggests that this garden path is a dead end, for deliberate ante hoc regulation is actual and ubiquitous.

2.4.2.6 Arguments for the D3 Corollary: Restrictions on Procedural Construction

Corollary Epistemic rationality norms restrict the construction of procedures; they determine how basic epistemic actions are to be sequenced, iterated, combined, chosen among, conditioned on tests, avoided, reversed, and cautiously executed.

The arguments for the D3 corollary derive from a number of sources: (i) direct examination of some classes of epistemic rationality norms, (ii) derivation from D3, (iii) derivation from the requisite structure for epistemic regulation and knowing adherence/application, (iv) derivation from the requisite structure for epistemic planning.

The fundamental assumption of the D3 Corollary is that epistemic processes can be analyzed into actions (or executions of actions) and sequences thereof. Actions, including epistemic actions, can be combined, sequenced, otherwise connected via procedural operations to generate processes. There is a basic set of procedural operations on basic actions. Relational algebra is a formalism for the study of procedural operations on actions, represented as binary relations [28, 30]. The procedural operations of relational algebra include sequential composition, iteration, parallel execution, choice, test, complement, reversal, and fixed point. Perhaps the most familiar procedural operation is sequential composition, which is an operation that orders actions in a linear sequence. In general, there are graph-theoretic and set-theoretic structures associated to each procedural operation. We will see more on this in later chapters. For now, procedural operations can be thought of as instructions. Sequential composition instructs one to carry out an action a and then, come what may, to carry out an action b. Other familiar procedural operations are associated with familiar instructions: do a some finite number of times, do a while doing b, choose among a and b, do a conditional on some condition C holding, avoid a, undo the effects of a, do a extra carefully, etc.. These are associated in Table 2.2.

It's tempting to think that the corollary follows trivially from the fact that epistemic norms restrict epistemic processes. After all, epistemic processes are sequences of epistemic actions, so if norms rule in sequence p and rule out sequence q, it's plausible that they do so on the basis of the ordering of the basic actions in p and q, respectively. But this is too fast. It might be the case, for all that has been shown, that norms only rule out sequences on the basis of distinguishing features of processes that are independent of the ordering of epistemic actions that comprise them. For instance, norms might only rule out processes because they result in prohibited transitions among states, because they require prohibited

Table 2.2: Procedural Operations

Instruction	**Procedural Operation**	**Relational Structure**
Do a, then do b!	Sequential composition "$a; b$"	Relational Composition
Do a some finite n times!	Kleene Iteration "$a*$"	Relational Composition
Comply with both the instructions for a and b!	Parallel Execution "$a \cap b$"	Intersection
Select either a or b (inclusive)!	Indeterministic Choice "$a \cup b$"	Union
Avoid a, do anything other than a!	Complement "$-a$"	Complement
Reverse the effects of a!	Reversal "$\curvearrowleft a$"	Inverse of Relation
Do a with minimal effects!	Fixed Point "∇a"	Fixed Point

individual steps (irrespective of order), or because they lead to certain kinds of epistemic states. So, there needs to be an argument for the claim that epistemic rationality norms really do rule in some procedural operations on epistemic actions and rule out others on the basis of the structure of the procedures. That some restriction on procedural construction emerges can't be merely incidental. Fortunately, several arguments are ready at hand.

Direct Examination of Procedural Norms First, one can simply point to both accepted and proposed normative and methodological standards that dictate or require particular procedural operations on epistemic actions. These are fairly numerous. Consider Table 2.3. To explicate just a handful of these examples, first, consider "argumentation methods". By this, I just mean the procedures that are explicated in logic and critical thinking texts for analyzing and evaluating arguments. In standard textbook presentations, the following processes typically have chapters or subchapters devoted to them: initial presentation of argument (excising it from its original spoken or written context; denote this with π_p), search for enthymemes (π_e), mapping out putative evidential or support relations (π_m), formalizing

Table 2.3: Norms Requiring Types of Procedural Operation

Norms Requiring Types of Procedural Operation	**Norm/Procedure/ Method**	**Field of Study**	
Norms that Require Sequential Composition	- Argumentation Methods - Hypothesis Testing Strategies (FLT) - Proof Construction and Model Checking - Mathematical Algorithms - Enabling and Prevention Procedures - Inductive Strategies (Mill, Bacon)	- Formal and Informal Logic - Formal Learning Theory - Formal Logic - Formal Methods - DEL, Lab Manuals - Philosophy of Science	- [Angell 1964, Walton 2013] - [Schulte 2017, Kelly 1988] - [Quine 1950, Kleene 1967] - [van Benthem 2007] - [van Ditmarsch et. al 2007] - [Bacon 1851, 1960, Mill 1843]
Norms that Require Iteration	- Argumentation Methods - Hypothesis Testing Strategies (Formal Learning Theory) - Proof Construction and Model Checking - Iterated Belief Revision - Iterated Credence Update	- Formal and Informal Logic - Formal Learning Theory - Formal Logic - AGM, Belief Revision - Bayesian Epistemology	- [Ibid] - [Kelly et. al RBR, Schulte 2017] - [Ibid] - [Gardenfors 2008, Hansson 2017] - [Hansson 2017, Gardenfors 1992]
Norms that Require Parallel Execution	- Every System of Multiple Norms	- Traditional Epistemology, Ethics of Belief	- [Littlejohn and Turri 2014, Clifford 1999, Adler and Rips 2008]
Norms that Require Choice	- Argumentation Methods - Logical Constraints - Proof Construction and Model Checking - Belief Revision	- Formal and Informal Logic - Beall, Field, Bayes - Formal Logic - AGM	- [Ibid] - [Beall 2013, Field 2009a,b,c, 2015, Christensen 2004] - [Ibid] - [Ibid]
Norms that Require Test	- All Conditional Norms	- Traditional Epistemology, Ethics of Belief	- [Ibid]
Norms that Require Complement	- Norms of Prohibition	- Traditional Epistemology, Ethics of Belief	- [Ibid]
Norms that Require Reversal	- Belief Revision - Backtracking Algorithms (Ariadne's Thread)	- AGM - Algorithms	- [Ibid] - [Ibid]
Norms that Require Fixed Point	- Efficient FLT Strategies	- Formal Learning Theory	- [Ibid]

the argument in some logical language (e.g., translating the argument into first-order logic; π_f), testing putative support relations using formal and informal methods (e.g., deploying semantic tableau methods to prove the arguments valid or invalid; π_{ft}), and evaluating the truth of premises (π_{pr}). The chapters usually involve an array of algorithms, informal procedures, or rules of thumb for structuring these processes.[26]

The norms that govern the overall process of analyzing and evaluating arguments dictate particular sequences of these subprocesses. For instance, the requirement that one interpret arguments charitably requires that one not rest one's case on the evaluation of the strength of the argument without having completed the search for enthymemes. That is, one can't complete the evaluation of argument well without searching for enthymemes. At the very least, this suggests a ranking of possible procedures formed out of the foregoing subprocesses that places the search for enthymemes prior to completion of testing support relations. That is, the principle of charity requires that the process of analysis and evaluation of argument take some procedural form that matches the sequence: $...; \pi_e; ...; \pi_{ft}; ....$[27] Thus, the norm of charity divides these ways of constructing epistemic processes into relatively acceptable and unacceptable subsets. Insofar as charity is a norm of epistemic rationality, the point is made. But perhaps this is too capacious a sense of epistemic norms. More direct examples are available.

The usual model-checking and proof construction procedures [82, 145, 200] involved in the formal testing phase of argument evaluation themselves require fairly straightforward sequential orderings of substeps. One must apply the non-branching rules to a tableaux prior to applying the branching rules; else, in systems that countenance infinite trees, one's attempt at proving the validity of an argument will be vulnerable to erroneous failure to close the tree [82]. Insofar as accurate verdicts on the validity of arguments shape require-

---

[26] Among many others, see: [9, 109, 145, 159, 137, 203, 253]

[27] Really, an algorithm for doing this would probably involve looping from formalization and testing procedures $\pi_f; \pi_{ft}$ back to π_e; a good test for whether or not one has charitably interpreted an argument is how obvious it seems that the argument is invalid. If it's trivially invalid, and your interlocutor is smart, you're probably being uncharitable!

ments of rationality - either because accuracy is itself a norm of epistemic rationality or because epistemic rationality conditions rationality of belief and inference on logical validity - epistemic rationality divides ways of ordering steps of proof construction into relatively acceptable and unacceptable subsets.

The main thought of both of these examples is that the results of an epistemic process can depend on the ordering of epistemic actions and other procedural features of the construction of a process. This is a familiar fact of epistemic life that is well-recognized in iterated belief revision and DEL. It's well-known that some ways of iterating belief revision functions are utterly untenable.[28] There are also dynamic paradoxes of belief revision, as Harman [119, 121] points out. Given a limited set of available actions, it may be impossible to revise one's belief states out of an inconsistent state once one ends up at one. One of the sets of driving observations for DEL is that some actions can obscure or reveal information. One needn't consider complex subterfuge to see this, as the revelation of information is readily apparent in the basic PAL example of the muddy children in Chapter 1. Though those cases were multi-agent cases, one can easily see the observation to single agent cases, as in The Detective case. Performing some actions, for instance, makes information available for inference. Obfuscation is also easy to see. One can perform actions that prevent later epistemic actions from being feasible. If the detective, for instance, had made insulting information about K available to both A and B, the deduction of K's homebase location wouldn't have been possible. These cases show that getting to the right terminal state can require the use of some subset of procedural sequences. Scientific practice reveals the ubiquity of this. In order to identify a substance, one might need to destroy it by a chemical identification process. This prevents other such processes from being carried out - at least on the same sample. Finally, the explicit ordering of epistemic actions can be used as a form of de-biasing strategy. There are cases in which, in order to suppress an unreliable heuristic process from triggering, one simply has to prevent oneself from accessing some

---

[28] Here, I have in mind Tennant's Degeneracy results from [244].

information. For instance, blind review of article submissions explicitly hides information that can trigger biases. Additionally, there are cases in which one has to sequence the ordering of tasks in order to avoid certain kinds of outcome. For instance, the sequencing of, first, registration of hypotheses, and then the discrete separation of data gathering from data analysis processes can be used to avoid some forms of p-hacking [126]. Insofar as epistemic rationality norms divide outcomes into rational and irrational subsets (et al), these general observations imply that epistemic rationality divides procedural constructions into rational and irrational subsets.

The foregoing table of norms and methodological rules and the subsequent discussion shows, if I'm right, that the procedural construction of epistemic processes is not a matter of indifference from the epistemic perspective. First, some of these norms explicitly divide processes into rational and irrational sets on the basis of the procedural operations out of which they built. Thus, from the definition of restrictions on epistemic dynamics, there are restrictions on procedural construction.[29] Second, even if the static outcomes of processes are the only features of normative relevance, the static outcomes often depend on the procedural structure of processes. Thus, in such cases of dependence, the existence of restrictions on static outcomes of processes implies the existence of restrictions on ways of constructing processes out of actions. More, processes constructed in certain ways may be related to normatively relevant outcomes in certain modal ways. As in the cases above, particular ways of (e.g.) procedurally operating on actions may be reliable or even necessary for producing (or avoiding) an outcome even if there are other processes that also generate that outcome. Norms frequently dictate the execution of the reliable processes over unreliable processes that converge, in some cases, on outcomes with the reliable process. Thus, even in the absence of norms that directly specify restrictions on procedural construction, the norms that specify restrictions on static outcomes indirectly specify restrictions on procedural construction, and the structural restriction is not incidental. The restriction on procedural construction is

---

[29] The next section spells this out.

based on a dependence relation between the procedural structure and the epistemic outcomes of the procedure.

Derivation from D3 In addition to what can be gleaned from the examples of procedural norms in the foregoing sections, there is a straightforward derivation from D3:

1. Every restriction on processes is associated with a set of restrictions on procedural construction.
2. Every epistemic rationality norm is associated with a set of restrictions on processes.
3. $\therefore$ Every epistemic rationality norm is associated with a set of restrictions on procedural construction.

The first premise is direct. Consider an arbitrary restriction on epistemic processes, $R_{\preceq}$. Note that epistemic processes generally break down into basic actions and sequences of basic actions. At the very least, epistemic processes can be built from sequences of more basic actions or processes, so all but indivisible processes break down this way. Sets of processes may instantiate patterns of construction. All of the processes, p, in the set might contain particular actions a or subprocesses q in a specific sequence, or they might be so formed that, at a given point in a sequence, every process contains one element, p_i, of a particular set of subprocesses, $\{p_1, ..., p_n\}$. Such sets of processes implicitly rule out any processes not constructed according to these patterns, and these patterns can be described in terms of procedural operations. In the first case, processes without a particular sequencing of subprocesses are ruled out. In the latter, any processes without a choice among $\{p_1, ..., p_n\}$ are ruled out. In the absence of any such procedural patterns, there is a general procedural construction defined by any set of processes: choice over the set of token processes in the set. Thus, for any set of rational processes in $R_{\preceq}$, R_{Pr}, there is a corresponding division of ways of building procedures out of actions and processes into rational and irrational sets. Thus, by the definition of restrictions, if we can identify sets, S, of procedural constructions, then every restriction on processes defines a unique set of restrictions on procedural constructions - the set of procedural constructions such that the processes in R_{Pr} match it.

Identify sets of procedural constructions with sequences of procedural operations on actions: sequencing, choices, iterations, etc.. A process, p, matches a procedural construction iff there is a set of subprocesses, $\{p_1, ..., p_n\}$, that, when interspersed with the sequence of procedural operations, is identical to p. That is, a process matches a procedural construction iff one can find a way of building p out of a set of subprocesses using the exact sequence of procedural operations. This definition yields numerous sets of restrictions on procedural constructions per restriction on processes.

The second premise is easier. Epistemic norms that take processes as objects straightforwardly define restrictions on processes. Moreover, the generation schemata from the D1 and D2 sections show how restrictions on processes can be generated from every restriction on states or transitions. Thus, the schemata show how every norm that restricts any epistemic factor is associated with a set of restrictions on processes. This point is strengthened in Chapter 4, where it is shown that for every restriction on states or transitions, there is a restriction on processes from which the state and transition restriction can be read off. From this fact, and the first premise, every norm can be seen to be associated with a set of restrictions on procedural construction.

Epistemic Regulation and Planning Finally, the D3 corollary has independent support from consideration of ante hoc regulation and epistemic planning in light of the examples of norms that restrict procedural construction. First, restrictions on procedural construction are sometimes necessary for ante hoc regulation. In cases where epistemic outcomes depend on procedures being constructed in certain ways (as in inductive rules, proof and model construction procedures, numerous other formal methods involving algorithms for the construction and manipulation of models, and de-biasing strategies), norms requiring those epistemic outcomes derivatively require those procedural constructions. For instance, whether a semantic tableaux procedure gives a correct answer about the validity of an argument depends on how one orders one's choice of rules to apply. The requirement that one get right answers entails a requirement that one use procedures constructed with certain sequential compositions of rule applications and not others. This point applies to nearly all of the examples in the table of procedural requirements above.

Second, connecting the foregoing to epistemic planning further strengthens the inference to the D3 corollary.

1. Ante hoc regulation requires epistemic planning.
2. Solutions to planning problems are restrictions on procedural construction.
3. ∴ Ante hoc regulation requires restrictions on procedural construction.

Ante hoc regulation is the process of solving epistemic planning problems. The solution to an epistemic planning problem is a set of plans that each suffice to achieve the goal state determined by some set of epistemic norms. A set of plans that achieve a goal state just is a restriction on procedural construction because plans are themselves sets of instructions built out of procedural operations of conditional choice, sequencing, etc. over a set of elements of the action library. Thus, insofar as ante hoc regulation requires solutions to epistemic planning problems, ante hoc regulation requires that there be restrictions on procedural construction. These restrictions will be specified by the norms that define the goal state. Typically, only a subset of possible plans constructible from the action library will reach the goal state. As noted above, this is often due to the way the plans are constructed, the ordering of the subtasks in the plan, the proper conditionalizing of choices in the plan, and so on. So, the goal state, in concert with the features of the constructible plans, divides the set of procedures into those that are effective for reaching the goal state and those not. Insofar as the goal state is defined by the requirements of some set of norms, those norms indirectly restrict procedural constructions.

2.4.3 Discussion

In arguing for D3 and its corollary, two of the motivating questions of this chapter were given tentative answers. First, one kind of information that epistemic rationality norms carry was identified as restrictions on epistemic dynamics. Certain kinds of restrictions are equivalent, in many respects, to programs. That is, these kinds of restrictions can, when put in the proper framework (i.e., thinking of epistemic dynamics with LTSs), be seen to be equivalent to rules or functions from conditions or states to instructions to perform an action. Given the insights from action logic and deontic logic [261, 174, 175, 89] that link imperatives and deontic operators to programs, it's frankly unclear how to answer the motivating question about what information norms carry. It seems quite clear, though, that norms carry some program-like information, and that this information can be translated back and forth from imperatives to procedural and imperatival claims.

Second, D3 found support in a certain conception of the cognitive role of epistemic norms that requires more than mere post hoc evaluation. The view defended was essentially a version of the idea that, whatever else epistemic norms do, they are to play a role in regulating epistemic cognition (as it is shaped by voluntary epistemic actions). Else, they are without any serious use or value..

Which is all to say that I side with the idea that epistemology is in the business of offering epistemic advice, and not merely describing regularities about epistemic practice. Of course, serious descriptions of epistemic practice are frequently exactly what is needed for devising epistemic advice. See means-ends epistemology [220], its formal output in formal learning theory [221, 183, 142, 140], or the "technological" side of naturalized epistemology [204] for interesting insights on this note. The basic thought is that, given regularities among (e.g.) the structure of the processes we use to generate our best models of the world and the features of those models, guidelines that references those regularities can be written for individual practice or performance improvement.

All of the foregoing is compatible with certain forms of normative anti-realism, like

Gibbard's plan [98] or norm [97] expressivism or Field's epistemic relativist-expressivism [77, 78, 79] as well as the forms of realism that act as a foil to either of these views. More precisely, these forms of normative anti-realism commit to at least the two theses that: (i) there is no uniquely best or uniquely correct set of epistemic norms, and (ii) epistemic norms, whatever they are, are not objective and mind-independent. There is really no reason - implicit or explicit in the account of epistemic rationality and its norms thus far developed - to rule out either (i) and (ii) or their negations.

There isn't sufficient space to explore this suite of questions. It's worth nothing, however, that the answers to the two motivation questions are highly complementary. Restrictions on epistemic dynamics are exactly the kind of information of information structure that is useful from the first-person perspective of epistemic regulation and epistemic planning. Some models of epistemic regulation require just the kind of information supplied by restrictions on epistemic dynamics. So, there is a sort of mutual reinforcement. That restrictions on epistemic dynamics are associated to epistemic norms makes it plausible that epistemic norms are useful in ante hoc regulation and epistemic planning. That regulation and planning are essential functions of epistemic norms makes it plausible that epistemic norms are associated with restrictions on epistemic dynamics.

Finally, D1 and D2 supply some clarifying information for D3; epistemic norms must specify restrictions on transitions and processes as well as on states. And this effects a bit of a simplification on the explanation of the Processes and Constraints problems. The thought is simply that, without tools to represent processes as distinct from transitions, the set of restrictions on processes cannot be distinct from the set of restrictions on transitions.

2.5 Conclusion

In this chapter, the second half of a core or minimal account of epistemic rationality and its norms was given; the first half was given in the form of the must-solve Processes and Constraints problems (and the knock-on problems they raise) for models of epistemic dynamics. The second half of the theory takes the form of principles D1-D3, a set of observations about the objects of epistemic rationality, the kinds of information that its norms convey, and the function of epistemic norms in epistemic cognition. As parts of a minimal theory, these observations were argued to characterize desiderata or even criteria of adequacy on accounts (esp. formal frameworks) of epistemic rationality. The following three chapters will argue that the protocol-theoretic account of epistemic rationality norms (via PLEN) uniquely satisfies these criteria.

Part II

PLEN: the Protocol-theoretic Logic of Epistemic Norms

Chapter 3

The Definition and Interpretation of PLEN

Introduction: PLEN

This chapter introduces PLEN, the protocol-theoretic logic of epistemic norms. PLEN is a formal framework for representing and reasoning about epistemic rationality norms based on propositional dynamic logic. Its primary philosophical innovation is the deployment of *protocols* to analyze procedural epistemic rationality norms, and its primary technical innovation is the deployment of paths and graphs to the interpretation of protocol expressions.

As a framework, PLEN isn't a complete or final model of epistemic norms. Rather, PLEN is a basic formalism that is defensible (re: solving the Framework Problem) and compatible with a wide array of modifications. Diverse theories, tailored to specific theoretical interests or commitments, can be built up from the basic formalism by adding restrictions or revisions to its semantic structures, extending its languages, or modifying its logic. These extensions of PLEN can then be studied and applied.[1]

---

[1] The developmental approach of PLEN is modeled after Logical Dynamics, an interdisciplinary research programme aimed at studying rational agency via logical tools and methods. The approach is focused on building a general technical framework that accommodates many different formalisms for thinking about the target systems rather than providing a specific theory to defend come what may. Dynamic logics are,

The core of the PLEN framework consists of (i) a jointly defined pair of languages; a syntactic system that represents procedural operations out of which protocols are composed and a modal propositional language encoding assertions about protocols, (ii) a semantics that simultaneous interprets both languages via abstract models of epistemic dynamics, and (iii) derivation systems that codify valid rules for reasoning about various aspects of protocols, including equivalence, execution, and the effects of execution. Informally, these things are, respectively, taken to represent the procedural contents of epistemic rationality norms and the fragment of natural language that embeds propositions about them, a model of epistemic dynamics that interprets those contents in terms of restrictions on dynamics, and a logic of epistemic rationality norms and propositions about them.

The plan for the chapter is straightforward. First, the languages and semantics of the PLEN framework will be defined. They are then discussed in the order in which they were introduced. Second, a derivation system for PLEN is defined, briefly explored, and shown to be sound. Third, the interpretation of PLEN as a formalism for representing epistemic rationality norms is explicitly outlined. The general theme of this part of the chapter is laying the groundwork for showing that PLEN is a philosophically defensible answer to the Framework Problem.

---

of course, central to the toolkit of Logical Dynamics, but they are just one kind of formalism deployed in the field. The programme is organized around a core, abstract model of dynamics that is perhaps best represented with labeled transition systems. Distinct formalisms, diverse in their expressive and logical properties and tailored to specific theoretical interests or commitments, are built around these core models, studied, and applied. One can, for instance, study classes of labeled transition systems with the languages of PDL, modal logics, deontic logics, temporal logics, or with fragments of first-order logic. Such study has unearthed interesting technical results such as, for instance, the fact that standard PDL can be translated into the two-variable fragment of FOL [30]. See [28, 30, 34] for the core statement of the programme and numerous important developments. See also [139, 114, 30, 74, 131] for seminal contributions, and [113] for a comprehensive textbook presentation on dynamic logic.

3.1 PLEN: The Formal Framework

PLEN has the structure of a logic.[2] It's fundamentally a triple of a syntax, semantics, and proof or derivation system. The syntactic system contains a language of *protocol expressions,* P_{PLEN}, and a language of propositional and modal, dynamic, and epistemic formulae, L_{PLEN}. P_{PLEN} defines the compositional structure of protocols, which encodes instructions for (inter alia) combining, iterating, sequencing, and choosing among basic epistemic actions and procedures. That is, P_{PLEN} codifies the procedural operations out of which complex procedures are formed out of a repertoire of basic actions [30]. The protocol expressions represent the procedural content of epistemic rationality norms - the instructions for correct epistemic action that those norms dictate. L_{PLEN} encodes a fragment of natural language - the fragment that expresses a partial theory of epistemic norms or at least, a partial theory of the procedural information that they carry. L_{PLEN} deploys a range of *dynamic epistemic operators* and *protocol connectives* to embed protocol expressions into propositional and modal expressions. The resulting dynamic epistemic expressions encode statements about the procedural content of norms, about the conditions of application or preconditions of norms, and about the outcomes, effects, or postconditions of complying with norms on the content of an agent's epistemic states as well as on the world.

The semantics is defined over the set of *epistemic directed graphs* (EDGs). EDGs are structures for modeling the possible epistemic trajectories or paths of particular epistemic agents through a space of possible epistemic states. They are basically labeled directed multigraphs, allowing multiple, distinct edges between the same pair of states or nodes. EDGs form the basis of interpretations for the syntax. Each protocol expression is associated with *execution conditions* that determine sets of paths through epistemic state space for each EDG, and, given that EDGs can be used as models of the dynamics of agents, each protocol is associated with particular restrictions on the dynamics of agents. The expressions of

---

[2] More precisely, PLEN has the structure that basically follows the blueprint of Propositional Dynamic Logic (PDL) [113, 30] but with elaborations that resemble the path models of Process Logic (PL) [114] and the basic protocols of Epistemic Protocol Logic (EPL) [184].

L_{PLEN} are given truth conditions that are jointly defined with the execution conditions for protocols.

The plan for this section is to define the formalism first, with relatively minimal discussion, and then explore it informally. First, the syntax is defined and very briefly discussed. Then, the semantics is defined and briefly discussed. These subsections are followed by a more elaborate discussion of the formalism and the philosophical interpretation of the formalism.

3.1.1 Syntax I: Simultaneously Defined Languages

3.1.1.1 Protocol Language

First, define the language of protocols, P_{PLEN}.

Definition: Protocol Language P_{PLEN} The expressions of P_{PLEN} are *protocols* or *protocol expressions*. Let $\Pi_0 = \{a_1, ..., a_n\}$ be a finite set of *basic/atomic procedures*, *basic/atomic actions*, or *atomic protocol expressions*. Let $Prot = \{;, \cap, \cup, *\}$ and $Prot{+} = \{(-)?, -, \curvearrowleft, \nabla\}$ be distinct sets of *control operators*. Let A be a metavariable for an expression of L_{PLEN}, which is defined simultaneously with P_{PLEN}. Let π with subscripts be metavariables for protocols; I will omit subscripts when unnecessary. Define P_{PLEN} in Backus-Naur form:

$$\Pi_0 \mid \pi_1; \pi_2 \mid \pi_1 \cap \pi_2 \mid \pi_1 \cup \pi_2 \mid \pi_i * \mid (A)? \mid -\pi_i \mid \curvearrowleft \pi_i \mid \nabla \pi_i$$

Let Π_C be the set of all *complex protocol expressions* formed by the syntax above. Then $P_{PLEN} = \Pi_0 \cup \Pi_C$. Complex protocol expressions correspond to natural ways of building complex procedures out of basic actions. *Prot* is the set of *regular control operators*. These take protocol expressions and generate complex protocols. A protocol formed only with regular control operators is a *regular protocol*. *Prot*+ is the set of *nonstandard control operators*. The nonstandard control operators generate complex protocols that, by the execution rules defined below, identify executions of their subprotocols (the elements of P_{PLEN} that are composed inside a protocol expression using control operators) or entirely distinct protocols on the basis of graph-theoretic features of EDGs. This will become clearer in discussing the execution rules.

The expression $\pi_1;\pi_2$ is the *sequential composition* of π_1 and π_2; the procedure that requires doing π_1 and then immediately doing π_2 from the state reached by π_1. The expression $\pi_1 \cap \pi_2$ is the *parallel execution* of π_1 and π_2; the procedure that requires doing π_1 while doing π_2. The expression $\pi_1 \cup \pi_2$ is the *indeterministic choice* among π_1 and π_2; the procedure that requires indeterministic selection from among π_1 and π_2 (including the execution of both). The expression $\pi *$ is the *iteration* of π any number of times; the procedure that requires doing π any n times sequentially.

The expression $(A)?$ is a test for A; a unique protocol that checks the current state for the truth of A. The expression $-\pi$ is the *complement* of π; the procedure comprising execution of the set of protocols other than π incompatible with doing π. The expression $\curvearrowleft \pi$ is the *reversal* of π; the procedure that reverses the effects that are brought about by doing π. The expression $\nabla\pi$ is the *fixed point of* π; the version of π that makes no changes to the world or to any epistemic states.

Sequential composition, parallel execution, and choice are binary, the rest are unary, and P_{PLEN} is recursively enumerable. Complex protocols can be composed to create more complex protocols. Parentheses can be used to clarify structure. For example, the following are well-formed protocols:

- $(\pi_1; \pi_2); (\pi_3; \pi_4)$
- $(\pi_1; \pi_2 *); (\pi_3; \pi_4) *$
- $\pi_1; (\pi_2; (\pi_3; \pi_4))$
- $((\pi_1; \pi_2); (\pi_3; \pi_4)) \cup (\pi_4 \cap (\pi_6 \cup \pi_1))$
- $-((\pi_1; \pi_2); (\pi_3; \pi_4))$

- $(A)?;((\pi_1;\pi_2);(\pi_3;\pi_4))$
- $\curvearrowleft((\pi_1;\pi_2);(\pi_3;\pi_4))$
- $-(\pi_1;\pi_2);(\pi_3;\pi_4)$
- $((A)?;(\pi_1;\pi_2));(\pi_3;\pi_4)$
- $\curvearrowleft(\pi_1;\pi_2);(\pi_3;\pi_4)$

The protocols, $\pi_1;\pi_2;\pi_3;\pi_4$, $\pi_1 \cup \pi_2 \cup \pi_3$, and $\pi_1 \cap \pi_2 \cup \pi_3$ are each ambiguous, but for any complex regular protocol formed from multiple iterations of a single control operator, this won't matter. For example, $(\pi_1;\pi_2);(\pi_3;\pi_4)$ and $\pi_1;(\pi_2;(\pi_3;\pi_4))$, while syntactically distinct protocols, are equivalent (in ways to be explicated), as are $(\pi_1 \cup \pi_2) \cup (\pi_3 \cup \pi_4)$ and $\pi_1 \cup (\pi_2 \cup (\pi_3 \cup \pi_4))$, and $(\pi_1 \cap \pi_2) \cap (\pi_3 \cap \pi_4)$ and $\pi_1 \cap (\pi_2 \cap (\pi_3 \cap \pi_4))$. In general, from a semantic perspective, each of the binary control operators is associative, while parallel execution and choice are commutative, as well. Permutations on the order of constituent protocols thus make no semantic difference for choice and parallel execution, though, obviously, they do for sequential composition.

As for intuitively reading off what each protocol expression says, protocols can be read as programs, recipes, or instructions for carrying out complex procedures that are formed out of basic actions. If an atomic protocol expression, π, is an instruction of the form "Do π!", then:

- $\pi_1;\pi_2$ reads "Do π_1, then do π_2!"
- $\pi_1 \cap \pi_2$ reads "Do π_1 concurrently with π_2!" or perhaps "Do π_1 while doing π_2!"
- $\pi_1 \cup \pi_2$ reads "Do π_1 or do π_2!"
- $\pi*$ reads "For some $n \geq 0$, do π n times!"
- $(A)?$ reads "Test for A!"

- $-\pi$ reads "Do anything but π, and don't do π!"

- $\curvearrowleft \pi$ reads "Undo the changes wrought by π!"

- $\nabla\pi$ reads "Do π in a way that doesn't change anything!"

Consequently, protocols formalize programs. In particular, the most important programming and decision structures can be expressed in PLEN. Here is a small sampling:

- *Conditional Choice* If A, do π_1; else, do π_2 :: $((A)?; \pi_1) \cup ((\neg A)?; \pi_2)$

- *Generalization of Conditional Choice* If A, do π_1; if B, do π_2 :: $((A)?; \pi_1) \cup ((B)?; \pi_2)$

- *Guarded Iteration* While A, do π :: $((A)?; \pi)*; (\neg A)?$

- *Guarded Repetition* Repeat π until A :: $(\pi; (\neg A)?)*; (A)?$

This will be a point that I return to in arguing for PLEN's theoretical adequacy with respect to epistemic norms.

3.1.1.2 Propositional Language

Now, define the propositional language, L_{PLEN}.

Definition: Propositional Language L_{PLEN} Let $AForm = \{P_i, Q_i, ..., \}$, $Conn = \{\wedge, \vee, \rightarrow, \equiv, \neg\}$, $Op = \{[x], \langle x \rangle, \Box, b, d\}$, $Pop\,\{\leftrightarrow, \Yup, \Ydown, M\}$. Define L_{PLEN} in Backus-Naur form:

- $AForm \mid \pi \leftrightarrow \pi' \mid \pi \Yup \pi' \mid \pi \Ydown \pi' \mid M(\pi)$

- $[\pi]\, A \mid \langle \pi \rangle\, A \mid \Box A \mid b(A) \mid d(A)$

- $A \wedge B \mid A \vee B \mid A \rightarrow B \mid A \equiv B \mid \neg A$

The set of propositional parameters or atomic formulae is $AForm$. The connectives in $Conn$ connect formulae of L_{PLEN} in a standard, propositional fashion. The operators in Op take formulae of L_{PLEN} as arguments. The operators $[x]$ and $\langle x \rangle$ embed protocol expressions and operate on any formulae of L_{PLEN} to form expressions about protocol executability and postconditions. The expression $[\pi]A$ says that A is a *strong postcondition* of π; A is true after any way of doing π. The expression, $\langle\pi\rangle A$, says that A is a *weak postcondition* of π; A is true after some way of doing π. The special case, $\langle\pi\rangle \top$, encodes the *executability* of π. That is, $\langle\pi\rangle \top$ says that π is *executable* (at the current state); it can be carried out from the current state.

The operators $\Box, b, d$ are pretty standard modal or epistemic operators, taking any formulae of L_{PLEN}. $\Box A$ stands in for "Necessarily, A." The epistemic operators b and d stand in for *belief* and *disbelief* or acceptance and rejection; $b(A)$ reading "A is fully believed/accepted" and $d(A)$ reading "A is fully disbelieved or rejected".

The connectives in *Pop* - $\leftrightarrow$, $\curlywedge$, $\curlyvee$, and M - take only protocol expressions as arguments. The first three connectives are binary and connect protocol expressions directly to form statements about protocols. The expression, $\pi \leftrightarrow \pi'$, tells us that the protocols π and π' are *procedurally equivalent at the current state*, i.e., that, from the current state, π and π' instruct us to carry out exactly the same processes of epistemic state change. This will be a notion more fully articulated below. The expression, $\pi \curlywedge \pi'$, tells us that π and π' *converge* in their dictates at the current state; there is some process that complies with both π and π', so one can comply with or execute both at once. The expression, $\pi \curlyvee \pi'$, tells us that the protocols *diverge* in their instructions at the current state; there is process that complies with one and not the other. The formulae $(\pi \leftrightarrow \pi') \wedge [\pi]A$ and $(\pi \curlywedge \pi') \wedge \langle\pi\rangle A$, respectively, read as "π and π' result in equivalent changes, and A always results from π" and "π and π' result in equivalent changes, and there is a way of carrying out π that results in A being true".

Finally, M names a temporary *methodology*, the set of norms that an agent endorses at a given point. $M(\pi)$ expresses *methodological inclusion*; π is part of the methodology endorsed at the current state. Formally, $M(\pi)$ is unary, M names a subset of P_{PLEN}, and the expression says that π is an element of M. Methodological expressions other than those of inclusion, expressions about the postconditions or preconditions of methodologies in their entirety, can be encoded by explicitly replacing M with the nondeterministic choice or parallel execution of each protocol in M, depending on the structure of the methodology and the structure of the claim.

3.1.1.3 Remarks on PLEN's Languages

PLEN's languages encode fragments of natural language that can be used to talk about epistemic dynamics and epistemic norms from both "internal" and "external" perspectives.[3] This distinction between these perspectives isn't quite rigorous, but, roughly speaking, the internal perspective on epistemic dynamics (respectively, norms) is focused on the evolution of the subjective part of epistemic dynamics. Internal languages exclusively represent what is going on inside an agent's epistemic states and how this content changes over time. AGM, for instance, takes an internal perspective; it doesn't describe what is going on in the world represented by the content of an epistemic state, it describes how idealized models of epistemic states (ideally) evolve under specific kinds of change. There are no actions in AGM that change the non-epistemic facts in the world. A model of epistemic dynamics takes an external perspective if it represents the state of the world co-evolving with the content of epistemic states. External perspectives can tell us about what's going on outside an agent's head. Languages that contain statements about the truth or falsity of an agent's beliefs or about the actual effects on the world of the agent's actions take external perspectives.

P_{PLEN} encodes external information about epistemic norms and epistemic dynamics. As remarked on above, the syntactic form of protocol expressions encodes the procedural information of epistemic norms. This procedural information specifies (i) the procedural commitments that an agent makes in employing a norm or complying with it, and (ii) the processes and sequences of actions that count as complying with a norm.

---

[3] This distinction is radically different from that between exogenous and endogenous logics of programs or actions. The latter distinction is strictly defined: if programs are explicit in the language of a logic, then it's exogenous. If programs are not explicit in the language, then it's endogenous.

L_{PLEN} encodes both internal and external information about norms and epistemic dynamics. On the internal side, the epistemic operators, b and d, describe the content of an epistemic state, and these scope over any other formulae of L_{PLEN}. Importantly, L_{PLEN} introduces no indices for epistemic operators - it's a single-agent system. The formulae of L_{PLEN} can only describe the internal states of a single agent. On the external side lie formulae that are not embedded in the scope of a b or d. Some of these may describe changes to the external world that are wrought by courses of epistemic action. Weak postconditions - formulae of the form $\langle\pi\rangle A$ - state the effects of complying with epistemic norms on the content of epistemic states and on the external world. Preconditions can be stated in the form of conditionals, $A \rightarrow \langle\pi\rangle\top$, $\langle\pi\rangle\top \rightarrow A$, that state conditions under which norms can be complied with. The various "protocol operators" that form complex expressions out of protocol expressions are strictly external. They describe structural and procedural relations between protocols. For instance, $\pi_1 \leftrightarrow \pi_2$ tells us that these two protocols can be carried out in exactly the same ways, and formulae of the form $[\pi_1]\langle\pi_2\rangle\top$ and $[\pi_1]\neg\langle\pi_2\rangle\top$ tell us, respectively that π_1 enables or disables π_2.

L_{PLEN} formulae can freely combine internal and external information. The expression $\langle\pi\rangle(A \wedge b(B))$ says that π can be carried out in such a way that A is true and B is believed. Given that contents of epistemic states can be expressed in L_{PLEN}, preconditions can also state what kinds of things an agent must know about the world or about the norms themselves in order to carry them out. The contents of an agent's beliefs about the preconditions, postconditions, equivalence, convergence, and divergence relations among, and even execution conditions of her own norms can be encoded by embedding these formulae of L_{PLEN} in expressions over which b and d have scope.

Importantly, the internal and the external in L_{PLEN}, as in life, can come apart. L_{PLEN} can describe the fact that an agent mistakenly thinks she can comply with a norm under some conditions; $b(\langle\pi\rangle\top)$ is not equivalent to $\langle\pi\rangle\top$. More, agents can, after executing protocols, also come to learn things about protocols: $(\pi_1 \leftrightarrow \pi_2) \wedge \langle\pi_1\rangle(b(\pi_1 \leftrightarrow \pi_2))$. Intuitively, this reads "Protocols π_1 and π_2 are equivalent and doing π_1 can result in believing this fact." Protocols can function as "hard information events", revealing information to agents infallibly: $[\pi]b(A)$. They can be softer, revealing less reliable information or less reliably revealing information to agents: $\langle\pi\rangle\, b(A \vee B \vee ...)$ or $\langle\pi\rangle\, b(A)$. They can also be obfuscating events, obscuring ($[\pi]\neg b(A)$) or preventing access to information (given a standardly defined dual to $\Box$, $\Diamond$, then $[\pi]\neg\Diamond b(A)$), or enabling or disabling possible informational actions: $[\pi]\neg\langle(A)?\rangle\top$, $[\pi_1]b(A) \rightarrow ([\pi_2]\neg\langle\pi_1\rangle\top)$, $[\pi_1]\langle\pi_2\rangle\, b(A)$. These well-known and important dynamic epistemic actions can thus be formalized in ways that mirror their familiar formalization (re: DEL [65]).

3.1.2 Structures: Epistemic Directed Graphs

The semantics of P_{PLEN} and L_{PLEN} are defined by simultaneous definition of execution rules for protocol expressions and truth conditions for propositional expressions. The key objects are graph-like structures called *epistemic directed graphs* (EDGs). These function in PLEN's semantics as Kripke Frames do in modal logics, labeled transition systems do in PDL, or path models do in Process Logic (PL). Epistemic directed graphs represent the possible epistemic states of a reasoner at a time along with all of the possible changes and trajectories of change of epistemic state that the reasoner's repertoire of actions, procedures, and processes could take her through. It is in relation to these that protocols, as objects representing the procedural information given by epistemic norms, are interpreted.

The class of EDGs is a class of structures isomorphic to labeled directed multi-graphs. EDGs are graphs, which is to say that they are sets of nodes connected by edges or n-ary relations. In a directed graph, the relations are ordered. Letting s_1 and s_2 be nodes, if an edge, e, links them, the idea is that you may be able to get from s_1 to s_2 but not from s_2 to s_1. The labeling tells you what action for moving from s_1 to s_2 the edge represents. Normally, edges in graphs are defined by their endpoints, the nodes they connect. In a multigraph, multiple edges may link the same pair of nodes. This is important for PLEN's philosophical prospects, especially regarding D1 through D3.

The correlate of semantic conditions for protocol expressions will be *execution conditions*. This chapter will introduce *procedural execution conditions* in order to interpret P_{PLEN} and L_{PLEN}. Execution conditions will be defined in terms of edges and paths. Edges are sets of ordered triples, $\langle s, \pi, s' \rangle$ where the middle term is the label. Paths are sequences of concatenated edges. For the purposes of PLEN, the middle terms of edges will be protocol expressions. The semantics of PLEN will correlate edges with each atomic protocol and n-ary relations or *paths* with each complex protocol.

Each EDG will act as a model or an interpretation for expressions of P_{PLEN}, on the basis of which truth conditions can be defined for expressions of L_{PLEN}. Before defining these conditions, I define the constituent parts of EDGs. As in all epistemological theories [91, 86], PLEN analyzes *epistemic states* and *epistemic dynamics*. Structures representing these are the essential components of EDGs.

3.1.2.1 Epistemic States

First, epistemic states. PLEN requires machinery to represent epistemic states, their content, and their relations to the world, and, eventually to epistemic dynamics.

Definition: Epistemic States Let W be a finite, non-empty set, $W = \{w_1, ..., w_n\}$. Each element of W is an *epistemic state*.

The most important part of an epistemic state is its content: the epistemic judgments an agent makes and the propositions or information toward which those attitudes are held. Formally:

Content States Define a set of ordered triples, $C \subseteq P(L_{PLEN}) \times P(L_{PLEN}) \times P(P_{PLEN})$. Each element of C is a *content state*.

Call the first term of a given such triple, X, the second term of the triple, Y, and call the third term Π_M. The *acceptance set*, X, is a syntactic representation of the statements of L_{PLEN} that an epistemic agent accepts or (fully) believes, these terms being used, for now, interchangeably. The *rejection set*, Y, is a syntactic representation of the statements of L_{PLEN} that an agent rejects or (fully) disbelieves. The intent of Π_M is to encode a *methodology*, the temporary epistemic rationality norms that an agent endorses or employs.

For each epistemic state, there is corresponding content. Define a function from W into C to assign content to each epistemic state. Formally:

Content Function Let all functions, $c : W \longrightarrow C$, be *content functions.*

Call $c(w)$ the content state of w. For brevity, epistemic states will be denoted with w where it isn't necessary to note $c(w)$. In discussing the graph-theoretic properties of EDGs, for instance, content states will frequently drop out of the picture.

To flesh out PLEN's external approach to epistemic dynamics, add valuations for each epistemic state:

Valuations Define *valuations* $v_w(P) : W \times AForm \longrightarrow \{0, 1\}$ as functions from states and propositional parameters (or atomic formulae) of L_{PLEN} to the values 0 and 1, with $v_w(\top) = 1$ for all $w \in W$.

Valuations are standard. Each pair of a valuation and an epistemic state combines, at a time, a representation, $c(w)$, of the key parts of an agent's epistemic judgments and methodology with a (basic) representation of the state of the world, v_w. Valuations are clearly useful for deploying EDGs as models of the languages of PLEN; we need some way to interpret the atomic sentences of L_{PLEN}. Moreover, they enable the representation of the co-evolution of epistemic state and state of the world. An epistemic action at w may transition a system to the epistemic state w' in a way that changes one, both, or neither the content, $c(w) = c(w')$, nor the non-epistemic state of the external world, $v_w = v_{w'}$.

3.1.2.2 Epistemic Dynamics

Second, epistemic dynamics. Actions, processes, and procedures can change epistemic states and states of the world. PLEN requires machinery to represent this. We start with two basic parts.

Transitions Each ordered pair of epistemic states, $\langle w, w' \rangle$, is a *transition.*

Edges Ordered triples, $\langle w, \pi, w' \rangle$, are *labeled transitions* or *(directed) edges.*

Sets of transitions are binary relations. Sets of edges are *labeled binary relations* or ternary relations. Note that each edge, e, is a triple, $\langle w, \pi, w' \rangle$. There is thus a first state, $first(e)$, and a last state, $last(e)$. Call $first(e)$ the *root state* of e and $last(e)$ the *terminal state* of e. Just so for transitions.

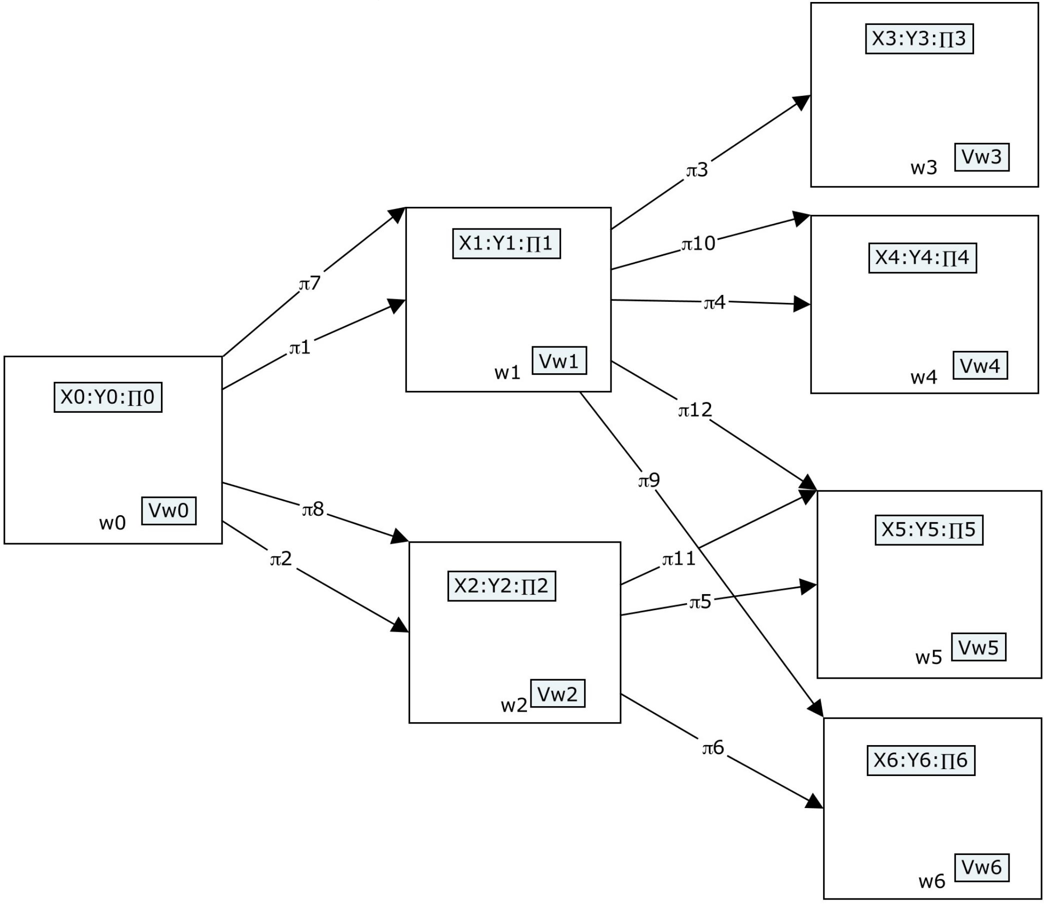
Transitions are possible changes of epistemic state that can be effected by actions, procedures, or processes. Edges link transitions with basic protocols; they represent the possible changes of both the world (the truths delivered by valuations) and epistemic states (content states) by execution of particular actions, procedures, or processes. In particular, a procedure or action π may take an agent from w to w'. Actions or procedures may be indeterministic: from w, π may take an epistemic agent to w' or any number of other states, w''.

These pieces all come together to form *epistemic directed graphs:*

Definition: Epistemic Directed Graphs (EDGs) Let W be a set of epistemic states. Let C be a set of triples of the form $\langle X, Y, \Pi_M \rangle$. Let c and v be a content function and a valuation function, respectively. Let $\Pi_g \subseteq \Pi_0$. For each $\pi \in \Pi_g$, map a set of edges, R_π, to it: $R_\pi \subseteq W \times \{\pi\} \times W$. Define the *atomic edge set* R_0 as the union of each set of edges mapped to each π in Π_g: $R = R_{\pi 1} \cup ... \cup R_{\pi n}$ for each $\pi_i \in \Pi_g$. The *atomic action set* in g is the set $R = \{R_\pi \mid \exists \pi(\pi \in \Pi_0 \wedge R_\pi \subseteq W \times \{\pi\} \times W)\}$. Then: *g is an EDG iff* $g = \langle W, C, \Pi_g, R, R_0, c, v \rangle$.

Occasionally, there will be the need to refer to graphs and parts of graphs by name. Subscripted names will pick out objects in relation to specific graphs, e.g., R_g, v_g. When we are

Figure 3.1: An Epistemic Directed Graph



dealing with two EDGs, g and g', rather than subscripts, the graphs and their parts may be "primed", e.g., W and W', R and R'. Generally, conventions that pick apart different graphs will be self-explanatory. Alternately, dealing with graphs g_1 and g_2, the subscripts directly transfer: W_1 and W_2 will be the obvious sets of epistemic states, for instance. Where there is no need for such devices, objects are defined with respect to some arbitrary graph, g, and the labels for g drops out of consideration. Figure 3.1 presents a simple EDG.

The languages P_{PLEN} and L_{PLEN} coupled with the set of EDGs form the basic formal framework of PLEN. This formal framework can generate different logics by using EDGs as the basis for semantic interpretations of L_{PLEN}. But interpreting L_{PLEN} requires interpreting protocols. Protocols are not standard semantic items; like imperatives, they require execution conditions rather than truth conditions, formal definitions of conditions under which the protocols are executed, carried out, or, in normative terms, complied with. There are different strategies for interpreting protocols. In defining the logic of PLEN, EDGs are deployed as path models [114]. The next section defines execution rules for protocols and the truth rules for propositions of L_{PLEN} by joint recursion.

3.1.3 Semantics: Procedural Execution Rules for P_{PLEN}

Execution of protocols will be the formal equivalent of following or complying with the dictates of epistemic norms. Execution conditions defined recursively for protocol expressions in P_{PLEN} function analogously to truth conditions for the expressions of L_{PLEN}. EDGs function as path models in PLEN. That is, they interpret protocol expressions and the operators of L_{PLEN} in terms of *procedural paths* - concatenated sequences of edges $\langle e_1, ..., e_n \rangle$ or, equivalently, alternating sequences of state and protocol expressions, e.g., $\langle s_1, \pi_1, ..., \pi_{n-1}, s_n \rangle$. These paths are those that follow or comply with the protocols to which they are assigned.[4]

EDGs can be deployed as path models after the definition of several useful preliminary devices. These include *procedural paths*, *state-sequences, correspondence* of paths, and *inverting functions.* Say that two edges, $e = \langle w, \pi, w' \rangle$ and $e' = \langle w'', \pi, w''' \rangle$, are parts of the same graph, g. Then e and e' are *concatenated* iff $w' = w''$. That is, if $w' = w''$, then there is a path in g, $ee' = \langle w, \pi, w', \pi, w''' \rangle$. In other words, e and e' are concatenated iff $last(e) = first(e')$. Of course, this entails that $v(w') = v(w'')$ and $c(w') = c(w'')$. Then define two kinds of paths on the basis of concatenation of edges:

Procedural Paths Finite sequences of edges, $\langle e_1, ..., e_n \rangle$, such that, for all e_i where $i < n$,

---

[4] As noted in earlier chapters, spelled out below with respect to solving the Processes and Constraints problems, this is merely one possible way to use EDGs to interpret protocol expressions. Alternatives can be defined in terms of binary relations or ternary relations. The decision to use paths is not justified in full detail in this dissertation. See [135] for details. The basic thought is that, in order to define execution conditions that make good philosophical sense, paths must be defined and employed in the definition of execution rules. Else, execution conditions for parallel execution are exceptionally hard to define without introducing EDGs with odd features that clash with the goal of parsimoniously representing epistemic dynamics. More, introducing paths and deploying EDGs as path models [114] for PLEN enables definitions of powerful equivalence relations among protocols, and this plays a role in PLEN's theoretical adequacy with respect to D1-D3.

$last(e_i) = first(e_{i+1})$ are *procedural paths*. Let P_g be the set of all procedural paths in an EDG, g.

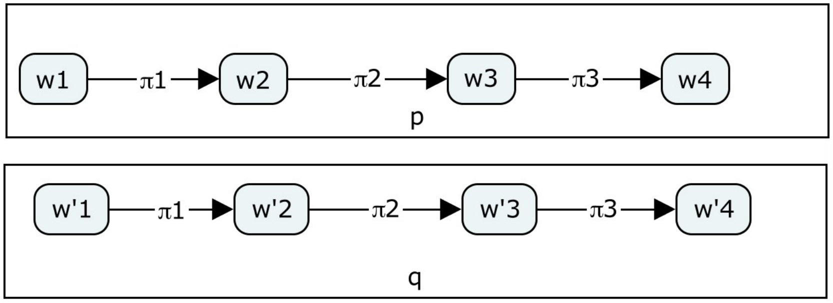
State-Sequences Finite sequences of transitions, $\langle t_1, ..., t_n \rangle$, such that, for all t_i where $i < n$, $last(t_i) = first(t_{i+1})$ are *state-sequences*. Let S_g* be the set of all state sequences in and EDG, g.

Denote procedural paths (hereon, "paths") with p, q, r, subscripted, as necessary. Equivalently, procedural paths are linearly ordered sets of edges s.t.: if $i < n$, $last(e_i) = first(e_{i+1})$. Thus, edges are simply the smallest paths in any graph. The length of a path p is the number of edges it contains: $length(p) = |\{e_i \mid e_i \in p\}|$. Denote state-sequences with $seq(p)$, $seq(q)$, etc., with $p, q, ...$ subscripted as necessary. This notation is suggestive for a reason that will be fleshed out below.

Paths are finite sequences of edges, and more complex paths can be formed by concatenations of paths. Each path, p, has, in virtue of being finite, a root state, $first(p)$ and a terminal state $last(p)$. That is, $first(p)$ is the first state of the first edge in any finite sequence of edges, p, and $last(p)$ is the last state of the last edge in p. Paths q and r are *concatenated* iff $last(q) = first(r)$. Thus, we can denote longer paths as concatenations of shorter paths. Say that $p = qr$ iff p is the concatenation of q and r. Letting the length of a path be the number of edges in it, if the length of q is n and the length of r is m, then the length of qr is $n + m$. A special case of finite path is a path that contains loops; p is a *looping path* iff p is such that, for some w_i, w_i appears more than once in the sequence of states - $p = \langle w_1, \pi_1, ..., w_n, ..., w_n, ..., \pi_m, w_m \rangle$.

An integrally useful notion is that of subpaths. Simply put, p is a subpath of q, $p \leq q$, iff $p = \langle e_1, ..., e_n \rangle$ and $q = \langle e_j, ..., e_k \rangle$ where $j \geq 1$ and $k \leq n$. Naturally, if $p = qr$, then both $q \leq p$ and $r \leq p$. Each edge composing a procedural path is a subpath of it. The subpath relation is just a special case of the subgraph relation discussed below. It's easy enough to see: each path, p, is a non-branching graph. Thus, each subpath, q_i, is just a subgraph of p because the set of nodes in q_i is a subset of the nodes in p and the set of edges in q_i is a

Figure 3.2: Procedural Paths and State Sequences



subset of edges in p. Using subpaths, looping paths can be defined as those paths such that some subpath has identical root and terminal states. Note that if there is even one looping path in a graph, then there is a countably infinite sequence, $\langle p_1, ...\rangle$, of looping paths such that for each p_i and p_{i+1}, p_i is a subpath of p_{i+1}.

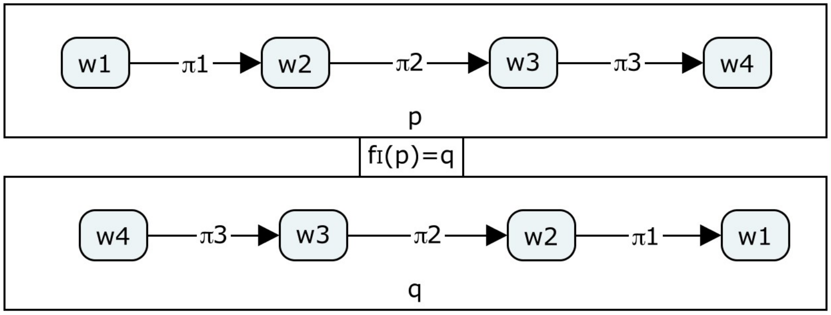
The edges of each procedural path can be mapped to any segment of the natural numbers. The resulting linear order on the edges induces an order on the states: $first(e_{i+1}) = last(e_i)$ for each e_i and e_{i+1} in any edge. Consider two paths of identical length that execute the same protocols, $p = \langle e_1...e_n\rangle$ and $q = \langle e'_1...e'_n\rangle$. There is an ordering of the states in each path, as shown in Figure 3.2. The sequencing of the edges in each path induces a sequencing of the states in each path. Thus, there is a state-sequence, $\langle w_1, w_2, w_3, w_2\rangle$, in the procedural path $p = \langle w_1, \pi_1, w_2, \pi_2, w_3, \pi_3, w_4\rangle$. Call this sequence of states $seq(p)$. Formally: $seq(p) = \langle w_1, ..., w_n\rangle$ iff (i) $p = \langle e_1...e_n\rangle$, and for all w_i, w_{i+1} in $seq(p)$, (ii) if $w_i = first(e_i)$, then $w_{i+1} = last(e_i)$, (iii) if $w_i = last(e_i)$, then $w_i = first(e_{i+1})$ and $w_{i+1} = last(e_{i+1})$. State-sequences will be useful tools for analyzing parallel execution. Note that paths may loop; for instance, in $p = \langle w_1, \pi_1, w_2, \pi_2, w_3, \pi_3, w_4\rangle$, it may be the case that $w_2 = w_4$.

An important concept for defining executions for parallel execution protocols is the notion of correspondence of paths.

Correspondence Let $p = \langle e_1, ..., e_n\rangle$ and $q = \langle e'_1, ..., e'_m\rangle$. Path p *corresponds* with q, ($p \rightleftharpoons q$), iff $n = m$, and $\forall i \leq n(first(e_i) = first(e'_i) \wedge last(e_i) = last(e'_i))$.

Let $Corr(p)$ be the paths in g that correspond to p. Abuse notation a bit to let $p \rightleftharpoons q \rightleftharpoons r$

Figure 3.3: Inverting Functions



stand for the conjunction of $p \rightleftharpoons q$ and $q \rightleftharpoons r$. Transitions and state-sequences can also correspond with procedural paths. A state-sequence, *seq*(p) corresponds to a procedural path, p, iff *seq*(p) is what results from removing all of the protocol labels from the edges that compose p. Which is just to say that the definition of correspondence applies just as well to state-sequences. A transition corresponds with a procedural path p or state-sequence *seq*(p) iff the transition is the pair consisting of the first and terminal states of p and *seq*(p), respectively.

To define the reversal control operator in a philosophically useful way, we will need *inverting functions*:

Inverting Functions Let $p \in P_g$. Let $\overline{e_i} = \langle w, \pi, w' \rangle$ iff $e_i = \langle w', \pi, w \rangle$. Define an *inverting function* $f_I : P_g \longrightarrow P_g$ such that $f_I(p) = \langle e_1, ..., e_n \rangle$ iff $p = \langle \overline{e_n}, ..., \overline{e_1} \rangle$.

The inverting functions output the reverse sequence of edges of their input path. See Figure 3.3 for an illustration. With the foregoing machinery, *procedural execution conditions* can be defined for each complex protocol for each EDG.

Definition: Procedural Execution Rules Assign *procedural execution sets*, P_π, to each complex protocol, π by the following rules:

(Atomics) $P_\pi = R_\pi$ for any π in Π_0.

(;) $P_{\pi 1;\pi 2} = \{p \mid p = qr \wedge q \in P_{\pi 1} \wedge r \in P_{\pi 2}\}$

$$(*)\ \ P_{\pi *} = \left\{ p \mid \left[\begin{array}{l} \exists n \geq 1 \\ \left[\begin{array}{l} \exists \langle x_1 ... x_n \rangle \\ \left[\begin{array}{l} \forall_{i \leq n}(x_i \in W), first(p) = x_1, last(p) = x_n, \\ \forall p', \forall_{i \leq n-1}[(x_i = first(p'), x_{i+1} = last(p')) \rightarrow p' \in P_\pi] \end{array} \right] \end{array} \right] \end{array} \right] \right\}$$

$$(\cap)\ \ P_{\pi 1 \cap \pi 2} = \left\{ p \mid \left[\begin{array}{l} p \in P_{\pi 1} \vee p \in P_{\pi 2}, \\ p \in P_{\pi 1} \rightarrow \left[\begin{array}{l} \exists q \in P_{\pi 2}, \\ p \rightleftharpoons q \end{array} \right], \\ p \in P_{\pi 2} \rightarrow \left[\begin{array}{l} \exists q \in P_{\pi 1}, \\ p \rightleftharpoons q \end{array} \right] \end{array} \right] \right\}$$

$$(\cup)\ \ P_{\pi 1 \cup \pi 2} = \{p \mid p \in P_{\pi 1} \vee p \in P_{\pi 2}\}$$

$$((\text{-})?)\ \ P_{(A)?} = \left\{ p \mid \left[\begin{array}{l} first(p) = last(p) \\ last(p) \Vdash A \end{array} \right] \right\}$$

$$(-)\ \ P_{-\pi} = \left\{ p \mid \left[p \notin P_\pi, \right] \right\} = Path(g)/P_\pi$$

$$(\curvearrowleft)\ \ P_{\curvearrowleft \pi} = \left\{ p \mid \left[\exists q \left[\begin{array}{l} q \in P_\pi, \\ p = f_I(q) \end{array} \right] \right] \right\}$$

$$(\nabla)\ \ P_{\nabla \pi} = \left\{ p \mid \left[\begin{array}{l} p \in P_\pi, \\ first(p) = last(p) \end{array} \right] \right\}$$

Say that a protocol π can be *procedurally executed* in g iff P_π is non-empty. Call P_π the *procedural executions of* π *in* g. Now, the truth conditions can be defined for L_{PLEN}.

3.1.4 Semantics: Procedural Truth Rules

Truth rules for the expressions of L_{PLEN} can now be defined on the basis of the execution rules.

Definition: Procedural Truth Rules (Base) $w \Vdash A$ iff $v_w(A) = 1$.

(Con) The standard rules for the truth-functions, $\vee$, $\wedge$, $\neg$, $\rightarrow$.

($\Box$) $w \Vdash \Box A$ iff $\forall w' \in W(w \Vdash A)$.

(b) $w \Vdash b(A)$ iff $c(w) = \langle X, Y, \Pi_M \rangle$, and $A \in X$.

(d) $w \Vdash d(A)$ iff $c(w) = \langle X, Y, \Pi_M \rangle$, and $A \in Y$.

($\langle \pi \rangle A$) $w \Vdash \langle \pi \rangle A$ iff $\exists p$ s.t.: $w = first(p)$, $p \in P_\pi$ and $w \Vdash A$.

($[\pi]A$) $w \Vdash [\pi]A$ iff $\forall p$ s.t.: $w = first(p)$, if $p \in P_\pi$, then $w \Vdash A$.

($\pi \leftrightarrow \pi'$) $w \Vdash \pi \leftrightarrow \pi'$ iff $\forall p$, if $w = first(p)$, $p \in P_\pi$ iff $p \in P_{\pi'}$.

($\pi \curlywedge \pi'$) $w \Vdash \pi \curlywedge \pi'$ iff $\exists p$ s.t.: $w = first(p)$, $p \in P_\pi$ and $p \in P_{\pi'}$.

($\pi \curlyvee \pi'$) $w \Vdash \pi \curlyvee \pi'$ iff $\exists p$ s.t.: $w = first(p)$ and either $p \in P_\pi / P_{\pi'}$ or $p \in P_{\pi'} / P_\pi$.

($M(\pi)$) $w \Vdash M(\pi)$ iff $c(w) = \langle X, Y, \Pi_M \rangle$, and $\pi \in \Pi_M$.

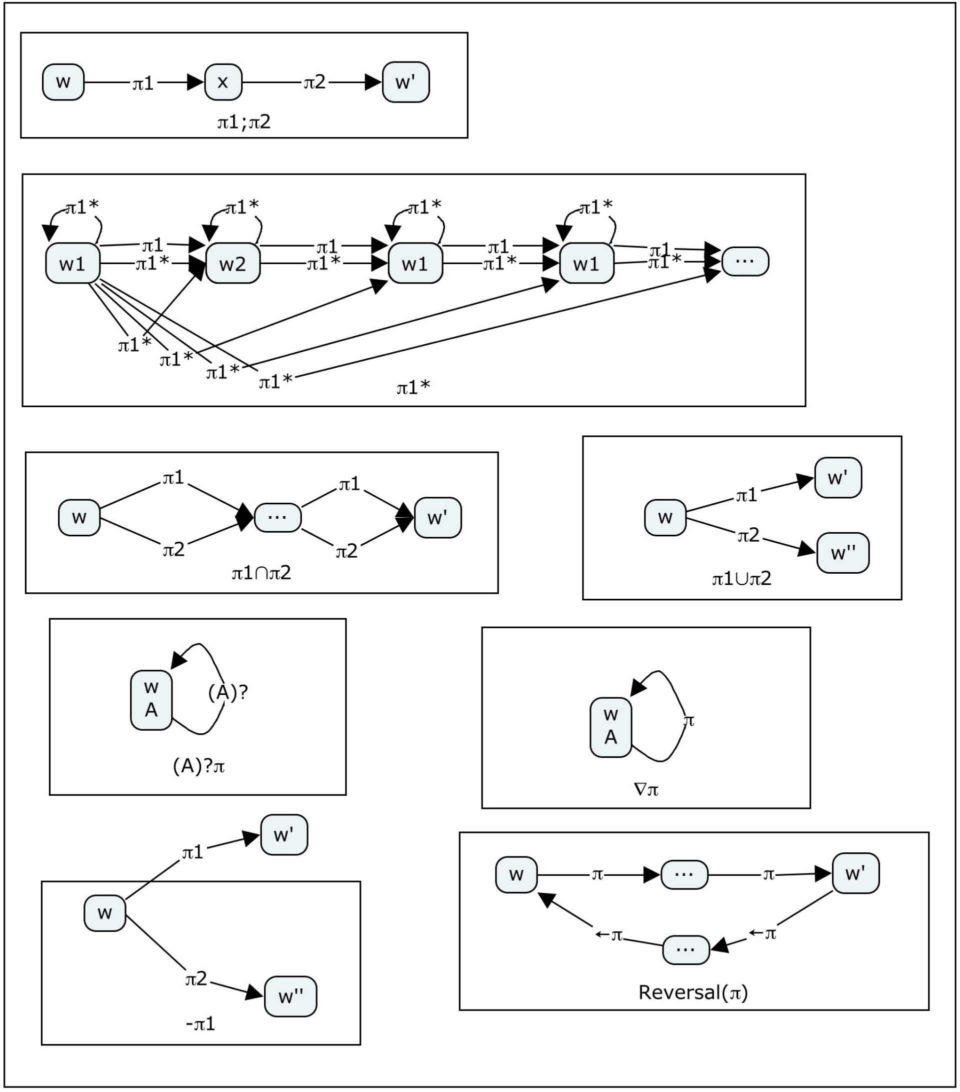
3.1.5 Remarks on PLEN

3.1.5.1 Execution and EDG Structure

The execution conditions determine the gross structures of EDGs in which protocols can be executed. Figure 3.4 provides simple diagrams of the executions of complex protocols and also illustrates this correlation of graph-theoretic structure and the syntactic form of protocol expressions. Think of the labeled arrows as paths that execute each π_i: The execution conditions for sequential composition identify paths that are formed by concatenation of executions of the sequentially composed protocols. Iterations provide the reflexive, transitive closure of P_π. More precisely, the definition of the execution rules for $\pi*$ imply that $P_{\pi*}$ is the reflexive, transitive closure of P_π. The definition implies that every execution of π is an execution of $P_{\pi*}$ and also that every execution of a sequence of π executions is an execution of $\pi*$. The $n = 1$ case ensures that every root state of any execution of π is reached by a looping execution of $\pi*$.

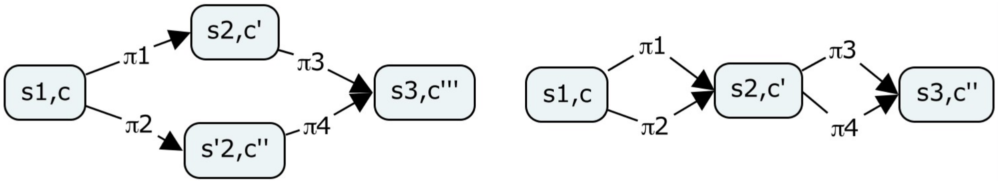
Parallel execution requires convergence not merely of the terminal states of executions of each subprotocol, but correspondence between procedural executions. See Figure 3.5. In

Figure 3.4: Induced Graph-theoretic Structure



The graph-theoretic structure of EDGs - the depth and width of the graphs, the vertex degree of each node, and other, standard features (connectedness, Eulerianness) - covary with their execution properties. From the fact that a protocol is executable at a state, we can, by means of the execution rules, infer the structure of the graph and vice versa.

Figure 3.5: Executability of Parallel Executions



Call the left graph g_1 and the right one g_2. In g_1, $(\pi_1;\pi_3) \cap (\pi_2;\pi_4)$ can not be executed at w. In g_2, it can be executed at w, and this fact determines a specific state-sequence for the agent to follow.

contrast, the conditions for indeterministic choice pick out executions of either subprotocol or may indicate a graph in which both are executable.

The form of test in PLEN is rather different from the form of test in PDL. The test in PLEN is procedural; such tests require invariance of A over paths rather than invariance over edges. Since they are defined over paths, they may end up being equivalent in some ways to complex protocols in particular EDGs. The remarks for fixed point are similar. The executions of a fixed point of π are simply the executions of π that make no difference to the world or to any epistemic contents. This is useful for avoiding conflicts, as a fixed point execution of π can never take one to a state at which π' is no longer executable due failure of its preconditions.

Complementation of a protocol, π, is seeking out any protocols the execution of which does not also suffice for any execution of π. That is, complementation doesn't generate a single protocol. Rather, as above, complementation generates an indeterministic choice among protocols the executions of which are incompatible with the execution of π.

Procedural executions complicate the semantics of reversal of π by adding the inversion function. As above, reversal requires information about the executions of π. The reversal of π is the indeterministic choice among protocols that invert the changes of truth and content wrought by π and do so, not simply by taking whatever path one can from the terminal state resulting from π to the root state, but by traversing the inversion of the path that π dictated.

3.1.5.2 Truth in L_{PLEN}

For each epistemic state, PLEN pairs representations of the content of epistemic states $c(w)$ and the true description of the state of the world $v(w)$. All of the formulae of L_{PLEN} are interpreted relative to these structures. The necessity operator is fairly standard. It says that A is true at every state in the EDG; every description of the world at any time that can be reached in any epistemic process is one in which A is true.

The operators $b(A)$ and $d(A)$ can be interpreted as belief or acceptance (b) and disbelief or rejection (d) at w. Note that the b and d operators are syntactic epistemic operators.[5] Their behavior is only determined by the properties of sets of formulae - X and Y - rather than the transition relations of EDGs, as might be the case in dynamic epistemic logics. This allows theorists to focus on subsets of EDGs in which assumptions or stipulations about the behavior of these operators hold. No assumptions about the behavior of these operators is built into the system at this stage, so, if, for instance, one distinguishes between belief and acceptance, conditions can be added to PLEN to sync the operators up to the (best philosophical accounts of the) epistemic judgments one wants to model. For instance, one might argue that the *exclusion* condition on acceptance holds: for all $A \in L_{PLEN}$, $\neg b(A) \equiv d(A)$, while it doesn't hold for belief. Similarly for the *exhaustion* condition: for all $A \in L_{PLEN}$, $(b(A) \vee d(A))$.

Note that the valuation at w may act as an external check on the epistemic content state. It may very well be the case that $b(A)$ at w but A is false. No factivity condition was built into the valuation of $b(A)$. This enables protocols to fold in reliability considerations. A protocol can, for instance, dictate increases or decreases in true belief.[6]

---

[5] As in Konolige and Jago [148, 134]. It has similar advantages and disadvantages, such as the absence of logical omniscience.

[6] Though such work is omitted from the main text, metrics for protocols that increase true belief and decrease error over time are easy enough to construct. Consider the easy case where valuations assign semantic values to finitely many formulae at states. Let states s_1 and s_2 be such that s_1 is the root state of a process executing π and s_2 is the terminal state. Then consider the formulae of L_{PLEN} accepted at s_1 or s_2. Let n be the cardinality of the set of all true accepted formulae at s_1, and let m be the cardinality of the set of all true accepted formulae at s_2. If $m > n$, π results in an absolute increase in true belief, if $n > m$, then π results in an absolute decrease in true belief. Define comparative increases and decreases

The first epistemic dynamic operator, $\langle\pi\rangle A$, encodes weak postconditions of π. The weak postconditions are the conditions that may hold after executing (or, more intuitively, complying with) π or, perhaps more intuitively, the conditions that hold after complying with π in some way. That is, one would say that A is a weak postcondition of π at w if some way of executing π takes you to w' and A holds at w'. Call the special case of weak postconditions, $\langle\pi\rangle\top$, *the executability* of π. The basic idea is simple; π is executable at a state iff there is a path through an EDG that executes π - some series of epistemic actions or procedures is such that it accords with π.

The second dynamic epistemic operator, $[\pi]A$, encodes *strong postconditions.* The strong postconditions are the conditions that must hold after every execution of π. In other words, the formula encodes the conditions that hold if one complies with π in any way. The special case of strong postconditions is *strong executability*, $[\pi]\top$. The truth conditions tell us that π is strongly executable at a state iff every path that begins at that state is an execution of π. Every epistemic process that begins at that state accords with π.

To sum: The *weak postcondition* operator, $\langle\pi\rangle A$, tells us that some epistemic process starting at the current state that accords with π results in a state where A is true. The strong postcondition operator, $[\pi]A$, tells us that every epistemic process starting at the current state that accords with π results in a state where A is true.

The *procedural equivalence* between formulae of L_{PLEN}, $\leftrightarrow$, is more informative about protocols than any truth-functional or even intensional equivalence between propositional formulae that embed protocol expressions. The expression $\pi \leftrightarrow \pi'$ tells us that, from the current state, any path executing π executes π' and vice versa. In other words, the idea is that, where π and π' represent norms, N and N', N and N' license exactly the same moves (actions, processes, changes of epistemic state) from the current state.

Procedural equivalence is stronger than equivalence between postcondition expressions

---

in terms of proportions between true accepted formulae and false rejected formulae. From there, more interesting properties can be defined by quantifying over all transitions associated with protocols within EDGs and across EDGs, as well as interactions with postconditions and other expressions in which protocols are embedded.

because focusing on inclusion relations among sets of executions starting at a state is more informative than focusing just on quantification over paths that execute protocols. Formulae formed with $\leftrightarrow$, for instance, say more about preservation of L_{PLEN} formulae than even combinations of modal and dynamic operators like $\Box([\pi]A \equiv [\pi']A)$. For instance, suppose that $\Box([\pi]A \equiv [\pi']A)$ and $\langle\pi\rangle\top$ are true at state w. From this, it does not follow that $\langle\pi'\rangle\top$ or that $\langle\pi'\rangle A$ are true at w. The countermodel is one in which π' is not executable at w. If so $[\pi']A$ is vacuously true.[7] Thus, there is no execution of π' starting at w, nor one that terminates with a state verifying A. Substituting π' for π in larger L_{PLEN} expressions does not preserve truth for these formulae. This point can be seen more easily from the fact that $\pi \curlywedge \pi'$ follows from $\pi \leftrightarrow \pi'$ and $\langle\pi\rangle\top$ but not from $\Box([\pi]A \equiv [\pi']A)$ and $\langle\pi\rangle\top$. However, from $\pi \leftrightarrow \pi'$ and $\langle\pi\rangle\top$, it does follow that $\langle\pi'\rangle\top$ because the relation holds only if the sets of executions of π and π' starting at w are identical. Any procedural path executing π also executes π'. Thus, if $\langle\pi\rangle A$, then $\langle\pi'\rangle\top$ and $\langle\pi'\rangle A$ must also hold. This equivalence shows that the embedded protocols obey a kind of substitutivity. Additionally, $\pi \leftrightarrow \pi'$ preserves strong postconditions; $[\pi]A \equiv [\pi']A$ follows. In general, for formulae involving only the dynamic operators, if $\pi \leftrightarrow \pi'$, then uniform substitution of protocols π and π' preserves truth. Not so for $\Box([\pi]A \equiv [\pi']A)$.

However, there are modal expressions that $\Box([\pi]A \equiv [\pi']A)$ preserves truth for that $\pi \leftrightarrow \pi'$ doesn't. Suppose that at w, $\langle\pi\rangle(\langle\pi\rangle\top \wedge [\pi]A)$ holds. Then $\langle\pi\rangle\top \wedge [\pi]A$ holds at some w' that π can bring the system to. If $\Box([\pi]A \equiv [\pi']A)$ also holds at w, $[\pi']A$ must hold at w', but $\langle\pi'\rangle\top$ and $\langle\pi'\rangle A$ may fail at that state. If $\pi \leftrightarrow \pi'$ holds at w, then all of the foregoing formulae may fail at w'. However, $\Box\pi \leftrightarrow \pi'$ ensures preservation of truth by substitution of π with π' for each of these expressions as well as any that $[\pi]A \equiv [\pi']A$ fails to preserve at w. So $\pi \leftrightarrow \pi'$ is strictly stronger than $[\pi]A \equiv [\pi]A$ but not $\Box([\pi]A \equiv [\pi']A)$, but $\Box\pi \leftrightarrow \pi'$ is strictly stronger than $\Box([\pi]A \equiv [\pi']A)$; in fact, it implies it, and the converse fails.

---

[7] Of course, this only follows if the truth condition is formulated with a material conditional.

The weaker $\pi \curlywedge \pi'$ does different work. It only guarantees (at w) that there is some path starting at w that executes both π and π'. If $[\pi]A$ holds at w, then, assuming convergence, so do $\langle\pi\rangle\top$, $\langle\pi'\rangle\top$, $\langle\pi\rangle A$, and $\langle\pi'\rangle A$. It ensures that substituting π for π' in the context of a dynamic operator preserves executability and weak postconditions, but not strong postconditions. It's tempting to think that this expression makes parallel execution redundant, but the notions are clearly distinct. That $\pi \curlywedge \pi'$ is true at w, it does follow that $\langle\pi \cap \pi'\rangle\top$ at w; given the former, there must be a p such that $first(p) = w$ and $p \in P_\pi \cap P_{\pi'}$, but then $p \in P_\pi$ and there is a $q \in P_{\pi'}$ such that $p \rightleftharpoons q$ because $p \in P_{\pi'}$ and $p \rightleftharpoons p$. However, the converse fails. Supposing that $\langle\pi \cap \pi'\rangle\top$ at w, it follows that there is a p such that $first(p) = w$, $p \in P_\pi$, there is a $q \in P_{\pi'}$ such that $p \rightleftharpoons q$. But that doesn't entail that there is any r such that $first(r) = w$ and $r \in P_\pi \cap P_{\pi'}$. So $\pi \curlywedge \pi'$ may fail.

The expression $\pi \curlyvee \pi'$ merely ensures the falsity of some formulae of L_{PLEN}; it is essentially the dual of $\pi \leftrightarrow \pi'$; if there is some path from w that executes π but not π', then clearly π and π' are not executed by exactly the same paths from w.

The Π_M in every content state represents the current methodology of the epistemic agent whose arena is g at w. The agent endorses or accepts the epistemic rationality norms in Π_M at w. $M(\pi)$ says that π *is part of the methodology (of the agent whose arena is g) at w*; that is, that $\pi \in \Pi_M$.

On that note, L_{PLEN} is a powerful language. Postconditions are represented by the dynamic epistemic operators. Preconditions can be formalized by conditionals: $A \rightarrow \langle\pi\rangle\top$, for instance, says that if A is true at s, then π is executable at s. That is, A is a *sufficient precondition* of π's execution. The converse conditional, $\langle\pi\rangle\top \rightarrow A$, encodes a *necessary precondition* of π's execution. Naturally, $A \equiv \langle\pi\rangle\top$ encodes a necessary and sufficient precondition of π's execution.

Note that L_{PLEN} contains expressions stating various kinds of interaction between methodologies, pre and postconditions, and execution conditions. A small sample:

- $M(\pi) \rightarrow [\pi]A$ says that if π is in the methodology, then A is a strong postcondition of

π.

- $M(\pi) \wedge \langle\pi\rangle\top$ says that some protocol in the methodology can be executed at the current state.
- $\langle\pi\rangle A \to M(\pi)$ says that if π can bring about A, then it's in the methodology.

The truth of these expressions at a given state imposes structural requirements on the EDG, including conditions on the edges and on the content functions. This can be easily seen; the truth conditions of the operators pair with the execution conditions of protocols to dictate graph-theoretic properties. Let w be a state at which any of the foregoing expressions is true. If π is a sequential composition, for instance, this dictates a minimum length for any path of which w is the root, and it dictates that the terminal state verifies A. If π is indeterministic at w, that is, if there is more than one path executing π at w, then there is a tree of minimum depth in which each leaf node verifies A. Each control operator induces other graph-theoretic features: corresponding paths, branching paths, etc.. It's trivial to see how conditions on content functions in an EDG are constrained by the truth of these claims at a state.

3.2 Axiomatizing PLEN

The final part of PLEN is a logic of protocols; a derivation system that codifies the valid rules governing protocol and process equivalence, execution of protocols, and the effects of executing protocols. There are numerous ways to build a logic with the syntactic and semantic resources of PLEN. In this section, valid reasoning in PLEN is partially axiomatized with a simple Hilbert-style system.

The system has two features worthy of note at the outset. First, the system is sound. There is, however, the caveat that some of the axiom schemata are syntactically restricted: Let $A(\pi)$ be any formula of L_{PLEN} that embeds π in either the dynamic operators or the connectives of L_{PLEN}, except that the formula variable $A(\pi)$ cannot embed π directly in

$\Box A(\pi)$, $M(\pi)$, or in a postcondition. Else, at least one schema fails. Second, the derivation system is strictly stronger than the standard axioms of propositional dynamic logic (PDL), with some qualifications. It validates the standard axioms of PDL [113] as well as novel axioms. This is not just due to the larger language of PLEN over that of PDL; PLEN validates axioms that, were they added to PDL in the natural ways, would fail in PDL. This demonstrates one of the ways in which PLEN can account for features of epistemic norms that is not recognized in other systems.

3.2.1 Syntax II: A Proof System for PLEN

Definition: (Rules) These are standard derivation rules for PDL or, indeed, for modal logic. Note that detaching the conditional is valid in that it preserves truth and necessitation is valid in that it preserves theoremhood; explicitly, necessitation is restricted to apply only to axioms and formulae derivable from axioms.[8]

1. $\left[\ \dfrac{A \to B,\, A}{B}(MP) \quad \dfrac{\vdash_{PLEN} A}{\Box A}(NEC)\ \right]$

Definition: (Axiom Schemata) PLEN is partially, soundly axiomatized by the standard propositional axiom schemata for the truth functions (assuming standard definitions of conjunction and disjunction), standard axioms of a suitable modal logic (for concreteness, let it be $S5$ - but note that alternative modal logics can integrated into PLEN easily), plus rules for the novel connectives and operators of PLEN.

---

[8] This feature makes the system compatible with additional machinery for reasoning from premises without the ugly - or at least perplexing - results that often accompany such an addition. In particular, this revision makes room for a valid deduction theorem and so makes it possible to have rules for introducing conditionals in the usual natural deduction style. See Hakli and Negri [116] for discussion of the matter and the machinery for adding a valid deduction theorem to normal modal logics. The box "$\Box$" of PLEN works like the box of $S5$, and the modalities for each protocol work as in K. The machinery that Hakli and Negri construct should make deduction for PLEN a simple matter.

$$\text{1.}\quad\left[\begin{array}{l}
A \rightarrow (B \rightarrow A) \\
(A \rightarrow (B \rightarrow C)) \rightarrow ((A \rightarrow B) \rightarrow (A \rightarrow C)) \\
(\neg B \rightarrow \neg A) \rightarrow (\neg B \rightarrow A) \rightarrow B)) \\
(\neg B \rightarrow \neg A) \rightarrow (A \rightarrow B) \\
\Box(A \rightarrow B) \rightarrow (\Box A \rightarrow \Box B) \\
\Box A \rightarrow A \\
\Diamond A \rightarrow \Box\Diamond A
\end{array}\right]$$

$$\text{2.}\quad\left[\begin{array}{ll}
[\pi 1][\pi 2]A \equiv [\pi 1; \pi 2]A & ([\pi]A \wedge \langle\pi\rangle \top) \rightarrow \langle\pi\rangle A \\
\langle\pi 1; \pi 2\rangle A \equiv \langle\pi 1\rangle \langle\pi 2\rangle A & \langle\pi\rangle A \rightarrow \langle\pi\rangle \top \\
\langle\pi 1 \cap \pi 2\rangle A \rightarrow (\langle\pi 1\rangle A \wedge \langle\pi 2\rangle A) & [\pi](A \wedge B) \equiv ([\pi]A) \wedge [\pi]B) \\
([\pi 1]A \wedge [\pi 2]B) \rightarrow [\pi 1 \cap \pi 2](A \wedge B) & ([\pi](A \rightarrow B) \wedge [\pi]A) \rightarrow [\pi]B \\
[\pi 1 \cup \pi 2]A \equiv ([\pi 1]A \wedge [\pi 2]A) & ([\pi](A \rightarrow B) \wedge \langle\pi\rangle A) \rightarrow \langle\pi\rangle B \\
\langle\pi 1 \cup \pi 2\rangle A \equiv (\langle\pi 1\rangle A \vee \langle\pi 2\rangle A) & (\langle\pi\rangle (A \rightarrow B) \wedge [\pi]A) \rightarrow \langle\pi\rangle B \\
([\pi][\pi *]A) \wedge A \equiv [\pi *]A & \Box A \rightarrow [\pi]A \\
\langle\pi *\rangle A \equiv (\langle\pi\rangle \langle\pi *\rangle A) \vee A & ((\pi \leftrightarrow \pi') \wedge A(\pi)) \rightarrow A(\pi') \\
(A \wedge [\pi *](A \rightarrow [\pi]A)) \rightarrow [\pi *]A & ((\pi \curlywedgeuparrow \pi') \wedge [\pi]A) \rightarrow \langle\pi'\rangle A \\
\langle(A)?\rangle B \rightarrow (A \wedge B) & (\pi \curlyveedownarrow \pi') \equiv \neg(\pi \leftrightarrow \pi') \\
\langle(A)?\rangle \top \rightarrow A & \\
A \equiv [\nabla\pi]A & \\
\Box(A \equiv \langle\pi\rangle \top) \rightarrow \Box[\curvearrowleft \pi]A &
\end{array}\right]$$

Note two important properties and a qualification to note about the axiomatization of PLEN.

First, the deductive system of PLEN is sound. The usual definitions related to soundness have counterparts in PLEN:

Derivation A derivation is any finite sequence of formulae of L_{PLEN}, $\langle A_1, ..., A_n\rangle$, such that each $A_{i\leq\text{n-1}}$ is either a premise, an axiom, or the conclusion of a derivation rule whose premises appear earlier in the sequence, and A_n is the conclusion of a derivation rule whose premises occur among $\langle A_1, ..., A_{n-1}\rangle$.

Derivability Let $A \in L_{PLEN}$, $X \subseteq L_{PLEN}$, and $\{B_1, ..., B_n\} = X$. A is *derivable* from X, $X \vdash_{PLEN} A$, iff there is a derivation $\langle B_1, ..., B_n, A \rangle$.

Theorem A is a *theorem* in PLEN, $\vdash_{PLEN} A$, iff A is an axiom or A can be derived from the axioms of PLEN via the system in Chapter 3.

Verification An EDG, $g = \langle W, C, c, v, \Pi_g, R \rangle$, *verifies* A *at* w, $w \models_{PLEN} A$, iff $w \in W$ and $w \Vdash A$.

Tautology A is *a tautology* in PLEN iff A is verified by all states in all EDGs.

Logical Consequence A is a *logical consequence in PLEN* of X, $X \models_{PLEN} A$, iff for every $g = \langle W, C, c, v, \Pi_g, R \rangle$ and every w in W, if $\forall B_i(B_i \subseteq X \rightarrow w \Vdash B_i)$, then $w \Vdash A$.

Validity A derivation rule $\dfrac{B_1, ..., B_n}{A}$ is *valid* in PLEN iff A is a logical consequence in PLEN of $\{B_1, ..., B_n\}$. A formula, A, of L_{PLEN} is valid iff A is a tautology in PLEN.

Definition: Soundness Syntax II is sound iff if $X \vdash_{PLEN} A$, then $X \models_{PLEN} A$, and if $\vdash_{PLEN} A$, then $\models_{PLEN} A$.

From these, it follows that the PLEN axiomatization is sound:

(Soundness Theorem) *If* $X \vdash_{PLEN} A$*, then* $X \models_{PLEN} A$*, and if* $\vdash_{PLEN} A$*, then* $\models_{PLEN} A$.

Proof-Plan It suffices to show that all of the rules in the derivation system are valid, and all of the axioms are tautologies. First, prove that the rules are valid - they preserve the truth of all expressions of L_{PLEN} at every state in every EDG. Second, prove that each of the axiom schemata are tautologous - every instance of each schema is a tautology. The details of the proof are entirely mechanical, and shown in detail in [135]. $\square$

This axiomatization of PLEN is very likely not complete. But this is beside the main theoretical point of PLEN's deductive system. The primary functions of the axiomatization of PDL are to: (a) show that formalizing reasoning can be easily carried out in the PLEN framework by (b) provide a basic framework for analyzing valid reasoning about epistemic rationality norms. This system is likely far from a final word on how best to model such reasoning. This is intentional; PLEN is a formal framework rather than a final model. As such, its full technical development is relegated to supplementary [135] and future work and likely to require modifications to the logical apparatus.

Second, note that the deductive system of PLEN is strictly stronger than propositional dynamic logic. The following list of axiom schemata, coupled with modus ponens and necessitation, forms the standard axiomatization of PDL:

$$\left[\begin{array}{l}
(\textbf{PDL}) \\
[\pi 1][\pi 2]A \equiv [\pi 1; \pi 2]A \\
[\pi 1 \cup \pi 2]A \equiv ([\pi 1]A \wedge [\pi 2]A) \\
[(A)?]B \equiv (A \to B) \\
([\pi][\pi *]A) \wedge A \equiv [\pi *]A \\
(A \wedge [\pi *](A \to [\pi]A)) \to [\pi *]A \\
[\pi](A \to B) \to ([\pi]A) \to [\pi]B) \\
[\pi](A \wedge B) \equiv ([\pi]A) \wedge [\pi]B)
\end{array}\right]$$

Then:

Proposition The axioms of PDL (without parallel execution) are all valid in PLEN, but the converse fails.

Proof-Plan Each (instance of each) PDL axiom schema is a tautology of PLEN, which is shown in the proof of soundness. However, the equivalence $((\pi_1; \pi_2) \cap (\pi_3; \pi_4) \leftrightarrow (\pi_1 \cap \pi_3); (\pi_2 \cap \pi_4))$ is valid in PLEN but not in PDL, assuming the natural execution conditions based on binary transition relations. Thus, PLEN is strictly stronger. □

In some sense, this means that PLEN gives an alternative semantics to PDL. Or, more precisely, the language of PDL is one of the logical languages that can be interpreted by the basic model of epistemic dynamics in PLEN. This point can be strengthened further:

Proposition Every PDL model is isomorphic to an EDG.

Proof-Plan This is trivial. For every binary transition $t = \langle w, w' \rangle$ that executes π in the PDL model g, construct an EDG, g', such that there is an edge $e = \langle w, \pi, w' \rangle$. Then there will be a bijection $f : W \longrightarrow W'$ such that $f(w) = w'$ iff $\langle w, w' \rangle$ executes π iff $\langle w, \pi, w' \rangle$ executes π. □

That is, PDL can be applied as a model of epistemic dynamics, and its logic can be applied as a logic of epistemic norms. PLEN is thus a proper extension of PDL; L_{PLEN} contains more expressions and the logic of PLEN can be used to reason about these expressions.

Note that, for purposes of modeling epistemic dynamics and epistemic rationality, many of the constructions in L_{PLEN} simply make no good philosophical sense if interpreted over PDL. The key idea is simply that PDL interprets its complex protocols over binary transition relations, where PLEN interprets them over procedural paths. This difference gives very different properties to expressions like $\pi \leftrightarrow \pi'$. In the style of PDL, we could interpret these expressions like so: $\pi \leftrightarrow \pi'$ is true at any state w in any EDG iff every pair of states, $\langle w, w' \rangle$ is such that it executes π iff it executes π'. This condition is very different from the truth conditions of the expression in PLEN, which requires sharing not just binary transition relations but procedural paths. The latter cannot even be interpreted in PDL without adding structure to the models of PDL that render those models essentially equivalent to EDGs. In models that apply both the execution conditions of PDL and those of EDG, the procedural equivalences imply syntactic properties of π and π' that the binary interpretation doesn't.

If π is a sequential composition, so must π' be, for instance. This is not so for the binary transition based interpretation.

Finally, the soundness result comes with a caveat. The axiom schema $((\pi \leftrightarrow \pi') \wedge A(\pi)) \rightarrow A(\pi')$ encodes a limited replacement property of $\pi \leftrightarrow \pi'$; it is only valid under certain syntactic restrictions. The schematic L_{PLEN} formula $A(\pi)$ can contain π up to any level of syntactic depth, but $A(\pi)$ cannot embed π in $\Box A(\pi)$, $M(\pi)$, or in a postcondition of a protocol. If we lift this syntactic restriction, some instances of the axiom schema fail. Consider the instance of the foregoing schema: $((\pi \leftrightarrow \pi') \wedge \Box[\pi]A) \rightarrow \Box[\pi']A$. Let the antecedent be true at w in EDG g. Then $[\pi]A$ is true at all w' in g. However, it doesn't follow that $[\pi']A$ holds at all w' in g, it only follows that $[\pi']A$ at w. The same holds for any M expressions. That execution of π always suffices for execution of π' and vice versa simply doesn't imply that π and π' are both in Π_M. For postconditions, the point is equally simple. Suppose that $A(\pi)$ embeds π in the postcondition of some protocol π'': $\langle \pi'' \rangle \langle \pi \rangle A$. It may be the case that π and π' share all executions starting at w, but they may not share executions at some w' reachable by π''.

However, even with this syntactic restriction, the set of valid replacements of procedural equivalents is fairly robust. If we let $A(\pi)$ be any expression meeting the syntactic restriction defined above, then replacement of π with any π' such that $\pi \leftrightarrow \pi'$ preserves truth. From this, various derivative rules governing the relation between procedural equivalence and the other operators and with pre and postconditions can be gotten:

- $(\pi \leftrightarrow \pi') \rightarrow ((\pi' \leftrightarrow \pi'') \rightarrow (\pi \leftrightarrow \pi''))$
- $(\pi \leftrightarrow \pi') \rightarrow (\langle \pi \rangle A \equiv \langle \pi' \rangle A)$
- $(\pi \leftrightarrow \pi') \rightarrow ([\pi]A \equiv [\pi']A)$
- $(\pi \leftrightarrow \pi') \rightarrow ([\pi]A \rightarrow \langle \pi' \rangle A)$
- $(\pi \leftrightarrow \pi' \wedge \langle \pi \rangle A) \rightarrow (\pi \Yup \pi' \wedge \langle \pi' \rangle A)$

- $(\pi \leftrightarrow \pi') \rightarrow \neg(\pi \Yup \pi')$
- $(\pi \leftrightarrow \pi') \rightarrow (([\pi 1][\pi 2]A \equiv [\pi]A) \equiv ([\pi 1][\pi 2]A \equiv [\pi']A))$
- And so on.

With the definition and exploration of the core system of PLEN out of the way, the next step is examining it as an account of epistemic dynamics and epistemic rationality norms. The final section of this chapter acts as a prolegomenon to Chapter 4 by outlining the philosophical interpretation of PLEN.

3.3 PLEN and Epistemic Rationality Norms

The core thought of PLEN is that, in PLEN, epistemic rationality norms are protocols. This requires a bit of clarification. Epistemic norms - really, norms in general - are simultaneously mysterious and ubiquitous objects. They clearly play important roles in cognitive life, and there is much that we can say about these roles and how norms function to perform them, but what norms exactly are is uncertain. Different fields take different approaches. The approach in epistemology [164] seems to follow in von Wright's footsteps [261], identifying norms with conditional propositions. Deontic logicians sometimes take them to be conditional rules, perhaps best identified with imperatives [89, 250]. The social sciences [59] takes another approach, identifying norms as regularities that emerge in models of populations or in the cognitive architectures of individual agents. There is a great deal of unclarity.

The point of the protocol-theoretic account and its formalization in PLEN is to account for what can be figured out about the functions and features of epistemic rationality norms with a simple, coherent formalism, so as to provide a clear model of them. This is a modeling task, with the metaphysics of norms set to the side. Chapter 2 went to great lengths to show that epistemic rationality norms are dynamic and procedural, and Chapter 1 argued that the underlying model of such norms must have certain features. PLEN is a formal theory in which protocols function with respect to the models of epistemic dynamics just as epistemic rationality norms function with respect to the epistemic dynamics of reasoners. Protocols represent procedural epistemic rationality norms by formalizing a kind of procedural or program-like structure defined over epistemic actions and processes. Procedural or program-like structure is directly encoded by the syntax of protocols defined above, and modeled by interpreting protocols over EDGs by means of the execution rules. This idea can be made somewhat more precise, now that the formalism has been introduced.

3.3.1 Preliminaries, Procedural Operations, and Dynamics

The formal features of PLEN are designed to represent epistemic rationality norms and epistemic dynamics. This can be seen by considering four ideas. The first idea is that epistemic rationality norms are conceptualized as abstract objects like policies, programs, or plans for restricting epistemic dynamics. This conceptualization was one of the main upshots of Chapters 1 and 2, especially considering D3 and its corollary. On this informal conceptualization, program-like structures are natural and useful for modeling procedural epistemic rationality norms because they encode instructions for procedural operations over actions, and this is the primary motivation for building the formalism around protocols, which are established logical representations of programs or plans.[9]

The second idea is that the protocols of P_{PLEN} are built with a subset of operations of relational algebra.[10] This syntactic feature reflects the idea that actions can be represented with transition relations (binary, ternary, or n-ary), and, consequently, that processes and

---

[9] The protocols of PLEN are basically just the programs of PDL, and, as PDL is the logic of programs, the point should be obvious. However, do note that the core model of temporal action logics (e.g. [130]) is also the core model of dynamic logics. If this point still evades one's intuitions, one need only recognize that protocols are constructions built out of some suite of basic actions with procedural operations that encode programs or instructions for complex action and that plans are instructions for complex action built out of some set of basic actions with procedural operations. More formally, see [195] for an explicit program-based account of planning and practical reasoning, see [95] and [115] for a variety of formalisms for representing plans, some of which bear obvious relations to the dynamic logics or program logics that provide the basic tools of PLEN, and, finally, note the features of planning observed in [43] that have obvious analyses with these tools.

[10] See Chapter 3 of [30] or the earlier establishment of the idea in [28].

procedures constructed out of procedural operations on actions can be represented by relational operations on representations of actions.[11] This leads to a natural way of interpreting protocols - assign relational algebraic operations on transitions (n-ary) to protocols in a way that parallels the assignment of operations on binary transitions to the symbols of relational algebra. That protocols encode instructions that specify procedures falls out of this idea. This is spelled out a bit below.

The third idea is that graph-theoretic structures provide useful models of (epistemic) dynamic processes, and they are built out of operations on relations. The first conjunct is the core insight of dynamic logics [28, 30], dynamic epistemic logics [65, 193], and belief revision theories [91, 92]. The second conjunct is a technical fact that has been well-established.[12]

The fourth idea is that a good formal framework should be general.[13] PLEN should not focus on any particular subset of epistemic actions (e.g., public announcements or any other kind of transparent updating actions, opaque actions involving kinds of subterfuge, etc.) but be entirely general concerning epistemic actions. Such generality allows a formal framework to generate a unified theory of epistemic norms and not just the plurality of theories of the epistemic norms that govern specific kinds of actions.

Putting these ideas together, the core thought of PLEN is that protocols encode the procedural content of epistemic rationality norms in two ways:

(i) Their syntax encodes *procedural operations*, abstract instructions for carrying out, sequencing, combining, choosing among, conditioning on tests, avoiding or complementing, reversing, and minimally executing *arbitrary* epistemic actions, processes, and procedures or, in other words, constructing complex epistemic procedures out of more basic epistemic actions.

(ii) Their semantics - defined by their execution conditions - model the restrictions

---

[11] See [261] for a classic source as well as [30, 28, 114].

[12] See, again, Chapter 3 of [30] and couple this with the basic observation that transition systems are useful for representing epistemic dynamics.

[13] This is the core project and insight of [18].

> on epistemic dynamics that norms specify. The execution conditions of protocols interact with these syntactically coded instructions to determine what epistemic actions are to be executed at what states (according to the instructions of a norm), what future courses of action must be carried out in order to comply with these instructions (as well as derivative and partial instructions), and what the resulting changes of epistemic state must be. In other words, the syntax of protocols defines the structure of the graphs that model rational epistemic dynamics.

The basic slogan of PLEN is that epistemic rationality norms are protocols. More literally, protocols are the objects in the theory that stand-in for epistemic rationality norms, and these two ways of encoding norms in PLEN spell out how. Basically, protocols function in PLEN with respect to EDGs - the models of epistemic dynamics - the way that epistemic rationality norms function with respect to epistemic dynamics. Let me elaborate on (i) and (ii).

The syntax of complex protocols encodes instructions for sequencing or iterating basic actions or more complex procedures, concurrently complying with multiple instructions or complex procedures, indeterministically choosing among actions or procedures, avoiding actions or procedures, reversing their effects, or ensuring that they are carried out with minimal effects. The control operators, in order of definition above, correspond to these instructions in ways obvious from the natural translations in section 3.1, but reiterated in Table 3.1.

The control operators of P_{PLEN} form the core of possible procedural instructions/operations; there are no other obvious ways in which actions can be combined in plans, programs, or instructions. That is, the control operators of PLEN represent the core possible procedural operations; the set of operations on actions that can be used to compose actions into more complex procedures. The elements in Π_0 encode basic actions. Epistemic actions are understood in a capacious sense, including not only uncontroversially voluntary actions like choosing which arguments to attend to, which cards to turn over, what transformations or

Table 3.1: Procedural Operations

Instruction	**Procedural Operation**	**Relational Structure**
Do a, then do b!	Sequential composition "$a; b$"	Relational Composition
Do a some finite n times!	Kleene Iteration "$a*$"	Iterated Relational Composition
Comply with both the instructions for a and b!	Parallel Execution "$a \cap b$"	Intersection, Correspondence
Select either a or b (inclusive)!	Indeterministic Choice "$a \cup b$"	Union
Avoid a, do anything other than a!	Complement "$-a$"	Complement
Reverse the effects of a!	Reversal "$\curvearrowleft a$"	Inverse of Relation
Do a with minimal effects!	Fixed Point "∇a"	Fixed Point

manipulations of a formal or physical model to carry out, or which experiments to carry out, but also arguably involuntary actions like inference. The complex protocols in Π_C encode the possible combinations of basic actions that form any possible epistemic procedures.

Coupled with the semantics of PLEN based on EDGs, this procedural information determines the properties of models of the ideally rational evolution of epistemic systems. First, distinguish acts and actions. Let acts be tokens of act-types, like a particular instance of my pressing a particular button at a time t with particular effects. In contrast, let actions be act-types, abstract particulars like pressing a button or even pressing *that* button, each of which can be instantiated in different locations in space-time. With EDGs, think of a particular *epistemic act* as the pair of a label for an action or act-type in Π_0 and an epistemic transition mapped to it by the execution rules. That is, an act, associated with a labeled transition, is the concrete instance of an action, which is associated with the label in Π_0. In PLEN, acts are edges. Actions are sets of acts - the execution sets of basic protocols. At this stage, PLEN aims for robust generality. The set of epistemic actions is not exclusive; epistemic actions are any actions that result in change of epistemic state or result in change of factors relevant, via effects on cognition, to change of state. The class of epistemic actions thus includes everything from direct expansions, contractions, and revisions of belief to the steps involved in the physical arrangement of representations of beliefs so as to make logical connections more perspicuous.

Second, define epistemic histories and epistemic processes. This is simple: an epistemic history is a state-sequence, and a token epistemic process is a procedural path.[14] That is, *epistemic processes* are chains of *epistemic act*s. Say that *epistemic procedures* are process types; they are sets of processes defined by procedural relations among their constituent acts. For instance, a procedure might be defined by sequences or choices of actions or the presence of convergences or divergences between processes. Process types are, as should be clear, mapped to complex protocols, and the defining relations correlate, via the execution rules, to the syntactic forms of protocols. Processes are instantiations or realizations of procedures; they're elements of the process type.

Third, note that each EDG has a graph-theoretic structure that combines with the execution conditions to represent the feasibilities and capabilities of the particular agent, i, whose dynamics are represented by g.[15] The thought is that acts and processes are concrete, particular parts of an epistemic agent's life, while actions and procedures are abstract, and execution rules link these concrete and abstract objects to each other. That is, the execution rule maps a procedure to every protocol expression - the set P_π is a procedure defined by the syntax of π.

In sum: think of edges (triples of the form $W \times \{\pi\} \times W$) in EDGs as instances or tokens

---

[14] A needlessly complex version can be given as follows: Let a map, h, from a segment, $\langle 1, ..., n\rangle$ of the natural numbers to W pick out part of the *epistemic history* of an epistemic agent; the resulting sequence of content states represents the evolution of the content of her epistemic states up to point n. Let Ah be a map from a segment, $\langle 1, ..., m\rangle$ of the natural numbers to Π_g. This picks out another part of the epistemic history of an epistemic agent; the sequence of actions that the agent actually carried out to evolve her epistemic state up to point m. Say that w, w' are connected by a in g iff $\langle w, a, w'\rangle \in R$. An *epistemic process* is the pair of a sequence of epistemic actions and a sequence of transitions mapped to it by the execution rules: $\langle h, Ah\rangle$, such that for every a_i, a_{i+1} mapped to $i \leq m-1$ by Ah, there is a triple of states, $\langle w_i, w_{i+1}, w_{i+2}\rangle$ mapped to $j \leq m-2$ such that a_i connects w_i to w_{i+1} and a_{i+1} connects w_{i+1} to w_{i+2}. Note that each basic action is mapped by the execution rules to a set of acts in each EDG, and each complex protocol is mapped by the execution rules to a set of processes in each EDG. The latter fact can be seen by considering the set of processes, Ah, such that the first and last elements of Ah form a transition that is an execution of π.

[15] All EDGs are labeled, directed multi-graphs. They may be deterministic (containing unique successors for each state according to R) or indeterministic. They may be cyclic or acyclic. They may be connected, Hamiltonian, Eulerian, and so on.

of actions, processes as concatenated sequences of edges (paths) and instances of procedures, histories as the results of processes, and protocols as the procedural instructions that define procedures, the limit case of which is that of basic actions. Each protocol $\pi \in P_{PLEN}$ is mapped to a procedure or process-type (set of processes) by the execution rules.[16]

3.3.2 Representation Rules: Interpreting Protocols as Epistemic Rationality Norms

Let a protocol, π, codify the procedural content of an epistemic rationality norm, N, in the sense that we interpret the executions of π to be those processes of evolution of i's epistemic state that are rational according to N. Explicitly, PLEN is to be deployed according to the following interpretive rules:

Representation A protocol, π, represents a procedural epistemic rationality norm, N, and g represents agent i's epistemic dynamics iff

(1) π is executable at w in g iff i can comply with N at the state represented by w.

(2) transition $\langle w, w' \rangle$ executes π iff the transition represented by $\langle w, w' \rangle$ is a rational change of epistemic state for i according to N.

(3) process $\langle w_0, a_0, ...a_{n-1}, w_n \rangle$ executes π iff the process represented by $\langle w_0, a_0, ...a_{n-1}, w_n \rangle$ is a rational epistemic process for i according to N.

These rules are meant to partially characterize the representation of norms with protocols. The idea is that these rules tell us that a protocol π represents a norm N, with respect to a specific agent, i, whose dynamics are modeled by g, iff (i) the possibility of i's compliance with N correlates to executability of π in g and (ii) the verdicts of rationality delivered by N with respect to i's possible transitions and processes correlate to the execution sets of π

---

[16] Under the rules for canonically representing epistemic rationality norms below, every protocol in PLEN represents the procedural content of some norm; if there is nothing to an epistemic rationality norm other than its procedural content, then PLEN identifies norms with procedures.

in g. From these rules, it follows that the set of ideally rational evolutions of i's epistemic state according to N contains any state change that is mapped to π by the execution rules. Let me flesh this out.

PLEN is a single agent system in contrast to the more familiar multi-agent systems in the literature [65, 74]. Each EDG is thought of as a possible model of the epistemic dynamics of an arbitrary, ideal-but-bounded epistemic agent; the class of EDGs is the class of possible models of the ideal epistemic agent.[17] Call the ideal epistemic agent, i. For EDG, g, each element, w, of W represents a possible epistemic state for i. The content state, $c(w)$ represents the possible content of one of i's epistemic states; a possible configuration of belief, X, disbelief, Y, and explicitly endorsed methodology, Π_M, for i. Expressions built with the acceptance and rejection operators like $b(A)$ and $d(A)$ represent i's acceptance and rejection of A. Each transition represents a possible change of epistemic state for i that would result from some basic epistemic action, a (an element of Π_g). As above, each edge or labeled transition represents a token execution of a basic epistemic action. Each sequence of states represents a trajectory of epistemic state change for i that results from an epistemic process. Each procedural path represents a token epistemic process composed of sequential or iterated token executions of basic and complex actions. Each corresponding pair of procedural paths represents a sequential composition of parallel token executions of actions and processes. The execution rules map these sets of possible transitions and processes of epistemic state change to each basic action and each complex protocol formed from the basic actions.

So, the idea is that if a protocol, π, represents a procedural epistemic norm, N, then the processes that are mapped to π are interpreted as those that are rational according to N. Every norm defines a process type, a rational procedure for controlling an agent's trajectory through epistemic state space. Those processes in the complement of π's processes are the irrational processes, the processes diverging from the rational procedure. So, PLEN

---

[17] "Bounded" in generally the sense of [50]. The thought is that, an arbitrary ideal epistemic agent may still be seriously limited in terms of capacity and feasibility. Here, I'm not assuming that all ideal bounded agents are finite in particular respects (lifespan, possible epistemic states, possible actions). Accordingly, I'm not assuming that EDGs are finite, though attention may be restricted to finite EDGs for many purposes.

represents the ideally rational epistemic dynamics according to N with the processes that are mapped to π by the execution rules. Equivalently, for any EDG, g, that represents the dynamics of agent i, each protocol is mapped by the execution rules to a subgraph, g', of g if π can be executed in g. The subgraph, g', captures some of the procedural structure of π and so also of N. This is in virtue of the relations that any such g' has to the syntax of π. So, if π represents N, then, given the interpretation of PLEN proposed in the form of the representation rules above, g' marks out the possible rational epistemic actions, processes, evolutions of i's epistemic states according to N. This thought is the basic idea of applying PLEN to the formalization of epistemic norms. One can visualize this thought by imagining the rational procedure defined by a norm (i.e., its procedural execution set) as an overlay on top of the paths of EDG, as in Figure 3.6.

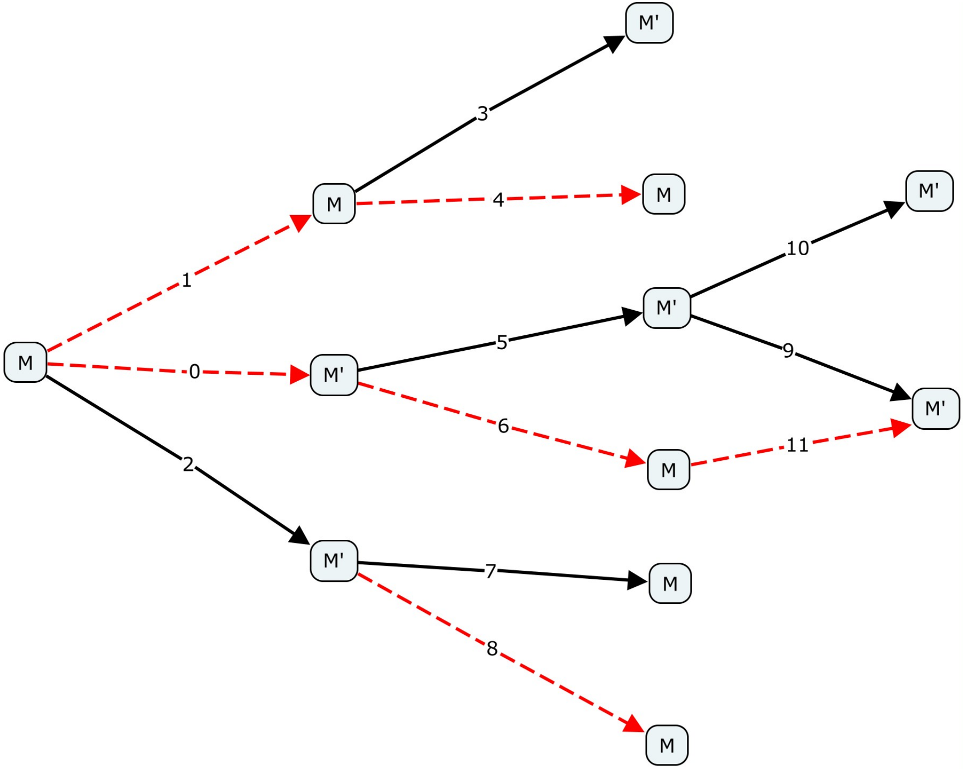
3.3.3 Preliminary Theoretical Advantages Over Other Theories

There is room for debate about how best to use the objects in PLEN to represent procedural epistemic rationality norms. The core representation rules form a plausible first stab that is expanded upon in Chapter 4. As such, the rules above codify the application of PLEN to the representation of epistemic rationality norms in the rest of this dissertation. Let me illustrate two of the theoretical advantages of taking this approach.

First, consider a very basic question about epistemic rationality norms: What are they? There are three plausible candidates. First, epistemic norms could be pieces of language or the propositions that encode their meaning; the set of epistemic norms just contains prescriptive expressions, normative statements, hypothetical imperatives, and norm-kernels. Second, epistemic norms could be regularities or patterns of actual or idealized epistemic behavior. Third, epistemic norms could be abstract objects like plans or programs. PLEN is a rigorous spelling out of third option, and it shows why the third option is superior to the other two.

Consider the first option. Thinking of epistemic norms as merely pieces of language

Figure 3.6: Restrictions as Overlays



Think of all of the arrows - the solid and dashed - as being part of the EDG above. The solid arrows execute the protocol $(\pi_2;\pi_7)\cup(\pi_5;\pi_{10}\cup\pi_9)\cup(\pi_3)$. Thus, the protocol defines a rational procedure displayed by the solid paths in this graph; the solid processes are the permissible processes. The rational procedure specified by a protocol under the interpretation rules is thus a subgraph of the EDG. The dashed arrows represent those paths that don't execute the protocol, and so are not within its designated permissible paths.

gives an impoverished and piecemeal account of them. It's impoverished because conceiving of epistemic norms merely as pieces of language overlooks their probative force. What is it about these pieces of language that matters so much for the regulation of epistemic behavior? It's piecemeal because the set of pieces of epistemically normative language is diverse, and lacks any obvious semantics that links them together. Thinking of norms as propositions might address the piecemeal objection, though it's hard to see how given the diversity of normative expressions (re: exactly this discussion in [89]). In any case, the probative force problem remains; it's mysterious what relation normative propositions like norm-kernels could have to rational epistemic behavior that wouldn't implicate restrictions on epistemic dynamics and so, in turn, implicate protocols. This point is more fully substantiated in the arguments from Chapter 2 connecting the function of epistemic norms to restrictions, and the arguments in Chapter 4 connecting restrictions to protocols.

Consider the second option. Thinking of epistemic norms as regularities in idealized epistemic behavior generates a vicious circle. In order to define an idealized model of epistemic behavior that can illuminate normative regularities, we can't just model actual behavior and then remove computational or temporal limits, we have to remove errors in performance. But epistemic norms are what define what counts as an error in performance. So, we need to have a way of determining what behaviors are epistemically normative prior to idealizing.

The third option is superior. Thinking of epistemic norms as protocols provides a systematic recipe for relating the normative pieces of language to regularities in epistemic behavior. This is one of the throughlines of Chapters 3 and 4. As argued in Chapter 2, epistemic normative language is semantically linked to restrictions on epistemic dynamics. As will be argued in Chapter 4, restrictions on dynamics are logically linked to protocols. Thus, normative expressions and propositions describe or convey some of the information that is associated with protocols. This can be shown in detail by crafting semantics for formal expressions encoding the pieces of languages in question. This is explicitly done for some normative expressions in Chapter 4, and hints as to how to do this are contained in the discussion above. This project is taken much further in the separate technical exploration of PLEN in [135]. Protocols are, in turn, linked to regularities in epistemic behavior, as described above. Every protocol, π, is associated with models of epistemic dynamics (P_π per EDG) that execute the protocol or otherwise carry out precisely the instructions contained in the protocol; errors are divergences from protocol. Thus, the idealization problem is solved; thinking of epistemic norms as abstract objects (protocols) that are systematically associated to some models of epistemic performance and not others gives us a way of defining errors in performance. These models are then associated with the pieces of language via the semantics gestured at above, giving a rich and unified account of epistemic norms in terms of protocols. Chapter 4 more fully spells out the advantages of the approach, but before moving on to PLEN's soundness, I consider one other advantage of PLEN's semantics.

Second, it's of fundamental importance to the dissertation's main arguments that, in applying PLEN to the formalization of epistemic norms, transitions are not mistaken for acts in PLEN, histories are not mistaken for processes, and feasibility is not mistaken for normativity. Given the foregoing definitions, and letting g represent i's epistemic dynamics, we can show how PLEN formalizes certain features of epistemic dynamics, constraints of feasibility and capacity:

- At any w in g, an action or procedure π is feasible for i iff it can be executed at w.
- An action or procedure π is within i's capabilities iff it can be executed at any w in g.

Thus, g represents (a) all of the possible epistemic actions and procedures that i could carry out (or undergo), (b) all of the possible changes of epistemic state that would result from them, and, consequently, the set of possible histories that i could instantiate, and (c) all of the acts and processes that comprise the agent's possible epistemic dynamics. Then, the protocols of PLEN carve these up in terms of the restrictions that they define. This enables the distinction of transitions from processes, and feasibilities and capacities from normative restrictions.

In PLEN, two basic actions, a and b, can be mapped to the same transitions in a graph, g, while being mapped to distinct sets of labeled transitions. The languages of PLEN can distinguish these facts. From the fact that transition $\langle w, w' \rangle$ executes a, it doesn't follow that a and b are equivalent in terms of executions, and there are thus classes of expressions in L_{PLEN} that distinguish them (e.g., $a \leftrightarrow b$ is false, $a \curlywedge b$ is false, etc.). This point gets stronger with complex protocols - a complex protocol π that is mapped to the same set of transitions as π' in g may have procedural characteristics (requiring the execution of a basic action c) that π' does not. Finally, with respect to any given protocol π, the feasibility of some process $\langle w_0, a_0, ...a_{n-1}, w_n \rangle$ can be distinguished from its rationality. After all, the feasibility of $\langle w_0, a_0, ...a_{n-1}, w_n \rangle$ at w_0 does not imply that it executes π at w_0, and thus it does not follow from mere feasibility that the process falls within the restriction defined by π. Thus, PLEN handles the Processes and Constraints problems, as will be shown in more detail in Ch.4.

In sum, PLEN's technical structures provide a model of epistemic rationality norms that analyzes their procedural content with protocols. These syntactical items are built using operations that map naturally onto procedural operations on actions, and are correlated to the graph-theoretic structure of EDGs which model epistemic dynamics, picking out submodels that execute the protocol. Protocols thus provide submodels that represent the rational dynamics with respect to a given norm. Finally, at this stage, we have some reason to think that PLEN is theoretically advantageous.

3.4 Conclusion

In sum, PLEN is composed of a language for codifying procedural operations on epistemic action, a language of modal, epistemic, and dynamic propositions about protocols, a semantics for both languages, and a sound deductive system for reasoning in the object language about protocols. The architecture of PLEN is that of a propositional dynamic logic with some additional machinery. Given PLEN's pedigree, it should be noted that PLEN provides ample tools for supporting and investigating (i) the process equivalence relations that are so frequently of note in the context of dynamic logics, i.e., graph isomorphism and bisimulation, and (ii) methods of devising new graph equivalence relations to analyze protocol equivalence relations. This work takes place in the forthcoming technical document [135].

The development of PLEN explicated in this chapter lays the groundwork for the solution of the Framework Problem. PLEN contains an exogenous language of protocols, and, interpreted as I've suggested in this chapter, thus an exogenous language of epistemic rationality norms. The semantics provide an array of set-theoretic constructions for modeling epistemic dynamics, protocols stand in for epistemic rationality norms in these models, and there are arrays of constructions for modeling the restrictions on epistemic dynamics that epistemic norms define. In the next chapter, these pieces are put together to solve the Framework Problem.

Chapter 4

PLEN as a Theory of Epistemic Rationality Norms

4.1 The Theoretical Adequacy of PLEN

This chapter shows that PLEN adequately formalizes the core theory of epistemic dynamics, epistemic rationality, and epistemic rationality norms defended in Part I, thereby solving the Framework Problem. Chapter 1 argued that, in order to adequately represent epistemic dynamics, a formal model must solve the Processes, Constraints, and First-Class Citizenship problems. A model of epistemic dynamics that fails to do so is incomplete or incorrect. Chapter 2 argued that principles D1 through D3 enumerate important features of epistemic rationality and its norms. If the arguments there are right, these principles are grounded in plausible considerations concerning the nature and structure of epistemic rationality norms, and must thus be accounted for by any serious account of epistemic rationality. These two lines of argument intertwine, thereby setting up the Framework Problem:

(Framework Problem) Any formal framework for reasoning about epistemic dynamics or the norms of epistemic rationality, F, that either fails to adequately represent epistemic dynamics or fails to adequately represent epistemic norms is either seriously incomplete

> or inaccurate with respect to its target domain. If a framework for epistemic dynamics fails to adequately formalize epistemic norms, it fails to be adequate with respect to epistemic dynamics. If a framework for norms fails to adequately represent epistemic dynamics, it fails to be adequate with respect to epistemic norms. Now, it is established in Chapter 1 that any framework that suffers the Processes, Constraints, or First-Class Citizenship problems (or related subproblems) fails to adequately formalize epistemic dynamics. It is established in Chapter 2 that any framework that fails to formalize desiderata/principles D1-D3 fails to adequately formalize epistemic norms. Thus, the fundamental problem is to devise a formal framework for representing and reasoning about epistemic dynamics, epistemic rationality, and epistemic rationality norms that (I) solves the Processes, Constraints, and First-Class Citizenship problems and related sub-problems, and (II) adequately formalizes principles D1-D3.

The clauses (I) and (II) represent important theoretical desiderata. Other things being equal, a theory that solves these problems and formalizes D1-D3 is preferable to a theory that doesn't. The main argument for accepting PLEN as a logic of epistemic rationality norms is, accordingly, a theory preference argument. The first step, carried out in this chapter, is arguing that PLEN is a solution to the Framework Problem.[1] More precisely, I attempt to prove the following proposition:

(The PLEN Proposition) **(i)** The semantics of PLEN contains structures that distinctly represent epistemic states, transitions, processes, actions, procedures, and constraints thereon.

(ii) The syntax and semantics of PLEN provide structures that explicitly represent epistemic rationality norms and their procedural content.

(iii) PLEN explicitly formalizes D1-D3; for each of D1-D3, there is a provable formal assertion at the meta-level of PLEN where formal assertions include: propositions

---

[1] Chapter 5 argues that PLEN is, among rival formalisms, the unique solution to the Framework Problem, thereby completing the theory preference argument.

> about the semantic structures of PLEN, sets of valid meta-level rules or object-level laws, theorems at the object-level, and object-level theories.

The plan of the chapter reflects the proof plan of this set of propositions. First, Section 2 discusses the devices with which PLEN represents states, transitions, processes, actions, procedures, and constraints; this is done with EDGs and their parts. This gives us clause (i) of the PLEN Proposition. Section 3 explores the devices with which PLEN represents epistemic rationality norms, thereby establishing clause (ii) of the PLEN Proposition. Here, two threads are pulled together. On the one hand, restrictions on epistemic dynamics are shown to unify norm-kernels and other forms of normative expression. On the other, restrictions on epistemic dynamics are shown to be reducible to protocols in that everything to know about restrictions can be derived from the syntactic and semantic features of protocols. Together, these threads bind protocols to epistemic rationality norms; these two lines of thought show that protocols are the devices in PLEN with which to represent and, by means of tools for reasoning about protocols, reason about epistemic norms. Section 4 walks through the formalization of D1-D3 in PLEN. This establishes clause (iii).

These arguments all assume the conception of representation of epistemic rationality norms by protocols defined in Chapter 3. That representation of the content of epistemic norms can be extended with an analysis of deontic operators in PLEN. Such an analysis strengthens the argument for clause (ii) of the PLEN Proposition, though the absence of such an analysis is not fatal to it. Accordingly, I've relegated a plausible first stab at formalizing deontic operators in PLEN to an appendix that follows the conclusion of the chapter; it's skippable but potentially interesting.

A final note of clarification: It is important to be explicit about exactly what the argument has been and what the rest of the argument is going to be. The thrust of the main argument for PLEN and the protocol-theoretic account of norms in this chapter is that the models of epistemic dynamics and epistemic norms provided by PLEN accurately captures the real features of epistemic rationality norms, which were adduced in Chapters 1 and 2. The thought is that Chapters 1 and 2 articulated at least some of the important features of epistemic dynamics and epistemic norms “out there in the world”. This chapter shows that, taking EDGs to represent epistemic dynamics, protocols do with respect to EDGs what epistemic norms do with respect to actual epistemic dynamics.

I will often write directly or loosely, and say that “epistemic norms are protocols”. But this is obviously false. Protocols are formal objects in dynamic logics. Whatever norms are, they aren’t those! This is like saying, in talking about gravity, that physical objects are basins of attraction in spacetime. That’s obviously false. Basins of attraction are just mathematical objects in topology. Whatever physical objects are, they aren’t just topological entities! Rather, basins of attraction *model* the gravitational properties of physical objects. The formal object and the real thing bear important structural relations with each other such that reasoning about the former can illuminate the properties of the latter. That is what I intend to show in this chapter, with respect to protocols and epistemic rationality norms.

4.2 Formalizing Epistemic Dynamics: Processes and Constraints

In this section, I argue that the EDGs model epistemic dynamics in a way that solves the Processes and Constraints problems by representing epistemic transitions in a manner distinct from that in which epistemic actions and processes are represented. As a bit of preliminary analysis, I want to explicitly lay out which parts of EDGs are intended to be the formal counterparts of which epistemic factors. This is done in Table 4.1.

The first six rows of relations are mapped out in Chapter 3 explicitly or implicitly. The essential thought being that EDGs are direct representations of the epistemic dynamics of an individual agent. The states are epistemic states with a content state explicitly capturing the epistemic judgments of acceptance and rejection. The transitions among epistemic states represent changes of epistemic state by the application of some basic or complex action -

Table 4.1: Representation of Epistemic Dynamics in PLEN

Epistemic States: Judgment and Content	$\Longleftarrow$ *Represents*	States: $W = \{w_1, ..., w_n\}$, Content States: $\langle X : Y : \Pi_M \rangle$, Content functions: $c(w)$
Epistemic State Transitions	$\Longleftarrow$ *Represents*	Transitions: $t = \langle w, w' \rangle$
Epistemic Action Tokens	$\Longleftarrow$ *Represents*	Edges: $e = \langle w, a, w' \rangle$
Epistemic Action Types	$\Longleftarrow$ *Represents*	Execution Sets for Atomic Actions: $R_a = \{e \in W \times \Pi_0 \times W \mid e \in W \times a \times W\}$
Epistemic Process Tokens	$\Longleftarrow$ *Represents*	Procedural Paths: $p = \langle e_1, ..., e_n \rangle$ where for each e_i and e_{i+1}, $last(e_i) = first(e_{i+1})$ and there are π_i and π_{i+1} such that $e_i \in R_{\pi i}$ and $e_{i+1} \in R_{\pi i+1}$
Epistemic Process Types (Epistemic Procedures)	$\Longleftarrow$ *Represents*	Execution Sets for Complex Protocols: P_π such that $p \in P_\pi$ iff p satisfies the relevant execution rules with respect to π
Feasibility of a Transition under C	$\Longleftarrow$ *Represents*	Let t be the transition. Then t is feasible under C iff $first(t) \Vdash C$.
Capability of a Transition	$\Longleftarrow$ *Represents*	Let t be the transition. Then t is within the capacities of agent i iff g represents i's epistemic dynamics and there is some w such that $w = first(t)$.
Feasibility of an Action under C	$\Longleftarrow$ *Represents*	Let e be the action token. Then e is feasible under C iff $first(e) \Vdash C$.
Feasibility of a Process under C	$\Longleftarrow$ *Represents*	Let p be the process token. Then p is feasible under C iff $first(p) \Vdash C$.
Capability of an Action	$\Longleftarrow$ *Represents*	Let e be the action token. Then e is within the capacities of agent i iff g represents i's epistemic dynamics and there is some w such that $w = first(e)$ (iff $e \in R$).
Capability of a Process	$\Longleftarrow$ *Represents*	Let p be the action token. Then p is within the capacities of agent i iff g represents i's epistemic dynamics and there is some w such that $w = first(p)$ (iff $e \in P_g$).

as each procedural path is ultimately bookended by a transition. Each basic action is an instruction to perform a basic epistemic action in the agent's repertoire or action library. Each complex protocol formed from these is a protocol, program, or plan of instructions to perform some procedural operations on basic epistemic actions. These are each mapped to parts of EDGs in ways that represent, respectively, action tokens and types and process tokens and types.

Since every EDG in PLEN is meant to represent a possible model of epistemic dynamics for a rational agent, capacity and feasibility come cheap. Each state is mapped to a propositional valuation. This valuation describes how the world is, so each node in an EDG represents a pairing of epistemic state and world state. The epistemic actions connect these in pairs. The feasibility of a state or condition relative property; a particular action is feasible at a state s or under a condition C just when that action can be carried out at s or under C. That is, given that EDGs are models of an agent's epistemic dynamics, the feasible actions for an agent at a state or under a condition are just those actions that can be carried out at that state or at the states verifying that condition.

The capacities of agents are relative to the entire space of possibilities in an agent's epistemic dynamics. That is, i is capable of carrying out an action (instruction) a iff there is some state that i could possibly get to at which a can be carried out.[2] That is, given that an EDG, g, is a model of an agent's epistemic dynamics, the agent's capacities are modeled by the possible actions that can be carried out anywhere in that model of possible epistemic actions.

Now, given this analysis, feasibility and capacity are equivalent to executability conditions. The action, $e = \langle w, a, w' \rangle$ is feasible for an agent iff that edge is part of the EDG

---

[2] Note that I exchange action symbols from the language of P_{PLEN} with what I'm calling actions, here, which are edges that execute action symbols. Really, symbols like a and π denote protocols and informally represent procedural epistemic norms whose content is essentially a plan or program of epistemic action formed out of procedural operations. The acceptability of these exchanges should be manifest; every action symbol is correlated to an action type (a set of edges) in every EDG, so when talking about one graph, the difference between a and R_a disappears. However, if talking more abstractly about action types that transcend the dynamics of a single agent, actions may have many concrete instantiations or tokens and what links these together is represented better by the instruction a than with a single action type in a graph.

that accurately represents that agent's epistemic dynamics. But this is equivalent to the executability of a at s in the model of the agent's epistemic dynamics. Similarly, the action is within an agent's capacities iff the action is executable somewhere in the model represents the agent's dynamics. Given that different EDGs may represent different agent's epistemic dynamics, these facts imply that the four distinctions between feasibility and capacity from Chapter 1 are formalized trivially:

- First, capacity and feasibility are distinct. An action a may be possible for an agent or within her capacities, but not feasible under a specified set of conditions.
- Second, capacity and feasibility may differ for distinct agents. Some agents can do under condition C what others cannot under C or at all.
- Third, the conditions that contribute to the feasibility of an action can be external facts in the world or they can be facts about what epistemic states the agents are in. That is, external and internal conditions can determine whether an action is feasible for an agent at a time.
- Finally, there are changes in what actions can be carried out by what agents over time. For example, given that epistemic states are a factor in determining what actions are feasible, the basic idea of epistemic dynamics implies that the set of feasible actions for an agent need not be constant from state to state. Learning and environmental changes can enable new actions or disable actions normally within one's repertoire.

Aside from the multi-agent aspect of these observations, these are all trivial features of EDGs![3] First, given g, what is feasible in g is not identical to what is in the capacities

---

[3] One can treat EDGs as multi agent systems in several easy ways. First, one could treat each EDG as an ideal model of a single agent's epistemic dynamics. Then, each graph gives the constraints of capacity and feasibility for an individual agent, and inter-graph differences encode inter-agent differences in capacity and feasibility. Simply compare states with identical valuations for executability of the same, syntactically defined procedures or actions. Second, index each action to a set of agents, Ag, relativize the content function so that each state gets a set of content states each indexed to agents in Ag. Then simply generalize the semantics of PLEN for each of these indexed actions and indexed program constructions. Then, for instance, π_i and π_j will be the same action for different agents, and their executability may come apart. The impact

defined by g; after all the set of protocols executable any particular s is not, in most EDGs, identical to the set of protocols executable at any state in g. Second, given g and g', what is executable at states verifying C in g need not be identical to the set of executable protocols at states verifying C in g', nor need the set of executable protocols in g be identical to that of g'. Third, because L_{PLEN} has both propositional expressions, A, and epistemic expressions, $b(A)$, the conditions that define the feasibility of any particular action may clearly be either internal facts about epistemic state or external facts in the world. Across EDGs that potentially model an agent's dynamics, we may even find (or stipulate, for the sake of drawing conclusions) regularities connecting internal and external conditions and the executability - at particular states or in general - of actions. Fourth and finally, it's clear that what is feasible need not be constant from state to state. If a is always feasible under c, it may very well be the case that for some transition, t, a is feasible at $first(t)$ and not feasible at $last(t)$ or vice versa. To be explicit, the observations about feasibility and capacity from Ch.1 are trivially formalized in PLEN.

From the foregoing, it can be shown that the Processes and Constraints problems are solved by PLEN. More precisely, it can be shown that none of the falsehoods that are validated by systems that conflate processes with transitions hold in PLEN. First, note the straightforward fact that edges in EDGs are distinct from transitions in EDGs. That is, EDGs are isomorphic to directed multi-graphs rather than directed graphs, and this implies that there may be more than one edge connecting any pair of states. Thus, the conflation is avoided.

Recall the Honest Twin case. Let the action of asking my twin for an answer to a validity problem be a, and let b be deploying my infallible validity detection method myself. Because PLEN correlates a and b with sets of triples with the middle term identical to a and b, more generally, it associates π and π' (where π and π' are any pair of complex procedural instructions) to paths, none of the problems raised in Ch.1 arise.

---

of agent actions them may impact both their own epistemic states, the world, and the epistemic states of others.

First, it was shown in Ch.1 that the conflation results in confusions about methodological inclusion. That is, b may be in my methodology at any particular content state while a is not, despite the fact that $R_a = R_b$ for all graphs, where R_x is the set of binary transition relations assigned to x as in PDL. And this fact is encoded in the semantics of the $M(\pi)$ formulae in L_{PLEN}. That is, L_{PLEN} can distinguish between one set of instruction's being included in a methodology at a state and another set of instruction's not being so included; there is a systematic recipe for determining the truth of these kinds of propositions. This is not so for systems with simpler languages or even for systems with the requisite language but that interpret propositions about actions over transition relations. Explicitly, because the expressions of L_{PLEN} that explicitly embed protocols are not interpreted over transition relations, the system can systematically distinguish $M(\pi)$ from $M(\pi')$ (one can be true at a state while the other isn't) even if the transition relations associated with π and π' are identical.

PLEN actually embodies two solutions to this problem. First, $M(a)$, which is the PLEN expression for what was denoted with $M_i(a)$ in Chapter 1, is interpreted in syntactic terms: $w \Vdash M(a)$ iff $c(w)$ is such that $a \in \Pi_M$. The syntactic non-identity of a and b does the work, though the methodological inclusion expression is interpreted over parts of the semantic structures. This immediately solves the problem articulated in Chapter 1, and does so in a way that is open to systems that do not distinguish edges from transitions, so long as those systems have content states. This emphasizes the point that actions and processes cannot be analyzed solely in terms of the transition relations associated with them. Second, if PLEN were to interpret methodological inclusion expressions in terms of edges rather than binary transition relations, the fact that $R_a = R_b$ would not imply $M(a) \equiv M(b)$. In fact, the uniqueness of edge labels implies an important connection between the syntactic solution and the ternary edge solution. In ad hoc notation, let R^e_a and R^e_b be the ternary edge sets for a and b, respectively. Then: $R^e_a \neq R^e_b$ iff $a \neq b$. This is due to the uniqueness of edge labels. Thus, $R^e_a \neq R^e_b$ implies that the syntactic solution is available, and the syntactic solution implies that the ternary edge solution is available. Thus, whether $M(a)$ and $M(b)$ are interpreted in terms of either syntax or edges, $M(a) \equiv M(b)$ might fail even if $R_a = R_b$.

As a result, the validation of falsehoods about methodological feasibility no longer holds. If a and b are mapped to the same transitions, then they are executable under equivalent circumstances. But this doesn't imply that the formula from Ch.1, $\langle M \rangle \top$, is true just because b can be executed and $a \in M$. This transparently fails in PLEN; let $w \Vdash \langle M \rangle \top$ iff some π such that $w \Vdash M(\pi)$ is executable at w. Then, the mere fact that $R_a = R_b$ does not imply that $b \in M$, and so it's perfectly possible that $w \nVdash \langle M \rangle \top$. The difference maker is that methodological inclusion is not interpreted in terms of R_a and R_b, but determined by syntax.

Second, where $A(\pi)$ and $A(\pi')$ encode procedural information about π and π', say convergence, divergence, dictation relations, or implementation relations among protocols, the mere fact that π and π' are associated with identical transition relations does not lead to confusion between $A(\pi)$ and $A(\pi')$. Again, this is simply because any such expressions will be interpreted not with respect to the transition relations associated with them but in terms of relations between syntactic features of π and π' themselves, the methodologies at states, and the graph-theoretic features of EDGs. Explicitly, $Com(\pi)$ and $R_\pi = R_{\pi'}$ do not imply $Com(\pi')$ for any of the procedural statements described above.

Finally, given the representation rules articulated in Chapter 3, the rationality and irrationality of actions and processes is determined not by transition relations associated with protocols, but with edge and procedural path sets associated with them. This means that, given the multi-graph structure of EDGs, even if π and π' are associated with the same transition relations, there may paths that execute π that do not execute π' and vice versa, thereby distinguishing the verdicts that π and π' deliver with respect to any particular process p. To sum, PLEN contains structures that explicitly represent epistemic processes as distinct from epistemic transitions and contains structures that represent constraints on epistemic processes (feasibility and capacity of processes) as distinct from constraints on epistemic transitions (feasibility and capacity of transitions).

Epistemic norms are represented with protocols in PLEN, and protocols are not interpreted with transition relations but with ternary or larger polyadic relations (procedural paths). This enables PLEN to formalize several concepts of equivalence much stronger than mere identity of transition relations. Conflation of processes with transitions resulted in validation of propositional equivalences, $A(\pi) \equiv A(\pi')$, between feasibility, procedural, methodological, and normative statements about the instructions or norms that protocols represent that were simply false. If the procedural content of protocols just comes down to transition relations, and propositions about protocols are interpreted in terms of their procedural content, then identity of transition relations yields about as strong an equivalence between π and π' as can be had. This protocol equivalence relation is what implied the equivalences above; in fact it's really a syntactic replacement result for a much simpler, more confused system than PLEN. However, if the procedural content of protocols is analyzed in the more complex terms of procedural paths, we get stronger protocol equivalence relations that imply quite a lot. However, because propositions about protocols will no longer be interpreted in terms of transition relations, the propositional equivalences noted above will all fail.

Thus, the Processes and Constraints problems are solved. The Constraint Unfolding problem can be solved simply by comparing EDGs, g and g', such that g can be properly embedded in g' (the states of g can be mapped to those in g' in such a way that all executability facts about g are preserved in g' and g' satisfies further executability facts). Between these two graphs, one can determine what it would look like if some constraints modeled in g did not hold. That is, at some state, s, in g, a might not be executable because its preconditions are not met. But in g', a might be executable at s's counterpart.

Finally, the Reasoning problems are trivially solved by the metalogical reasoning about epistemic dynamics and protocols possible in PLEN and also in the object level deductive system of Syntax II. The latter provides an explicit encoding of valid forms of reasoning about protocols (hence, epistemic norms) and the features of epistemic dynamics that (e.g.) compliance with norms determines. The differences between formulae like $\pi_1 \leftrightarrow \pi_2$ and $[\pi_1]A \equiv [\pi_2]A$. The former provides explicit mathematical structures for reasoning rigorously about processes as distinct from transitions as well as constraints on both of these and different constraints across EDGs. I attack these problems in more detail in [135].

4.3 Formalizing Epistemic Norms: First-Class Citizenship

In this section, I argue that the syntax and semantics of PLEN provide structures that explicitly represent epistemic rationality norms and their procedural content, thereby solving the First-Class Citizenship problem. To phrase it bluntly: epistemic rationality norms are protocols in PLEN. To phrase it a bit more delicately, protocols are the formal structures in PLEN that represent epistemic rationality norms. The thought is that norms straightforwardly instruct the selection and construction of courses of epistemic action the way that programs instruct the selection and construction of courses of computational action. It takes a bit of work to see this.

The argument that follows essentially shows that protocols can serve as constructions from which the verdict information of epistemic norms can be read off. These structures are restrictions on epistemic dynamics. Each protocol defines, for each EDG, a division of its processes into rationally permissible and rationally impermissible sets, these restrictions in turn specify restrictions on paths of states, transitions, and terminal states. More important, for every such restriction on states, transitions, and paths, there is a protocol such that the verdicts of these restrictions can be inferred, by means of the execution rules, from the syntax of the protocol. Protocols are thus PLEN's stand-ins for epistemic norms.

Additionally, the structures in PLEN with which protocols are associated represent indeterministic compliance with epistemic rationality norms. The contrast between these structures and EDGs in which they can be embedded represents how epistemic dynamics unfolds under normative restriction as contrasted to how it would unfold without the normative restriction. This solves the Unfolding problems noted in Chapter 1.

Finally, the foregoing arguments, spelled out and defended in greater detail below, are strengthened by certain formal results in PLEN. These results might be thought to lack intuitive force or philosophical interest in the case that PLEN doesn't offer a plausible analysis of deontic notions, and, though far from decisive, this might weaken the overall solution that PLEN offers to the First-Class Citizenship problem. To the end of resolving this worry, there is a technical appendix giving a plausible - if preliminary - analysis of epistemic deontics in PLEN.

4.3.1 How PLEN Solves the Problem

The basic thought of the solution is that protocols can be used to reduce a large class of restrictions on procedural construction, which in turn reduces wide classes of other kinds of restrictions on epistemic dynamics, which in turn reduce all of the important features of a wide class of expressions that state the content of epistemic rationality norms. The "important features" in this argument concern only the categorical and deontic verdicts of the norms and the assignment of objects to rational and irrational sets by restrictions. The following argument states the main moves of the solution more explicitly:

1. For all members of a large class of expressions that plausibly state the content of epistemic norms, all of the most important things to derive about that content - the categorical and deontic verdicts assigned to objects by epistemic rationality norms under their conditions of application - can be derived from restrictions on dynamics of the following kinds:

    (a) Restrictions on states

    (b) Restrictions on transitions

   (c) Restrictions on processes

2. Epistemic norms (also) restrict the construction of procedures; there are restrictions on how epistemic actions can be sequenced, iterated, combined, chosen among, etc., and these are defined by some epistemic norms.

3. For each member of the class of restrictions on dynamics that account for the content of the norm expressions considered, there is a class of restrictions on procedural construction such that the verdicts of each such restriction can be used to derive the categorical and deontic verdicts of the other forms of restriction on dynamics. More precisely:

   (a) For every restriction on processes, there is a (class of) corresponding restriction(s) on procedural constructions. The converse also holds.

   (b) For every restriction on transitions, there is a (class of) corresponding restriction(s) on procedural constructions.

   (c) For every restriction on states, there is a corresponding (class of) restriction(s) on procedural constructions.

4. For each member of the foregoing class of restrictions on procedural construction, there is a protocol in PLEN from which the verdicts of that restriction on procedural construction can be read off syntactically.

5. ∴ For each member of the class of restrictions on dynamics that account for the content of the norm expressions considered, there is a protocol from which the categorical and deontic verdicts of that restriction on dynamics can be derived (even if only by means of reading off the procedural restriction).

6. ∴ For all members of a large class of expressions that plausibly state the content of epistemic norms, there is a protocol from which the categorical and deontic verdicts

> of that normative content can be derived (even if only by means of reading off the procedural restriction).

In simpler terms, protocols unify many sorts of normative expression by doing exactly what we would expect epistemic rationality norms to do with respect to epistemic dynamics; specify restrictions on them, thereby dividing the good from the bad (the permissible from the prohibited, the rational from the irrational).

4.3.2 Details of the Solution: Premise 1

The thought behind the first premise is that the central kinds of epistemic norms (categorical and deontic norms) and the different pieces of natural language that encode them (including norm-kernels) are all unified - in a conceptual, informational, or pragmatic sense - by restrictions on dynamics. Where I talk about the "content" of epistemic norms, I am speaking loosely. I want to make no claims about what the contents of thoughts expressed by normative claims are or what the best semantics of those expressions is. Accounting for the contents of thoughts or best semantics is, at this stage of the project, simply too big a task to take on. Rather, in speaking of the "content" of norms, I intend to denote the information that is associated with or conveyed by epistemic norms that is useful for complying with, applying, or responding appropriately to them. This is a point I raised in Chapter 2, in discussing restrictions on dynamics and the cognitive role of epistemic norms. There, I pointed out that in order to comply with norms, we need to be able to figure out what verdicts they assign to their objects under their conditions of application. This information is the "content" at which I am aiming.

There are several grammatically and/or semantically distinct types of expressions that are called epistemic norms. Perhaps the standard account of norms, following von Wright, identifies norms with norm kernels: conditionals that link normative expressions (assignments of valences to objects) to conditions of application, as described in Chapters 1 and 2. But it's far from clear that these kinds of conditional deontic expressions are actually the only or the best candidates for representing the content of epistemic norms. There are, after all, numerous types of norms and numerous kinds of natural language construction that are plausible candidates for encoding each of these types of norms.[4] While von Wright argues that norm-kernels are the core form of norms, this is disputed by imperativalists [250, 251], who argue that norms are, at their core, imperatives or commands. The spiraling apart and occasional reconvergence of deontic and imperative logic [89] is driven by the plausibility of both of these thoughts. More, natural language is straightforwardly noncommittal about how norms are represented. We state norms with constructions including imperatives, straightforward declaratives, and other kinds of rules or requirements, and it's not clear that any has any privileged claim to truly representing norms.

All of these types of normative expression (for the epistemic case) can be understood as stating properties of restrictions on epistemic dynamics; they state what kinds of epistemic factors go into what parts of the restriction. A restriction on dynamics is, informally, a sorting of the possible changes in the world into different normative categories like rational or irrational, legitimate or illegitimate, moral or immoral. The thought, then, is that epistemic norms - however they are expressed - define restrictions on epistemic dynamics; they tell us about the properties and contours of restrictions. Restrictions on epistemic dynamics thereby unify all of these types of normative expression (for the epistemic case). Explicitly, all epistemic normative expressions are associated with a restriction on epistemic dynamics that, from the arguments of Ch.2, serve the normative expression's essential cognitive roles. This is what I mean when I say that restrictions on dynamics unify these expressions. There are a few direct arguments for this premise.

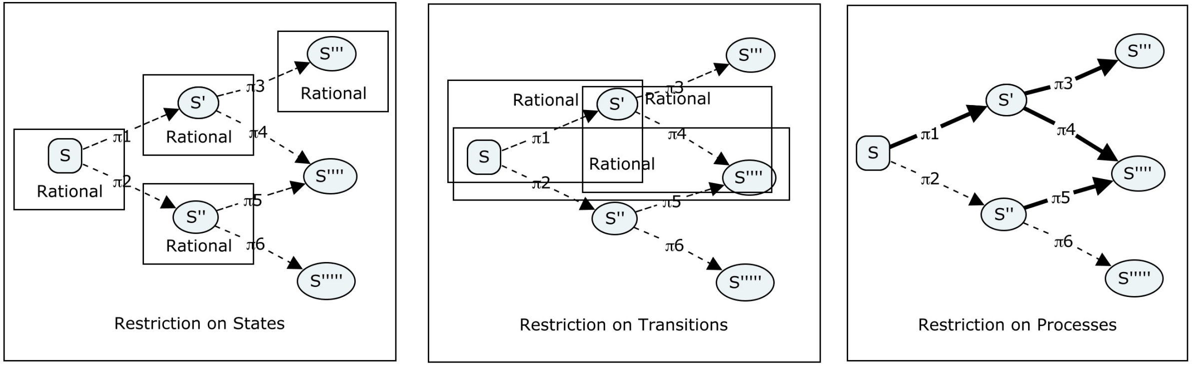
First, recall restrictions on epistemic dynamics:

(Restrictions on Epistemic Dynamics) Consider some set S of components of epistemic dynamics; these could be states, transitions, processes, or some specified subset thereof. Let $P(S)$ be the powerset of S. Consider some partial order relation $\prec$ on the elements,

---

[4] See Chapter 1 of von Wright [261] for a much finer typology than I'll be giving, here.

Figure 4.1: Restrictions on Epistemic Dynamics



> P_i, of $P(S)$. Call some designated element of $P(S)$, T_I, the standard of irrationality. Now, let R_S and I_S be subsets of S such that: $\forall P_i \in P(S) \forall s \in P_i(P_i \preceq T_I \text{ iff } s \in I_S)$, and $\forall P_i \in P(S) \forall s \in SP_i(P_i \succ T_I \text{ iff } s \in R_S)$ Then: every $R_{\prec} = (P(S), \prec, T_I, R_S, I_S)$ is a *restriction* on epistemic dynamics.

Formally, the basic idea of restrictions is that restrictions involve rankings over the powerset of a set of elements of epistemic dynamics, and these can be thought of, diagrammatically, as boundaries separating different classes of epistemic factors into rational or irrational, permissible or impermissible sets. Recall the information conveyed in Figure 4.1.

Restrictions on dynamics are complex tools for capturing the content of norms. The core thought is that every epistemic rationality norm, in order to function in ante hoc regulation or epistemic planning (re: Chapter 2), is associated with a set of restrictions on epistemic dynamics that encode its verdicts. The powerset over the components of dynamics, S, breaks up the set of components by property. Each P_i in $P(S)$ can be thought of as the extension of some property of these components, e.g., consistency of states or reliability of processes. The subset R_S is the set of rational components, and I_S is the set of irrational components. The ranking may often do no work, but it defines something like a relation of "more rational than", where this is relevant to the verdict assignments of norms.

Generally speaking, every expression that employs a normative term or a deontic expression obviously defines a set of restrictions: the objects of the normative expressions are directly or conditionally assigned either to (e.g.) rational or irrational sets (R_S or I_S), or placed somewhere in the ordering of subsets of the domain. In the central case of deontics, the objects of norms are assigned to deontic categories, and this can be captured by restrictions. All of those actions that are rationally permissible compose the rational set, those that are prohibited are in the irrational set. The obligatory actions are simply those actions such that either (i) they are the uniquely permissible actions under some condition or (ii) they must be executed in addition to any other permissible action. All one needs to analyze obligation in terms of restrictions is machinery for thinking about these kinds of features of action. All restrictions whose orderings match these assignments to rational and irrational sets are models of such deontic norms.

Taking the analysis of normative expressions in terms of restrictions further, normative terms like “right” and “wrong”, “legitimate” and “illegitimate”, “rational” and “irrational” obviously define simple sets of restrictions. Conditional rules straightforwardly define restrictions, placing those contents that satisfy C into whatever category is employed in their consequents. Normative declarations identify elements as members of the rational or irrational sets in a restriction. The deployment of evaluative terms rather than categorical normative terms or deontics is also analyzable in terms of restrictions; positive evaluative terms imply that, under conditions of parity, categorical and deontic restrictions apply. To say that x is correct or that one has reason to a is just to say that if x or a is just like everything else among one’s available options but that it satisfies the evaluative property in question, one is rationally obligated to attain x or do a (one is irrational not to). Normative claims that employ comparative terms are similarly analyzable in terms of restrictions; where x is more F than y, one is ceteris paribus rationally obligated to choose x over y (it’s irrational not to). Thus, given that deontics are appropriately represented by restrictions on dynamics, even these kinds of normative claims are unified by restrictions.

The case for premise 1, then, is that sets of restrictions on dynamics can be given that capture the content of any familiar class of epistemic norm expressions. See the Table 4.2 for a loose typology of normative expressions and the categories that correspond to R_S and

I_S in any restriction associated with each.

All of the natural language constructions that look like norms in the typology given in Table 4.2 can be understood - at least insofar as the information necessary to comply with them or respond appropriately to them is concerned - entirely in terms of restrictions on epistemic dynamics. All of the foregoing, I think, makes the first premise highly plausible. But a stronger argument can be given.

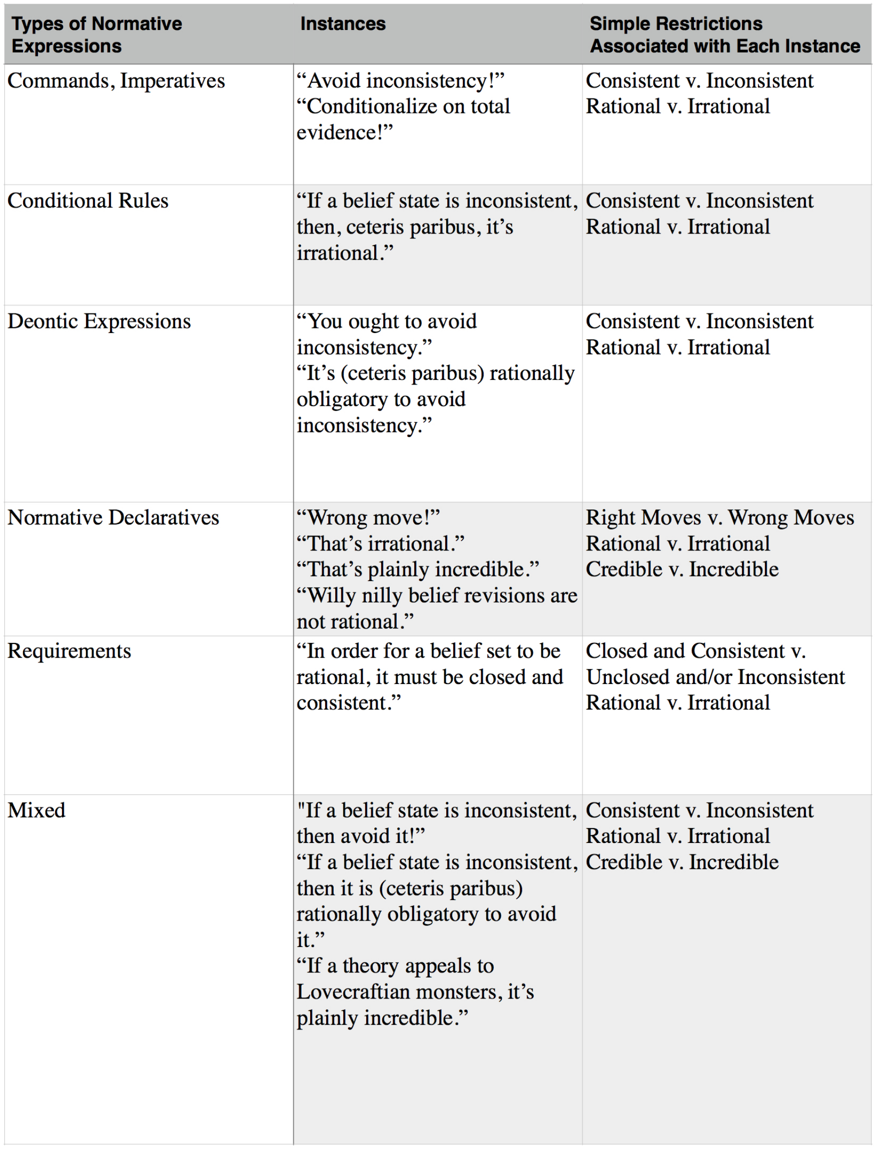
4.3.2.1 Norm-kernels, Norm Functions, and Protocol Functions

The argument focuses on norm-kernels. It will show that, insofar as the most ubiquitous classes of epistemic rationality norms have their most important features - what factors they assign to the rational and irrational sets under what conditions - captured by norm-kernels, these features of these classes of norms are captured by restrictions on processes, transitions, and states. That gives us premise 1. From there, the rest of the argument in section 4.3.1 tells us that the most important properties of these restrictions can be read off of restrictions of procedural construction, and that the most important features of procedural restrictions can ultimately be read off of protocols themselves.

So as to properly recognize its scope, I want to note explicitly that the argument for premise 1 quantifies over only those restrictions on epistemic dynamics that capture the content of norms that are (i) straightforwardly categorical (i.e., those norms that explicitly divide epistemic factors into rational or irrational subsets) or (ii) deontic. The partial ordering on subsets of epistemic factors cannot be captured by the properties of the constructions in PLEN, as it stands. However, as suggested in Chapter 6, a modification of PLEN with weighted paths will plausibly do the trick. With that caveat out of the way, let's move onto the argument itself.

First, suppose, along with [261] and many others [164], that norm-kernels express the core content of normative expressions. Every norm-kernel (hereafter, this term is interchangeable with "norm") maps conditions of application and objects (inter alia) to a valence. Thus, norms

Table 4.2: Preliminary Typology of Epistemic Norms and Correlated Restrictions on Dynamics



Types of Normative Expressions	**Instances**	**Simple Restrictions Associated with Each Instance**
Commands, Imperatives	“Avoid inconsistency!” “Conditionalize on total evidence!”	Consistent v. Inconsistent Rational v. Irrational
Conditional Rules	“If a belief state is inconsistent, then, ceteris paribus, it’s irrational.”	Consistent v. Inconsistent Rational v. Irrational
Deontic Expressions	“You ought to avoid inconsistency.” “It’s (ceteris paribus) rationally obligatory to avoid inconsistency.”	Consistent v. Inconsistent Rational v. Irrational
Normative Declaratives	“Wrong move!” “That’s irrational.” “That’s plainly incredible.” “Willy nilly belief revisions are not rational.”	Right Moves v. Wrong Moves Rational v. Irrational Credible v. Incredible
Requirements	“In order for a belief set to be rational, it must be closed and consistent.”	Closed and Consistent v. Unclosed and/or Inconsistent Rational v. Irrational
Mixed	"If a belief state is inconsistent, then avoid it!” “If a belief state is inconsistent, then it is (ceteris paribus) rationally obligatory to avoid it.” “If a theory appeals to Lovecraftian monsters, it’s plainly incredible.”	Consistent v. Inconsistent Rational v. Irrational Credible v. Incredible

relate objects of epistemic appraisal to conditions and normative properties or valences. Because of their conditional form, they can be thought of as functions from pairs of conditions and objects to valences or normative properties of objects.

Let n be a norm. Let Val be a set of valences of epistemic rational appraisal, the normative properties attributed to the objects of epistemic appraisal. In other words, valences encode the verdicts of norms, and so indirectly encode their requirements. For concreteness, consider only the norms that apply deontic operators like rational obligation, rational forbiddance, and rational permissibility to their objects. Again, it's a very plausible conjecture that all of the remaining argument can be extended to more complicated norm-kernels, but this is left to future work.

Let $Dom(R)$ be the domain of rationality, the objects that can be appraised for the valences in Val. The domain of norms of rationality contains actions, conditions, and epistemic states. Let Con be the set of conditions expressible in PLEN. That is, assume Con is a subset of L_{PLEN}. Let $nDom(R) \subseteq Dom(R)$ be the objects that n applies to. Let $nCon \subseteq Con$ be the conditions of application of n. Then:

Norm Functions Norm-kernels have the functional form: $f_n : (nCon, nDom(R)) \longrightarrow Val$

The general reason to think that norm functions are suitable tools for approximating norm-kernels is the basic form of norm-kernels. They associate conditions of application with assignments of deontic valences to their objects. There must thus be some norm function that agrees with each norm-kernel. For instance, consider the norm-kernel:

Indiscriminate Consistency If your beliefs are inconsistent, then, if possible, it's rationally obligatory to revise them in a way that removes the inconsistency by any procedure, plan, or program that effectively does so.

This rule can be formalized by means of protocol functions. First, define the following constructions:

- Let $nCon = \{s \mid (b(A) \wedge b(\neg A))\}$

- $nDom(R) = \{\pi \in P_{PLEN} \mid ([\pi]\neg b(A)) \vee ([\pi]\neg b(\neg A))\}$
- $Val = \{O_M, P_M, F_M\}$

Read the elements of Val as deontic operators with the flavor of epistemic rationality, but relativized to a methodology. For instance, "rationally obligatory", O_M, is read as obligatory with respect to the union of all norms in the agent i's methodology. Then our toy norm - inadequate as it may be - can be formalized with the following norm function:

Indiscriminate Consistency Function $f_n : (nCon, nDom(R)) \longrightarrow Val$ such that:

> $f_n(s, \pi) = P_M$ iff $s \in nCon$, $\pi \in nDom(R)$, and $s \Vdash \langle\pi\rangle \top$, or $s \notin nCon$ and $s \Vdash \langle\pi\rangle \top$, or $\neg\exists\pi'(\pi' \in nDom(R) \wedge s \Vdash \langle\pi'\rangle \top)$
>
> $f_n(s, \pi) = O_M$ iff $s \in nCon$, for all $\pi' \in nDom(R)$, it holds that $s \Vdash \pi \curlywedge \pi'$, and $s \Vdash \langle\pi\rangle \top$
>
> $f_n(s, \pi) = F_M$ iff $s \in nCon$, $s \Vdash \langle\pi\rangle \top$, and $\neg\exists\pi'(\pi' \in nDom(R) \wedge \pi \curlywedge \pi')$, and $\exists\pi'(\pi' \in nDom(R) \wedge s \Vdash \langle\pi'\rangle \top)$,

At states at which inconsistencies are believed, this function maps executable protocols that result in the elimination of the inconsistency to rational obligation (with respect to methodology M). It maps executable protocols that don't remove the inconsistency to rational permissibility where there are no ways to remove the inconsistency. It maps executable protocols to rational prohibition or forbiddance where they do not remove the inconsistency and there are available methods to do so. Note that the function is entirely indiscriminate with respect to what protocol one uses to eliminate the inconsistency if there are many. It arbitrarily selects some executable inconsistency-removing procedure. It's a mechanical task to extend this kind of analysis to norms that states, transitions, actions, or processes - under arbitrary conditions - to other valences.

Now, the crux of the argument is the following thought:

Norm-Kernel Analysis For any norm-kernel, N, if all conditions of application relevant to N are expressible in L_{PLEN} and all of N's objects are actions, expressible conditions, or states in PLEN, and all valences with respect to N are expressible in PLEN, then there is a norm-function, f_N, that represents that norm-kernel by mapping N's conditions of application and objects to its valences. For every norm-function, f_N, there is a protocol, π, whose universal protocol function, $f_\pi^\forall$, agrees with f_N for every pair of condition of application and object.

The idea behind Norm-Kernel Analysis is that protocols do the same thing that norm-kernels do; they map permutations of pairs of (i) conditions of application expressible in L_{PLEN} and (ii) objects (elements of epistemic dynamics) to valences, which are analyzed in terms of execution-theoretic properties of protocols. This thought actually goes further than we need it to; it goes all the way to the fundamental argument for the solution of First-Class Citizenship by showing that the verdicts of norm-kernels can all be read off of some protocol. I take this up in subsection 4.3.6. Here, however, I want to focus on a part of this thought to establish the first premise above.

The protocol functions of π, as I define protocol functions below, are explicitly defined in terms of the restrictions on dynamics that π is mapped to by the execution rules of PLEN: the execution set of π. Thus, because the universal protocol function of π completely agrees with the norm function of norm-kernel N, the restrictions on dynamics that are associated with π by the semantics of PLEN are themselves constructions from which all of the verdicts of N can be inferred. I now want to try to prove Norm-Kernel Analysis.

To prove Norm-Kernel Analysis, first define protocol functions. For each graph g in which a protocol, π, is executable, π defines a set of conditions, $\pi Con(g)$, a set of objects, $\pi Dom(R)(g)$, and a set of valences $\pi Val(g)$. The idea is that $\pi Con(g) \subseteq L_{PLEN}$ and $\pi Con(g) = \{A \mid \vdash_g (A \equiv \langle \pi \rangle \top)\}$. It's the union of the preconditions of π in g. These are the conditions of application of π, a subset of all definable conditions in L_{PLEN}.

Now, let $\pi Dom(R)(g) = \pi Dom(R)Act(g) \cup \pi Dom(R)Con(g) \cup \pi Dom(R)Stat(g)$ be the domain of π. The general idea is that protocols, in virtue of the mechanisms of PLEN, bear all sorts of relations to protocols, states, formulae of L_{PLEN}, and so on. For instance, let $\pi Dom(R)Act(g) \subseteq P_{PLEN}$, let $\pi \equiv_p^g \pi'$ mean that π and π' have identical sets of procedural paths in graph g, and $\pi Dom(R)Act(g) = \{\pi' \mid \exists \pi'' \in Sub(\pi)(\pi' \equiv_p^g \pi'')\}$. Explicitly, $\pi Dom(R)Act(g)$ is the set of protocols that are procedurally equivalent with some subprotocol of π; it's the set of actions or procedures implicated in executions of π. Every execution of π in g executes some subset of $\pi Dom(R)Act(g)$. These actions are all in the purview of the restriction of π. $\pi Dom(R)Con \subseteq L_{PLEN}$ and $\pi Dom(R)Con(g) = \{B \Vdash_g ([\pi]B)\}$. That is, $\pi Dom(R)Con(g)$ is the set of strong postconditions of π in g; they're the conditions that are brought about at the states reachable by π. $\pi Dom(R)Stat(g) = \{s \in S \mid \exists p \in P_\pi(s = last(p))\}$. This is the set of states that π can bring a system to via its executions. In other words, these are the states reachable by π. In sum, $\pi Dom(R)(g)$ is the set of objects of π in g; the things that the protocol can "assign" deontic valences to.

The deontic properties that an action, postcondition, or state can have relative to π are contained in $\pi Val = \{Ob, Per, Forb\}$.[5] These are the valences of the objects of π. The idea is that actions, conditions, and states can have various relations to the restrictions defined by π that encode deontic operators. The analysis of deontic operators in terms of protocols and restrictions on dynamics will occupy a brief appendix to this chapter. The key, however, is that the analysis of the deontic operators in terms of protocols and restrictions on dynamics is plausible, and so each protocol function maps the conditions of application and the objects in the domain of a protocol to the protocol-theoretic analysis of deontic operators. Please grant me that for the moment.

Given the foregoing definitions, every protocol π generates a function, f_π^g, for every EDG g in which π is executable:

Protocol Functions $f_\pi^g : (\pi Con, \pi Dom(R)) \longrightarrow \pi Val$.

---

[5] Again, only focusing on the basic deontic operators. With weighted graphs, any comparative property relative to π, might be added to πVal, but this is a conjecture to be spelled out in future work.

I omit the details of the specification of such functions; they are spelled out in the appendix to this chapter and in the technical document [135]. Given this fact, for each set of $g_1, ..., g_n$ in which π is executable, a universal protocol function for that set can be defined that agrees with each f_π^{gi} for each g_i:

Universal Protocol Functions $f_\pi^\forall : (\pi UCon, \pi UDom(R)) \longrightarrow \pi Val$ where $\pi UCon = \pi Con(g_1) \cup ... \cup \pi Con(g_n)$ and $\pi UDom(R) = \pi Dom(R)(g_1) \cup ... \cup \pi Dom(R)(g_n)$ for all g_i in which π is executable at any state.

Define $f_{Val} : Val \longrightarrow \pi Val$ as a bijection from the objects in Val to the objects in πVal. Then, the following proposition holds:

(PFT) *For any f_n, with any $C \in nCon$ and any $x \in nDom(R)$, if $C \in L_{PLEN}$ and x is an action or procedure in PLEN, a condition expressible in L_{PLEN}, or a state or class of states in the semantics of PLEN, then there is a protocol π such that the universal protocol function $f_\pi^\forall$ for arbitrary $g_1, ..., g_n$ in which π is executable is such that $f_\pi^\forall(C, x) = f_{Val}(f_n(C, x))$.*

The proof of PFT is not difficult, and it's left to the technical document [135]. This can easily be seen to suffice for Norm-Kernel Analysis. Given that, as argued above, Norm-Kernel Analysis suffices to show that restrictions on dynamics can be used to infer everything we need to know about norm-kernels (i.e., their assignments of valences to objects under conditions of application), we have the first premise.

4.3.3 Details of the Solution: Premise 2

The various forms of restriction on epistemic dynamics in Premise 1 are established by D1 and D2. What's interesting is the possibility of a further category of restriction on epistemic dynamics. The observation of plausible sets of methodological norms from various fields of logic and epistemology suggest a further category of restriction: restrictions on the construction of epistemic procedures. More precisely, as the second premise states, there are numerous plausible methodological norms that define restrictions on procedural operations on epistemic actions.

Informally, a restriction on procedural constructions is simply a restriction on the possible ways that epistemic actions can be sequenced, iterated, chosen among, executed, conditioned on tests, etc.. Formally:

(Restrictions on Procedural Construction) Let S be the set of all possible procedures constructible out of any set of basic actions and any set of procedural operations. Then every restriction on S, $R_{\prec} = (P(S), \prec, T_I, R_S, I_S)$, is a *restriction on procedural constructions.*

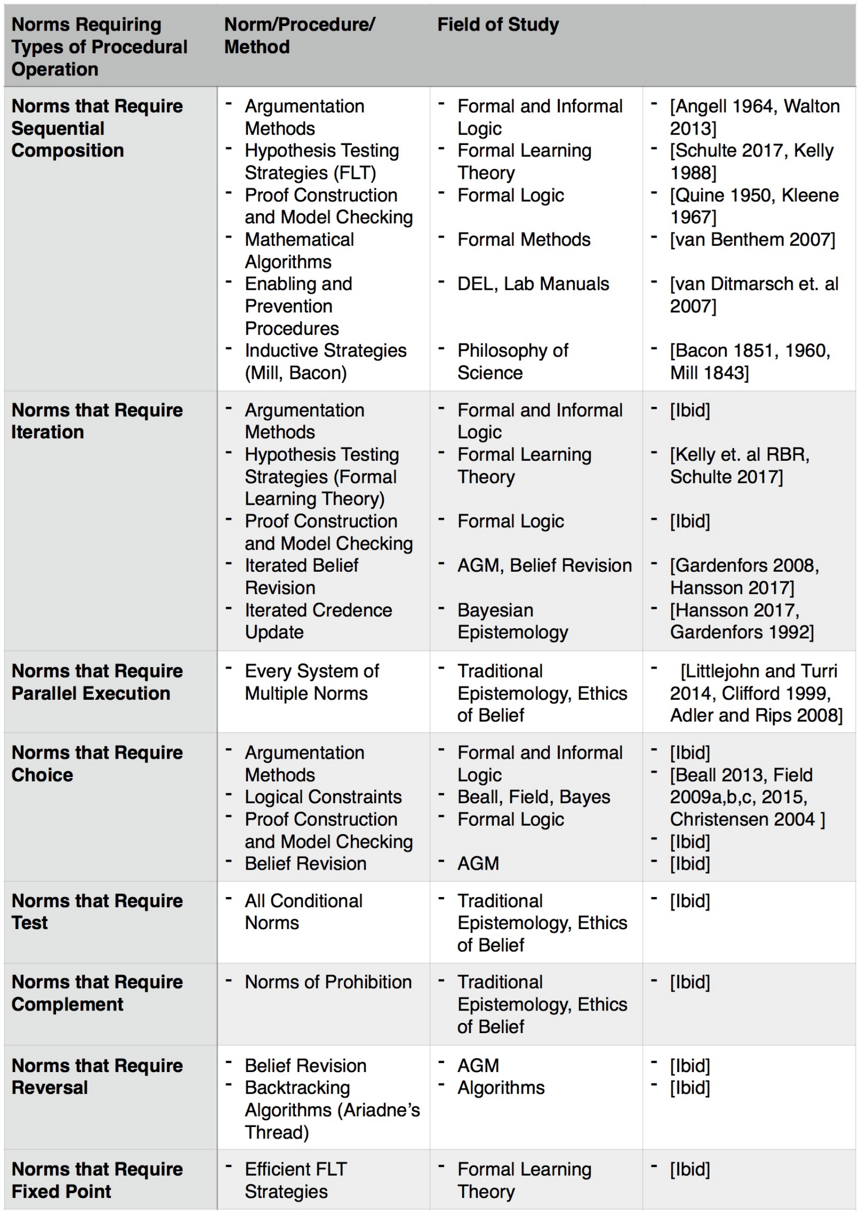
Any rule that specifies that some ways of building a procedure out of constituent actions are rational and others are irrational defines such a restriction. Every such rule defines a set of procedures that "match" the rule and a set of procedures that fails to match it. Let R_S be the set of matching procedures, I_S be the non-matching procedures, and $\prec$ be some suitable similarity relation on the structure of the procedures or on the holding of the conditions of application of the rule. The result is transparently a restriction on procedural constructions.

There are multiple ways of working out the details of this argument. For now, consider this simple way of working it out. Suppose that a rule has the form of a norm-kernel, and simply takes as its objects the sequential compositions of a set of actions, $\{a, b, c, d\}$. Suppose a rule says that, if a sequential composition, $\pi; ...; \pi'$, doesn't place a prior to c, then that sequential composition is irrational. The procedures that match the rule are those that sequentially compose a and c in the specified way. Generally, rules like this can be mapped to a restriction such that if $S1$ is the set of procedures such that C holds of them, and $S2$ is the set of procedures that fails C, then $R_S = S1 \succ T_S = S2 = I_S$. Any binary partition of a set can be matched with an ordering relation, as pointed out above with respect to categorical norms. That norm-kernel thus defines a restriction on procedural constructions.

The norms in Table 4.3 define restrictions on ways in which epistemic actions can be sequenced, iterated, concurrently executed, chosen among, combined with tests, complemented, reversed, or executed with minimal changes.

That each of these norms imposes a restriction on how rational procedures are built out of more basic actions implies the second premise. This is straightforward. Some actions make it impossible to execute other actions. In chemistry lab procedures, for instance, some tests that detect properties of a substance destroy that substance, so other tests must be performed first (assuming that tests must be performed on the same sample). In the more purely epistemic case, the Detective example from the first chapter is one in which the ordering of the epistemic actions matters. If the detective had drawn conclusions earlier in the process or failed to make some information available or failed to utilize certain sources of information in inference, she'd have been unable to perform the deduction at the end. Very many algorithms for manipulating technical constructions require some sequential orderings of tasks and not others, semantic tableaux being a familiar philosophical example. Iterated belief revision provides other examples. Proof construction provides yet more; in tableaux methods, one ought to apply all applicable non-branching rules prior to applying branching rules in order to manage complexity of proofs. Proof systems also require iteration of the same step, as do belief revision systems, and projection rule testing systems in FLT. Systems incorporating multiple norms require concurrent compliance. Procedures for proof construction give indeterministic choice among applicable steps. Norms like "If belief set K is inconsistent, then revise it to consistency" require the performance of a test for consistency sequentially followed by a revision action that may involve reversing the effects of a prior process. All prohibitionary norms require performing actions other than those prohibited. The controversial recovery rule of AGM requires reversal, as do all backtracking algorithms. Finally, FLT procedures require finding projection rules that reach a fixed point. These examples will be familiar from Chapter 2.

Table 4.3: Norms Requiring Types of Procedural Operation



Norms Requiring Types of Procedural Operation	**Norm/Procedure/ Method**	**Field of Study**	
Norms that Require Sequential Composition	- Argumentation Methods - Hypothesis Testing Strategies (FLT) - Proof Construction and Model Checking - Mathematical Algorithms - Enabling and Prevention Procedures - Inductive Strategies (Mill, Bacon)	- Formal and Informal Logic - Formal Learning Theory - Formal Logic - Formal Methods - DEL, Lab Manuals - Philosophy of Science	- [Angell 1964, Walton 2013] - [Schulte 2017, Kelly 1988] - [Quine 1950, Kleene 1967] - [van Benthem 2007] - [van Ditmarsch et. al 2007] - [Bacon 1851, 1960, Mill 1843]
Norms that Require Iteration	- Argumentation Methods - Hypothesis Testing Strategies (Formal Learning Theory) - Proof Construction and Model Checking - Iterated Belief Revision - Iterated Credence Update	- Formal and Informal Logic - Formal Learning Theory - Formal Logic - AGM, Belief Revision - Bayesian Epistemology	- [Ibid] - [Kelly et. al RBR, Schulte 2017] - [Ibid] - [Gardenfors 2008, Hansson 2017] - [Hansson 2017, Gardenfors 1992]
Norms that Require Parallel Execution	- Every System of Multiple Norms	- Traditional Epistemology, Ethics of Belief	- [Littlejohn and Turri 2014, Clifford 1999, Adler and Rips 2008]
Norms that Require Choice	- Argumentation Methods - Logical Constraints - Proof Construction and Model Checking - Belief Revision	- Formal and Informal Logic - Beall, Field, Bayes - Formal Logic - AGM	- [Ibid] - [Beall 2013, Field 2009a,b,c, 2015, Christensen 2004] - [Ibid] - [Ibid]
Norms that Require Test	- All Conditional Norms	- Traditional Epistemology, Ethics of Belief	- [Ibid]
Norms that Require Complement	- Norms of Prohibition	- Traditional Epistemology, Ethics of Belief	- [Ibid]
Norms that Require Reversal	- Belief Revision - Backtracking Algorithms (Ariadne's Thread)	- AGM - Algorithms	- [Ibid] - [Ibid]
Norms that Require Fixed Point	- Efficient FLT Strategies	- Formal Learning Theory	- [Ibid]

4.3.4 Details of the Solution: Premise 3

The third premise asserts that restrictions on procedural construction effectively reduce all other restrictions on epistemic dynamics. The details of this reduction are relegated to the technical document [135], but the reduction can be verified relatively easily. First, the idea of restrictions on procedural construction comes to the thought that, if we identify a set or space of possible ways of building procedures out of a repertoire of basic actions, a restriction is simply a division of that set/space into subsets into the correct or legitimate ways of building procedures and the incorrect or illegitimate ways of building procedures. We can define restrictions on procedural constructions more formally by defining syntactic schemata:

Def.: Syntactic Schemata Let a syntactic schema be a sequence of symbols satisfying the rules of P_{PLEN} but defined over a set of variables (Π_V) for elements of Π_0. Formally:

$$\Pi_V \mid \pi_1;\pi_2 \mid \pi_1 \cap \pi_2 \mid \pi_1 \cup \pi_2 \mid \pi_i* \mid (A)? \mid -\pi_i \mid\curvearrowleft \pi_i \mid \nabla\pi_i$$

- where Π_V is a set of subscripted variables, $\langle x_1, ..., x_n \rangle$.

Call the set of syntactic schemata Π_{Sch}. Explicitly define restrictions on procedural construction like so:

(Restrictions on Procedural Construction) Let $S = \Pi_{Sch}$. Then every restriction on S, $R_{\prec} = (P(S), \prec, T_I, R_S, I_S)$, is a *restriction on procedural constructions* or a *syntactic restriction.*

By replacing the subprotocols in a protocol with variables that themselves can be replaced with complex schemata, we identify ways of building procedures. Take $x_1; x_2$. This schema identifies the way of building procedures by sequential composition of atomic actions. The more complex schema $(x_1; x_2) \cap (x_3; (x_1 \cup x_4))$ identifies the way of building procedures by parallel execution on a sequential composition of atomics and a sequential composition of an atomic and a choice formed out of the first component of the first sequential composition and a new component. Given finitely many atomic protocols, there are finitely many protocols built according to this pattern, among them: $(a; b) \cap (c; (a \cup d))$ and $(a; a) \cap (a; (a \cup a))$.

Given the set of all syntactic schemata, a restriction on procedural constructions is just a way of partitioning the possible ways to build protocols into rational and irrational sets. A restriction on procedural constructions might, for instance, single out all sequential compositions as the rational ways of building procedures and the rest as irrational. The rational sets of such a restriction would contain $x_1; x_2$, $x_1; (x_2; x_3)$, $(x_1; x_2); x_3$, $(x_1 \cup x_2); x_3$, and so on for all of the possible ways of building a protocol such that sequential composition is the main control operator.

Second, define correspondence among restrictions:

Def.: Correspondence Let R_Π be the rational set of syntactic schemata in a restriction where $S = \Pi_{Sch}$. Let R_{PX} be the rational set of processes in a restriction where $S = P_g$ for some g. Let R_{TX} be the rational set of transitions in a restriction where $S = T_g$. Restrictions on different sets of elements, $R_\prec$ and $R'_\prec$, *correspond* iff:

1. For any π_v, π that matches π_v, and p such that $p \in P_\pi$, π_v is in $R_{\Pi Sch}$ iff p is in R_{PX}.
2. For any p and any t such that $t = \langle first(p), last(p) \rangle$, p is in R_{PX} iff t is in R_{TX}.
3. For any t and any s such that $s = last(t)$, t is in R_{TX} iff s is in R_{SX}.
4. Same for I_Π, I_{PX}, I_{TX}.

Now, the Reduction Lemma, proven in [135], is as follows:

Reduction Lemma For every connected EDG:

(1) For every restriction on the construction of procedures, there is a corresponding restriction on processes.

(2) For every restriction on processes, there is a corresponding restriction on procedural constructions.

(3) For every restriction on transitions, there is a corresponding restriction on processes.

(4) For every restriction on states, there is a corresponding restriction on transitions.

(5) It is not the case that, for every restriction on processes, there is a corresponding restriction on transitions nor a corresponding restriction on states.

(6) It is not the case that, for every restriction on transitions, there is a corresponding restriction on states.

From the Reduction Lemma, premise 3 follows; for every restriction on epistemic dynamics, $R_{\prec E}$, there is some restriction on procedural constructions, $R_{\prec C}$, from which all categorical and deontic verdicts of $R_{\prec E}$ can be derived via the relevant correspondence rule. From another angle, for any restriction on epistemic dynamics, $R_{\prec E}$, there is a restriction on procedural constructions, $R_{\prec C}$, that provides sufficient information for inferring the verdicts of $R_{\prec E}$ for any object in its domain. Whether or not a process, transition, or state is rational according to a restriction of the relevant kind can be determined by examining the corresponding restriction on procedural constructions. The converses of these fail; there aren't restrictions on transitions or states that can be used to derive the verdicts of restrictions on processes or procedural constructions. Thus, restrictions on procedural construction provide structures for representing and reasoning about all other restrictions on epistemic dynamics.

4.3.5 Details of the Solution: Premise 4

Premise 4, which asserts that protocols can be used to read off the verdicts of restrictions on procedural constructions, can be gotten rather simply from the notion of restrictions on procedural construction. It turns out that every protocol can be associated with a restriction on procedural constructions by means of the following notion of a matching syntactic schema:

1M) Every protocol contains a sequence of atomic subprotocols, $\pi_1, ..., \pi_n$. Consider an arbitrary function, $f : \langle \pi_1, ..., \pi_n \rangle \longrightarrow \Pi_V$. A protocol schema, π_V, *matches* a protocol, π, iff there is a sequence of variables $\langle x_1, ..., x_n \rangle$, such that $\langle x_1, ..., x_n \rangle = \langle f(\pi_1), ..., f(\pi_n) \rangle$.

By 1M, every protocol, π, can straightforwardly be mapped to a set of syntactic schemata such that uniformly substituting the sequence of atomic protocols in π for the variables in the schema yields π. For instance, by 1M, $(a \cap b); (c \cup d)$ is matched by schema $(x_1 \cap x_2); (x_3 \cup x_4)$ but is not matched by schema $x_1; x_2$. Thus, it's easy to see that each protocol is mapped to a restriction on procedural constructions by 1M. The restriction on procedural constructions such that $R_{\Pi Sch}$ is the set of matching schemata and $I_{\Pi Sch}$ is the set of non-matching schemata suffices to show that there is a set of restrictions on procedural constructions (i.e., syntactic schemata) that can be read off of the syntax of π. The elements of the set of matching and non-matching schemata can be straightforwardly enumerated by checking to see whether replacing the atomics in π with the variables in π_v generates π_v. Thus, the elements of $R_{\Pi Sch}$ (resp. $I_{\Pi Sch}$) can be enumerated by checking to see whether replacing the atomics in π with the variables in π_v generates π_v. If so, then it goes into $R_{\Pi Sch}$. If not, then it goes into $I_{\Pi Sch}$.

With the notion of matching schemata, the following theorem and its proof give premise 4:

Reading Off Theorem For every restriction, $R_{\preceq Sch}$, on procedural constructions, there are sets of protocols, $\{\pi_1, ..., \pi_n\}$, and individual protocols, π, such that a set, R, of restrictions containing $R_{\preceq Sch}$ can be read off of the set (resp. protocol), and every restriction that R contains is identical to $R_{\preceq Sch}$ with respect to its rational, irrational, and threshold sets.

Proof-plan Define procedures for reading off restrictions on construction from protocols:

R1 Take the set of protocols, $\{\pi_1, ..., \pi_n\}$, such that each π_i is matched by at least one π_v in $R_{\preceq Sch}$ by 1M. Eliminate redundancies. From this set of protocols,

$R_{\preceq Sch}$ can be read off by simply adding a matching π_v to the rational set for each π_i. Do the same for each other component of the restriction.

R2 Let $\pi = \pi_1 \cup (...(... \cup \pi_n))$ where each π_i is in $\{\pi_1, ..., \pi_n\}$ from R1. Then read off the set of schemata such that they match the π_i by 1M. Do the same for each other component of the restriction.

For any such restriction, each element of the rational set is matched by a set of protocols that differ only in the sequence of atomics that they contain. The set of protocols thus only requires one of these protocols for each schema in order to read off exactly the rational set of schemata. Just so for each of the other components. From this set of protocols we can thus read off the set of restrictions containing exactly that restriction. Every other restriction in that set that can be so read off will be identical to $R_{\preceq Sch}$ with respect to its categorical and deontic verdicts; they will contain all and only schemata that match each π_i. Thus, these restrictions can differ by only their ordering relations. Thus, everything to be derived about the verdicts of the restriction can be derived from the set of restrictions that can be read off of $\{\pi_1, ..., \pi_n\}$ (resp. π). □

This theorem, coupled with the Reduction Lemma, implies that all restrictions on epistemic dynamics can be read off of the syntax of protocols in verdict-equivalent sets, and this enables us to derive everything about the verdicts of the norms represented by restrictions.

4.3.6 Details of the Solution: Putting it All Together

Premise 2 worked to motivate the examination of restrictions on procedural construction. From there, the rest of the argument falls into place. From premises 1, 3, and 4, it follows that protocols are the proper devices in PLEN with which to represent epistemic rationality norms. To summarize: Protocols are devices from which the most important information about restrictions on epistemic dynamics can be inferred, and restrictions on epistemic dynamics unify the information necessary to comply with or respond appropriately to epistemic rationality norms.

To recapitulate detail: For each restriction on processes that determines the properties of the corresponding restrictions on transitions and states, there is a corresponding restriction on procedural constructions. Correspondence between two restrictions on dynamics ensures that we can derive information about the assignment of objects to rational and irrational sets of one restriction from the assignment of the other. This verdict information about restrictions on procedural constructions can be read off of the syntax of some protocol, π, or set of protocols, $\{\pi_1, ..., \pi_n\}$. Given that restrictions on processes unify the classes of normative expressions adumbrated above (whether by means of defining protocol functions that agree with the norm functions of the underlying norm-kernels or simply by codifying the verdicts of those expressions), procedural constructions unify these classes just as well due to the Reduction Lemma; anything inferable about the verdicts of the normative expressions from the restriction on processes can be inferred from the restrictions on procedural constructions.

Thus, by the Reading Off Theorem, these restrictions can all be read off of the syntax of protocols; protocols themselves are constructions from which anything inferable about the verdicts of the normative expressions can be inferred. Hence, protocols are the best possible candidate out of the formal devices available in PLEN for analyzing epistemic rationality norms themselves, insofar as epistemic rationality norms are objects that underly the different kinds of epistemic norm expression.[6]

4.3.7 PFT Solves the First-Class Citizenship Problem

Now, as I hinted above, the solution can actually just be derived from PFT itself. The basic idea is that for every norm under consideration, there is a formal object - a norm function - that corresponds to it that mirrors its mapping of objects to valences under conditions

---

[6] It's also possible to run nearly same argument while bypassing the restrictions on procedural construction. Simply note that each protocol is directly associated with a restriction on processes by the execution rules. From the relevant clauses of the reduction lemma, the restrictions that capture the content of norm-kernels can all be read off of protocols anyway.

of application. Norm functions can all be mirrored in the same way by protocol functions. Thus, for every norm, n, with content fully analyzed by a norm function, there is a protocol that represents n by defining an agreeing function from the conditions of application and objects of n to the protocol-theoretic analysans of n's valences. More precisely, assuming that each valence has a formal counterpart in PLEN, there is a protocol that maps every pair of condition C and object x of any norm n to the formal counterpart of the valence of x under C according to n. Thus, for instance, the Indiscriminate Consistency Function has protocol-theoretic partner - the protocol π such that $f_{\pi}^{\forall}$ agrees with the Indiscriminate Consistency Function for every condition of application and object.

If the foregoing arguments are right, and, for every norm-kernel that can be captured by the norm functions considered, there is a protocol that can be used to read off all of the verdicts (assignments of valence to object under condition of application) of the norm-kernel, then it follows that that protocols perform a theoretical unification of the normative expressions in our taxonomy above. To see the foregoing claim, recall that protocol functions are defined in terms of the execution sets of protocols, which are defined in terms of the syntactic forms of protocols. Execution sets of protocols are basic forms of restrictions on epistemic dynamics and define derivative restrictions. Let $S = P_g$, let the execution set be R_S, and let $T_I = I_S = P_g/R_S$. Each execution set defines restrictions on states: let $S = W$, and let the set of terminal states be R_S. For transitions: let S be the set of transitions possible in a graph, let R_S be the set of all pairs of first and last states of any execution of π. Thus, from PFT, every norm-kernel can be represented, via its norm function, by a set of restrictions on epistemic dynamics defined by the syntactic form of some protocol, π, in that the restrictions on dynamics can be "read off" of the syntax by means of the execution rules and the relations between execution sets and restrictions on dynamics.

Thus, from PFT, it follows that if norm-kernels unify or underly the informational content of all of the other forms of expression commonly called "epistemic norms", then the restrictions on epistemic dynamics defined by protocols provide a unifying formal construction for all of these expressions[7]. Protocols thus provide a formal analysans of the unifying content of norm expressions. From properties of the syntax or semantics of protocols, properties of norm kernels can be inferred, and thus properties about the unifying content of epistemic norm expressions can be inferred. More explicitly, let N be a norm-kernel such that there is a norm-function, f_n, expressible in PLEN that captures the assignment of valences in Val to its objects under any of its conditions of application. The PFT results show that there is a universal protocol function, $f^{\forall}_{\pi}$, from which all such assignments of N's norm function can be inferred. The foregoing results also show that there is a protocol, π, such that the assignments of $f^{\forall}_{\pi}$ can be inferred from the syntax of π by means of first figuring out π's execution sets, and then deriving the restrictions on processes, transitions, states, and formulae of L_{PLEN} from that. Thus, protocols are constructions that represent the unifying content of epistemic rationality norms.

4.4 Formalizing D1-D3

Section 4.2 showed how PLEN solved the Processes and Constraints problems. This section puts all of the foregoing pieces together to complete the formalization of the core theory defended in Part I. The formalization of D1-D3 is central to this project. In particular, I show that for each of D1-D3, there is a provable formal assertion at the meta-level of PLEN, a derivable (and thus valid) formula at the object-level of PLEN, a set of valid rules, or an object-level theory that corresponds to it.

Prior to getting on with the formalizations, I want to lay groundwork concerning the theoretical adequacy of the formalization of D1-D3. The basic conception of adequacy for formalizations is adapted from the formal adequacy theses of [18]:

---

[7] Again, I wish not to wade into the battles concerning content, but merely to argue that (a) norms can be complied with or applied, and (b) normative expressions can be appropriately or inappropriately responded to. The idea of this argument is merely that, if we suppose that the information necessary to comply with or apply norms and respond appropriately to normative expressions can be read off of norm-kernels, then, because we can read off all of the features of norm-kernels from restrictions on dynamics and the features of restrictions can, in turn, be read off of protocols, then we can read all of this information off of protocols.

> THESIS I. Let s be a social situation involving the intuitive concepts of knowledge, justifiable beliefs and common knowledge among a group of agents. Assume that s is presented in such a way that all the relevant features of s pertaining to knowledge, beliefs and common knowledge are completely determined. Then we may associate to s a mathematical model S. (S is a multi-agent Kripke model; we call these epistemic state models.) The point of the association is that all intuitive judgements concerning s correspond to formal assertions concerning S, and vice-versa.
>
> THESIS II. Let σ be a social "action" involving and affecting the knowledge (beliefs, common knowledge) of agents. This naturally induces a change of situation; i.e., an operation o taking situations s into situations $o(s)$. Assume that o is presented by assertions concerning knowledge, beliefs and common knowledge facts about s and $o(s)$, and that o is completely determined by these assertions. Then
>
>> (a) We may associate to the action σ a mathematical model Σ which we call an epistemic action model. (Σ is also a multi-agent Kripke model.) The point again is that all the intuitive features of, and judgments about, σ correspond to formal properties of Σ.
>>
>> (b) There is an operation $\otimes$ taking a state model S and an action model and returning a new state model $S \otimes \Sigma$. So each induces an update operation O on state models: $O(S) = S \otimes \Sigma$.
>>
>> (c) The update O is a faithful model of the situation change o, in the sense that for all s: if s corresponds to S as in Thesis I, then again $o(s)$ corresponds to $O(S)$ in the same way; i.e. all intuitive judgements concerning $o(s)$ correspond to formal assertions concerning $O(S)$, and vice-versa. (pp. 166-167)

For this project: A formalism, F, is theoretically adequate to the analysis of epistemic dynamics, epistemic rationality, and epistemic rationality norms iff it contains elements satisfying the following theses:

THESIS I. Let s be an epistemic state. Assume that s is presented in such a way that all the relevant features of s pertaining to the content of s are completely determined. Then we may associate to s a mathematical model S. The point of the association is that all intuitive judgements concerning s correspond to formal assertions concerning S, and vice-versa.

THESIS II. Let σ be an epistemic "action" involving and affecting the epistemic states of agents. This naturally induces a change of epistemic state; i.e., an operation o taking situations s into states, $o(s)$. Assume that o is presented by assertions concerning the content of s, the postconditions, preconditions, and alternatives of $o(s)$ (including the content of the new state), and about the procedural properties of σ and their relation to s and $o(s)$, and that o is completely determined by these assertions. Then

(a) We may associate to the action σ a mathematical model g which we call an epistemic directed graph. The point again is that all the intuitive features of, and judgments about, σ correspond to formal properties of g.

(b) There is an operation $\otimes$ taking a state model S and an action σ and returning an EDG, $S \otimes \sigma$. So each induces an executability assignment O on state models: $O(S) = S \otimes \sigma$.

(c) The assignment O is a faithful model of the state change o, in the sense that for all s: if s corresponds to S as in Thesis I, then again $o(s)$ corresponds to $O(S)$ in the same way; i.e. all intuitive judgements concerning $o(s)$ correspond to formal assertions concerning $O(S)$, and vice-versa.

THESIS III. Let N be an epistemic norm concerning the rationality of states, transitions, and actions, processes, or procedures of epistemic state change. This naturally induces a restriction on states, transitions, and actions/processes/procedures of epistemic state change; i.e., an operation r taking norms N into sets of states, transitions, and actions/processes/procedures, $r(N)$. Assume that r is presented by assertions concerning procedural operations that construct complex actions and procedures, the procedural content of N, the postconditions, preconditions, and alternatives of complying with N or controlling one's dynamics in accordance with $r(N)$, and about the procedural properties of N and their relation to s and $o(s)$, and that r is completely determined by these assertions. Then

(a) We may associate to the norm N a mathematical model π which we call an epistemic protocol. The point again is that all the intuitive features of, and judgments about, N correspond to formal properties of π.

(b) There is an operation $\otimes$ taking an EDG, g, and a norm N and returning an EDG $g \otimes N$. So each induces a restriction operation $R_{\preceq}$ on EDGs: $R_{\preceq}(\pi) = g \otimes N$.

(c) The restriction operation $R_{\preceq}$ is a faithful model of the restriction r, in the sense that for all N: if N corresponds to π as in Thesis I, then again $r(N)$ corresponds to $R_{\preceq}(\pi)$ in the same way; i.e. all intuitive judgements concerning $r(N)$ correspond to formal assertions concerning $R_{\preceq}(\pi)$, and vice-versa.

This defines the fundamental argumentative goals of the section. Note that the desiderata, D1-D3, defended in Chapter 2 are or imply "intuitive" or, rather, semi-formal but philosophically defensible judgments concerning epistemic dynamics and epistemic rationality norms. Thus, adequacy of PLEN as a formal framework requires providing formal assertions to correspond to the objects of these judgments. Now, if section 3 is right, then protocols are the best devices in PLEN for representing epistemic rationality norms. It is the properties of protocols, then, that we must search for assertions that correlate to D1-D3. More precisely, adequacy requires that for every principle about epistemic rationality and its norms, D1-D3, there is a provable formal assertion at the meta-level of PLEN about protocols, a derivable (and thus valid) formula at the object-level of PLEN, a set of valid rules governing the properties of protocols, or an object-level theory that corresponds to it. I argue in this section that this is the case by spelling out the formal assertions that correspond to each of D1-D3.

Before presenting those results, I must emphasize that none of the analyses below is taken to be final; each of them is merely a demonstration that PLEN contains the resources to provide formal assertions corresponding to each of D1-D3. This is not to say that the analyses don't still need work. The goal is to make it plausible that good work can be done with the tools of PLEN.

4.4.1 D1

D1 Epistemic rationality is dynamic; there are epistemic restrictions on transitions among epistemic states, not just states or contents of states, and there are characteristic relations among restrictions on states and restrictions on transitions.

This principle contains the core assertion of epistemic dynamicism: there are epistemic restrictions on transitions among epistemic states. This principle was defended in Chapter 2. However, in arguing against epistemic dynamicism, Chapter 2 defended more than just this core assertion; the following were also defended, spelling out D1's hint of "characteristic relations" among types of restrictions:

(D1.1) There are norms of epistemic rationality that take epistemic transitions as their objects.

(D1.2) There are norms of epistemic rationality that take properties of epistemic transitions as their conditions.

(D1.3) Norms that restrict epistemic states can generate norms that restrict epistemic transitions.

(D1.4) Norms that restrict epistemic transitions can generate norms that restrict epistemic states.

Now, how are all of these claims formalized in PLEN? Much of the work was done above in section 3. First, the execution rules of PLEN map sets of transitions to each protocol - the set, T_π, of all transitions composed of the first and last states of some path in P_π. Defining a restriction on transitions that is mapped to each protocol is thus trivial; $T_\pi = R_{TX}$ and $T_I = I_{TX} = T_g/T_\pi$. Since epistemic rationality norms are represented in PLEN by protocols, it's transparent that there are epistemic restrictions on transitions: the formal counterparts of epistemic restrictions on transitions are the T_π mapped to each protocol. That is, a formal assertion corresponding to the first clause of D1 is this:

(RET) For every protocol, π, for every EDG, g, there is a set of transitions, $T_\pi = \{\langle w, w' \rangle \mid \exists p(p \in P_\pi, w = first(p), w' = last(p)\}$, in g. There is thus, trivially, a restriction on transitions corresponding to π such that $R_{TX} = T_\pi$, and T_I is the complement of T_π.

In other words, that w' can be reached from w by execution of π implies that the transition is in the restriction defined by the norm that π represents. Switching to intuitive epistemological language, every execution of π is a rational or correct epistemic process according to π, so the rational transitions according to π are those that bookend the rational processes. Another way of putting this is that the rational transitions are those that can be achieved by implementing π. So, there are epistemic restrictions on transitions.

Now, it might be objected that the presence of restrictions on transitions does not imply that there are restrictions that are not just restrictions among states or contents of states. Of course, an epistemic staticist would argue that restrictions on transitions are, in principle, reducible to restrictions on states. However, EDGs contain restrictions on transitions that cannot be reduced to restrictions on contents of states.[8] First, there is a technical reason:

---

[8] Staticists would have to defend restrictions on states irrespective of content in order to reduce these

states are represented with nodes that are then associated with contents via the content function, there are restrictions on states that rule in $\langle w, w' \rangle$ and rule out $\langle w, w'' \rangle$ even if w' has identical content with w''. Identity of content does not imply identity of state in PLEN. Second, there is a philosophical reason: even if we focus on restrictions on content states alone, there are restrictions that rule in $\langle c(w), c(w') \rangle$ and rule out $\langle c(w''), c(w') \rangle$ while also ruling in $\langle c(w), c(w') \rangle$ and ruling out $\langle c(w), c(w'') \rangle$ that, if the arguments of Chapter 2 are correct, capture the content of plausible epistemic norms. These restrictions cannot be reduced to restrictions on states by any function defined by a rule that considers only the properties of content states nor, for the reasons in Chapter 2, by any function that meets the staticist's aims.

Second, PFT implies all of D1.1-D1.4.

(PFT) *For any* f_n*, with any* $C \in nCon$ *and any* $x \in nDom(R)$*, if* $C \in L_{PLEN}$ *and* x *is an action or procedure in PLEN, a condition expressible in* L_{PLEN}*, or a state or class of states in the semantics of PLEN, then there is a protocol* π *such that the universal protocol function* $f_\pi^\forall$ *for arbitrary* $g_1, ..., g_n$ *in which* π *is executable is such that* $f_\pi^\forall(C, x) = f_{Val}(f_n(C, x))$*.*

Simply put, PFT shows that, for any epistemic norm with conditions of application expressible in L_{PLEN}, domain contained somewhere in the set of structures of PLEN, and valences encoded in execution-theoretic facts, there is a universal protocol function for some protocol that agrees with that norm. It follows trivially from PFT that:

PFTC1 If there is an epistemic norm, N, such that (i) its condition of application C is expressible in L_{PLEN} (e.g., any postcondition statements, any precondition statements), (ii) its objects are represented by objects in the structures of PLEN (e.g., states,

---

kinds of restrictions on transitions, and for them, this is unintelligible. Only the static features of states are normatively salient, and individuating properties other than content (e.g., time index, relations to processes, relations to predecessor and successor states) are dynamic. However, the protocol-theoretic account via PLEN makes perfectly good sense of transitions among content-identical states being rationally distinct; there might be a rational process that takes you from w to w' but not rational process that takes you from s to s''.

transitions, paths, syntactic schemata), and (iii) its valences are encoded in execution-theoretic properties of PLEN, then there is a protocol π_N such that the domain of the universal protocol function of π takes the expression that corresponds to C and any object, x, of N to the valence that N attributes to x.

PFTC2 If a transition, t, is in $nDom(R)$, then there is a π such that t is in $\pi Dom(R)$.

PFTC3 If a condition, C, is in $nDom(R)$, then there is a π such that C is in $\pi Dom(R)$.

Thus, if there are epistemic norms that take transitions (that are, in principle, representable in PLEN) as their objects, then there are protocols that take transitions as their objects - there are protocols the protocol functions of which take transitions. PFT also implies that if there are epistemic norms that take properties of transitions expressible in L_{PLEN} as their conditions of application, then there are protocols that take those expressions as conditions of application. Again, for any such norm, there is a protocol whose protocol function takes those expressions as inputs. Taking an object or condition as an input, the protocol function maps condition and object pairs to valences, thus delivering valences about those objects. Valences are encoded in the execution sets and derivative constructions. Thus, protocol functions place objects of norms into restrictions on dynamics that correspond to the protocols. Thus, PLEN has a systematic way of saying that there are epistemic restrictions on transitions and that these are conditioned on properties of transitions. This, I think, suffices to show that PFTC2 and PFTC3 correspond to D1.1 and D1.2

Now, for D1.3 and D1.4, much simpler features of PLEN correspond:

(RST) For every connected EDG, g, and every restriction on states, $R_{\prec S}$, defined on the states of g, there is a restriction on transitions, $R_{\prec T}$, defined on the set of transitions in g such that a transition, t, is in R_{TX} iff $last(t) \in R_{SX}$.

(RTS) For every connected EDG, g, (i) there is some restriction on transitions, $R_{\prec T}$, defined on the transitions of g, such that there is a restriction on states, $R_{\prec S}$, defined on the set of states in g such that a transition, t, is in R_{TX} iff $last(t) \in R_{SX}$, and (ii) there is some restriction on transitions such that there is no restriction on states meeting the foregoing condition.

These are mechanically derivable from the definition of restrictions, as shown in the proof of the Reduction Lemma in [135]. Thus, PLEN has systematic means for thinking about how transition restrictions can generate state restrictions and vice versa. Note that this coheres with the objections to staticism in Chapter 2. It cannot be the case that all restrictions on transitions are reducible to restrictions on states in virtue of their terminal states, as there may be transitions t and t' such that $last(t) = last(t')$. These cases are, of course, exemplified by the cases discussed in Chapters 1 and 2.

4.4.2 D2

D2 Epistemic rationality is procedural; there are epistemic restrictions on epistemic actions, processes, and procedures, not just states or transitions among states, and there are characteristic relations among restrictions on transitions and restrictions on processes.

D2 expands on D1. Not only is epistemic rationality dynamic, it's procedural. In particular, Chapter 2 argues that:

(D2.1) There are norms of epistemic rationality that take epistemic processes as their objects.

(D2.2) There are norms of epistemic rationality that take epistemic procedures as their objects.

(D2.3) The norms of epistemic rationality take properties of epistemic processes and procedures as their conditions.

(D2.4) Norms that restrict epistemic transitions can generate norms that restrict epistemic processes.

(D2.5) Norms that restrict epistemic transitions can generate norms that restrict epistemic procedures.

(D2.6) Norms that restrict epistemic processes and procedures can both generate norms that restrict epistemic transitions.

Again, formalization is largely carried out by the machinery already introduced. First, the execution rules of PLEN map restrictions on processes to each protocol. The execution rules map sets of processes to each protocol:P_π. Defining a restriction on processes that is mapped to each protocol is trivial: the formal counterparts of epistemic restrictions on processes are the executions mapped to each protocol: $R_{PX} = P_\pi$ and $T_I = I_{PX} = P_g/P_\pi$. That is, a formal assertion corresponding to the first clause of D2 is this:

(REP) For every protocol, π, for every EDG, g, there is a set of processes, P_π, in g that is mapped to π by the execution rules. There is thus, trivially, a restriction on processes corresponding to π such that $R_{PX} = P_\pi$ and T_I is the complement of P_π.

In other words, the processes that execute π are the rational processes according to π. Switching to intuitive epistemological language, every execution of π is a rational or correct epistemic process according to π. Another way of putting this is that the rational processes are those that can be achieved by implementing π.

Second, each of D2.1-D2.3 correspond to the following trivial corollaries of PFT:

PFTC4 If a process, p, is in $nDom(R)$, then there is a π such that p is in $\pi Dom(R)$.

PFTC5 If a procedure, P_π, is in $nDom(R)$, then there is a π' such that any π'' s.t. that $P_{\pi''} = P_\pi$ is in $\pi' Dom(R)$

PFTC6 If a condition, C, is in $nDom(R)$, then there is a π such that C is in $\pi Dom(R)$.

Thus, if there are epistemic norms that take processes (that are, in principle, representable in PLEN) as their objects, then there are protocols that take processes as their objects - there are protocols the protocol functions of which take processes as inputs. PFT also implies that if there are epistemic norms that take procedures or sets of processes (that are, in principle, representable in PLEN) as their objects, then there are protocols such that their protocol functions take protocols for which those procedures are execution sets as inputs. Simply put, protocols can deliver rationality verdicts about procedures by mapping them conditionally to valences. Finally, PFT implies that if there are epistemic norms that take properties of processes expressible in L_{PLEN} as their conditions of application, then there are protocols that take those expressions as conditions of application. Again, for any such norm, there is a protocol whose protocol function takes those expressions as inputs. Taking an object or condition as an input, the protocol function maps condition and object pairs to valences, thus delivering valences about those objects. Valences are encoded in the execution sets and derivative constructions. Thus, protocol functions place objects of norms into restrictions on dynamics that correspond to the protocols. Thus, PFTC4, PFTC5, and PFTC6 correspond, respectively, to D2.1, D2.2, and D2.3.

Finally, D2.4-D2.6 correspond to the simple propositions:

(RTP) For every EDG, g, and every restriction on transitions, $R_{\prec T}$, defined on the transitions of g, there is a restriction on processes, $R_{\prec P}$, defined on the set of processes in g such that a transition, t, is in R_{TX} iff $p \in R_{PX}$.

(RTPr) For every EDG, g, and every restriction on transitions, $R_{\prec T}$, defined on the transitions of g, there is a restriction on procedures, $R_{\prec Pr}$, defined on the set of procedures (the set of sets of processes, the powerset of P_g) in g such that a transition, t, is in R_{TX} iff $p \in R_{PrX}$.

(RPT) For every EDG, g, (i) there is some restriction on processes, $R_{\prec P}$, defined on g, such that there is a restriction on transitions, $R_{\prec T}$, defined on g such that for all transitions, t, such that $t = \langle first(p), last(p) \rangle$, $p \in R_{PX}$ iff $t \in R_{TX}$, (ii) there is some restriction on processes such that there is no restriction on transitions meeting the foregoing condition.

(RPrT) For every EDG, g, (i) there is a restriction on procedures, $R_{\prec Pr}$ defined on the set of procedures (the set of sets of processes, the powerset of P_g) in g, such that there is a restriction on transitions, $R_{\prec T}$, defined on g such that a transition, t, $t = \langle first(p), last(p) \rangle$, for some $p \in P_\pi$ for some π that matches a π_V, $\pi \in R_{Sch}$ iff $t \in R_{TX}$, and (ii) there is some restriction on procedures such that there is no restriction on transitions meeting the foregoing condition.

These propositions - trivial consequences of the Reduction Lemma - each show simple ways of constructing restrictions on elements of epistemic dynamics from restrictions on other elements of epistemic dynamics. The latter two formalizations note the limited reducibility of restrictions on processes and procedures to other restrictions. There can't be, as shown in [135], for all arbitrary restrictions on processes or procedures, restrictions on transitions from which the restrictions on processes (resp. procedures) can be derived. Suffice it to say that these assertions show how to formalize claims about the generation of restrictions from other kinds of restrictions in PLEN.

4.4.3 D3

D3 Epistemic rationality norms define restrictions on epistemic dynamics that model (a) ideally rational evolution of epistemic state, (b) ideally rational changes of epistemic state, (c) ideally rational procedures and processes of epistemic state change, and that (d) can be distinguished from constraints of capacity and feasibility.

Corollary Epistemic rationality norms restrict the construction of procedures; they determine how basic epistemic actions are (rationally) to be sequenced, iterated, combined, chosen among, conditioned on tests, avoided, reversed, and cautiously executed.

As argued in Chapter 2 and earlier in Chapter 4, D3 characterizes the unifying model of the content of epistemic rationality norms. Epistemic rationality norms define restrictions on epistemic dynamics in the sense that the norms determine what elements of epistemic dynamics occupy what parts of the restriction. The extension of the restriction then characterizes ideally rational epistemic dynamics. If a norm determines that a class of epistemic states is rational (or at least maximal in the ordering of the restriction), then ideally rational (with respect to the norm in question) agents will end up in those states. If a norm determines that a transition is rational (or higher in the ordering than other available transitions), then an ideal agent, when in the root state, will transition accordingly given the option. If a norm determines that a process or procedure is rational (or higher in the ordering than other available processes), then, at a state at which that process (or a process in the procedure) can be carried out, an ideally rational agent will carry it out. These are trivial consequences of the notion of ideal rationality - an ideally rational agent complies with epistemic norms. But compliance with a norm, in PLEN, is cashed out in terms of membership in the rational subset of the restriction mapped to the norm or comparative superiority in the ranking in the restriction. Thus, an ideal agent's epistemic dynamics are in the rational subsets or are comparatively superior in the ranking of the restrictions defined by epistemic norms.

The corollary of D3 extends this basic idea to the construction of procedures out of basic actions. As argued in Chapter 2, epistemic norms frequently tell us that certain sequences, combinations, choices, iterations, conditioning on tests, reversals, and other procedural operations on epistemic actions are rational, irrational, or more or less rational. A procedure's compliance with an epistemic norm, in PLEN, is characterized by restrictions. Restrictions divide and order ways of building procedures. A way of building a procedure is rational according to a norm if that way of building the procedure is in the right place with respect to the restriction. An ideally rational agent will structure procedures accordingly.

As in Chapter 2, some details of D3(d) are spelled out by these further assertions:

(D3.1) Normative restrictions are distinct from constraints of capacity and feasibility.

(D3.2) Particular models of the actual evolution of a reasoner's epistemic states can be compared with the correct evolution as defined by the norms of rationality.

As argued, capacity and feasibility impose bounds on epistemic dynamics; there are formally, physically, or logically possible states, transitions, and processes that are ruled out by various factors. These bounds are distinct from those imposed by epistemic normativity. There are perfectly feasible (e.g.) processes that are not rationally permissible. A model of the feasible courses of epistemic development is thus very different from a model of the normative courses of epistemic development.

The D3 corollary was spelled out in Chapter 2 with the following assertions:

(D3c.1) Epistemic rationality requires the execution of sequential and iterative procedures.

(D3c.2) Epistemic rationality requires parallel execution of procedures.

(D3c.3) Epistemic rationality requires nondeterministic choice among procedures.

(D3c.4) Sequential and iterative procedures determine sequentially composed paths through epistemic state space.

(D3c.5) Parallel execution procedures determine converging and corresponding paths through epistemic state space.

(D3c.6) Procedures instructing nondeterministic choice determine branching paths through epistemic state space.

D3c.1-D3c.3 simply outline some of the procedural operations that epistemic norms rule in. The rest correlate features of epistemic dynamics to the procedural operations that are ruled in by various norms thereon. For instance, if a norm rules in - as rational - a sequential composition of actions, then, given that processes sequences of actions that transform epistemic states, there is a corresponding rational process.

The formal assertion corresponding to D3 and its corollary has not been introduced above, but has been implicated by it. Naturally, all of D3 can be gotten in PLEN by means of restrictions. The proper assertion to which D3 (and its corollary) corresponds is this:

((RRT) Restriction Representation Theorem) Protocols define restrictions on every aspect of epistemic dynamics. More precisely:

- **(RRT.1)** Protocols define restrictions on the construction of epistemic procedures: for every protocol, π, there is a restriction on procedural construction, $R^{\pi}_{\leq}(\Pi_{Sch})$, that can be read off of the syntax of π in the sense that there are simple rules relating the syntax of π to a determination of whether a given protocol schema is in the rational or irrational set in $R^{\pi}_{\leq}(\Pi_{Sch})$.
- **(RRT.2)** Protocols define restrictions on epistemic processes: for every protocol, π, there is a set of restrictions on processes per EDG, g, $R^{\pi}_{\leq}(P_g)$, that can be read off of the syntax of π in the sense that there are simple rules relating the syntax of π to a determination of whether a given process is in the rational or irrational set in $R^{\pi}_{\leq}(P_g)$.
- **(RRT.3)** Protocols define restrictions on epistemic transitions: for every protocol, π, there is a set of restrictions on transitions per EDG, g, $R^{\pi}_{\leq}(T_g)$, that can be read off of the syntax of π in the sense that there are simple rules relating the syntax of π to a determination of whether a given transition is in the rational or irrational set in $R^{\pi}_{\leq}(T_g)$.
- **(RRT.4)** Protocols define restrictions on epistemic states: for every protocol, π, there is a set of restrictions on states per EDG, g, $R^{\pi}_{\leq}(W_g)$, that can be read off of the syntax of π in the sense that there are simple rules relating the syntax of π to a determination of whether a given state is in the rational or irrational set in $R^{\pi}_{\leq}(W_g)$.

This theorem is a short step from the syntactic schemata, matching criteria, restrictions on procedural construction, and Reduction Lemma of section 4.3.4. A proof is available in the separate technical document [135].

Given the canonical interpretation of PLEN such that protocols are the formal objects representing epistemic rationality norms, this theorem straightforwardly encodes D3. Every protocol is mapped by the execution rules to restrictions on procedural construction, processes, transitions, and states by means of its syntax. The syntax of the protocols determines - via the rules - the extension of the rational subsets in these restrictions. The D3 corollary corresponds to RRT.1 and the theorem itself suffices to capture D3. Every protocol (hence, norm) is mapped to a restriction on dynamics such that the syntax (hence, procedural instructions) of the protocol determines what elements of epistemic dynamics or what procedural constructions are in what parts of the restriction mapped to it. It is straightforward to see how this works, and the Reduction Lemma and Reading Off Theorem above essentially provide the details.

As for assertions D3.1, D3.2, and D3(a) through D3(d), now that we have seen RRT, it's easy to see how sections 4.4.2 and 4.4.3 above essentially took care of these. The formal assertions that correspond are trivial corollaries of RRT and graph-theoretic properties of EDGs. RRT.2-RRT.3 correspond to D3(a)-D3(c). D3(d), fleshed out with D3.1 and D3.2, corresponds to the trivial assertion that, for each protocol, π, there is a state, w in some g, such that the feasible states, transitions, and processes at e may not identical to the rational elements per the restrictions of π. These sets can be directly compared for basic set theoretic relations.

D3c.1-D3c.3 follow simply from the fact that epistemic protocols represent epistemic norms and are built out of control operators that represent procedural operations of sequential composition, parallel execution, and indeterministic choice. D3c.4-D3c.6 correspond to GTP1 and GTP2:

GTP1 For all protocols, π, and EDGs, g, if $g \in G^{\pi}$, where $g \in G^{\pi}$ iff all subprotocols of π and all computation sequences [113] of π are executable in g, and g is the smallest graph that does so while also meeting other structural conditions [135], then the syntactic properties of π correspond to the graph-theoretic properties of g.

GTP2 For all protocols, π, and EDGs, g, if π is executable in g, then the syntactic properties of π correspond to the graph-theoretic properties of subgraphs of g.

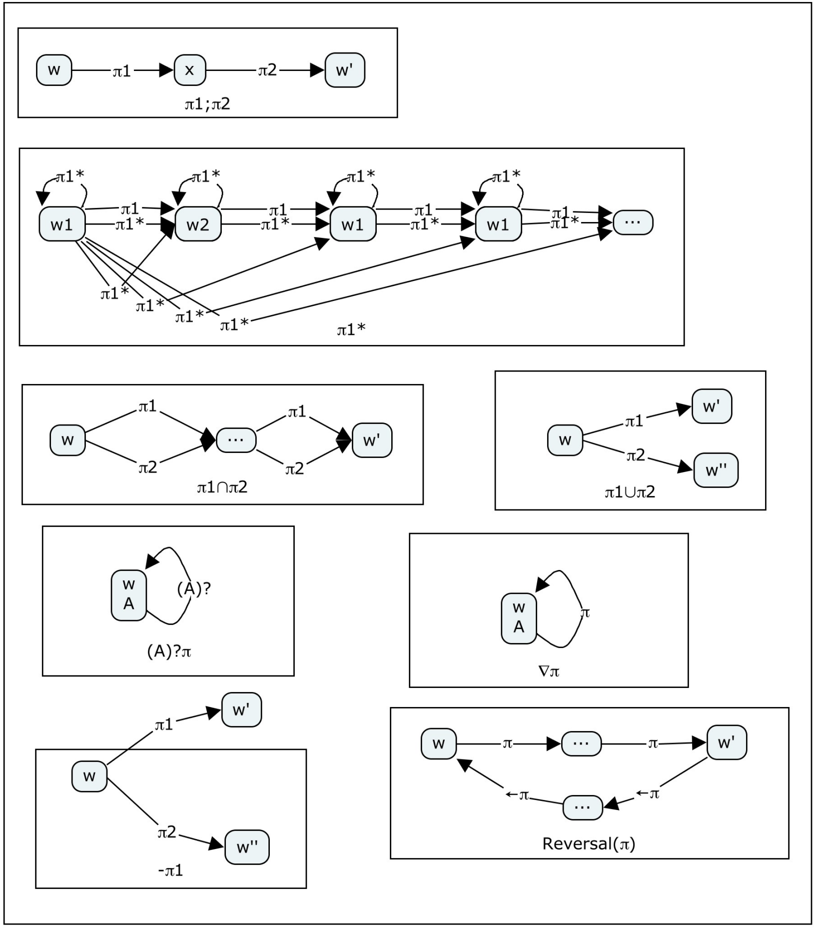
These are proven in [135]. To spell these claims out, I simply note some features that are shown there. For example, if there is a $\pi \in P_{PLEN}$ such that $g \in G^{\pi}$, then the number of nested sequential compositions in π is proportional to the depth of g where depth of g is determined by the length of maximal paths in g (the finite paths that are not subpaths of any other paths) if g is non-looping. That is, for a regular protocol π, every sequential composition that is not nested in a parallel execution or choice corresponds to a pair of concatenated paths. This property extends to any graphs that execute protocols. For any g, $p \in P_g$, for any protocol, π, if p executes π, appending an atomic to π via ; adds 1 to the depth, n, of p. In general, the executability of a protocol at a given state imposes structural requirements on the EDG, including conditions on the edges and on the content functions. This can be easily seen; the truth conditions of the operators pair with the execution conditions of protocols to dictate graph-theoretic properties. Let w be a state at which π is executable and $[\pi]A$ is true. If π is a sequential composition, for instance, this dictates a minimum length for any path of which w is the root, and it dictates that the terminal state of that path verifies A. If π is indeterministic at w, that is, if there is more than one path executing π at w, then there is a tree of minimum width (and, depending on π's structure, depth) in which each leaf node verifies A.

Each control operator induces other graph-theoretic features: corresponding paths, branching paths, etc.. It's trivial to see how conditions on content functions in an EDG are constrained by the truth of these claims at a state.[9] In general, recall how the simple diagram of the executions of complex protocols illustrates this in Figure 4.2. Properties of EDGs like depth and width, as well other features. For some examples: The length of the executions of

---

[9] However, some these claims are nonetheless fleshed out in [135].

Figure 4.2: Induced Graph-theoretic Structure



π are proportional to the number of X_i in the immediate subprotocol hierarchies of implementations of π with main operator ;. It follows that the subgraphs, g', of any g in which π is executable such that the maximal paths of g' are executions of implementations of π are such that the depth of g' is proportional to the number of sequential compositions in π. If the main operator of π is $*$, then any execution of π is in the reflexive, transitive closure of the edge relation of π. If the main operator of π is $\cap$, then, for any EDG, g, in which π is executable, there is an $(s, g') \leq g$ that is a lattice. I'll let the illustration of this point rest with these facts.

4.5 Conclusion

This chapter is an extended argument for the PLEN Proposition. The first section argued that PLEN's semantics contain structures that adequately represent epistemic dynamics and its various parts. Epistemic states were modeled with content states that encode attitudes of acceptance, rejection, and endorsement of methodologies while imposing no substantive restrictions on how these sub-states interact. Epistemic dynamics were modeled by transitions, edges representing actions, and procedural paths representing processes. It was then shown that PLEN's representation of all of these elements of epistemic dynamics was theoretically adequate insofar as it solves or evades the Processes and Constraints problems (and the knock-on problems that followed them).

I then spelled out the slogan of PLEN: epistemic rationality norms are protocols. I showed that for a large and representative class of epistemic rationality norms, all of the most important things to derive about norms can be captured by restrictions on dynamics, and all of the important things to derive from restrictions on epistemic dynamics can be captured by restrictions on procedural construction. I then showed in outline that these could be read off of the syntax of protocols. This shows that PLEN solves or evades the First-Class Citizenship problem.

Finally, on the basis of the representation of epistemic norms with protocols, each of D1-D3 were shown to be correlated with formal assertions in the object language of PLEN or features of the PLEN formalism itself. In particular, each of D1-D3 were shown to correlate to the properties of protocols. D3, in particular, was shown to be analyzed in terms of restrictions on epistemic dynamics which, by the theorem called RRT, was reduced to properties of protocols.

The PLEN Proposition fits into a theory preference argument for PLEN. Given that solving the Processes, Constraints, and First-Class Citizenship problems and formalization of each of D1-D3 are criteria or desiderata of theoretical adequacy, the PLEN Proposition shows - in a qualitative fashion - that PLEN is a theoretically adequate formal framework for analyzing epistemic rationality and its norms. This takes us some way to completion of the central part of the dissertation: defining PLEN and displaying its virtues. The results of the argument, thus far, are summarized in the following, partially empty table:

(Key) Systems in this family are generally:

- + :: Adequate
- - :: Somewhat Adequate, Some Systems Adequate, or Unclear
- X :: Inadequate

	Norm Exogeny: Broad or Narrow	Modeling Epistemic Dynamics: Constraints Problem, Process-Transition Distinction, Restrictions	Desiderata: D1-D3	Paradoxes and Other Tradeoffs: Paradox, Generality, Simplicity, ...
Belief Revision Systems				
Dynamic Logics				
Dynamic Epistemic Logics				
Deontic Logics				
PLEN	+	+	+	-

	D1	D2	D3
Belief Revision Systems			
Dynamic Logics			
Dynamic Epistemic Logics			
Deontic Logics			
PLEN	+	+	+

The rest of this table is filled out in Chapter 5. For now, it's enough to note that the table records the following facts. PLEN is "narrowly norm exogenous".[10] Where an action logic is exogenous for having program expressions that are interpreted as names for actions in its logic, PLEN's protocols are interpreted as norm expressions. Each protocol is an expression that encodes the instruction of an epistemic norm. So much for First-Class Citizenship. As argued in Section 2, PLEN's underlying model of epistemic dynamics handles the Processes and Constraints problems readily. Section 3 shows how PLEN represents restrictions on epistemic dynamics, and associates each protocol (i.e., norm expression) with restrictions on epistemic dynamics. Section 4 provides formalizations of D1-D3 by showing how formal assertions in PLEN's object languages or about PLEN's constructions correspond to the

---

[10] See section 5.1 of Chapter 5 for explication of this term and its importance.

features of epistemic rationality and its norms asserted in each. The Paradoxes and Tradeoffs column is set to "-" by default, as I've yet to stumble on any paradoxes or other serious theoretical tradeoffs for PLEN's positive features.[11] PLEN is extremely general; it imposes no constraints on what action names are admissible in its protocol languages or representable by its models. It's also a simple formalism with relatively few moving parts and a great deal of flexibility and room for modification.

Chapter 4 has filled in the PLEN row on these evaluative matrices. The second step of the theory preference argument is moving from PLEN's virtues to PLEN's overall preferability among its rivals. That is the project of Part III. Chapter fills in the remaining rows; I argue that, among the alternative formal frameworks for epistemic norms, PLEN comes out best.

Appendix: Analyzing Deontics in PLEN

This section merely closes lacunae in the main argument of section 4.3. As such, it can be skipped by those already convinced or looking for a better flowing transition into Part III. The goal is to deepen the connection between protocols and epistemic rationality norms, by showing how deontic operators can be re-conceptualized in PLEN. This gives an interpretation of the valences, Ob, Per, and $Forb$ deployed in the argument for Norm-Kernel Analysis. More explicitly, the idea, here, is that the mapping from Val to πVal in the definition of the Protocol Function Theorem (PFT) is actually principled; each intuitive deontic operator in Val is mapped to a plausible formal analysis of the operator in PLEN. Now, strictly speaking, the mapping alone gives us the ability to read off verdicts from protocols, but the following material makes the notion of agreement in the PFT more substantive and provides a bit of justification for the mapping used in the technical result.

To analyze deontics in PLEN, it's useful to first have the technical notion of dictation.

Def.) Dictation Norms dictate actions, conditions, or states; this is the most basic notion

---

[11] There isn't, for instance, any obvious disunity in the ways that PLEN solves the Constraints, Processes, or First-Class Citizenship problems.

of restriction or requirement. Norms tell us that we must execute some action, bring about some condition, or be in some state. Protocols do the same work. Protocols even dictate other protocols, whether these specify basic actions or complex procedures.

(1) π *dictates* π', $\pi \Rightarrow \pi'$, iff for all g, every execution of π contains an execution of π' or corresponds to an execution of π'.

(2) π *dictates* π' in g, $\pi \Rightarrow_g \pi'$, iff every execution of π in g contains an execution of π' or corresponds to an execution of π'.

(3) π *dictates* π' in g, $\pi \Rightarrow_{gw} \pi'$, iff for all g, every execution starting at w of π contains an execution of π' or corresponds to an execution of π'.

(4) π *dictates* $\pi_1, ..., \pi_n$ *exclusively* at w in g, $\pi \Rightarrow_{gwx} \pi_1, ..., \pi_n$, iff every execution of π starting at w in g contains or corresponds to executions of $\pi_1, ..., \pi_n$ and no other protocols.

(5) π *dictates* $\pi_1, ..., \pi_n$ *exclusively* in g, $\pi \Rightarrow_{gx} \pi_1, ..., \pi_n$, iff every execution of π in g contains or corresponds to executions of $\pi_1, ..., \pi_n$ and no other protocols.

(6) π *dictates* $\pi_1, ..., \pi_n$ *exclusively*, $\pi \Rightarrow_x \pi_1, ..., \pi_n$, iff every execution of π in any g contains or corresponds to executions of $\pi_1, ..., \pi_n$ and no other protocols.

If π dictates π' at w in g, this tells us that executing π at w in g implies that π' is carried out at some point reachable from s in the restriction of π. The dictation of π' by π in g tells us that every execution of π implies the execution, at some point, of π'. The general dictation of π' by π tells us that π dictates π' for every g. These distinctions are useful for thinking about various levels of particularity in the requirements of norms; general dictation encodes a protocol's imposing uniform requirements on all agents. Dictation in g encodes a protocols requirements on the agent whose arena is g. A protocol's dictating π' at w in g encodes the protocol's imposition of requirements unfolding from a particular state.

Now, the basic plan for formalizing deontic notions in PLEN is the following set of rules, where we let p be a subpath of q, $p \preceq q$, iff $p = \langle e_1, ..., e_n \rangle$ and $q = \langle e_i, ..., e_j \rangle$ where $i \geq 1$ and $j \leq n$:

Requirement/Obligation π requires/obligates π', $\pi R \pi'$, iff $\forall p \exists q(p \in P_\pi \rightarrow (q \in P_{\pi'} \land q \preceq p))$ (iff π dictates π')

Permission π permits π', $\pi P \pi'$, iff $\exists p \exists q(p \in P_\pi \land q \in P_{\pi'} \land q \preceq p)$ (iff π doesn't dictate π')

Prohibition/Forbiddance π prohibits/forbids π', $\pi F \pi'$, iff $\forall p \forall q((p \in P_\pi \land q \in P_{\pi'}) \rightarrow q \not\preceq p))$ (iff π dictates $-\pi'$)

We can punch these formalizations into the object language of PLEN by defining variants that are true over states rather than paths per graph:

State Requirement π requires π', $\pi R \pi'$, at w iff $\forall p \exists q(w = first(p) \land p \in P_\pi \rightarrow (q \in P_{\pi'} \land q \preceq p))$

State Permission π permits π', $\pi P \pi'$, at w iff $\exists p \exists q(w = first(p) \land p \in P_\pi \land q \in P_{\pi'} \land q \preceq p)$

State Prohibition π prohibits π', $\pi F \pi'$, at w iff $\forall p \forall q((w = first(p) \land p \in P_\pi \land q \preceq p) \rightarrow q \in P_{-\pi'})$

Finally, here is the full-on definition of the deontic operators in PLEN:

(Rational Obligation wrt π) $w \Vdash RO_\pi(\pi')$ iff $\pi R \pi'$ at w

(Rational Permission wrt π) $w \Vdash RP_\pi(\pi')$ iff $\pi P \pi'$ at w

(Rational Forbiddance wrt π) $w \Vdash RF_\pi(\pi')$ iff $\pi F \pi'$ at w

These are defined for application to protocols, but the extension to formulae of L_{PLEN} is straightforward.

These basic translations of deontics are encoded in the definition of protocol functions for the proof of PFT. Protocol functions basically just map Ob, Per, and $Forb$ to the foregoing translations of deontic notions.

Individual Protocol Functions Let g be an EDG. Let $\pi Val = \{Ob, Per, Forb,\}$ and define the duals as usual. Define the domains and conditions of application of protocol functions as in the main text. Then: $f^g_\pi : (\pi Con(g), \pi Dom(R)(g)) \longrightarrow \pi Val$ is such that:

(i) If $A \in \pi Con(g)$ and $x \in \pi Dom(R)\Pi(g)$, then f^g_π maps (A, x) to Ob iff for all w, if $w \Vdash A$, then $\pi \Rightarrow_{gw} x$.

(ii) If $A \in \pi Con(g)$ and $x \in \pi Dom(R)\Pi(g)$, then f^g_π maps (A, x) to Per iff for some w, $w \Vdash A$, and $\pi \Rightarrow_{gw} x$.

(iii) If $A \in \pi Con(g)$ and $x \in \pi Dom(R)\Pi(g)$, then f^g_π maps (A, x) to $Forb$ iff for all w, if $w \Vdash A$, then $\pi \Rightarrow_{gwx} -x$.

(iv) If $A \in \pi Con(g)$ and $x \in \pi Dom(R)Con(g)$, then f^g_π maps (A, x) to Ob iff for all w, if $w \Vdash A$, then $w \Vdash [\pi]x$.

(v) If $A \in \pi Con(g)$ and$x \in \pi Dom(R)Con(g)$, then f^g_π maps (A, x) to Per iff for some w, $w \Vdash A$, and $w \Vdash \langle\pi\rangle\, x$.

(vi) If $A \in \pi Con(g)$ and $x \in \pi Dom(R)Con(g)$, then f^g_π maps (A, x) to $Forb$ iff for all w, if $w \Vdash A$, then $w \Vdash ([\pi] - x$.

(vii) If $A \in \pi Con(g)$ and $x \in \pi Dom(R)Stat(g)$, then f^g_π maps (A, x) to Ob iff for all $p \in P_\pi$ such that $w = first(p)$, if $w \Vdash A$, then $x = last(p)$.

(viii) If $A \in \pi Con(g)$ and $x \in \pi Dom(R)Stat(g)$, then f^g_π maps (A, x) to Per iff for some $p \in P_\pi$ such that $w = first(p)$, $w \Vdash A$, and $x = last(p)$.

(ix) If $A \in \pi Con(g)$ and $x \in \pi Dom(R)Stat(g)$, then π maps (A, x) to $Forb$ iff for all $p \in P_\pi$ such that $w = first(p)$, if $w \Vdash A$, then $-x = last(p)$.

(x) If $A \in \pi Con(g)$ and $x \in \pi Dom(R)Tran(g)$, then f^g_π maps (A, x) to Ob iff for all $p \in P_\pi$ such that $w = first(p)$, if $w \Vdash A$, then $x = \langle w, last(p)\rangle$.

- **(xi)** If $A \in \pi Con(g)$ and $x \in \pi Dom(R)Tran(g)$, then f^g_π maps (A, x) to Per iff for some $p \in P_\pi$ such that $w = first(p)$, $w \Vdash A$, and $x = \langle w, last(p) \rangle$.
- **(xii)** If $A \in \pi Con(g)$ and $x \in \pi Dom(R)Tran(g)$, then f^g_π maps (A, x) to $Forb$ iff for all $p \in P_\pi$ such that $w = first(p)$, if $w \Vdash A$, then $x = \langle w, last(p) \rangle$.
- **(xiii)** If $A \in \pi Con(g)$ and $x \in \pi Dom(R)Proc(g)$, then f^g_π maps (A, x) to Ob iff for all $p \in P_\pi$ such that $w = first(p)$, if $w \Vdash A$, then x is a subpath of p.
- **(xiv)** If $A \in \pi Con(g)$ and $x \in \pi Dom(R)Proc(g)$, then f^g_π maps (A, x) to Per iff for some $p \in P_\pi$ such that $w = first(p)$, $w \Vdash A$, and x is a subpath of p.
- **(xv)** If $A \in \pi Con(g)$ and $x \in \pi Dom(R)Proc(g)$, then f^g_π maps (A, x) to $Forb$ iff for all $p \in P_\pi$ such that $w = first(p)$, if $w \Vdash A$, then x is a subpath of p.

All of the foregoing analyzes deontics with respect to specific protocols/norms. To encode what is generally rationally obligatory (etc.), replace π in $RO_\pi(\pi')$ with an entire methodology: $RO_M(\pi')$. This isn't strictly speaking well-formed in PLEN. So, strictly speaking, add formulae to L_{PLEN} by the following rules:

(Methodological Deontics) **(O)** If $\pi \in P_{PLEN}$, then $RO_M(\pi) \in L_{PLEN}$, and if $A \in L_{PLEN}$, then $RO_M(A) \in L_{PLEN}$.

- **(P)** If $\pi \in P_{PLEN}$, then $RP_M(\pi) \in L_{PLEN}$, and if $A \in L_{PLEN}$, then $RP_M(A) \in L_{PLEN}$.
- **(F)** If $\pi \in P_{PLEN}$, then $RF_M(\pi) \in L_{PLEN}$, and if $A \in L_{PLEN}$, then $RF_M(A) \in L_{PLEN}$.

Then, define truth rules for them:

(Rational Obligation wrt M) $w \Vdash RO_M(\pi')$ iff for every p in the intersection of $P_{\pi i}$ for all π_i such that $w \Vdash M(\pi_i)$ is such that if $w = first(p)$, then there is a $q \in P_{\pi'}$ such that $q \preceq p$.

(Rational Permission wrt M) $w \Vdash RP_M(\pi')$ iff for some p in the intersection of $P_{\pi i}$ for all π_i such that $w \Vdash M(\pi_i)$ is such that if $w = first(p)$, then there is a $q \in P_{\pi'}$ such that $q \preceq p$.

(Rational Forbiddance wrt M) $w \Vdash RF_M(\pi')$ iff for no p in the intersection of $P_{\pi i}$ for all π_i such that $w \Vdash M(\pi_i)$ is such that if $w = first(p)$, then there is a $q \in P_{\pi'}$ such that $q \preceq p$.

Again, extension to objects other than protocols in the language of L_{PLEN} is straightforward.

To the foregoing deontic operators, we can add the analysis of defeasible reasons. One has a defeasible reason to believe A if one has a reason to believe it, but that reason is inconclusive because it may be outweighed or undermined on some general balance of reasons. There are two key conditions to consider in formalizing some defeasible reason R for belief in A: (a) If R makes no difference to the permissibility or obligation of belief in A under any norms of rationality under any conditions, then R isn't a reason for belief in A at all. (b) If R cannot fail to obligate belief in A (under norms of rationality under any conditions), then R can't be defeasible. There is machinery for formalizing defeasible reasons in PLEN. I'll introduce one apparatus.

We want to make sense of how R can make a difference to the permissibility or obligation of belief in A. The translations above provide a general framework. The reader can fill in the obvious duals. The key idea for formalizing defeasible reasons is that R must encode the difference between conditions under which a methodology, M, permits or obligates belief in A. Let R be a putative reason for A. Let $R*$ be some set of propositions excluding R but compatible with R. Let $R**$ be some set of propositions including R. Then:

Defeasible Reason Say that $RD_\pi(A)$ iff one has a defeasible reason, R, for accepting A A proposition, $R \in L_{PLEN}$ (just a propositional variable with a suggestive name, not to be confused with other uses of capital "R" in the notation), is a *defeasible reason* for accepting A (relative to π) iff there are $R*$ and $R**$ (also propositional variables of

L_{PLEN}) such that:

(O) (i) $RP_\pi(b(R*) \wedge d(A))$, (ii) $RO_\pi(b(R * \wedge R) \rightarrow b(A))$, and (iii) $RP_\pi(b(R**) \wedge d(A))$.

OR:

(P) (iv) $RO_\pi(b(R*) \rightarrow d(A))$, (v) $RP_\pi(b(R * \wedge R) \wedge b(A))$, and (vi) $RO_\pi(b(R**) \rightarrow d(A))$.

Extend the definitions in the natural way to methodologies M, then add $RD_\pi(A)$ and $RD_M(A)$ to the language as in Chapter 4. The first set of conditions tells us that, according to π, any state at which $R*$ is believed is one in which one can adhere to π and end up rejecting A, but if one adds belief in R, then adhering to π requires accepting A. However, the final condition tells us that belief in R alone does not guarantee that adhering to π requires acceptance of A because there is a full set of reasons (possibly including R via, monotonicity of $\rightarrow$ in classical logic, which, for the moment, we use in PLEN) that doesn't require belief in A. The second set of conditions tells us that, according to π, any state at which $R*$ is believed is one in which one cannot adhere to π and end up accepting A, but if one adds belief in R, then adhering to π permits accepting A. However, the final condition tells us that belief in R alone does not guarantee that adhering to π permits acceptance of A because there is a full set of reasons including R (possibly including R via, monotonicity of $\rightarrow$) that doesn't permit acceptance of A.

Given the usual dual definitions of obligation and permission, the intuitive thought of the definition comes to the idea that defeasible reasons for A make a difference to the permissibility and impermissibility of accepting or rejection A, but not a difference that can't be neutralized by other reasons. Defeasible reasons to reject A can be formalized in the obvious, complementary way. Encode the righthand side of the definition of defeasible reasons in PLEN with respect to π, acceptance, and A with the expression: $D_\pi(b(A))$. Encode the complementary defeasible reasons with respect to π, rejection, and A: $D_\pi(d(A))$. From here, the analysis of defeasible reasons is ignored because it is reducible to the analysis of the deontic operators.

Finally, we can extend the suite of deontic operators to contain optionality and omissibility which are, arguably, not analyzable in terms of the others [89]. They can be formalized in PLEN semantically like so:

Optionality π' is optional under π, $\neg(\pi P \pi') \wedge \neg(\pi R \pi')$, iff $\exists p \exists q (p \in P_\pi \wedge q \in P_{\pi'} \wedge q \preceq p)$ and $\exists p \forall q (p \in P_\pi \wedge q \preceq p \rightarrow q \notin P_{\pi'})$

State $\neg(\pi P \pi') \wedge \neg(\pi R \pi')$ at w iff $\exists p \exists q (w = first(p) \wedge p \in P_\pi \wedge q \in P_{\pi'} \wedge q \preceq p)$ and $\exists p \forall q (w = first(p) \wedge p \in P_\pi \wedge q \preceq p \rightarrow q \notin P_{\pi'})$

Omissibility π permits omission of π', $\neg(\pi R \pi')$, iff $\exists p \forall q (p \in P_\pi \wedge q \preceq p \rightarrow q \notin P_{\pi'})$

State $\neg(\pi R \pi')$ at w iff $\exists p \forall q (w = first(p) \wedge p \in P_\pi \wedge q \preceq p \rightarrow q \notin P_{\pi'})$

These notions can be added to the object language of PLEN easily enough, and from there, added to the definition of protocol functions. With this machinery, we can extend the arguments in Chapter 4 beyond the limited range of deontic norm-kernels that were targeted in Chapter 4.

Part III

Defending PLEN

Chapter 5

PLEN vs. Rival Formalisms

5.1 Recapitulation: Modeling Epistemic Dynamics and Formalizing Norms

The protocol theoretic account of epistemic rationality norms is defended by means of a theory preference argument. The account, formalized by PLEN, satisfies a set of critically important theoretical desiderata on formal accounts of epistemic rationality norms. As shown in Chapter 4, PLEN achieves all of the following goals:

- (i) provision of devices for *explicitly* representing and reasoning about (protocols that encode) norms of epistemic rationality,
- (ii) capturing of the distinction between processes of epistemic change and the changes they bring about when completed, and
- (iii) formalization of the features of epistemic rationality articulated in D1-D3 that are inadequately recognized in both the formal and traditional epistemology literatures.

These goals characterize what I take to be necessary conditions - or at least important desiderata - for the project of giving an adequate account of epistemic norms with an adequate underlying model of epistemic dynamics.

The first goal requires what I'll call "norm exogeny". Normally, a logic of action is exogenous if the logic's languages contain expressions that explicitly name or describe actions. As a result of the success of program logics in the logic of action, this is often cashed out in terms of containing a language of program expressions [113, 230]. Norm exogeny analogously requires that a logic's languages contain norm expressions: formal expressions that explicitly name or describe the structure or content of norms. This is what I mean by devices for "explicitly" representing and reasoning about norms.

A broader condition of exogeny could be satisfied by there just being some features of the formalism that correspond in a perspicuous way to actions. If there are some explicit statements, theorems, or objects stipulated to represent actions or that satisfy (formal assertions that correspond to) the intuitive features of actions, say that the formalism has the feature of *broad action exogeny.* Hence, *broad norm exogeny* requires that there be some features of the formalism that perspicuously correspond to norms. If there are some explicit meta-language statements, theorems, or objects stipulated to represent norms or that satisfy (formal assertions that correspond to) the intuitive features of norms, the formalism has the feature of broad norm exogeny.

Now, a formalism can very well be deeply theoretically interesting and useful with only broad norm exogeny, but it's preferable for a formalism to be narrowly norm exogenous, as this enables the possibility of a proof system for encoding explicit reasoning about norms. The soundness of such a system shows the soundness of the formalized reasoning. More, there are philosophically important forms of argument concerning equivalence, dependence, and independence relationships among epistemic norms that only an exogenous logic can formalize [135]. Exogenous logics are thus preferable to non-exogenous logics, and narrow exogeny is preferable to broad exogeny. Or so I'll assume going forward. It's shown that PLEN formalizes the aforementioned forms of reasoning about epistemic norms in the technical document [135] in a far more robust way than is shown in this thesis. For the purposes of the thesis, simply recall that the axiomatization of PLEN, Syntax II, is sound, and that it provides machinery for deriving and proving expressions encoding statements about various properties of epistemic methodologies M, procedural features of protocols (i.e., epistemic norms), and verdict relations among objects in L_{PLEN}.

The protocols of P_{PLEN} account for the norm exogeny of PLEN; they explicitly represent epistemic rationality norms and are embedded in formulae of L_{PLEN}. The semantics of PLEN provide models that formally distinguish between processes of epistemic state change and mere transitions among epistemic states via the difference between procedural paths and transitions. The laws governing protocols thus provide theories governing various properties of epistemic rationality norms. The derivation systems of PLEN thus formalize explicit, valid reasoning about these properties, including valid reasoning about equivalence relations among epistemic rationality norms. This enables PLEN to formalize D1-D3. Formally, this is all captured by the following proposition, proven over the course of Chapter 4:

(The PLEN Proposition) **(i)** The semantics of PLEN contains structures that distinctly represent epistemic states, transitions, processes, actions, procedures, and constraints thereon.

(ii) The syntax and semantics of PLEN provide structures that explicitly represent epistemic rationality norms and their procedural content.

(iii) PLEN explicitly formalizes D1-D3; for each of D1-D3, there is a provable formal assertion at the meta-level of PLEN where formal assertions include: propositions about the semantic structures of PLEN, sets of valid meta-level rules or object-level laws, theorems at the object-level, and object-level theories.

Now, of course, alternative accounts of epistemic rationality norms might also achieve the same goals, and so furnish no preferential argument for the protocol theoretic account. This chapter aims to show that no rival formalism, and so no rival precise accounts of epistemic rationality, does as well as PLEN without either significant theoretical costs or without simply being equivalent to a fragment of PLEN. This is done by considering a set of families of systems as formal frameworks for epistemic rationality and its norms.

Now, the field of formalisms is too vast to survey. As such, I want to focus on what I take to be *rivals* of PLEN with respect to the twin goals of adequately modeling epistemic dynamics and giving an account of epistemic norms. What I mean by a *rival formalism* is this: a formalism F rivals PLEN with respect to the project of this thesis iff:

(i) F has structures (in its syntax or model-theory) that: (a) are explicitly taken to characterize the elements of epistemic dynamics epistemic (e.g., epistemic states, transitions among epistemic states, epistemic actions and processes), (b) could easily be modified to do so, (c) are explicitly taken to encode epistemic norms, or (d) can be easily modified to encode epistemic norms, and

(ii) F isn't identical to any fragment of PLEN, F can't be embedded in PLEN, or F isn't identical to an easy extension of PLEN.

This definition carries an element of relativity: ease of technical modification is hardly a rigorously understood concept. Nonetheless, for the purposes of this thesis, there is a reliable indicator of ease: if I can come up with ways in which a formalism can be modified for certain ends, and it really can be, then it can be easily modified for those ends.

A key feature of this notion of rivalry is, even where a system was not expressly designed to be used to think about epistemic rationality, there may be interesting machinery in the theory for doing just that, so, in what follows, I make extremely modest modifications to the systems (or I suppose them without much detail because they can be safely ignored while seeing the problems to be raised) or to their informal interpretations to see how they fare as accounts of epistemic norms with underlying models of epistemic dynamics. To be clear, the following systems are not all designed to model epistemic dynamics and provide a formal account of epistemic norms. They all do, however, contain machinery that might be useful to this project, and I interpret them as such in the examinations that follow.

For the sake of efficiency, many groups of rival systems will be taken on together in the examination on the basis of their shared features. For instance, offshoots of dynamic logics often build in basically the same action theory - actions are represented with binary relations and programs are encoded with operations on those relations. If this underlying program theory is shown to suffer (e.g.) the Processes problem, then, barring some special additional features, all of the systems with the same basic program theory will suffer the Processes problem. As such, I omit detailed examination of these systems beyond what suffices to show the conditions or desiderata under consideration.[1]

The groups of systems examined with respect to the desiderata above include belief revision theories, dynamic logics, dynamic epistemic logics, and deontic logics.[2] The chapter follows this listing of groups. First, I consider belief revision frameworks, including Bayesian frameworks, and show how they fall short with respect to the desiderata above. Next up for the same treatment are dynamic logics, including PDL, Process Logic (PL), and EPL. The next family are dynamic epistemic logics including epistemic action logics (EALs) and logics of epistemic planning (LEPs). Following these two sections, deontic logics - including von Wright's logic of norms and actions (VWNA), STIT theory (STIT), and dynamic deontic logics (PDeL) are considered and subjected to the same sort of criticisms. Finally, I close by considering PLEN's modularity - one can add, subtract, and restrict elements of PLEN to generate formalisms that go beyond the expressive and representational power of PLEN. I omit detailed explication and/or definition of systems for which I provided explication or definition in prior chapters (e.g., AGM, PDL, EPL, DEL, ...).

---

[1] Which is to say that it may seem that I am ignoring the technical and philosophical merits of the systems. I certainly don't want to downplay any of the technical or philosophical merits of these systems outside of the project of modeling epistemic dynamics and accounting for epistemic norms; rather, I merely want to show how these systems fall short with respect to precisely that problem. Their value for other projects is unquestioned, here.

[2] The focus on these systems is not meant to ignore the enormous amount of work in the intersection of AI, deontic and imperative logic, epistemic logic, and non-monotonic logic. See, e.g., [89]. Much of this work focuses on finding new formal devices for representing normative discourse from dynamic perspectives and resolving the old problems of deontic logic. This work often formalizes important conceptual distinctions that are relevant to this project, such as that between norms (themselves) and normative propositions, and the distinctions remarked on above. Rather, the focus of this project is meant to throw a spotlight on the formalisms that are PLEN's nearest philosophical neighbors.

5.1.1 Schematic Problems

There are just too many formalisms to evaluate in any depth in a single chapter; there are dozens upon dozens of variants and modifications of modifications of doxastic dynamic logic built out of PDL [228, 229, 227, 34] or action-specific variants of DEL [65, 193, 99] or twists on AGM [224, 211, 256]. As such, I break the project into general evaluations of each of four groups that are based on their general features. For instance, I articulate general arguments for the theoretical inadequacy of PDL and PDL-based theories. Rather than assessing the technical details of each particular formalism that has spun out of PDL and might plausibly be used to think about epistemic dynamics, I identify a feature of PDL-based theories that supports a general strategy of objection. This results in identifying a number of schematic arguments that can be applied against the groups of formalism in virtue of their shared features.

Some groups simply fail to have machinery that represents epistemic norms. Other systems fail to achieve a high enough degree of "model validity" by failing to provide models of epistemic dynamics that capture important truths about dynamics.[3] Some groups fail to formalize some of D1-D3. Still others are plagued by paradoxes or general theoretical costs that are too steep to pay. Some groups suffer more than one of these problems. Accordingly, the rest of the chapter unfolds by applying the following schematic problems alone or in groups to rival formalisms:

1. Failure of Epistemic Norm Exogeny: System S doesn't have any (syntactic or semantic) constructions that encode the structure or anatomy of epistemic norms.
    (a) Special Case 1: System S is a logic and doesn't have anything in its languages to encode directives, instructions, plans, or programs/protocols.

---

[3] I mean this in the senses implicit in [178, 12]. Roughly, models are valid to the extent that the mathematical structures that make up the model suitably resemble and/or correlate with the parts of the systems one is trying to give an account of, and the relations among the structures suitably resemble and/or correlate with the relations among those parts.

   (b) Special Case 2: System S is a logic and doesn't have anything in its languages to encode norm-kernels, i.e., conditions, deontics, or formulae that encode suitable conditions of application.

   (c) Special Case 3: System S is a logic and doesn't have an action language; i.e., there's nothing to encode epistemic actions that change the world, epistemic actions that are not uniform across all epistemic states, epistemic actions that are not uniform across preconditions, or epistemic programs that are formed from mixed world-changing and information changing actions. Without encodings of actions, plausible expressions that encode norms like instructions, plans, or programs will not be encoded in the language.

2. Failure in the Model Epistemic Dynamics: System S doesn't have (syntactic or semantic) constructions that represent feature F of epistemic dynamics.

   (a) Special Case 1: The Constraints Problem

   (b) Special Case 2: The Processes Problem

   (c) Special Case 3: System S doesn't have anything in its constructions to represent: epistemic actions that change the world, epistemic actions that are not uniform across all epistemic states, epistemic actions that are not uniform across preconditions, or epistemic programs that are formed from mixed world-changing and information changing actions. Without representations of these actions, epistemic dynamics is inaccurately modeled and epistemic norms that restrict such actions are unanalyzable.

3. Failure of Desideratum in D1-D3: For some of D1-D3, there is no corresponding formal assertion of S, where formal assertions include: propositions about the (semantic - where S is a logic) structures of S, sets of valid meta-level rules or object-level laws of S, theorems at the object-level of S, and object-level theories in S.

4. Paradox or Other Theoretical Tradeoff: System S suffers from paradoxes or other theoretical tradeoffs (explicated where applied):
    (a) Paradox
    (b) Lack of Generality
    (c) Lack of Uniformity of Problem Solutions

The idea is that every rival formal framework below will be shown to suffer at least one of these schematic problems. The simplest case will be when every member of the family of formalisms in the framework simply lacks the machinery for norm exogeny, adequate modeling of epistemic dynamics, the desiderata, or suffers some unresolved paradoxes or tradeoffs. A more interesting case that sometimes arises is when the formalisms are incompatible with adding machinery that would represent or formalize the relevant notions; some formalisms can't be amended to handle the correct representation of epistemic dynamics or epistemic norms without paradox or breakdown of some kinds. These will, of course, be spelled out in some detail below. Of course, if the arguments of Chapter 4 worked, PLEN survives all of these schematic problems. And that's the thrust of the defense of PLEN.

5.2 Belief Revision

There are multiple paradigms for models of dynamic belief revision. The most well-known and developed are AGM [5, 91] and Bayesian theories [182]. Both of these paradigms encompass multiple exact formalisms built around the same basic notions. AGM is built around belief sets, deductively closed sets of sentences of a formal language. There are, for instance, non-classical versions of AGM [216, 118] that close belief sets under non-classical consequence relations, variants of AGM that reject some or other rationality postulates (rules characterizing rational belief set changes) [211, 52, 256], and extensions of AGM to iterated belief revision [237, 118], in addition to variants based on possible worlds representations of belief sets [108] and predecessor [4] formalisms.

Bayesian theories are centered on probability functions into sets of sentences of a formal language. These are taken to encode credences or partial beliefs. Alternative Bayesian theories arise under distinct probability theories as well as alternative rules for characterizing rational updates of credence/partial belief or theories that build in both partial beliefs and partial disbeliefs as separate functions [75, 80].

A number of alternative belief revision formalisms have been developed in response to the perceived theoretical inadequacies of these frameworks, including Spohn's Ordinal Conditional Functions [233], Tennant's Finite Dependency Networks [245], and Pollock's Defeasible Reasoning Semantics [197, 195].

The foregoing are the rival systems I will focus on in this section.

5.2.1 AGM Theories

In many ways, AGM is the most natural model of epistemic dynamics. Epistemic states are represented simply and intuitively with sets of sentences of a propositional language. The core constructions of AGM are belief sets, deductively closed sets of sentences K (i.e., $K = Cn(K)$). This gives them rather transparent appeal as models of the content of beliefs. This modeling decision also makes it easy to apply classical logic to the analysis.

There are numerous belief revision theories based on the AGM framework or, if you will, versions of AGM that can be built by modifying the basic formalism [256, 118, 90]. These belief revision formalisms can be taken out in swathes. Every version of AGM - in virtue of being based on the same basic formalism - suffers the same problems. The core problem among AGM theories is that epistemic dynamics reduces to transitions among epistemic states. Thus, AGM theories suffer at least the Processes problems, and this spreads to the failure to formalize several of D2-D3.

Explicitly, dynamics in AGM theories is defined by a set of epistemic commitment functions that take belief sets and inputs (particular sentences in the base formalism) to belief sets. As Gardenfors shows [91], these functions are equivalent to relations on belief sets. Hence, the basic actions of belief change in AGM theories are analyzed with transition relations among belief sets. This is as natural a model of epistemic dynamics as the model of epistemic dynamics; changes of epistemic state consist of transitions among representations of the contents of belief. Unfortunately, this generates a couple serious problems with respect to the goal of providing an underlying model of dynamics for an account of epistemic norms. Before considering these problems, however, there is the indirectly related failure of norm exogeny in AGM theories.

First, AGM theories or AGM-based theories all fail the narrow condition of norm exogeny. Strictly speaking, AGM is a logic-based analysis of epistemic dynamics and doesn't itself provide a logic of epistemic dynamics. As such, it doesn't have an object language into which to embed norm expressions. Thus, AGM theories fail to provide a logical analysis of valid reasoning about epistemic rationality norms.

There are, however, constructions in AGM theories that arguably represent or encode epistemic rationality norms: the rationality and entrenchment postulates.[4]

Rationality Postulates for Contraction For propositional formula A, a belief set K, an an independently defined expansion function $K{+_A}$, a rational contraction function satisfies the following postulates:

- **(-1)** $K{-_A} = Cn(K{-_A})$
- **(-2)** $K{-_A} \subseteq K$
- **(-3)** If $A \notin K$, then $K{-_A} = K$.
- **(-4)** If $A \in K{-_A}$, then $\vdash A$.
- **(-5)** $K \subseteq (K{-_A}){+_A}$
- **(-6)** If A is equivalent to B, then $K{-_A} = K{-_B}$.
- **(-7)** $(K\, {-_A} \cap K{-_B}) \subseteq K{-_{A \wedge B}}$

---

[4] See, for instance, [5] and [118]. I follow the presentation in [91] pretty closely.

(-8) If $A \notin K-_{A\wedge B}$, then $K-_{A\wedge B} \subseteq K-_{A}$.

These are motivated independently in AGM, but can also be shown to be equivalent to the entrenchment postulates.[5] Let $A \leq B$ say that either both A and B are logical truths, that A is not believed, or that the need to discard one of A or B leads to discarding A. The intuitive though is that $\leq$ encodes entrenchment; it says that B is at least as entrenched as A. Then entrenchment is thought to satisfy these conditions:

Entrenchment Postulates (EE1) If $A \leq B$ and $B \leq C$, then $A \leq C$.

(EE2) If $A \vdash B$, then $A \leq B$.

(EE3) $A \leq A \wedge B$ or $B \leq A \wedge B$.

(EE4) If K is consistent, then $A \notin K$ iff $A \leq B$ for all B.

(EE5) If $A \leq B$ for all A, then $\vdash B$.

Given a rational contraction function, the following rule defines epistemic entrenchment:

(CtoE) $A \leq B$ iff $A \notin K-_{A\wedge B}$ or $\vdash A \wedge B$.

Given entrenchment, the following rule defines a rational contraction function:

(EtoC) $K-_{A} = K \cap \{B \mid A < A \vee B\}$ if $\nvdash A$, and $K-_{A} = K$ otherwise.

The rationality and entrenchment postulates are the rules in AGM that specify what the requirements of rational belief revision are. This suffices for broad norm exogeny: entrenchment and rationality do perspicuously correspond to norms of epistemic rationality. Those are the features of AGM that characterize rational belief dynamics. As a result, there is a simple model of compliance with epistemic rationality norms: the epistemic commitment functions. Dynamics in AGM theories just are models of ideally rational belief change with respect to these norms. The entrenchment and rationality norms are not, however, really theoretically adequate.

---

[5] An epistemic commitment function is a partial meet contraction function iff it satisfies (-1) - (-6).

The entrenchment and rationality postulates fail to encode crucial properties of epistemic rationality norms. For one, AGM theories say absolutely nothing about epistemic actions that change the world or epistemic programs that are formed from mixed world-changing and information changing actions. As such, AGM theories cannot escape Special Case 3 of the failure of broad norm exogeny; there are no ways in AGM to represent these kinds of norms. More generally, AGM theories don't say anything about actions or processes as distinct from transitions, and this leaves AGM theories entirely powerless to represent normative restrictions on action or processes where they come apart from restrictions on transitions. This takes us to the second major problem for AGM theories, which, after a bit of discussion, connects back up to this first problem.

AGM theories or AGM-based theories all suffer from the Processes and Constraints problems. Epistemic dynamics, even if we translate into the iterated case, entirely reduces to transitions among belief sets, however those are defined.[6] In other words, no AGM theory can represent the basic distinction between processes and transitions; every process of epistemic change in AGM theories just is a sequence of transitions. Additionally, no AGM theory can represent the constraints on epistemic processes due to feasibility and capacity; there aren't even preconditions in AGM. AGM theories simply provide a model of the space of all formally possible transitions among belief sets and the subset that are defined by the epistemic commitment functions. Thus, there is no distinction between processes and epistemic transitions in AGM theories, and this leaves AGM theories powerless to represent the numerous distinctions and properties of epistemic dynamics articulated in Chapters 1 and 2.

As a result of these problems, no AGM theory can accurately represent properties like methodological inclusion, procedural information, or epistemic distinctions between processes. The methodological inclusion point can be seen as a result of the failure of broad norm exogeny. While AGM theories do have rationality postulates or entrenchment rules

---

[6] This needn't be stepped through for each case; just look at the explicit definitions of the formalisms, re: [91, 92, 256, 216, 211, 52].

that, via representation theorems, explicitly characterize rational epistemic dynamics, there is nothing in them that represents the endorsement or commitment to epistemic norms.[7]

Procedural information is entirely invisible in AGM theories except where it corresponds to differences in belief set transitions. In AGM theories, if two processes (think of the Honest Twin case) define identical transition relations (epistemic commitment functions), they are identical in AGM. Expanding with A by inference is entirely indistinguishable from expanding by observation. Revising as a result of applying some inconsistency or consequence detecting algorithm is indistinguishable from revising as a result of paying undue credence to psychic testimony. And so on. This straightforwardly implies that AGM theories can say nothing of the procedural differences in cases like the Honest Twin.

Normative information is also obscured or neglected. Despite the features of AGM theories that represent epistemic norms (the rationality postulates, entrenchment conditions, etc.) and compliance with them (epistemic commitment functions), nothing in them represents all or even any of the same features of procedural epistemic norms that the arguments of Chapters 1 and 2 suggest are important. The rational distinctions between epistemic processes in the Detective case (the deduction vs the psychic testimony), Bad Reasoning case, or the Honest Twin case are invisible in AGM. Thus, AGM theories cannot adequately represent epistemic dynamics nor can they adequately represent procedural epistemic rationality norms.

As a result of all of this, AGM theories uniformly straightforwardly fail to formalize D2 and D3. AGM theories don't recognize procedural norms at all, let alone the restrictions that procedural norms specify. They thus also incompletely characterize the restrictions on epistemic dynamics associated with epistemic normativity in general as a result of their failure to represent processes at all. Thus, D3 is mischaracterized.

Now, to its credit, AGM is broadly norm exogenous. Some of the primary results of AGM are equivalence results among the features of AGM [91] (pp. 211-241). Further, one

---

[7] This is somewhat ironic as one of AGM's founders, Carlos Alchourron, was explicitly concerned with modeling the revision of norms in the work he did that lead to his contributions to AGM. See [3, 4].

Table 5.1: Belief Revision Theories

AGM	[5, 91]
Multiple Language or Relevance Sensitive AGM	[52, 216]
Non-Classical AGM	[211, 256]
Neighborhoods AGM	[108]
Iterated AGM	[237, 118]
Special Operator AGM	[92, 118]
Offshoots and Hybrids	[67, 69, 84, 91]

can reason about how epistemic dynamics looks under compliance with epistemic norms and about the verdicts of various norms; again, this is illustrated by the fundamental results of AGM (pp. 217-226). However, there is no machinery for reasoning about the features of epistemic dynamics and epistemic rationality that AGM ignores, e.g., procedural requirements, process requirements, procedural structure, and knowing execution. And, of course, none of the available reasoning can be done internally in some logical object language. As Segerberg [226] remarks, "AGM is not really logic; it's a theory about theories." (p. 136)

All of these problems apply equally to all or most of the many versions of AGM in Table 5.1. The application of the central problems generalize to the variants of AGM because the underlying model of dynamics in all of these theories reduces to content-changing transitions among states.

Additionally, for many versions of AGM, there are theoretical problems and tradeoffs that are independent of the desiderata above. The absence of expressions for higher order belief [118], Tennant's Degeneracy Theorem [244], the amnesiac revision function [141], and the controversiality of the Recovery postulate in basic AGM [118] also mar AGM theories - at least the basic theories for which these results were initially proven and any systems to which the results can be extended. More, AGM theories are committed only to the analysis of a specific typology of epistemic commitment functions even in variants modified with operators additional to expansion, contraction, and revision. Thus, AGM theories fail to be general theories of epistemic dynamics, let alone theories with any claim to general analyses of epistemic rationality. The addition of new operators still limits AGM in its generality

- there is no general analysis of epistemic action or process in AGM, and so no analysis of epistemic norms in general given that some norms are procedural. Thus, AGM theories suffer with respect to general theoretical tradeoffs in addition to the failure of the desiderata above.

Finally, AGM theories can be embedded in PLEN satisfactorily.[8] One need only let there be a set of atomic actions e_A, c_A, and r_A that represent the expansion, contraction, and revision functions with respect to the formula A for each propositional formula (i.e., not formulae that embed protocol expressions) of L_{PLEN} (barring application of each of these actions to formulae embedding themselves). Then, one can define subsets of EDGs such that the binary execution sets of these actions satisfy the rationality postulates and/or entrenchment conditions. The easiest way to do this would be to first identify the subset of EDGs in which all acceptance sets in all content states are belief sets (for now, ignoring rejection and methodology sets) or closed sets of formulae of L_{PLEN}. Then, from there, one can identify the transition relations that respect the rationality postulates. From there, Gardenfors' representation theorems [91] (pp. 217-226) equate these actions to commitment functions that satisfy the entrenchment postulates. If this all works out, PLEN satisfies the desiderata that AGM theories don't without the theoretical tradeoffs, and PLEN can subsume AGM.

5.2.2 Bayesian Theories

My comments on Bayesian theories will be brief - perhaps moreso than Bayesianism deserves. Extremely basic explication of the core parts of Bayesian theories of epistemic dynamics [182] is contained in Chapter 1. The reason for the brevity of this treatment is primarily that Bayesian theories seem susceptible to being taken out in general by some of the same arguments that applied to AGM theories. First, an emphasis on the positive. It's clear

---

[8] I won't delve into this in detail, but dynamic logics for belief revision are well-known and well-developed. See, for instance, Dynamic Doxastic Logic [224, 228], which has actions that satisfy the AGM postulates. There are also results showing that the AGM postulates are satisfied by actions in certain dynamic epistemic logics [65]. These cash out earlier conjectures made by van Benthem [28, 30].

that Bayesian theories of rationality, insofar as they are concerned with credence update rules like conditionalization are fundamentally dynamic. Conditionalization is, at the very least, a diachronic norm of rational credence change such that only some transitions count as rational credence updates: those transitions from credal state w to credal state w' such that, if $Pr_{w'}(E) = 1$, then $Pr_{w'}(H) = Pr_w(H \mid E)$. More, it's quite plausible that conditionalization is a procedural norm - so long as the grain at which we individuate procedures is coarser than, for instance, causal processes.[9] This can all be granted, and it can nonetheless be shown that Bayesian models of epistemic rationality fail to solve the Constraints problem and fail to formalize D3.

Now, the criticisms are straightforward. First, Bayesian models of dynamics just don't say anything about constraints of feasibility and capacity. Like AGM, they concern only an idealized model of rational processes and are unconcerned with idealized models of agent capabilities with which to contrast the normative processes. If one can conceive of probabilistic versions of the Muddy Children or hidden envelope cases [65, 131], for instance, there is nothing - in general - in Bayesian models to tell us anything about them. Thus, for instance, whether or not one can carry out an update that complies with the rule of conditionalization - at all - is not something that a Bayesian model is interested in. This straightforwardly results in a failure of completeness with respect to modeling epistemic dynamics. Second, a key part of the representation of restrictions on epistemic dynamics is providing machinery that can distinguish the normative processes from those that are merely feasible processes. Without machinery to represent constraints of feasibility and capacity, Bayesian models provide no account of epistemic dynamics into which to embed the normative paths and thereby carry out important cognitive work, such as determining whether, as Chapter 2 argues is essential, conditionalization can even be complied with by an agent, let alone what specific processes or actions to carry out in order to conditionalize.[10] In other words, one has to augment the

---

[9] If this is not the case, then D2 cannot be formalized with Bayesian models. This would be frosting on the cake that I am baking.

[10] Aside, of course, from the action of conditionalizing on one's prior credences. But this is a particularly capacious action type that one can realize by numerous different action subtypes. It's not at all clear how,

Bayesian model of normative credence update with a model of epistemic dynamics in order to formalize D3. Without such machinery, there is just nothing in Bayesian models to tell us about D3(d).

This all results in parallel problems to those suffered by other belief revision theories. Let $Pr_1(H \mid E) = k$ be one's conditional credence at t_1. Let π_L be the process of subjecting oneself to a series of electric shocks that cause one randomly to take on the credences: $Pr_2(E) = 1$ and $Pr_2(H) = k$. Let π_R be the process of coming, by hard won argument and observation to have credences such that $Pr_2(E) = 1$ and $Pr_2(H) = k$. There is no difference between π_L and π_R in Bayesian theories from either the perspective of modeling dynamics or from the perspective of rationality.[11] Nonetheless, there is something distinctly epistemically troubling about any causal process that achieves conditionalization by chance, and Bayesian theories, at their core, are powerless to capture it.

5.2.3 Ordinal Conditional Rankings

As with Bayesian theories, my commentary will be minimal - in fact, much more minimal! The arguments applied to AGM theories apply nearly as well to Spohn's Ordinal Conditional Ranking theories. I omit detailed exposition of the system for the sake of brevity. See [233] for the full explication. Spohn's theory reduces epistemic dynamics to transitions among ranking functions. That is, only content changes of epistemic states are represented in the theory - just like AGM theories. As such, they are powerless to represent, with any real granularity, epistemic actions or processes. Thus, at least the Processes problem follows, as does failure of D2 and D3 for parallel reasons as those above.

---

at any given state, one is actually supposed to bring these credal changes about.

[11] Even if we assume that the update of belief in E is rational in both processes and restrict our attention to the update of the unconditional probability of H, the randomness with which one updates one's unconditional probability in H is objectionable. Here, as elsewhere, I'm channeling Goldman [105].

5.2.4 Finite Dependency Networks

The arguments applied to AGM and Bayesian theories apply nearly as well to Tennant's Finite Dependency Networks [245]. I omit detailed exposition of the system for the sake of brevity. See chapter 2 of [245] for the basics of it, and the rest of the book for exploration, defense, and application of it. Tennant's theory reduces epistemic dynamics to transitions among models of epistemic states that, insightfully, contain structures that track reasons for belief via color-coded graph-theoretic structures representing acceptance and non-acceptance and support and defeat relations. However, only content changes of epistemic states are represented in the theory. As such, they are powerless to represent, wth any real granularity, epistemic actions or processes. Thus, at least the Processes problem follows, as does failure of D2 and D3 for parallel reasons as those above.

5.2.5 Defeasible Reasoning Semantics

The arguments applied to the foregoing families of theories apply nearly as well to Pollock's Defeasible Reasoning Semantics [197, 195]. I omit detailed exposition of the system for the sake of brevity. Pollock's theory reduces epistemic dynamics to transitions among models of epistemic states that, insightfully, contain structures that track reasons for belief via graph-like structures representing support and defeat relations. However, content changes of epistemic states are represented in the theory. As such, they are powerless to represent, wth any real granularity, epistemic actions or processes. Thus, at least the Processes problem follows, as does failure of D2 and D3 for parallel reasons as those above.[12]

---

[12] However, Pollock's complete theory of rational agency, articulated in [195], builds upon this semantics and places it in a dual epistemic-practical model of a rational agent. The practical reasoning module of the model is built around a plan synthesis system that is, broadly speaking, basically a program logic. The practical reasoning module is fed beliefs from the epistemic module (the defeasible reasoning semantics system), and it feeds information back into it. Interestingly, Pollock defends the view that epistemic norms are internalized plans of cognitive behavior [194]. If these internalized plans can be called up deliberately under some conditions, and the planning module can be used to construct plans with epistemic actions, then Pollock's theory might have something much more robust to say about D2 and D3. I leave these considerations to future work.

5.3 Dynamic Logics

Dynamic logics come in a wide variety [113, 65, 247, 28, 31, 34, 35, 175], as well as "dynamified" versions of static logics [30] (Ch. 2). The core constructions of dynamic logic are program expressions that encode sequences of actions, indeterministic choices among actions, and iterations of actions [113]. Attaching the language of program expressions to an underlying model of dynamics has become conventional [30], and the most common form for the underlying model of dynamics is that of a labeled transition system or multi-modal Kripke frame.

Propositional Dynamic Logic or PDL is the central formalism in dynamic logic. It is the modification or extension of PDL to myriad applications that comprises much of the field of logical dynamics [34], though the field also relies quite heavily on epistemic logics and game theory [34]. PDL supports useful combinations of both of these latter frameworks [34].

There are a number of different ways of adding additional procedural constructions to dynamic logic such as concurrency operators and things like negated action or complement and reversals of actions [247, 113], but I'll omit discussion of these. More interesting variants arise by modifying the underlying model of dynamics by interpreting program expressions over different structures. Interpreting program expressions over processes generates Process Logic (PL) [114]. Interpreting program expressions over protocols (indeterministic trees that model sequences of decision problems) and then embedding these in a model of agent uncertainty about actions generates Epistemic Protocol Logics (EPL) [184]. Treating the nodes of an LTS as belief sets and zeroing in on the subset of basic actions that represents rule-based revisions of belief set generates Epistemic Logics for Rule-based Agents (ELRA) [134]. Extending the language of PDL with operators about knowledge, ability, opportunity, belief, and more and suitably upgrading the model of dynamics to provide semantics for these additional notions generates KARO logics [128, 124].

The foregoing are the rival systems I will focus on in this section.

5.3.1 PDL

The promise of PDL as a model of epistemic dynamics and epistemic norms is one of the primary motivators of this entire project. In examining PDL, treat it the way PLEN was interpreted as a theory of epistemic norms: interpret the states in PDL's model theory as epistemic states, the propositional language as one that contains epistemic or doxastic expressions, and the program expressions as encoding the content of procedural norms. Keep this adaptation of PDL in mind in what follows.

PDL does well with respect to recognizing constraints on dynamics if programs are taken to encode capacity and feasibility. However, as has already been argued, PDL suffers the Processes problem and, as a result, cannot handle methodological inclusion, procedural information, or epistemic distinctions in cases like the Honest Twin. Thus, PDL fails to formalize D2. Because PDL can't represent processes as distinct from transitions, it cannot accurately represent normative restrictions on processes. PDL mischaracterizes D3 as a result - the restrictions on processes cannot be analyzed in PDL.

PDL does provide machinery for exogeny. There are explicit protocol expressions in PDL and operators encoding the dynamic properties of protocols. Taking protocols to be the analysans of epistemic norms, PDL is narrowly exogenous and its deductive systems encode valid reasoning about the dynamic properties of epistemic norms. Given the general strategy of interpreting deontics and verdicts in terms of dynamic properties (as in Chapter 4), these features of norms can be explicitly reasoned about in PDL. However, the exogeny of PDL pales in comparison to that of PLEN.

There are specific kinds of arguments concerning the equivalence and dependence relations among norms (as spelled out in terms of relations between restrictions on epistemic dynamics and normative verdicts). These are articulated in [135], as are their formalizations in PLEN. PDL, however, cannot adequately formalize these forms of argument for the simple reason that PDL suffers the Processes problem and, as a result, fails D2 and partially fails D3. PDL cannot analyze restrictions on processes and actions in a way that distinguishes them from transitions, and so the forms of reasoning that PDL can formalize are limited to arguments about properties that can be analyzed in terms of transitions only. PDL cannot, for instance, formalize an argument establishing that a norm that dictates constructing proofs myself is inequivalent in an important way to a norm that dictates either asking my honest twin or doing the proofs myself.[13] More, PDL cannot analyze any useful equivalence relations among epistemic norms stricter than transitional equivalence without simply adding machinery from PLEN.[14] Thus, it follows that if PDL's languages are extended to contain expressions that encode statements about (e.g.) procedural convergence and divergence, parallel execution, and path-theoretic equivalence, the logic of PDL will mischaracterize these notions or only partially characterize them (by codifying how they interact with transitional properties only). This is easy to see: given the notion encoded in $\pi \curlywedge \pi'$, PDL will have no way of showing that $\pi \curlywedge \pi'$ doesn't follow from $R_\pi = R_{\pi'}$.

Now, the foregoing implicitly assumes a re-interpretation of PDL for the epistemic case, but this reinterpretation does no work in the argument against PDL's adequacy. PDL is a logic of actions and programs, first and foremost, and has no native features for representing epistemic states. However, let us generally assume that, along with propositional valuations, the models of PDL associate to each state a structure that interprets epistemic operators like belief b and knowledge k - content states as in PLEN would be fine, replacing each states with an epistemic model as in DEL would do the trick, as well. There are numerous possible ways to do this; see [162, 219, 229, 247, 30, 34, 36, 68, 65] for developments of PDL into dynamic doxastic logics.

Additionally, PDL does well with respect to other theoretical tradeoffs. PDL makes no commitment to actions that result exclusively in changes of epistemic state. That is, there are actions, a, in PDL such that a may change the state of the world in addition

---

[13] Formally, one norm dictates carrying out a transition in transition relation R_d - deducing for myself. The other dictates carrying out a transition in $R_t \cup R_d$ - the union of the deduction and twin-asking transition relations. Of course, if $R_t = R_d$, then $R_t \cup R_d = R_t = R_d$. The two norms have the same requirements under PDL, but they are distinct. One can comply with the latter norm by carrying out an action that doesn't comply with the other.

[14] See the next section for some details

to epistemic states. So, PDL nicely evades Special Case 3 of the failure of norm exogeny and of the failure to model epistemic dynamics. PDL is also fully general, committing to the analysis of no specific kinds or subsets of epistemic actions like public announcements, subterfuge, expansions and contractions, or any other specific subset of epistemic actions. Finally, there are, as far as I know, few, if any, standard paradoxes plaguing PDL that aren't merely artifacts of the propositional, epistemic, or deontic logic that one attaches to it.[15] And these can be switched out while retaining the action and program logic.

As a final bit of comparison between PDL and PLEN, note that PLEN is more expressive than PDL. Denote bisimulation with $\sim$. Formula A is bisimulation invariant iff for any graphs, g and g' such that $s \in S$ and $s' \in S'$, if $s \sim s'$, then $s \Vdash A$ iff $s' \Vdash A$; A is bisimulation variant iff it's not bisimulation invariant. It is well-known that the control operators of PDL are bisimulation invariant. As a result, the language of PDL contains no expressions that are bisimulation variant.[16] However, PLEN contains a control operator and several connectives out of which bisimulation variant expressions can be formed in the object language. It also has plenty of room for ad hoc meta-language operators that are bisimulation variant.

(EXP) L_{PLEN} can distinguish between bisimilar EDGs; $\cap$, $\equiv_t$, $\curlywedge_p$ $\leftrightarrow_p$, and $\equiv^s_G$ generate bisimulation variant formulae.

Corollary L_{PLEN} contains formulae that are not bisimulation invariant, but are isomorphism invariant.

Corollary Graph-theoretic Protocol Equivalence Relations are bisimulation variant but isomorphism invariant. This accounts for the difference in strength among graph-theoretic protocol equivalence and the rest.

---

[15] To be clear, when applied to both the doxastic [158, 162, 224, 225, 229] or deontic [174, 228] cases, things are different; there are well-known paradoxes in PDL-based doxastic [219] and deontic [10] theories, despite how well the application of a PDL-based framework does at removing other paradoxes.

[16] Both of these technical results are proved and explored in [29, 30].

The foregoing arguments make no reference to how epistemic states are represented. All of the foregoing complaints have to do with the program logic that PDL employs. Fundamentally, PDL is constructed by attaching a bit of relational algebra to a modal logic. Actions are all interpreted with binary transition relations, and so are programs. It's this feature that generates the problems for PDL. This is an important point because many formalisms that are plausible rivals of PLEN simply build the action and program logic of PDL on top of some innovation at the level of states or logic or they offer mild modifications of the action and program logic of PDL. This extends to dynamic epistemic logics as well.

5.3.1.1 Modifying PDL

It might be thought that the main problems of PDL can be removed by giving a semantics of program expressions in terms of computation sequences [113]. Let a computation sequence be a sequence of atomic action symbols, e.g., $\langle a_1, ..., a_n \rangle$. Then, letting $CS(\pi)$ be the set of computation sequences of π, an elementary result of PDL is that $CS(\pi)$ is syntactically determined. For atomic, a, $CS(a) = \{a\}$. Letting p_i be an arbitrary computation sequence, and $p_i p_j$ be a concatenated computation sequence: $CS(\pi_1; \pi_2) = \{p_i p_j \mid p_i \in CS(\pi_1), p_j \in CS(\pi_2)\}$, $CS(\pi_1 \cup \pi_2) = \{p_i \mid p_i \in CS(\pi_1) \cup CS(\pi_2)\}$, and $CS(\pi*)$ is the reflexive, transitive closure of $CS(\pi)$. These can be associated, via the basic execution rules, with transition relations.

The solution of some of PDL's problems (with respect to modeling epistemic dynamics) might lie in analyzing epistemic norms in terms of computation sequences. One way to do this would be to either interpret the propositional language of PDL with the computation sequences. First, note the sets of computation sequences mapped to each program, and let the execution sets of programs be the computation sequences. Then let a function (relativized to LTS, g), $f_\pi : CS(\pi) \longrightarrow R_\pi \subseteq W \times W$, associate each computation sequence of π with a transition relation so that, basically, the execution rules of PDL are approximated. For atomics, $CS(a) = \{a\}$. Let $f_a(a) = R_a$ with R_a defined as in PDL or the transitional fragment of PLEN. Then, for example, define f_π for complex protocols so that (inter alia): if $p_1p_2 \in CS(\pi_1;\pi_2)$, then $t \in f_{\pi 1;\pi 2}(p_1p_2)$ iff $t = \langle first(f_{\pi 1}(p_1)), last(f_{\pi 1}(p_2))\rangle$. Then, the semantics of the operator $[\pi]A$ can be defined like so:

(1) $w \models [\pi]A$ iff for all $p \in CS(\pi)$, if $t \in (f_\pi(p))$ and $w = first(t)$, then $last(t) \models A$.

(2) $w \models \langle\pi\rangle A$ iff there is some $p \in CS(\pi)$ such that $t \in (f_\pi(p))$ and $w = first(t)$, then $last(t) \models A$.

For example, let $\pi = a;b$ where a and b are atomics. Then $CS(\pi) = \{ab\}$. Consider an LTS g such that P_g is defined in g as it is in PLEN. Consider a state s such that $\langle\pi\rangle\top$ at w. Then there is some t such that $t = \langle w, x\rangle \in f_{a;b}(ab)$. But then there must be some $t' = \langle w, x'\rangle \in f_a(a)$ and $t'' = \langle x', x\rangle \in f(b)$. The steps just go further with more complex protocols.

Now, with this kind of semantics, PDL might be capable of distinguishing processes from transitions. Simply put, computation sequences represent processes and transitions are simply correlated with them. Two distinct computation sequences (i.e., processes), p and q, that execute π can be bookended by the same transitions via f_π. So, it's possible that $R_\pi = R_\pi$, $\langle\pi\rangle\top$ and $\langle\pi'\rangle\top$ at w, but that $p \in CS(\pi)$ and $p \notin CS(\pi')$. For example, $CS(a;b) \neq CS(c;d)$, but, in g, let $f_a(a) = \{t\}$, $f_b(b) = \{t'\}$, $f_c(c) = \{t\}$, $f_d(d) = \{t'\}$, $f_{a;b}(ab) = tt' = f_{c;d}(cd)$. In this case, $R_a = R_c$, $R_b = R_d$, $R_{a;b} = R_{c;d}$, and $\langle a;b\rangle\top$ and $\langle c;d\rangle\top$ at $first(t)$. Thus, two distinct sets of processes might nevertheless define, via f_π, identical transition relations. Now, with this amendment, PDL would be capable of supporting formulae, $A(\pi)$, about π's methodological inclusion, procedural features, and normative status that distinguish programs π_1 and π_2 such that $R_{\pi 1} = R_{\pi 2}$. For instance, if we take programs to encode the requirements of epistemic norms, then process (computation sequence) p may be in the execution set of π without being in the execution set of π', even if for all p in $CS(\pi)$, $f_\pi(p) = f_{\pi'}(p')$ for some p' in $CS(\pi')$ and the converse also holds. Hence, if there are formulae $A(\pi)$ with semantics defined in terms of computation sequences, these will distinguish programs with identical transition relations and non-identical computation sequences. Or, in other words, $A(\pi) \equiv A(\pi')$ will not follow from $R_\pi = R_{\pi'}$ if the truth conditions of $A(\pi)$ and $A(\pi')$ are defined in terms of computation sequences rather than transition relations. Thus, modifying PDL so that programs assertions are interpreted over computation sequences is a marked improvement.

Unfortunately, this modification of PDL accomplishes - at best - no more than what is accomplished by the procedural fragment of PLEN without machinery for encoding epistemic and doxastic assertions (e.g., content states that interpret $b(A)$). Really, this modification accomplishes less. Correlating protocols/programs with computation sequences and pairing these with relation functions, f_π, delivers no more information than associating each program with procedural paths.

First, in PLEN, it can be shown that $P_\pi = P_{\pi'}$ for all structures, g, iff $CS(\pi) = CS(\pi')$.[17] All there is to know about equivalence defined via computation sequences can be gotten from procedural paths because procedural paths minus states are computation sequences. The set of computation sequences can be read off of the set of paths. Anything we need to know about computation sequences can thus be derived from some facts about procedural paths. More, computation sequences do not give us the same information about protocols that procedural paths do. For instance, the computation sequences of each protocol π are determined exclusively by its syntax. Thus, the extension of $CS(\pi)$ is not relative to any semantic structures. However, the procedural paths of π are determined partially by the syntax and partially by the graph-theoretic structure of each EDG. What this means is that, taking the structure of a model of epistemic dynamics to encodes (e.g.) constraints of capacity and feasibility, different conditions of feasibility and capacity with respect to epistemic actions can change what it means to comply with a norm. P_π in g may be distinct from P_π in g'. The paths in one set of executions may be structurally distinct from those in the other, and this means that, in one graph, different computation sequences will be enacted

---

[17] See section 5.4.2 of [135].

than in the other. But computation sequences alone don't tell us this - they can't tell us that two agents with different capacities will have to comply with the same norm differently. We need to exploit the function f_π in order to get information about what computation sequences are enacted in a particular graph. PLEN gives us this information for free via the same constructions that give us computation sequences.

Second, there is no obvious way of using this kind of semantics to define graph-theoretic relations among programs/protocols, and this will result in a blindspot in the equivalence relations in modified PDL. Let $\pi \equiv_G \pi'$ iff $G^\pi = G^{\pi'}$ (as defined in GTP1). Then $\pi \equiv_G \pi'$ implies $CS(\pi) = CS(\pi')$ while the converse fails. Thus, computation sequence identity between protocols can't establish "graph-theoretic equivalence". Unless the modified form of PDL, ultimately, just defines the procedural execution rules and implementations of PLEN, even a version of PDL defined over computation sequences is without any clear means of defining the sets of graphs exploited in defining $\equiv_G$. As a result, all of the useful results about graph-theoretic equivalence disappear in the modified system. These results are explicitly developed in the technical document [135].

Finally, if R_π is the set of transitions composed of $first(p)$ and $last(p)$ for each $p \in P_\pi$, then $f_\pi(p) = R_\pi$, and the converse holds, as can be easily shown. So computation sequences tell us no more about the transitions specified by protocols/programs than can be gotten from procedural paths.[18] And all of this ignores parallel execution or, in the PDL literature, concurrency, which at least is given an analysis in PLEN that is theoretically adequate in certain ways: e.g., parallel execution in PLEN delivers unique sequences of states for complex protocols - especially sequentially composed protocols - that are executed concurrently.[19] The failure of this property implies that the total epistemic states of agents can simultaneously evolve along two distinct paths, which is absurd.

Consequently, this modification of PDL is simply a step toward building PLEN that still

---

[18] All pf these results only hold if $f_\pi(\pi) = R_\pi$ for all atomics, where R_π is defined as in PLEN or PDL. Else, f_π might be a completely arbitrary transition relation.

[19] A well known and difficult problem. See chapter 10 of [113] as well as [247].

omits some information about protocols, rather than a genuine rival.[20] For the remaining PDL based theories, this sort of modification is available, and, in each case, it improves the performance of the formalism with respect to the desiderata at stake. However, in each case, the underlying model of epistemic dynamics is just a step towards PLEN, and the additional machinery in each theory can easily be approximated or added to PLEN. Thus, in none of the following cases does the formalism, even modified as above, amount to a step forward from PLEN with respect to solving the Processes, Constraints, and First-Class Citizenship problems or formalization of D1-D3.

5.3.2 Process Logic (PL)

Process Logic (PL) [114] reinterprets the logic of PDL over path-models. The path-models of PL are essentially ETL models. The innovations involve (a) the interpretations of actions and programs by assigning them sets of paths, and (b) the machinery defined on top of the path-models for analyzing features of processes like what states are reached over the course of the process. The path-theoretic execution conditions of PLEN are directly inspired by PL, as were many of the basic formal contrivances.[21]

There is much innovative treatment of these basic models, but the details can, without loss, be overlooked for my purposes, here. Accordingly, I omit detailed exposition of PL for the sake of brevity. However, the problems that plagued PDL as a model of epistemic dynamics and epistemic rationality norms apply just as well to PL. The key problem is that paths in PL are countable sequences of states. While this enables PL to analyze a number of interesting properties of processes [114], it ultimately implies that PL suffers the Processes problem. Actions, and sequence thereof, that correspond to identical sequences of states are indistinguishable in PL, and so the Honest Twin case and Bad Reasoning throw up barriers

---

[20] There are quite a few steps between this amendment to PDL and PLEN, however, among which lies semantic decisions about how to execute parallel execution. This is taken up in the document [135].

[21] The $first(p)$, $last(p)$ notation, the definition of execution conditions for sequential compositions over concatenated paths, and the general thought that programs ought to be interpreted by assigning them sets of paths rather than binary relations are all innovations of PL.

to the adequate formalization of D2 and D3 in PL. As a result, cannot handle methodological inclusion or epistemic distinctions in cases like the Honest Twin. Thus, PL fails to formalize D2 and mischaracterizes D3, while failing to match PLEN's norm exogeny just as PDL did.

5.3.3 Epistemic Protocol Logic (EPL)

EPL played a significant role in the design of PLEN. EPL protocols themselves have more than a passing resemblance to the restrictions on an EDG's paths that the execution rules assign to each protocol in PLEN. EPL's protocol expressions are are constructed from a very similar syntax. More, the protocols of EPL [184] were the initial inspiration for the "protocol graphs" used extensively in [135], though these constructions play little role in this thesis.

EPL protocols have the benefit of providing useful models of rules for selecting action that captures certain intuitive features of what commitment to a protocol requires.[22] More important, EPL builds these protocols into a model of dynamics that explicitly displays the actual courses of action according to a protocol (the protocol itself) as distinct from the entire set of possible courses of action (the arena) and distinct from those courses of action the agent thinks accord with a protocol (the subjectively enabled paths in the arena). Essentially, every protocol in EPL is a restriction on paths through arenas that is assigned to each protocol expression by the EPL semantics.

Unfortunately, EPL, being based fundamentally on PDL, interprets atomic actions with binary transition relations, and basic protocols are simply trees. This reinstates a form of the Processes and Constraints problems and, along with them, versions of the same failures that befell PDL. First, any two atomics correlated with the same binary relations are interchangeable in all EPL contexts. This follows from the fact that the transition relations in protocols are binary. Second, basic protocols with identical accessibility relations are interchangeable in all EPL contexts. Consider EPL protocols t and t', such that the basic action sets are,

---

[22] For instance, EPL's protocols provide means of distinguishing between different sequences of decision problems that PDL programs can't make sense of. The aforementioned [184] demonstrates this well. (pp. 4-6)

respectively, Σ and Σ', for every $a \in \Sigma$, there is an $a' \in \Sigma'$ such that $\Rightarrow_a = \Rightarrow_{a'}$ and vice versa. Then, at any state in any arena, t is enabled iff t' is enabled.[23] Any equivalence relation among protocols in EPL defined in terms of enabling or embedding, then, will hold between t and t' because they have equivalent transition or reachability relations. Finally, because complex protocols are generated by composition functions on trees, these interchangeabilities extend to complex protocols (the protocols in EPL that correspond to complex protocol expressions in EPL). Thus, there simply can't be EPL protocols that represent the kind of restriction on procedures that is required by epistemic rationality with respect to norms that distinguish transition-equivalent actions. This implies that some of the problems of PDL transfer directly over to each basic protocol of EPL.

Let $t = \langle S, \{R_a \mid a \in Act\}, s_0\rangle$ be a basic protocol. Let $t' = \langle S', \{R_a \mid a \in Act\}, s'_0\rangle$ be another basic protocol. Let b and c be atomic actions that are associated, for all basic protocols, with identical binary transition relations, $R_b = R_c$, but, as in the Honest Twin case, have distinct methodological, procedural, and normative properties. Then, every edge labeled in t with b is also labeled with c and the same holds for t'. Thus, at every state at which t dictates doing b, t also dictates doing c. Same for t'. This extends to every basic protocol. From this, a number of troubles follow. Insofar as basic protocol expressions are to be deployed to capture the content of norms, they fail. There can be no basic protocols that represent restrictions on action that distinguish b from c; this just gets restrictions on action wrong in cases like the Honest Twin. More, it follows that every basic protocol in one's methodology that dictates doing b also dictates doing c. This gets methodological inclusion wrong. Finally, this feature of protocols in EPL also obscures procedural facts, as a procedure that requires doing b needn't, intuitively, require doing c, and yet any basic protocol that requires doing one also requires doing the other.[24] Because complex protocols

---

[23] I am using Pacuit and Simon's notation for protocols rather than the PLEN notation; in their system, protocols are trees rather than syntactic items, so t is a convenient term for naming them. Keep in mind that t may but does not always denote a transition in EPL as it does in the PLEN notation. What I call protocols would be protocol expressions in EPL.

[24] Certain kinds of procedural facts will be fine. For instance, a protocol t that is bisimilar to a t' that requires a sequencing of atomics $a; b$ will also require the sequencing. This because every ordered pair is

are built by various composition functions out of basic protocols, these results extend to complex protocols; any pair of complex protocols will conflate b and c just as badly as any pair of basic protocols. The foregoing arguments merely assumed that atomic actions were interchangeable. These problems are exacerbated by the extension of interchangeability to basic and complex protocols.

All of the foregoing implies several failings of EPL with respect to D1-D3. First, EPL gets D2 wrong and D3 at least partially wrong by conflating actions with transitions (and protocols with trees rather than multi-graphs). As for norm exogeny, there are explicit protocol expressions in EPL and operators encoding the dynamic properties of protocols. Taking protocols to be the analysans of epistemic norms, EPL is narrowly norm exogenous and its deductive systems encode valid reasoning about the dynamic properties of epistemic norms. Given the general strategy of interpreting deontics and verdicts in terms of dynamic properties, these features of norms can be explicitly reasoned about in EPL. However, as with PDL, there are philosophically important kinds of arguments concerning the properties of epistemic rationality norms that cannot be analyzed correctly in EPL due to the interchangeability of transition-equivalent actions.

5.3.4 Epistemic Logic for Rule-based Agents (ELRA)

ELRA [134] was explicated in Chapter 1. The key insight of ELRA is that one can apply a dynamic logic to the analysis of the actual process of deductive reasoning rather than simply closing belief sets or knowledge under deduction. ELRA also shows us how to use dynamic logic to characterize the inferential features of actions as represented by transition relations.

The fundamental problem with ELRA is generality; ELRA is restricted to the analysis of

---

also a tree, as is every unconnected node, so there will be a basic protocol that interprets every action label in every basic protocol. So, many basic protocols are analyzable into more basic protocols, down to the atomic actions. Thus, if a basic protocol t contains a subtree t', and if t' is related to another subtree t'' in a way that forms t by means of one of the composition relations in EPL [184], then t' is a "more basic" protocol contained in t, and there will be a complex protocol expression corresponding to t that contains one corresponding to t'. Thus, if procedural facts like those at hand are encoded by some formulae of EPL, these procedural facts can, at least, be represented.

particular kinds of rule-based epistemic change, and these rule-based changes are all defined by specific content state changes. While ELRA can distinguish the processes in Bad Reasoning, for example, it has no machinery for representing epistemic actions in general or the Honest Twin case in particular. ELRA doesn't suffer from the Processes problem as severely as other formalisms, but it cannot handle methodological inclusion, procedural information, or epistemic distinctions in cases like the Honest Twin. Thus, ELRA mischaracterizes D3. Finally, because ELRA is restricted to reasoning about the specific rule-based processes it is designed for, it cannot provide machinery for analyzing reasoning about the properties of epistemic norms generally.

5.3.5 KARO Logics (KARO)

KARO logics, or logics of Knowledge, Action, and Opportunity (among many other things), are a complication of the logical analysis of agency based roughly on the combination of propositional dynamic logic and epistemic logic [128, 124]. The innovation of KARO is in the representation of much more than merely the structure of epistemic states and the structure of procedures, plans, or actions; KARO is a more robust logic for practical reasoning, including reasoning about motivational states, opportunities, and commitments (whether due to plans or norms) to action. A KARO model is a set of states over which several transition relations and equivalence relations are defined along with a function from states to actions. The transition relations interpret actions, the equivalence relations interpret knowledge, belief, and desire, and the function from states to actions functions as an agent's agenda - the set of actions per state that an agent is committed to.

The main failure of theoretical adequacy for KARO is a result of the incorporation of the action and program logics of PDL. Accordingly, I omit detailed exposition of PL for the sake of brevity. The same problems inherent in the action and program logics there transmit directly to KARO. In fact, the problems are multiplied because the language of KARO enable the interpretation of new assertions about programs and actions, and this increase in expressive power increases the possible errors that conflating processes with transitions result in.

5.4 Dynamic Epistemic Logics

The general form of Dynamic Epistemic Logic (DEL) [65] emerged out of the study of Plaza's Public Announcement Logic (PAL) [193], which itself was a natural, dynamic extension of Hintikka's classic framework for Epistemic Logic (EL) [129]. PAL took the multi-agent epistemic models of EL and added a dynamic action of public announcement, which, intuitively, has less to do with verbal announcements as it has to do with hard information updates for entire groups of epistemic agents. PAL defines a public announcement as an action that takes epistemic models to new epistemic models. The basic idea being that the transition relations in an epistemic model encodes the properties of knowledge for each agent, and the transition from one model to the next via public announcement updates these transition relations, and so encodes changes in what's known by each agent. DEL generalizes this picture by adding a general class of product updates, one for each event and/or action model, which encodes what agents know about events/actions.

The underlying model of dynamics is defined in terms of product updates. As such, DEL is an input assimilating model of dynamics, and so suffers the standard problems with these (e.g., inability to model constraints [131]). DEL can be upgraded by incorporating Epistemic Action Logics (EAL) and Logics of Epistemic Programs (LEP) [18] to overcome some of these limitations. It can alternately be combined with Epistemic Temporal Logic (ETL), resulting in a branching time model of epistemic dynamics, and a Temporal Dynamic Epistemic Logic (TDEL) [131]. Epistemic Planning Logics (EPPL) apply DEL to the analysis of deliberate action or the "reasoning" phase of action [41].

The foregoing are the rival systems I will focus on in this section.

5.4.1 DEL, Epistemic Action Logic (EAL), and Logics of Epistemic Programs (LEP)

Dynamic epistemic logic (DEL) has the best chances of being a logic of epistemic norms out of the rival formalisms of PLEN. Basic forms of DEL can be modified to incorporate action logics and even programs in addition to their well known and fruitful structural representations of epistemic states and the dynamics thereof. As such, such modifications of DEL seem plausible as formalisms with structures for representing both epistemic dynamics and epistemic norms.

Epistemic models represent epistemic uncertainty relations (per agent), and these are used to read off properties and contents of epistemic states. That is, for agent i, R_i in model M tells us everything we need to know about what i knows (or, in a doxastic logic, what i believes), e.g., what's true at the worlds connected by R_i in M gives a snapshot of the contents of i's knowledge and beliefs. Event models add to the content of an agent's epistemic state by representing uncertainty relations among events as well as some "external" information about the preconditions of events.

Product updates represent transitions from one set of uncertainty relations to another that are effected by particular events. Formally, product updates are just functions from pairs, (M, E), of epistemic and event models to epistemic models that combine the uncertainty relations among states in M and events in E while preserving valuations for atomics at all states. That is, epistemic dynamics are conceptualized in terms of transitions among epistemic states associated with events. The properties of epistemic dynamics can be read off of these updates.

Call a dynamic epistemic logic with structures analyzing actions an Epistemic Action Logic (EAL).[25] To analyze the epistemic features of actions, DEL deploys pointed models. First, we consider pointed epistemic models (M, w) where w is one of the worlds in M. Second, we consider pointed action models, (E, e), where E is an event model and e is a

---

[25] See [65] for technical details.

particular event in the set of events in E. Product update generates a pointed epistemic model $(M', (w, e))$ in a straightforward way that restricts or, in more fully generalized versions (re: [18, 65]), expands the uncertainty relation. “In this way, agent [i]’s uncertainty in the resultant model [...] comes from two sources: her initial uncertainty in M [...] as to which is the actual world and her uncertainty in E [...] as to which is the actual event.” [19] The thought is that action is a species of event, and to see the epistemic effects of an action, one must see how the action updates the uncertainty relations at the current world.

An action algebra can be added to DEL with EAL to generate a Logic of Epistemic Programs (LEP) [18]. Here are two ways to do this. First, simply take the foregoing machinery and add program control operators to the language so that the basic actions can be combined via procedural operations. At a minimum, consider choice $\cup$ and sequential composition ;. Then define functions on product updates in order to assign product updates to epistemic programs. For an action a, we know that there is an action model (E, e) and a product update function that identifies all transitions among epistemic models for a. The same, mutatis mutandis, works for b. For the program $a; b$, the product update of the program is just the composition of the product updates of a and b. The choice is the union of them. A second way to do this is to assign to each program a program model, (E, π), which is an action model such that π is a designated set of actions in E. Then the update for (E, π) is the update for E restricted to the designated actions, i.e., we update any (M, s) to $(M', (s, e))$ where $e \in \pi$. In either case, there is a way in which complex procedures built out of epistemic actions can be analyzed, and the underlying model of epistemic dynamics moves beyond input assimilation to tree-like representations of possible unfoldings of a system.

LEP is a dynamic epistemic logic with a program logic on top. As such, it provides a prima facie plausible underlying model of epistemic dynamics and a plausible looking model of epistemic normativity in just the same way that PDL or PLEN do. Let the programs in LEP represent procedural epistemic norms that dictate operations on epistemic actions. The product updates represent the transitions that result from executing the instructions of these programs and so represent the rational courses of action according to the norms the programs represent.

The fundamental problems of DEL (generalized with EAL) are three-fold, and they apply equally well to the base form of DEL as to EAL and LEP. First, there are the problems of DEL raised by Hoshi [131]. I quote him at length:

> [...] DEL does not provide a machinery that is suitable for representing protocol information. There are two senses in which it does not. First, in DEL, there is no restriction on which event models can be applied to given epistemic models. Any event model can be applied to any epistemic model, and the epistemic model obtained by the process can be described by using corresponding event operators. As we saw above, the informational state after I peep into the card on the desk can be represented (in Figure 1.4) by an epistemic model. No matter what communication constraints we think of for the situation after my peeping, e.g. I put the card into the deck after peeping and leave the office, DEL does not forbid us from applying the public announcement of $!p$ (The card is the ace of Diamonds), which will yield the truth of $\langle \varphi ! \rangle [2]p$ (p can be publicly announced after which you know p).
>
> Second, one may try to adjust the precondition functions of event models to represent communication constraints. One component of event models is a precondition function *pre*. The function is interpreted in such a way that an event e can happen iff $pre(e) = \varphi$. Thus, in the above example, we may introduce a new propositional letter, say d, to represent whatever communicational or observational constraints there will be after my peeping, and say that the public announcement of p is in fact the public announcement of $p \wedge d$, since the public announcement of p can happen only if p is true and the condition d is satisfied. For instance, we may interpret d here as “the card is still on the desk in the office in front of us", and make d false to represent the situation after I put the card

> back into the deck of cards and leave the office. $\langle !(p \wedge d)\rangle\, [2]p$ then become false, since $p \wedge d$ is false in that case. Thus, after I put the card into the deck and leave the office, you cannot know what the card is by turning the card over or asking me what the card was.
>
> Even though such an adjustment of precondition functions may yield satisfactory representations for certain cases of intelligent interaction, the strategy cannot be applied generally to represent protocol information. The main obstacle is that the informational events that can happen may change over time. In many interaction scenarios, the information about what can happen at a given moment depends on the information about what has happened earlier. For instance, our conversation may obey implicit rules such as "Do not repeat yourself", "Say p after q", etc. Protocol information of this kind cannot be captured by the above maneuver, since the preconditions of events are encoded by propositional letters, whose truth values are constant in DEL at a given world. A given world can evolve in various possible ways, depending on what event happens, and propositional letters cannot do the job of tracking how the world has evolved. For this reason, DEL is not suitable for capturing the temporality in protocol information. (pp. 24-25)

What Hoshi calls protocol information, I call information about the constraints of capacity and feasibility. In sum, DEL is inadequate to accurately representing the properties of constraints discussed in Chapter 1. DEL thus suffers the Constraints problem.

Second, the use of Kripke models as the basis for structures that represent epistemic states generates the well-known problem of logical omniscience and closure of knowledge [134, 148]. This is a result of defining the semantics of knowledge and belief operators in terms of transition relations among worlds. I won't comment on this more, except to note that syntactic epistemic logics do not have this bug; one can stipulate whatever conditions on belief and knowledge one wants. Hence, PLEN is free of logical omniscience and closure -

so long as one wants to be free of them. They can easily be reinstated by imposing conditions of closure under validity relations on the acceptance set.

Third, and most serious from the perspective of comparing DEL to PLEN, DEL suffers the Processes problem. In contrast to PDL, however, DEL conflates processes and transitions not by omitting machinery for representing processes, but by omitting machinery for representing transitions. No two actions represented in DEL can be transition equivalent. The epistemic features of actions in DEL are represented by event models, but the dynamic features of actions are defined by product updates. Consider actions a and b. It's tempting to take the phrase "action model" on its face, and think that a can be represented with pointed event models. But this is a mistake; actions are encoded syntactically, as in PLEN, by means of the symbols in action models, and their dynamic features are modeled indirectly by the mapping provided by product updates. Recall that product update is a function from pairs of pointed epistemic models and pointed action models, $((M, w), (E, e))$, to pointed epistemic models, (M', w'), where the states in the resulting epistemic model are pairs, $w' \in W \times E$. The transition relations that define execution of actions in DEL, then, are defined like so: $(M, w) \mapsto_{(E,a)} (M', w')$ iff the preconditions of a are all satisfied at w, M' is the update product of $((M, w), (E, a))$, and $w' = (w, a)$.

This rules out transition equivalent yet distinct actions. Consider the following action models. $(E, a) = \langle E, \rightarrow_i, pre\rangle$ where $a \in E$, and $(E', b) = \langle E', \rightarrow'_i, pre'\rangle$ where $b \in E'$. Suppose that a and b define identical product updates. Thus, for any pointed epistemic model $(M, w) = \langle S, R, V, w\rangle$, the product update from $((M, w), (E, a)) = (M', (w, a)) = ((M, w), (E, b))$. But this just means that $a = b$, that $\longrightarrow_i = \longrightarrow'_i$, and that $pre = pre'$. Thus, $(E, a) = (E, b)$. Thus, two distinct actions, a and b, cannot be transition equivalent; the terminal states in transitions assigned to actions via the product update carry a "record" of which action brought about the change, and so no two actions, DEL, can possibly be associated with equivalent transition relations.

The foregoing property is a serious failure to accurately model epistemic dynamics as well as the content of epistemic rationality norms. First, as argued in Chapters 1 and 2, it's plausible that there really are transition equivalent epistemic actions. This means that DEL straightforwardly mischaracterizes epistemic actions by ruling out an entire class of potentially - and, if Chapters 1 and 2 are right - epistemologically important actions by technical fiat. Second, the argument above shows that there cannot even be transition convergent actions in DEL. For no pair of actions, a and b, is it possible for there to be any (M, w) such that (E, a) and (E, b) to lead to the same terminal state from (M, w) while $a \neq b$. Explicitly, the execution rules of DEL rule this out. $(M, w) \mapsto_{(E,a)} (M', w')$ iff the preconditions of a are all satisfied at w, M' is the update product of $((M, w), (E, a))$, and $w' = (w, a)$, and $(M, w) \mapsto_{(E,b)} (M', w')$ iff the preconditions of b are all satisfied at w, M' is the update product of $((M, w), (E, b))$, and $w' = (w, b)$. If a and b converged, for even one transition, $((M, w), (M', w')$, then $w' = (w, a) = (w, b)$, and so $a = b$. But, again, it's quite clear that there are distinct epistemic actions and complex epistemic processes that are transition convergent. Thus, DEL gets epistemic dynamics wrong.

All of this adds up to DEL failing to adequately formalize D2 because epistemic rationality can distinguish transition equivalent and transition convergent actions and processes, and nothing in DEL can approximate such rational distinctions because DEL cannot even recognize these categories of actions. As a result, the restrictions on epistemic dynamics defined by procedural norms cannot be accurately represented (though dynamic restrictions might be well represented).

The problem is that the distinction between processes is built directly into the machinery for representing transitions. What is necessary is machinery that can represent transitions and processes separately while also capturing their relations. By building machinery that distinguishes processes at a fine grain (the syntactic level) into the terminal states of transitions, this assimilates all of the properties of transitions to those of processes rather than, as PDL does, the other way around. Consequently, D3 can't be formalized in DEL (extended to LEP). Given Hoshi's point above, D3(d) isn't going to come out right, either. The absence of machinery for representing constraints of capacity and feasibility results in the failure to distinguish them from normative restrictions. Now, even if that can be resolved by some maneuver with preconditions, D3(b) and D3(c) cannot be accurately formalized due to the failure to even recognize restrictions on transitions.

Finally, DEL (extended with EAL and LEP) supports logics that govern the expressions that encode the dynamic properties of epistemic programs. Taking epistemic programs to be the analysans for epistemic norms as we have above, these logics explicitly encode valid reasoning about (the dynamic properties of) epistemic norms. The logics are thus narrowly exogenous. However, just like PDL, its exogeny is lacking. As argued in [135], there are specific kinds of arguments concerning the equivalence and dependence relations among norms (as spelled out in terms of relations between restrictions on epistemic dynamics and normative verdicts). LEP cannot adequately formalize these forms of argument for the simple reason that LEP fails D2 and partially fails D3. LEP cannot analyze restrictions on processes and actions in a way that distinguishes them from transitions, and so the forms of reasoning that LEP can formalize are limited to arguments about properties that can be analyzed in terms of transitions only. LEP cannot, for instance, formalize an argument establishing that that a norm that prohibits asking my honest twin and dictates constructing proofs myself is inequivalent with a norm that dictates either asking my honest twin or doing the proofs myself. More, LEP cannot analyze any equivalence relations among epistemic norms stricter than transitional equivalence. Thus, it follows that if LEP's languages are extended to contain expressions that encode statements about (e.g.) procedural convergence and divergence, parallel execution, and path-theoretic equivalence, the logic of LEP will mischaracterize these notions or only partially characterize them (by codifying how they interact with transitional properties only). This is easy to see: given the notion encoded in $\pi \curlywedgeuparrow \pi'$, LEP will have no way of showing that $\pi \curlywedgeuparrow \pi'$ doesn't follow from $R_\pi = R_{\pi'}$.

These final problems apply across the board to all variants of DEL for the same reason that the arguments against PDL applied to all variants. The problems just articulated depend only on the underlying model of dynamics in DEL that analyzes actions in terms of transitions only. The resulting model of dynamics fails to distinguish actions from transitions, and so a cascade of failures of desiderata ensues.

5.4.2 Epistemic Planning Logics (EPLL)

Epistemic Planning Logics (EPLL) [41] extend DEL to consider epistemic planning problems. Explicitly, the idea is that we have a language of planning constructions:

$$\pi ::= E|skip|if\ K\varphi,\ then\ \pi\ else\ \pi|\pi;\pi'$$

with E a set of events or elementary actions, and $if\ K\varphi,\ then\ \pi,\ else\ \pi'$ and $\pi;\pi'$ being complex constructions out of action. Essentially, the former is conditional choice and the latter is sequential composition. Take a set of epistemic models, $(S, \rightsquigarrow, V)$, and define an equivalence class over them with respect to bisimulation. Call such classes *information cells*, and label them with M_0. Then, define an action library, A, which is a finite set of event models $(E, \sim\subseteq E \times E, Pre, Post)$. Intuitively, plan constructions combine elements of the action library. To what end? The end defined by the goal state, ϕ, which is a propositional parameter in the language of EPLL. These propositions, of course, define sets of worlds in epistemic models. Finally, a *planning problem* is a triple: (M_0, A, ϕ). A planning problem is, intuitively, the problem of figuring out how to get from the initial state M_0 to a state that satisfies the goal state via plan constructions over the action library, A.

Solutions to planning problems are *planning trees* or trees with nodes labeled with epistemic models and edges labeled with event models, such that the root node is labeled with the initial state and the leaf nodes are such that the goal state is satisfied at all worlds in the epistemic model that labels the leaf node. Define the tree expansion function intuitively so that it takes a node n, adds node m, and connects them with some $\varepsilon \in A$ such that ε

is applicable at $M(n)$ and $M(m) = M(n) \otimes \varepsilon$ via product update. Intuitively, the idea is just this: you start at an initial state, have a state you want to reach, and search the action library for actions that you can put together to take you from the initial state to the goal state.

The basic tree-like model of dynamics that underlies planning trees have a natural interpretation as models of epistemic dynamics. Epistemic models model epistemic states, actions in the action library are epistemic actions, and the goal state is defined by epistemic norms. For instance, given one's initial state, one might have the normatively informed goal of coming to accept the element of a set of theories that maximizes some set of desiderata or coming to a consistent belief state, etc.. One then tries to figure out what plans of epistemic action would achieve this goal. Ideally, then, one can characterize restrictions on actions, processes, and even plans that achieve one's normative goals. First, restrictions on states are straightforwardly defined by all goal states - only the states that satisfy one's goal state are to be reached. Second, only transitions that are mapped to the actions in a plan that succeeds at achieving the goal state are acceptable. Third, only processes that instantiate a plan that succeeds in achieving the goal state are acceptable. Finally, only actions that are parts of a plan that succeeds in achieving the goal state are acceptable. That is, the features of the planning problem determine which states, transitions, processes, and actions are in and which are out. Taking the requirements of epistemic rationality to specify our goal states, these requirements define restrictions on states, transitions, processes, and actions.

The main problems of EPLL are that (i) it fails norm exogeny, and (ii) actions are analyzed in terms of product update. Concerning the first, the norms that define epistemic goal states are entirely left out of the picture. No structure in EPLL represents them, and nothing in EPLL will formalize reasoning about norms directly. The key problem is that plans, which might plausibly encode norms, are constructed in EPLL in order to arrive at goal states, which are themselves defined by some underlying norms. The norms that determine the set of goal states are left implicit. Even if one can interpret deontics over

EPLL and so interpret translations of norm-kernels (e.g., if one can only reach ϕ by plan π, then it's obligatory to carry out π), it's clear that the properties of these norm kernels do not analyze the properties of epistemic norms. Rather they would be derived from some assumptions about what epistemic rationality requires that are encoded in the goal states that define a given planning problem.

The second problem is simply that the problems with DEL carry over to EPLL. There are no transition equivalent actions at all, and so the Processes problem appears in an interesting form, in which, rather than ignoring processes and actions, it is transitions that are neglected. As a result D2 and D3 - at the least - cannot be adequately formalized.

5.4.3 Temporal Dynamic Epistemic Logic (TDEL)

Hoshi [131] and van Benthem et al [37] develop Temporal Dynamic Epistemic Logic by combining ETL and DEL to construct Temporal Dynamic Epistemic Logic (TDEL). The fundamental goal of this project is to improve upon DEL's ability to represent constraints on epistemic dynamics in a manner similar to that of ETL. To that end, things like propositional stability ($V(w', e') = V(w')$) may be dropped from the definition of product update, and devices are given that relate histories in ETL to sequences of action models to sequences of product updates.[26] The result is a set of *DEL generated ETL models*, $(M, E*)$ where M is an initial epistemic model and $E*$ is a protocol or set of sequences of pointed action models. Call these TDEL models, for short. TDEL models are, intuitively, trees that represent all of the possible evolutions of M (sequences of updated epistemic models) that can be generated by applying any sequence of actions in $E*$.

I omit the interesting and innovative technical details, but do see the foregoing references

---

[26] Though, in the basic logic that Hoshi devises, it's left in, which gives the implausible theorem: for arbitrary action e and a proposition, A, $\langle e \rangle A \leftrightarrow \langle e \rangle \top \wedge A$. Which is to say, if you can apply e at the current state of the current model, then e can be applied (it's executable), and A is already true. This rules out the possibility that one can perform an action that brings it about that A, or, in von Wright's transition calculus, the theorem rules out the possibility of e being the sort of action that results in a transition like $\neg ATA$. That is a terrible expressive limitation for any model of epistemic dynamics that countenances a capacious notion of epistemic actions.

for them. As is implicit in the intuitive description of TDEL models, the primary objection to TDEL is simple: the exact problem that plagued the basic model of dynamics in PDL and DEL persists. While there are protocols in TDEL, and these nicely represent constraints on action, and, suitably interpreted, perhaps also epistemic norms, the resulting model of dynamics conflates sequences of transitions for processes. The protocols in TDEL are analogous to the computation sequences of PDL. These are extremely interesting structures in their own right, and they encode instructions for the performance of action. But the pieces of dynamics with which they are associated are not complex enough to carry information that allows the distinction of (e.g.) the actions in the Honest Twin case. To be a bit more explicit, the foundation being essentially that of DEL, the execution conditions of actions and processes make it impossible for there to be transition equivalent or convergent actions and processes, and so there cannot be machinery that represents epistemic rationality norms that distinguish these kinds of actions and processes. D2 and D3 are thus not formalized adequately.

I close consideration of TDEL with a conjecture. The fundamental representation theorem of the formalism establishes that the models generated by combining TDEL are isomorphic to a subset of ETL models that meets certain specified conditions. But, it's plausible that if one sets up a correspondence between pointed event models or pointed action models and elementary actions in PLEN, then each ETL model is isomorphic to a set of computation sequences in PLEN. From there, one can map these computation sequences to sets of sequences of states in a way that generates PLEN's state sequences. If one can map epistemic models and product updates on epistemic models to nodes in EDGs, then EDGs isomorphic to TDEL models can be generated. If all of that works, then TDEL can be embedded in PLEN by just redefining the truth rules.

5.5 Deontic Logics

The development of deontic logics has unfolded in a number of different, occasionally intersecting directions.[27] However, only some of the directions in deontic logic are relevant to the general project for which a formalism might rival PLEN. Standard Deontic Logic, for instance, despite having a model-theory that is built on Kripke frames just like PDL (and, ultimately, a lot like PLEN), is miles away from an analysis of epistemic norms that meets the desiderata in Chapters 1 and 2. SDL is not oriented toward action in the first place, and its model-theory uses Kripke frames not to model any form of dynamics but to capture the modality-ish behavior of deontic operators. As such, there is no clear way to apply SDL as a logic of epistemic norms that solves any of the problems with modeling epistemic dynamics, even if it's a nice theory of the logical behavior of deontic expressions. As such, I'll be ignoring most of the field of deontic logic in order to zero in on a number of formal frameworks that are plausible rivals of PLEN with respect to the project at hand.

The rival formalisms I will focus on are those that build deontic logics on top of an underlying dynamic model of action. These systems tend to be independent, but broadly parallel in many of their insights. First, I consider von Wright's analysis of norms and actions (VWNA), which provides a deep and insightful analysis of, predictably, norms and actions [261]. I then move on to the STIT paradigm for action and deontic logic (STIT) [27, 130]. Third, I consider Deontic Interpreted Systems (DIS) [166], which are built on top of the classic Interpreted Systems paradigm from [74] (RAKIS). The Interpreted Systems approach is, essentially, an alternative approach to dynamic epistemic logic that takes a somewhat simplified approach to epistemic dynamics as compared to DEL, building both actions and protocols for controlling the development of an epistemic system. DIS then builds a deontic logic on top of this underlying model of dynamics. Finally, I consider the application of PDL to deontic logic [174] that generates Propositional Dynamic Deontic Logic (PDeL). Finally, I turn to Deontic Update Logics (DUL), that builds a deontic logic on top of a logic of updates

---

[27] See [89] for an overview.

[246].

5.5.1 Norm and Action

The seminal [261] is credited by many with the establishing the foundations of deontic logic [89]. There, von Wright develops a transition relation based theory of actions and then imposes a deontic logic of action over this basic theory of world dynamics. Call the system VWNA.

Actions in VWNA are represented by transitions in von Wright's system. The basic piece of the puzzle is the idea that each proposition, p, describes a set of possible situations or states of the world. Actions are, fundamentally, things that can effect changes in the state of the world. Any such change can be represented by some proposition changing from true to false or vice versa. A change in the state of the world can thus be represented by a transition, $(...)T(---)$, where ... and $---$ are replaced by some proposition, p, and its negation in some order. For every proposition p, then, there are four basic possible transitions: pTp, $\neg pTp$, $pT\neg p$, and $\neg pT\neg p$. Transition expressions can be conjoined by truth functions.

Actions stand in a $1-1$ relation to transitions [261], on von Wright's pre-formal assumptions. Thus, the basic action expression schemata with respect to a proposition, p, are $d(pTp)$, $d(\neg pTp)$, $d(pT\neg p)$, and $d(\neg pT\neg p)$. The actions $d(\neg pTp)$ and $d(pT\neg p)$ are actions that effect a change in the world, transforming it from one where p fails to one where p holds or vice versa. The actions $d(pTp)$ and $d(\neg pT\neg p)$ are preservative actions, where the world would have transitioned from a state where p holds to one where it fails, $d(pTp)$ prevents the world from changing in this way. Mutatis mutandis for $d(\neg pT\neg p)$. Forbearing $f(d(...))$ is an operator encoding the omission of action. One can forbear to preserve p just as much as one can forbear changing the state of the world so that p fails or forbear to bring p about. Complex actions are formed in two ways. First, by applying the truth functions to the transition expressions within a $d-$expression. Thus, $d(pTp \vee qT\neg q)$ is the complex action expression of either preserving p or bringing it about that q fails. Second, by applying truth functions to $d-$expressions and $f-$expressions: $d(pTp) \vee f(qT\neg q)$. The operators d and f are conjunctively and disjunctively distributive over perfect normal form complex transition expressions.[28]

Norm expressions are applications of deontic operators O and P to action expressions. Norm-kernels are directly formalized by introducing a novel conditional operator: / that takes action expressions to the left and transition expressions to the right. In von Wright's terms, norm-kernels break into conditions of application, contents, and characters. The content of a norm-kernel is the object, and the character is the valence. The conditional operator / is introduced to formalize norm-kernels directly: $D(d(...))/C$ where $d(...)$ is an action expression, D is a deontic operator, and C is the condition of application. Thus, $O(d(pT\neg p \wedge f(qT\neg q))/rTs$ is a norm-kernel informing us that, if the world is about to transition from an r world to an s world, it's obligatory to bring it about that p fails while also forbearing to bring it about that q fails. The formalism makes room for both wide scope and narrow scope norms.

In examining VWNA, the most natural adaptation to the epistemic case is to treat the propositional language as one that contains epistemic or doxastic expressions and to interpret the deontic operators as epistemic deontic operators - $O(...)$ means that it is rationally obligatory that ..., etc.. That is, a norm-kernel, $O(d(pTp))$, may say that it is rationally obligatory that one maintain the fact p, where p encodes a fact about belief. Thus, the T-calculus is a calculus of epistemic transitions, the $df-$calculus is a calculus of epistemic actions, and the $OP-$calculus is a calculus of epistemic norm-kernels. Keep this adaptation of VWNA in mind in what follows.

VWNA does somewhat well with respect to capacity and feasibility. A key feature of the analysis is the fragment devoted to the analysis of agent abilities. Many of von Wright's motivating observations are direct parallels with the features of constraints on action identified in Chapter 1. For instance, von Wright's "can-do of success" corresponds to feasibility,

---

[28] I omit most of the detail of VWNA. Much of it won't be relevant to the forthcoming objections. There are very many subtleties of the system that are worth studying, but they won't be material in what follows.

and his “can-do of ability” corresponds to capacity. The only complaint to be made about von Wright’s analysis of constraints is that the formalism doesn’t account for changes -over time or due to action - in the preconditions of actions. This is due to a key limitation of the formalism.

VWNA is demonstrably norm exogenous given that norm-kernels have direct formal counterparts. These objects can thus be reasoned about straightforwardly; one can establish a number of dependence, independence, and equivalence relations among them easily because they are embedded in a truth functional logic.[29] For a quick instance, von Wright contends that, given the principles articulated in his system, that $Pd(\neg pTp) \wedge O(d(\neg pTp) \wedge d(\neg qTq) \vee f(\neg pTp) \wedge d(\neg qTq) \vee f(\neg pTp) \wedge f(\neg qTq)) \rightarrow P(d(\neg qTq)/\neg pT\neg p)$ is a tautology. That is, “If one is unconditionally permitted to produce a certain state of affairs but unconditionally forbidden to produce this state and forbear to produce a certain other state, then one is also permitted to produce this second state under circumstances which constitute an opportunity for producing the first state.” (p. 182)

The main problems with VWNA are straightforward and threefold: von Wright identifies acts with transitions and gives essentially an input assimilating model of dynamics. Thus, the system falls afoul of the Processes problem in two dimensions. First, the system conflates actions with transitions and so can not distinguish between the actions a and b in the Honest Twin case with respect to any of the properties identified. This, of course, results in a failure to accurately represent procedural norms - von Wright’s norm-kernels cannot require that one carry out an action or process a rather than b where a and b result in identical token transitions. The $df-$calculus that expresses the contents of norms simply cannot distinguish such actions. Explicitly, consider a transition expression in von Wright’s T-calculus, ATB, where A is a complex propositional expression capturing the complete state description of the world prior to a change and B is a complete state description after. Let two actions a and b result in identical token changes of the world described by $A'TB'$ (such that A implies

---

[29] See for instance the various principles that von Wright defends concerning preconditions of action, conditions of application, categorical and hypothetical norms.

A' and B implies B'). Explicitly, for every p in A', if a results in pTp (mutatis mutandis all other basic action expressions), then so does b and the converse also holds. Then the $df-$calculus analyzes the tokens of both a and b with $d(A'TB')$. Thus, given von Wright's $OP-$calculus, every norm N assigns a and b the same valence. This is a syntactic result: no OP expression can assign the token of a resulting in $A'TB'$ a deontic D without also assigning D to the token of b resulting in $A'TB'$ because the language can only assign deontics to *df* expressions like $A'TB'$. The same problem results if A is logically equivalent to A' and B is logically equivalent to B'.

Thus, VWNA (suitably modified to the epistemic case) fundamentally fails to adequately represent epistemic dynamics. VWNA suffers the Processes problem. As a result, it fundamentally mischaracterizes the content of epistemic norms by misconstruing the extensions of restrictions on epistemic dynamics. From there, despite its other logical successes, VWNA cannot give a fully adequate logic for epistemic norms simply because it omits so many of their essential properties. For instance, only transitional forms of norm equivalence, dependence, or sufficiency can be recognized in VWNA. If N and N' are procedurally inequivalent, VWNA has no way of capturing this or the kind of reasoning that can be used to establish it. So, VWNA's adequacy as a logic of norms is lower than a rival that can capture such distinctions.

Second, VWNA does not recognize the program-like aspects of procedural norms. There are no explicit procedural operations in VWNA, and such operations would ill fit the system. The logic is a logic of single opportunities; it has no temporal or dynamic dimension beyond the basic notion of transition. The underlying "T-calculus" or calculus of transitions has no mechanisms at all for iterated changes or extended processes of change. As such, there is no clear sense in which VWNA represents processes at all, and so, a fortiori, no way in which VWNA represents the program-like features of norms and their restrictions on dynamics.

Finally, VWNA does not contain direct epistemic operators, and, in its classical presentation, doesn't have a model theory that suggests a way to represent knowing or deliberate execution. It's not clear whether anything stands in the way of adding this machinery.

5.5.2 STIT Theory

The STIT theory analyzes action over a multi-agent, branching time model of dynamics [27, 130]. A STIT frame is a triple, $\langle T, <, V, C, I\rangle$ where T is a set of moments with a tree ordering $<$ over them, maximal sequences of moments in a frame are histories, V is a valuation function mapping pairs of histories and moments to propositions of the language of STIT, and I is a set of agents. The unique bit is C: a set of choice functions per agent i that partitions paths of moments into choice cells, C^i_m. A choice cell, C^i_m, is a sets of paths through a moment m in the tree ordering. The choice cell $C^i_m(h)$ is the choice cell that contains h. Say that $h, m \models \phi$ iff $(h, m) \subseteq V(\phi)$. Intuitively, the idea is that each agent's choices at any given moment, m, restricts the set of histories that can extend forward in time from m.

The basic form of action in STIT is $iSTIT\phi$, which means, roughly, that i sees to it that ϕ. Doing so, at any given moment, brings it about that ϕ is true and also restricts the future. Omitting quite a lot of detail, STIT focuses on the notion of Deliberative STIT: $h, m \models iSTIT\phi$ iff $C^i_m(h) \subseteq V(\phi)$ and there is at least one history that is not in $V(\phi)$.

In examining STIT, the most natural adaptation to the epistemic case is to treat the propositional language as one that contains epistemic or doxastic expressions, to interpret the STIT operators as epistemic actions, and to trade moments out for epistemic states. Doing so leaves STIT with much in common with ETL. STIT by itself is a logic of action. Building in action modals, one can encode postconditions with formulae like $[iSTIT\phi]\psi$. One can build deontic operators over STIT that function in rather obvious ways (like temporal or dynamic operators) in order to build a normative language. One can then construct norm-kernel like formulae like $C \rightarrow O(iSTIT\phi)$. An easy example is the epistemic consequentialist norm: $[iSTITb_M(A)]b(A) \wedge Tr(A) \rightarrow P(iSTITb_M(A))$. If i's seeing to it that she believes A by method M always results in a state where A is believed and true, then it's rationally permissible for her to see to it that she believes A by method M.

The major problem with STIT should be obvious given that STIT explicitly analyzes actions in terms of what effects they bring about: histories are merely series of states, and the actions in STIT are mere transitions. As such, they cannot distinguish transitions from actions.[30] If $iSTIT\phi$ and $iSTIT\psi$ correspond to identical transitions (e.g., if ϕ is logically equivalent to ψ), then $D(iSTIT\phi)$ iff $D(iSTIT\psi)$, where D is a deontic operator. As such, "extensionally equivalent" actions like those in the Honest Twin case will be normatively equivalent under all conditions, and this is, of course, false. Thus, D3 is inadequately formalized, as is norm equivalence; the restrictions on epistemic dynamics of procedural norms are ignored and STIT validates incorrect reasoning about norm verdicts. If the propositional language STIT encodes methodological or procedural statements, the Processes problem implies that STIT will fail to adequately represent dynamics by suffering the problems articulated in Chapter 1.

5.5.3 Propositional Dynamic Deontic Logic (PDeL)

Meyer [174] combines PDL and deontic logics with a unique semantics to generate a dynamic deontic logic in the same mold as von Wright's system. Which is to say, Meyer builds a deontic logic on top of a (dynamic) logic of action. He defines a set of actions *Act* and assertions A. Over *Act*, he deploys the sequential composition ;, choice $\cup$, parallel execution $\cap$, and negated action (or complement) $-$, which map onto the same intuitive procedural operations that they do in PLEN. Meyer adds a special control operator for conditional choice $A \rightarrow a_1/a_2$, which intuitively maps onto the instruction "If ϕ, then a_1; else a_2". The semantics of actions are defined in terms of (the finite prefixes of) s-traces, which are

---

[30] I omit detailed consideration, here, but Xu [262] combines STIT with an action theory in a way that is quite interesting, and gets at many of the same philosophical judgments about action that PLEN does. For instance, actions in the combined STIT and Action theory (call it ASTIT) are sets of transitions such that, in every history, each transition in an action occurs only once. This enables ASTIT to distinguish act tokens from act types in much the same way that PLEN and PDL do. However, the dynamics are still histories as in STIT, and there is seemingly no recognition in ASTIT that actions might be procedurally or normatively distinct even if they define identical sets of transitions.

sequences of elementary action symbols. These sequences of elementary action symbols are then associated with state transition relations, and the semantics of assertions is then built on top of this action logic. I omit the details, here, but do see the appendix of [174].

Unlike PLEN, PDeL does not represent epistemic norms via constructions in the action logic. Rather, norms are implicitly characterized via the propositional language and the deontic operators. In PDeL, A is a straightforward propositional language with a special violation constant V, which indicates that one has violated the requirements of whatever norms define the deontic notions being analyzed via PDeL. The violation constant is randomly assigned to states in the models of PDeL. The assertion $[a]A$ is central to PDeL, as the semantics of normative forbiddance or prohibition is defined in terms of the violation constant and this construction: letting s be a state, $s \models Fa$ iff $s \models [a]V$. The rest of the deontic operators are defined in terms of Fa.

PDeL can be appropriated for epistemic normativity by straightforward means. Think of the states as epistemic states. Suppose that A contains expressions encoding assertions about (e.g.) epistemic states. The facts about the deontics characterize the requirements of epistemic rationality. That is, where $[a]V$ is true, epistemic rationality tells us that execution of a is forbidden; the execution of a results in violation of the requirements of epistemic rationality. Then, by encoding normative assertions involving deontic operators, and conditionals with deontic assertions in their consequences translate norm-kernels.

The action logic of PDeL is superior to that of the base form of PDL because, essentially, s-traces are computation sequences. The action logic of PDeL can distinguish between processes and transitions, in principle, because it associates each program to a set of computation sequences which are then associated with transition relations. Then, all of the benefits of the modified form of PDL above accrue for PDeL. So, when applying PDeL to the epistemic case, the underlying model of epistemic dynamics is as adequate to the Constraints and Processes problems as PLEN. However, PDeL falls short of some of D1-D3.

D1 is as well taken care of in PDeL as it is in PLEN. As argued above, the basic program logic in PDeL gives a model of epistemic dynamics equivalent to that in a fragment of the path-theoretic fragment of PLEN. The programs that are associated with identical transition relations are not associated with identical s-traces. Taking s-traces to encode epistemic processes, this solves the Processes problem. Given that different s-traces are associated with different starting states, this handles the Constraints problem.

As for D2 and D3, the template for argument in this chapter tends to move from failure of Processes to failure of D2 and D3. PDeL bucks this trend by failing to formalize D2 and D3 without suffering the Processes problem. Now, on the one hand, given that s-traces can be rational or irrational, the restrictions on epistemic dynamics representable in PDeL include restrictions on processes, putting PDeL a step ahead of many other formalisms. Interpreting the violation constant as the violation of the requirements of epistemic rationality, the semantics of $F\pi$ for all programs enables the definition restrictions on processes by assigning the s-traces of forbidden programs to the irrational sets. So, it seems that the restrictions on dynamics that epistemic rationality specifies can be represented as well in PDeL as they can be in PLEN.[31] But this is too fast.

By deriving restrictions on s-traces from the violation constant, and associating the violation constant only with states, we have failed to properly represent restrictions on procedures and on processes. Consider a pair of programs, π_1 and π_2 such that the set of state transition relations associated with π_1's s-straces is identical to the set of state transition relations associated with π_2's s-traces. Then $F\pi_1$ iff $F\pi_2$ because the final states of those transition relations are identical despite the non-identity of the s-traces of π_1 and π_2. Thus, PDeL will mischaracterize the restrictions on procedures, actions, and processes that norm-kernels define because all normative verdicts will be equivalent for any transition equivalent programs. This directly fails to correctly represent restrictions on processes correctly because PDeL will obscure the restrictions that norms like Bad Reasoning from Ch.2 specify. Thus, PDeL will fail to adequately characterize restrictions on dynamics for the reason that it runs together

---

[31] This isn't to say that either PDeL or PLEN characterize the extensions of these restrictions accurately, but only that both formalisms can characterize restrictions on procedures.

the restrictions on processes where the processes map to the same transitions. As shown in Chapter 2, there will be states at which reasoning by modus ponens and by affirming the consequent will both be forbidden or both be permitted because one of them is, and they converge on a terminal state that does (or doesn't) satisfy the violation constant. This is a result of a more central problem with PDeL.

The primary failure of PDeL, even deployed to the epistemic case, is that it fails norm exogeny. Programs in PDeL are simply procedures or plans, and their place relative to the restrictions on dynamics that epistemic rationality determine is given, ultimately, by the violation constant. But the violation constant encodes a single, unchanging protocol underlying each model of PDeL. While one can draw inferences to or from specific norm-kernels, these are all defined with respect to the violation constant. This conflicts with a prima facie plausible conceptual principle: the definition of violation of the requirements of epistemic rationality should flow, in an account of epistemic rationality, from the norms that specify those requirements.[32] PDeL gets this exactly backwards, and this is what generates the failure of P2 and P3 above.

In PLEN, restrictions on states and transitions are derived from the restrictions on procedures (defined by execution sets of protocols), which encode the instructions of procedural norms, mirroring the conceptual priorities argued for in Chapter 2. In PDeL, restrictions on procedures are derived from restrictions on states formulated via the violation constant, reversing the order argued for in Chapter 2. The violation constant is assigned to states randomly, thereby characterizing a kind of ur-requirement of epistemic rationality that then determines what norm-kernels hold at what states. Then, assigning restrictions on processes on the basis of restrictions on actions defined by appeal to the violation constant gives us restrictions on processes that have to be linked extensionally to the verdicts of the norm-kernels - and these are dependent on a restriction on states defined by the satisfaction of the

---

[32] This can be seen as a consequence of the asymmetric "generation facts" argued for in Chapter 2. Every restriction on states or transitions can be read off of some restriction on procedures, but not the other way around.

violation constant.

From there, however, PDeL can formalize arguments to and from the dependence and equivalence relations among norm-kernel translations. However, there is no machinery in PDeL that can formalize reasoning about what it is that defines the violation constant. Which is to say that PDeL can formalize no epistemic norms nor reasoning about them. Norm-kernels then, do not have the conceptual primacy that epistemic norms have, so the formalization of reasoning about norm-kernels isn't a formalization of reasoning about epistemic norms. This, I think, suffices for a direct failure of narrow exogeny, and, without norm exogeny to embed norms in assertions about what agents know about various norms, a direct failure to formalize arguments about norms. Worse, as above, we derive equivalence of restrictions on processes from equivalence of restrictions on actions from equivalence of restrictions on states, and this mischaracterizes the equivalence relations among these kinds of restrictions by imposing a connection that was argued, intuitively, to not hold in Chapter 2: transitionally equivalent processes needn't be normatively equivalent! Just see the Honest Twin case.

Another pressing problem with the epistemic take on PDeL is that, as with the modified version of PDL, the program logic is just an alternative way of pursuing the basic underlying model of epistemic dynamics as in PLEN's path-theoretic fragment. From there, some machinery is added to account for epistemic normativity via the violation constant and deontic operators. All of this can just be added to PLEN without loss. Or, from another angle, one could replace the underlying action logic in PDeL with that of PLEN, and the formalism would come cheaper, in that PLEN has simpler and more diagrammatically intuitive structures. As such, PDeL represents an extension or modification of a fragment of PLEN with an alternative take on the primacy of requirements on states, transitions, and processes as and a suitably alternative account of deontic notions. Thus, even if none of the problems above are compelling, PLEN can assimilate PDeL's virtues as it could with the modified PDL.

5.5.4 Deontic Interpreted Systems (DIS)

Interpreted systems model dynamics in a straightforward way. First, given a set of agents $A = \{1, ..., n, e\}$, one of which is the environment, we define a set of *local states* for each agent $S_1, ..., S_n, S_e$. Then, we define a *global state* as a tuple of local states, $g \in S_1 \times ... \times S_n \times S_e$. So, the set of global states is $G \subseteq S_1 \times ... \times S_n \times S_e$. Define *runs over* G as sequences of global states or, alternately, maps from sequences of natural numbers into G. An *interpreted system* is a set of runs over G.

For each $i \in A$, there is a set of *actions*, $Act_i = \{a_1, ..., a_m\}$. Each action is associated with a transition among global states meeting certain constraints that I omit, here [166]. The essential thought being that actions change the states of agents, some actions change an agent's own states, other actions change the states of other agents. *Protocols*, then, are defined as rules that map local states to actions. Thus, *protocols define sets of runs over* G. Thus, each protocol determines an interpreted system.

A deontic logic can be built on top of interpreted systems in a straightforward way. Partition the sets of local states, L_i, for each agent into *green states* Gr_i (and, by complementation, red states Rd_i), including the environment. A *deontic system of global states* is any system of global states defined over the Gr_i; that is, each global state is defined as a tuple of green states and a deontic system of global states is a set of runs over this set of global states. A *deontic interpreted system* is simply a pair (DS, v) where DS is a deontic system of global states and v is a valuation function assigning the atoms of a logical language to sets of global states.

Define *generated deontic frames* as multi-model Kripke frames $\langle W, R_1, ..., R_k \rangle$ for each DS such that: $W = DS$ and for any $i \in A$, $\langle l_e, l_1, ..., l_n \rangle \, R_i \, \langle l'_e, l'_1, ..., l'_n \rangle$ if $l'_i \in Gr_i$. Deontic operators are interpreted over generated deontic frames:

(Obligation) A global state, g, in DS satisfies $O_i\phi$ iff, for all g' such that gR_ig', g' satisfies ϕ.

The thought is that, because each protocol defines a set of runs over G, there is a set of protocols that defines each DS. Thus, each g in every DS is a world that accords with such a protocol - it describes what happens if the protocol is perfectly followed by all parties (and the complement of the sets of runs over G in which global states are not green describes what happens in the worlds where the protocol isn't followed by all parties). Obligations, then, are defined in terms of what complies with the underlying protocols that define the green runs. Derivatively, the permissible actions are those that correspond to the R_i in generated deontic frames.

An alternative approach is to directly consider green runs and transitions: "In applications to specific examples it is often useful to classify as red the states that result from the failure of an agent to follow its functioning protocol. In these cases one can consider a finer-grained notion of interpreted systems in which the concepts of protocols and transitions are introduced. Moreover, a rather different and interesting approach is to label runs of the system as 'red' or 'green' instead of states, enabling us to reason about allowed/acceptable as opposed to disallowed/unacceptable/faulty runs." [74, 166] (p. 6)

The problems for DIS with respect to the project of modeling epistemic dynamics and epistemic norms should be obvious, now. Let interpreted systems encode epistemic states in any of a number of ways - by simply verifying doxastic and epistemic expressions added to the language, by replacing them with epistemic models, etc.. Epistemic norms are then the protocols that define the green runs defined over the green states. The problem is similar to that which plagued PDeL. There, the problem was that the violation constant essentially defined a restriction on states first, and then restrictions on transitions, actions, and processes was derived from that. In either version of DIS, the green states or green runs define a restriction on states or a restriction on runs. This, in turn, can be used to define a restriction on transitions, actions, and processes. But these are just extensional notions. So, however restrictions on transitions, actions, and processes are derived from green states or green runs, they will be limited in ways that get D2 and D3 wrong. It won't be possible for there to be restrictions on (e.g.) actions and processes that distinguish actions/processes with identical transition relations or identical runs.

Let's go with green states, first. Consider a pair of agent i's potential actions, a_i and b_i. Let a_i and b_i be associated with identical transition relations on global states. Say that Pa_i holds at g iff a_i is executable at g or at some green run that extends from g in a green run. It follows that Pa_i iff Pb_i. But this just gets the Honest Twin case wrong; let a_i be asking my twin and b_i be carrying out the deduction myself. These will not be permissible at all of the same states. Just so for reasoning by modus ponens and reasoning by modus morons, and so DIS can't formalize procedural norms like Bad Reasoning. This means that DIS fails D2 and D3.

Now, lets consider green runs. The set of green runs defines a set of green protocols - exactly those protocols that result in green runs. First, taking protocols to analyze norms, this gets norms wrong. Let π_1 and π_2 be such that the set of runs associated with π_1 is identical to the set of runs associated with π_2. Then every property of π_1 that holds strictly in virtue of π_1's runs holds of π_2 and vice versa. But, as was argued in Chapter 2, two norms might generate the same restriction on transitions and even sequences of transitions without specifying the same restrictions on actions. A protocol N that always requires asking my Honest Twin and always forbids deducing things for myself will always converge on runs with the protocol N' that always requires deducing for myself and forbids asking my twin. To conclude that these protocols (hence, norms) are equivalent would be to simply mischaracterize them. They ought to have different execution or compliance conditions, supply different verdicts about actions, and specify distinct restrictions on action.

Second, if distinct actions (by distinct agents!) may be assigned to the same transitions, then protocols π_1 and π_2 may be rather distinct, assigning different actions to at least some states, while defining the same runs. Let Pa_i at g iff a green protocol π maps a_i to g or a successor state of g. Then, let π and π' be protocols with distinct action to state mappings that are both green because their actions converge on a run. Thus, let a_1 be asking my honest twin and b_1 be deducing for myself. It follows that Pa_1 iff Pb_1. But this is false in the Honest Twin case.

5.6 PLEN as a Supersystem

It might be objected that, while PLEN satisfies the desiderata defended here, it fails the desiderata that animate its rivals. If this is right, then despite the foregoing arguments, it doesn't follow that PLEN is (ceteris paribus) rationally preferable to its rivals. For instance, PLEN might be argued to fail to provide plausible models of the classes of epistemic actions that effect expansion, contraction, or revision because it offers no rules or laws that correspond to or analyze the epistemic commitment functions that represent these things or the rationality postulates governing them. For another, PLEN contains no structures that represent credence and so, from the perspective of Bayesians that don't countenance full acceptance or rejection, fundamentally assumes a misbegotten model of epistemic states. More, there is no way to characterize (e.g.) conditionalization in PLEN - so how can it possibly provide a model of rational belief change? For yet another, PLEN's syntactic account of the content of epistemic states lacks the structure to ground the truths about epistemic states that epistemic models provide via their graph-theoretic properties; the analysis of knowledge in PLEN must thus be mistaken or ungrounded in the features of the formalism. Given that, PLEN's model of dynamics can't properly analyze the relations of dynamics to changes in knowledge.

This objection can be answered straightforwardly: the benefits of the other formalisms can, without loss, be retained in PLEN because PLEN is abstract enough to act as a *supersystem* for all of them. The core of the thought is just that the model of dynamics at the core of PLEN is so abstract that one can replace the model of states with just about anything while retaining the important features of the dynamics and the logic of protocols. PLEN would be a "supersystem" for all of these other systems of dynamics insofar as the features of dynamics represented by any model in the other theories would be represented by some subset of models of PLEN. PLEN would thus be a system in which other theories of dynamics appear as fragments. If this is right, it wouldn't entail that none of the foregoing systems are actually rivals to PLEN by the definition of rivalry in section 5.1. Rather, the fragments of a PLEN-based supersystem would not be formally identical to these other systems; they would just provide logical resources equivalent to them.

The procedure for embedding the foregoing systems in PLEN is this: First, modify PLEN's content states to mirror whatever epistemic structures one likes. Replace content states with belief sets, credence functions, or epistemic models. Second, introduce axioms governing them (e.g., deductive closure, exhaustion and exclusion of membership in belief set with respect to a propositional language, probabilistic coherence, arbitrary structural conditions on epistemic models, etc.). From there, we can characterize the protocols that correspond to the axioms as those such that all executions are bookended with transitions that meet the conditions on transition relations in the other formalisms. For instance, we can treat X in PLEN's content states as belief states in AGM and then reason from the stipulation that $\pi{+}_A$ always results in the expansion of the X at the root state of its executions with A. That is, expansion can be characterized by axioms that govern protocols in PLEN. Just the same, the content states can be replaced with credence functions, and we can reason from the stipulation that π effects coherent conditionalizations. Or again, replace the content states of PLEN with pairs of epistemic models and arbitrary (possibly empty) sequences of protocols, associate each protocol to action models per graph, and stipulate that π's executions are bookended by the inputs and outputs of a product update.

Given these kinds of modifications, the properties of epistemic commitment functions, updating procedures, and DEL style informational events can be embedded in PLEN as properties of protocols or their executions. For instance, let $\pi{+}_A$ and $\pi{-}_A$ encode expanding with A and contracting A, respectively, and let X be treated like a belief state ($X = Cn(X)$), and let $b(A)$ mean that A is in X ($w \Vdash b(A)$ iff $a \in X$). The Recovery postulate of AGM can be encoded with the schema: $b(A) \rightarrow [\pi -_A; \pi +_A]b(A)$.[33] Now, this is one case, and it's far from clear that this is an adequate approximation of Recovery or the rest of the analysis of commitment functions in PLEN, but (a) it's fairly well known that AGM can be embedded in dynamic and epistemic logics, and (b) there isn't space to fully carry out this embedding of the other formalisms in PLEN; establishing this is itself an interesting and fairly major project. The point is that the claim stands as a plausible conjecture for now.

Now, the important feature of this conjecture is that all of the good modeling work done by these alternative formalisms can be done in PLEN while retaining the features of PLEN; we don't lose any of the structures that made PLEN so useful at solving the Processes or First-Class Citizenship problems. All of the characterizations of (e.g.) the epistemic commitment functions can be done without conflating processes for transitions and while analyzing epistemic norms as protocols in such a way that all of the results of PLEN hold. This is because the foregoing embedding procedure just requires the modification of the epistemic states and perhaps accompanying revisions of the epistemic part of the language and not the protocol-theoretic part. None of the primary results of PLEN (the analysis of equivalence, the distinction of processes from transitions, the analysis of restrictions on epistemic dynamics, etc..) depends on these features of PLEN. Thus, we can zero in on (e.g.) the AGM-like fragment of PLEN focusing on the atomic (or complex) protocols (or classes of protocols), $\pi +_A$ and $\pi -_A$, which are characterized by axioms encoding the rationality postulates without the loss of the structural features of the formalism that (e.g.) distinguished processes from transitions. One could, after all, just think of all protocols that abide by the rationality postulates in a given EDG and still distinguish between the protocols themselves - as well as the paths that execute them - while not obscuring anything about the transitions that AGM wants to focus on. Two protocols, $\pi +_A$ and $\pi' +_A$ might be inequivalent, have different executability conditions, or different methodological inclusion conditions even if they both

---

[33] For Bayesianism, replace X with credence functions and introduce (a very large) expansion to the language that encodes credences on formulae, $Cr(A) = n$, and conditional credences $Cr(A \mid B) = m$. Then one can characterize conditionalization protocols, π_C, by means of formulae like: $Cr(A \mid B) = n \rightarrow [\pi_C](Cr(E) = 1 \rightarrow Cr(A) = n)$.

basically achieve expansions with A.

If PLEN truly is a supersystem in which all of the rival formalisms can be embedded as some variant of one fragment or another, then the theory preference argument is truly powerful. Accepting PLEN as a formal framework in which to represent and reasoning about epistemic dynamics and epistemic rationality is both a theoretical step forward, in that the Framework Problem is solved, and comes at no theoretical loss because all of the rival formalisms can be preserved as fragments or variants of fragments of PLEN. Rational theory preference thus dictates acceptance of PLEN over any rival assessed here.[34]

5.7 Conclusion

This chapter completes the primary argument for the protocol-theoretic account of epistemic rationality norms. PLEN formalizes the protocol-theoretic account by building its main theses directly into the semantic and syntactic features of PLEN. The primary argument consists of showing that PLEN is theoretically preferable to other formal frameworks for thinking about epistemic dynamics and epistemic rationality norms.

Chapter 4 showed that PLEN (i) solves the Processes, Constraints, and First-Class Citizenship problems, and (ii) formalizes D1-D3. More, PLEN has these properties, is narrowly exogenous, and it doesn't have any comparable theoretical tradeoffs. PLEN is completely general with respect to action, it's a formalism with relatively few moving parts that demands a minimum of mathematical sophistication, and it is at worst unclear whether PLEN contains any other theoretical tradeoffs like paradox.

Where Chapter 4 showed how PLEN satisfied the foregoing desiderata, this chapter

---

[34] Notably absent from the foregoing remarks are deontic logics. This is for rhetorical reasons. First, I attack the problem of formalizing epistemic deontics in the main text at some length. Second, it's difficult to see how there are any desiderata on deontic logics as applied to the analysis of epistemic rationality that aren't countenanced in the various problems and desiderata PLEN is shown to handle. Any further desiderata will really be a matter of adjusting the logic to account for the intuitive validities and invalidities concerning deontic reasoning, and the logic of PLEN is also fully adjustable without sacrificing (e.g.) the process-transition distinction: just add new operators or change the semantics of the operators or truth functions as one sees fit. This can be done safely within the general framework of PLEN, even if this changes the exact logic of PLEN.

showed how PLEN's rivals failed to satisfy those same desiderata. The recurring discovery was that rival systems almost ubiquitously built their underlying models of dynamics on the thought that actions are best represented by transitions or transition relations. This feature of dynamics results in the Processes problem and the failure to formalize D2 and D3. Given that the tools available for reasoning about epistemic norms in a system are dependent on the underlying model of dynamics, this also generates a general failure to provide adequate logics of epistemic norms. The addition of formal machinery in rival systems sometimes added to this underlying model, but the additions either did little to alleviate the main problems or just made the system more like PLEN.

The results of the primary argument are summarized in the following tables:

(Key) Systems in this family are generally:

- + :: Adequate
- - :: Somewhat Adequate, Some Systems Adequate, or Unclear
- X :: Inadequate

	Norm Exogeny: Broad or Narrow	Modeling Epistemic Dynamics: Constraints Problem, Process-Transition Distinction, Restrictions	Desiderata: D1-D3	Paradoxes and Other Tradeoffs: Paradox, Generality, Simplicity, ...
Belief Revision Systems	X	X	X	X
Dynamic Logics	+	X	-	-
Dynamic Epistemic Logics	+	X	-	-
Deontic Logics	+	X	X	X
PLEN	+	+	+	-

	D1	D2	D3
Belief Revision Systems	+	X	X
Dynamic Logics	+	X	X
Dynamic Epistemic Logics	+	X	X
Deontic Logics	-	X	-
PLEN	+	+	+

Even allowing that PLEN has general theoretical weaknesses that are as yet undetected, every rival framework does worse, especially with respect to modeling epistemic dynamics. As a result, PLEN's informal rationality index is (ceteris paribus) higher than that of any rival formalism.

Finally, I conjectured that PLEN has the interesting property of being able to embed approximations of the other formalisms thereby retaining the theoretical benefits of these systems without sacrificing anything that would cause a desideratum to fail. If this is correct, the theory preference argument need not be qualified as carefully as it has been.The argument contained in Chapters 3-5 shows that PLEN is demonstrably superior to its rivals with respect to the problems and desiderata focused on, here. The conjecture shows, if correct, that PLEN is already at least as good as its rivals at whatever they do. In other words, PLEN does better with respect to these desiderata and ceteris is paribus! So PLEN is flat out rationally preferable to its rivals as an account of epistemic dynamics and epistemic rationality.

Chapter 6

Recapitulation and The Road Ahead

6.0.1 The Entire Project in Focus

Epistemic norms are implicated in all of the most important epistemological projects. It's epistemic norms that define the contours of epistemic rationality and epistemic justification. These things feed epistemic norms into the rest of epistemology in some way or another. Traditional epistemology either builds assumptions about epistemic norms into accounts of knowledge and justification or builds them into the derivation of norms for action, assertion, and belief built on accounts of knowledge. Formal epistemology essentially builds models of what different kinds of reasoning or cognition look like according to assumed systems of epistemic norms. On procedural approaches to epistemology or the ethics of belief, epistemic norms transparently take central importance. Epistemic norms thus lie at the conceptual foundations of epistemology. In all cases, errors and oversights about the basic properties of norms may introduce barriers to getting things right in the end. The central focus of this project has been developing and defending a dynamic framework for thinking about epistemic norms that resolves some of the central errors and oversights in thinking about epistemic norms.

In this dissertation, I have defended a "protocol-theoretic" account of epistemic norms. The main idea of the account is that epistemic norms ought to be thought of as protocols or instructions that map epistemic actions and combinations of epistemic actions to conditions under which they can correctly be carried out. In order to study norms from this perspective, norms should also be correlated with the effects they would have on (among other things) an agent's belief states if followed. Thus, the account requires an underlying model of epistemic dynamics – a model of how things like beliefs and commitments change as a result of actions and processes. To make the account precise, and to defend it as a fruitful theory, I developed, explored, and applied a "protocol-theoretic" logical framework for analyzing epistemic rationality norms. I called it PLEN. Formally, PLEN is built out of pieces of propositional dynamic logic, epistemic dynamic logic, and formal belief revision theories like AGM.

The basic insights of dynamic logic are that they have formal languages that encode the structure of instructions that are given by programs, and semantics that – with some modifications - are excellent models of epistemic dynamics. The basic insights of epistemic logics and belief revision theories is that doxastic states or other contentful epistemic states can be represented by set-theoretic structures (in the simplest case, sets of sentences), and reasoning can be represented by transition relations among the models of epistemic states. Combining these parts - along with minor innovations in the technical machinery - we get PLEN, which basically provides a model of how normative instructions guide the behaviors of epistemic agents and result in evolutions of contentful states of epistemic agents.

The main results of the dissertation show that, though still in the early stages of its development, PLEN not only analyzes and justifies theoretical features of epistemic normativity that are frequently ignored or denied (desiderata D1-D3), but also overcomes some of the major drawbacks of the existing approaches to epistemic dynamics and belief revision. As I argued in Chapters 1 and 2, there are deep problems with any formal account of epistemic dynamics that doesn't distinguish epistemic actions, processes, and procedures from the transitions among epistemic states that they bring about. These problems result in fundamental failures to analyze the properties, especially logical properties like equivalence and complete accounting of normative verdicts, of any account of epistemic norms built on top of this underlying model of dynamics.

The innovation of PLEN, as defined in Chapter 3 and applied in Chapter 4, is the simple modification of the structures of PDL into EDGs and the accompanying redefinition of the execution conditions of protocols. This lead to a reconceptualization of restrictions on epistemic dynamics that enabled the idea to account for restrictions on processes as well as restrictions on procedural construction. More precisely, I showed that the protocols of PLEN are structures from which all of the theoretically important features of restrictions on dynamics can be derived by means of several technical results. These moving parts together made for the success of the analysis of epistemic rationality norms with protocols; the Framework Problem was solved. Chapter 5 showed that PLEN is unique in solving the problems raised in Chapters 1 and 2. None of the rival theories of PLEN can solve the Framework Problem as currently formulated, and it's plausible that their theoretical fruits can be assimilated into PLEN.

This project is intended to form the kernel of a “hard core” style research program. In future work, I want to take the protocol-theoretic account of norms and the PLEN framework and explore what can be done with them. Let me comment on some possible applications.

First, the problem of the epistemic role of logic is arguably the core problem in the intersection of epistemology and philosophy of logic. Reasoning is one of our primary sources of rational belief and knowledge. Logic is often conceptualized as the study of reasoning, but this can't be the case. Logic, as a field, focuses on the construction and exploration of logics, which are formal systems of not indeterminate but also not uncontroversial form [88]. Logics are not generally formal systems with special utility in analyzing reasoning, and what they do analyze - logical consequence or validity or proof rules - isn't indisputably related to the dynamic process of reasoning in any special way. Harman [121, 119, 122] showed us this, and thereby inaugurated the debate about the epistemic role of logic.

PLEN can be used to formalize parts of the dispute about the epistemic role of logic

[23, 79, 80, 121, 241] that are helpful for resolving it. Explicitly, PLEN can formalize the different accounts of the epistemic role of logic. More precisely, PLEN provides a formalization of epistemic rationality norms, and, as a result, it can formalize different possible bridge principles between logical validity and epistemic rationality. PLEN can also formalize desiderata on such accounts of logic's epistemic role. Thus, theoretical preferences among bridge principles may be clearly articulated and rigorously defended in the logical platform offered by PLEN. But this formalization is partly revisionary; it reconceives epistemic norms from a dynamic and procedural perspective. The application of PLEN might thus result in a reconceptualization of the debate about the epistemic role of logic, and may have surprising results. While theoretically interesting in its own right, such a reconceptualization attains added interest given that different accounts of the epistemic role of logic interact with other disputes in philosophy of logic, such as those about logical theory preference [8, 110, 202, 206, 22] and revision of logic [38, 56, 138].

Second, PLEN can almost certainly - in virtue of both its similarities to and differences from existing multi-agent settings [65, 74] - be extended fruitfully to the logical formalization of the EMIL-A model of norm immergence and compliance [59]. More work can also be one with respect to Chapter 2's gestures concerning how normative alternatives are supplied in default interventionist accounts of cognition, and it may be able to supply a model of the content that is taken on by individual agents in the development and adoption of norms. Beyond that, there is room to develop connections to the broader study of normativity and action, as the formalism can be used as a basis for deontic logic and the semantics of imperatives [89, 250, 251].

Of course, there is always room for further development of the underlying philosophy of the model as well as further technical development of the basic formalism. While I've argued fairly extensively for the philosophical account of norms to which the formalism is anchored, it's a wide-ranging discussion, launching criticisms of timeslice epistemology, devising arguments from robustness analysis, proposing accounts of the purpose of epistemic norms, and so on. Each of these lines can be further developed and defended. On the technical side, and I have only proven soundness results and a number of other results that were especially useful for the philosophical purposes of the dissertation. Considerations of completeness, decidability, alternative kinds of proof system, and extensions with new devices are all in the offing. Let me close with a superficial glance at potential future technical extensions of PLEN.

6.1 Extensions of PLEN for Future Work

PLEN is a modular framework. There are many points at which one can impose restrictions on the constructions in PLEN in order to generate distinct logics for distinct purposes; modifications practically suggest themselves for various purposes. First, and foremost, there is much room for the rigorous exploration of the modifications of PLEN that were briefly developed in Chapters 4 and 6. Here, I restrict myself to some potential avenues of modification that have heretofore gone unremarked upon.

6.1.1 Alternative Parallel Executions

There are several ways in which parallel executions are unsatisfactory. One way is that they require parallel executions to be composed of actions that decompose in to the same number of steps. A better way to analyze parallel executions would be to build in machinery that represents the temporal durations of actions, and defines parallel execution in terms of durations. One could read $\pi_1 \cap \pi_2$ as instructing the execution of π_1 while executing π_2 (and vice versa). One way to do this would be to treat states as temporal indices and to define the executions of parallel executions as those paths p such that $first(p)$ is either $first(q)$ for some execution of π_1 or $first(r)$ for some execution of π_2 and $last(p)$ is either $last(q)$ or $last(r)$ and the processes are interleaved in some way. A version of this is carried out in the technical appendix of Meyer's PDeL [174]. The analysis of concurrency is a well-developed subfield of dynamic logic [113], and a full analysis of the options for revising parallel execution is reserved to future work.

6.1.2 Alternative Tests

In modeling epistemic dynamics, there are a number of interesting ways to think of tests. One question to answer about tests is, independent of what epistemic action they are to represent, whether to think of them as atomic actions or processes that might have durations that vary. In the procedural fragment of PLEN, tests are treated in the latter way:

$$\textbf{((-)?)}\quad P_{(A)?} = \left\{ p \mid \bigwedge \begin{bmatrix} first(p) = last(p) \\ last(p) \Vdash A \end{bmatrix} \right\}$$

This definition allows tests to be executed in parallel or to be process or transition equivalent with complex processes. This allows them to be executed in parallel with other protocols. The treatment of tests in the ternary fragment gives us "basic tests":

Basic Test $P_{(A)?} = R_{(A)?} = \{\langle s, (A)?, s\rangle \mid s \Vdash A\}$

While this form of tests doesn't work as well with the considerations about parallel execution, it has the feature that the replacement theorems apply in an unrestricted fashion to them.

With the conception of tests as processes of indeterminate length, tests end up having some counterintuitive features. For one, in a test that takes one across many different states, we get what we might think of as "Gettierized" tests. It might be the case that A holds at $first(p) = last(p)$ but that it fails to hold at some or all of the intermediate states. One might want to think of tests as useful for detecting invariances:

Invariance Test

$$P_{(A)?} = \left\{ p \mid \bigwedge \begin{bmatrix} \forall s_{i \leq n} \in w(p)(w(p) = \langle s_1, ..., s_n\rangle \rightarrow A \in (v(s_i) \cap v(s_{i+1})) \\ last(p) \Vdash A \end{bmatrix} \right\}$$

Invariance tests require A to hold at the first state of the process and to carry through to all other states in the process. Now, invariance tests might have a number of features we don't want them to have. We might think that tests should change as little as possible about the world while revealing whether A holds:

Minimal Test

$$P_{m(A)?} = \left\{ p \mid \bigwedge \left[\begin{array}{l} \forall s_{i \leq n-1} \in w(p)(w(p) = \langle s_1, ..., s_n \rangle \rightarrow (v(s_i) = v(s_{i+1})) \\ v(last(p)) = v(first(p)) \cup \{A\} \end{array} \right] \right\}$$

Minimal tests are minimally mutilating to the world; the test procedure uncovers A without changing anything else in the world, but might mislead by bringing it about that A. A better minimal test:

Minimal Invariance Test

$$P_{mi(A)?} = \left\{ p \mid \bigwedge \left[\begin{array}{l} \forall s_{i \leq n} \in w(p)(w(p) = \langle s_1, ..., s_n \rangle \rightarrow (v(s_i) = v(s_{i+1})) \\ v(last(p)) = v(first(p)) \end{array} \right] \right\}$$

An alternative way to think of tests is as processes (whether basic actions or not) of introspection. Where an EDG represents the dynamics of an agent, this suggests that, rather than thinking of tests as reflexive transitions that always result in (and begin with) states where A holds, they might be restricted to those where doxastic or epistemic propositions hold:

Doxastic Test $P_{e(A)?} = \left\{ p \mid \bigwedge \left[\begin{array}{l} first(p) = last(p) \\ last(p) \Vdash b(A) \end{array} \right] \right\}$

Epistemic Test $P_{e(A)?} = \left\{ p \mid \bigwedge \left[\begin{array}{l} first(p) = last(p) \\ last(p) \Vdash K(A) \end{array} \right] \right\}$

The key differences between the two being those due to the semantics (in Chapter 4) for K; doxastic test does not require that one's beliefs end up true, where the Epistemic test does. Again, these tests might differ from intuition in the ways in which tests always do, so we might want to combine the forgoing pieces into more complex tests:

Minimal Doxastic and Epistemic Tests

(Doxastic)

$$P_{md(A)?} = \left\{ p \mid \bigwedge \left. \left[\begin{array}{l} \forall s_{i\leq n-1} \in w(p)(w(p) = \langle s_1, ..., s_n\rangle \rightarrow (v(s_i) = v(s_{i+1})) \\ c(first(p)) = [X : Y : \Pi_M] \rightarrow c(last(p)) = [X \cup \{A\} : Y : \Pi_M] \\ v(last(p)) = v(first(p)) \cup \{b(A)\} \end{array} \right] \right] \right\}$$

(Epistemic)

$$P_{me(A)?} = \left\{ p \mid \bigwedge \left. \left[\begin{array}{l} \forall s_{i\leq n-1} \in w(p)(w(p) = \langle s_1, ..., s_n\rangle \rightarrow (v(s_i) = v(s_{i+1})) \\ c(first(p)) = [X : Y : \Pi_M] \rightarrow c(last(p)) = [X \cup \{A\} : Y : \Pi_M] \\ v(last(p)) = v(first(p)) \cup \{K(A)\} \end{array} \right] \right] \right\}$$

Of course, epistemic actions - even those that are aimed at merely detecting things - are often not minimal, resulting in all sorts of changes to the world:

Non-minimal Doxastic and Epistemic Tests

(Doxastic)

$$P_{pd(A)?} = \left\{ p \mid \bigwedge \left. \left[\begin{array}{l} \forall s_{i\leq n} \in w(p)(w(p) = \langle s_1, ..., s_n\rangle \rightarrow (v(s_i, A) = v(s_{i+1}, A)) \\ c(first(p)) = [X_1 : Y_1 : \Pi_{M1}] \rightarrow c(last(p)) = [X_1 \cup \{A\} \subseteq X : Y_n : \Pi_{Mn}] \\ v(last(p)) = v(s_{n-1}) \cup \{b(A)\} \end{array} \right] \right] \right\}$$

(Epistemic)

$$P_{pe(A)?} = \left\{ p \mid \bigwedge \left. \left[\begin{array}{l} \forall s_{i\leq n} \in w(p)(w(p) = \langle s_1, ..., s_n\rangle \rightarrow (v(s_i, A) = v(s_{i+1}, A)) \\ c(first(p)) = [X_1 : Y_1 : \Pi_{M1}] \rightarrow c(last(p)) = [X_1 \cup \{A\} \subseteq X : Y_n : \Pi_{Mn}] \\ v(last(p)) = v(s_{n-1}) \cup \{K(A)\} \end{array} \right] \right] \right\}$$

All of the foregoing tests are invariance tests. The minimal epistemic test is mutilating to the world; the test procedure uncovers invariant A and pushes it into the acceptance set without changing anything else in the world. The non-minimal epistemic test is not minimally mutilating to the world; the test procedure uncovers invariant A and pushes it into acceptance set while allowing changes in the world. The addition of these forms of test into PLEN might increase the utility for representing epistemic instructions that require introspection or awareness.

6.1.3 Weighted Paths

In chapter 2, restrictions on epistemic dynamics were discussed in a somewhat limited way. I introduced a construction that plausibly (or so I argued) modeled the categorical and deontic content of norms as well as a kind of ranking or ordering on components of epistemic dynamics. In Chapter 4, it was shown that protocols, in virtue of the execution rules, can all be mapped to restrictions on epistemic dynamics, and that each restriction on epistemic dynamics defined by any norm-kernel (whose elements are representable in PLEN) correlates to those defined by some protocol. But this correlation was defined in terms of agreement between the rational and irrational subsets and the thresholds of irrationality. PLEN couldn't represent the actual ordering of the elements of dynamics in the restrictions.

One way to overcome this limitation is by modifying the execution rules so that they define not only sets of executions but ordering relations on all paths in each graph. For each protocol π, define a function, $Pref_{\pi}$, that, for each graph, g, arbitrarily maps each path in g to a value in $[0,1]$. Then, letting $t \in [0,1]$, for the regular operators, define the execution sets of π, P^t_{π}, for each threshold, t:

$$(;)\ \ P^t_{\pi 1;\pi 2} = \{p \mid p = qr \wedge q \in P_{\pi 1} \wedge r \in P_{\pi 2} \wedge Pref_{\pi 1;\pi 2}(p) > t\}$$

$$(*)\ \ P^t_{\pi *} = \left\{ p \mid \bigwedge \left[\begin{array}{l} \exists n \geq 1 \\ \bigwedge \left[\begin{array}{l} \exists \langle x_1 ... x_n \rangle \\ \bigwedge \left[\begin{array}{l} \forall_{i \leq n}(x_i \in W), first(p) = x_1, last(p) = x_n, \\ \forall p', \forall_{i \leq n-1}[(x_i = first(p'), x_{i+1} = last(p')) \rightarrow p' \in P_{\pi}] \end{array} \right] \end{array} \right] \\ Pref_{\pi *}(p) > t \end{array} \right] \right\}$$

$$(\cap)\quad P^{t}{}_{\pi1\cap\pi2} = \left\{ p \mid \bigwedge \begin{bmatrix} p \in P_{\pi1} \vee p \in P_{\pi2}, \\ p \in P_{\pi1} \rightarrow \bigwedge \begin{bmatrix} \exists q \in P_{\pi2}, \\ p \rightleftharpoons q \end{bmatrix}, \\ p \in P_{\pi2} \rightarrow \bigwedge \begin{bmatrix} \exists q \in P_{\pi1}, \\ p \rightleftharpoons q \end{bmatrix} \\ Pref_{\pi1\cap\pi2}(p) > t \end{bmatrix} \right\}$$

$$(\cup)\quad P^{t}_{\pi1\cup\pi2} = \{p \mid (p \in P_{\pi1} \vee p \in P_{\pi2}) \wedge Pref_{\pi1\cup\pi2}(p) > t\}$$

I conjecture that this will enable us to map every restriction on epistemic dynamics defined over the set of EDGs to some protocol in P_{PLEN} in such a way that not only corresponds to the rational and irrational subsets of elements, but also to the orderings they define over paths. Adding preference functions on the basis of states and transitions should be straightforward. Perhaps a function $Pref_{\pi}$ for transitions would work by requiring that: $Pref_{\pi}(t) > Pref_{\pi}(t')$ iff $t = \langle first(p), last(p)\rangle$, and $t' = \langle first(p'), last(p')\rangle$, and $Pref_{\pi}(p) > Pref_{\pi}(p')$ would work.

Of course, there are now infinitely many sets of executions of each protocol, and this will make proving anything meaningful about the formalism more difficult as a result. There are infinitely many sets of executions of each protocol - one for each $t \in [0, 1]$. Protocol graphs [135], for instance, seem apt to become massively more complex. For instance, if *every* subprotocol and implementation of π must be executed in a protocol graph of π, there will be infinitely many executions of each subprotocol and implementation.

6.2 PLEN and Nonclassical Logics

There are two very interesting ways of combining PLEN with non-classical logics. The first - and more straightforward way - is to make the underlying logic of PLEN non-classical. PLEN makes much room for non-classical modifications. After all, the truth functions in PLEN were interpreted in strictly classical ways; one could always (for instance) replace the conditional of PLEN by giving the PLEN conditional a different truth table or by replacing it with or adding an intensional conditional or a non-classical negation. It would be interesting to see what follows from a non-classical PLEN.

The second way is actually more interesting. It would be worthwhile to explore PLEN as theoretical basis for analyzing logical consequence relations. One way of doing this is to think about the dynamic consequence relations analyzable via PDL: test-to-test consequence, update-to-test consequence, update-to-update consequence, and so on.[1] These notions have already been deployed to analyze logical consequence [39, 258]. Another way of doing this would be to try to build a relevant logic with PLEN as its semantics, following [208] or [243]. The basic idea being that some set of constraints on epistemic dynamics might provide a natural basis for the semantics of relevant logic in much the same way that information dynamics [208] or function application do [201]. This is also similar to Gardenfors' basing logical consequence on rational belief dynamics [91].

6.2.1 Alternative Epistemic Logics in PLEN

PLEN built its representation of epistemic dynamics on a syntactic epistemic logic [148] that was inspired by Restall's "states" [210] and Jago's ELRA [134]. This enabled PLEN to simply bypass any of the usual paradoxes or counterintuitive features of other epistemic logics, i.e., logical omniscience. This made PLEN suitable to representing epistemic dynamics at a lower level of idealization. But this also left the properties of knowledge in PLEN relatively incomplete and unsettled.

Probability and Credence in PLEN First, credence had no place in PLEN. This can remedied easily enough by by amending content states with credence functions. Respecting the threshold view of partial belief, let $Cr(A) \in [0,1]$ such that, for some threshold t, $A \in X$ iff $Cr(A) > t$ and $A \in Y$ iff $Cr(A) < t$. Then throw credence statements into the language

---

[1] See Chapter 7 of [30] for an outline of these "dynamic styles of inference".

of L_{PLEN} and protocols can be characterized in terms of their effects on credences. This seems to be a way of injecting subjective probabilities into PLEN; I didn't specify that Cr had to obey any probability axioms, but it would be simple enough to require it at either the meta-level or to characterize protocols that satisfy this or that set of axioms.

Epistemic Models Second, we could follow the central thought of research on DEL and replace the states in EDGs with epistemic models, thereby moving from a syntactic representation of epistemic states to an epistemic models based representation. From there, we can identify those pairs of states that essentially result in the same changes of transition relation that would result from a product update, and associate them with the atomic protocols to which they're mapped. This gives us a way of embedding DEL into PLEN.

Norm Revision plus Conceptual Innovation Finally, we could extend PLEN to not only represent changes of epistemic state and changes of methodology, but even more radical epistemic changes like conceptual innovation. One can combine Multiple Language Architectures (MLAs) and dynamic logics to create Descriptive Dynamic Logics [231], in which protocols connect "units" over which distinct languages with distinct logics are interpreted.[2] Each unit is itself a graph, and the protocols pair states in one graph with states in another.

To treat this in PLEN, think of hyper-EDGs, in which the nodes are themselves EDGs. Then, after adding a meta-protocol language to PLEN, one could assign these new protocols transitions (or even paths) among EDGs within hyper-EDGs. This could represent transitioning to an entirely new way of thinking. The language of one unit could be strictly truth-functional while another has intensional connectives.

What would this technical modification represent? Well, to take one example, suppose that in one unit, the logical language contains modals and in another unit, the language doesn't. Then there are contents for doxastic states in one unit with contents inexpressible in the other. The expression $b(\Box A)$ is in one language and not in the other. The transition

---

[2] See also graph modifier logics [15].

from the modal-free unit to the modal unit looks like acquiring the concept of modality. Alternately, in one unit, there may be no protocol expressions, and in another there may be all of the protocol expressions of L_{PLEN}. The transition from the first to the second unit represents the acquisition of protocol-theoretic concepts.

With that, I end my brief excursion into additions and extensions of PLEN. And with *that*, I end my initial articulation, defense, and case for the continued development of the protocol-theoretic account of epistemic norms. Hopefully, it has added something to the project of doing better epistemically.

Bibliography

[1] Adler, Jonathan E. 2009. Another argument for the knowledge Norm. *Analysis*. 69 (3):407-411.

[2] Adler, Jonathan E. & Rips, Lance J. (eds.). 2008. *Reasoning: Studies of Human Inference and its Foundations*. Cambridge University Press.

[3] Alchourron, Carlos E. 1991. Conflicts of norms and the revision of normative systems. *Law and Philosophy*. 10 (4):413 - 425.

[4] Alchourron, Carlos & Bulygin, Eugenio. 1971. Normative Systems. *CrÃtica: Revista Hispanoamericana de FilosofÃa*. 5 (14):120-123.

[5] Alchourron, C., Gardenfors, P., and Makinson, D. 1985. On the Logic of Theory Change: Partial Meet Contraction and Revision Functions. *The Journal of Symbolic Logic*. Vol. 50, No. 2. pp. 510-530

[6] Alechina, Natasha; Jago, Mark & Logan, Brian. 2008. Preference-based belief revision for rule-based agents. *Synthese*. 165 (2):159-177.

[7] Alston, William (1989). _Epistemic Justification_. Cornell University Press.

[8] Anderson, A.R. and Nuel Belnap. 1975. *Entailment: The Logic of Relevance and Necessity*. Princeton: Princeton University Press, Volume I.

[9] Angell, Richard B. 1964. *Reasoning and Logic*. New York: Appleton-Century-Crofts.

[10] Anglberger, Albert J. J. 2008. Dynamic Deontic Logic and its Paradoxes. *Studia Logica*. 89 (3):427-435.

[11] Arieli, O., Avron, A. 2000. General patterns for nonmonotonic reasoning: from basic entailments to plausible relations. *Logic Journal of the IGPL*. 8 (2):119-148.

[12] Aris, R. 2012. *Mathematical Modeling Techniques*. Dover: Dover Books.

[13] Aristotle. 2012. *Organon: The Works of Aristotle on Logic*. Bishop, R.B., (ed.). CreateSpace Independent Publishing Platform.

[14] Aucher, Guillaume. 2012. DEL-sequents for regression and epistemic planning. *Journal of Applied Non-Classical Logics*. 22 (4):337 - 367.

[15] Guillaume Aucher, Philippe Balbiani, Luis Farinas del Cerro, Andreas Herzig, Global and Local Graph Modifiers. *Electronic Notes in Theoretical Computer Science*, Volume 231, 2009, Pages 293-307, ISSN 1571-0661, https://doi.org/10.1016/j.entcs.2009.02.042.

[16] Bacon Francis. 1851. *The Advancement of Learning*. Modern Library.

[17] ——. 1960. *The New Organon, and Related Writings*. New York: Liberal Arts Press.

[18] Baltag, A., & Moss, L.S. 2004. Logics for Epistemic Programs. *Synthese*. 139, 165-224.

[19] Baltag, Alexandru and Renne, Bryan,. Dynamic Epistemic Logic. The Stanford Encyclopedia of Philosophy (Winter 2016 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/win2016/entries/dynamic-epistemic/>.

[20] Baltag, Alexandru & Smets, Sonja. 2008. Probabilistic dynamic belief revision. *Synthese*. 165 (2):179 - 202.

[21] Baumgartner, Michael & Lampert, Timm. 2008. Adequate formalization. *Synthese*. 164 (1):93-115.

[22] Beall, J. 2011. Multiple-conclusion LP and Default Classicality. *Review of Symbolic Logic.* 4(2): 326–336.

[23] ——. 2013. “Free of Detachment: Logic, Rationality, and Gluts.” *NoÃ»s* 47 (3): n/a-n/a.

[24] Beall, Jc & Restall, Greg. 2005. *Logical Pluralism.* Oxford University Press.

[25] Beall, Jc and Restall, Greg. 2013. Logical Consequence. *The Stanford Encyclopedia of Philosophy* (Winter 2013 Edition), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/win2013/entries/logical-consequence/>.

[26] Beauchamp, Tom L, & Childress, James. 1994. P*rinciples of Biomedical Ethics.* Oxford University Press.

[27] Belnap, N., Perloff, M., and Xu, M. 2001. *Facing the future.* Oxford: Oxford University Press.

[28] van Benthem, Johan. 1990. General Dynamics. *ITLI Prepublication Series for Logic, Semantics, and Philosophy of Language.* University of Amsterdam.

[29] ——. 1994. Programming operations that are safe for bisimulation. *Studia Logica.* 60 (2):311-330.

[30] ——. 1996. *Exploring Logical Dynamics.* Stanford: CSLI Publications.

[31] ——. 1998. Program constructions that are safe for bisimulation. *Studia Logica.* 60 (2):311-330.

[32] ——. 2007. Dynamic logic for belief revision. *Journal of Applied Non-Classical Logics.* 17 (2):129-155.

[33] ——. 2010. *Modal Logic for Open Minds.* Center for the Study of Language and Information.

[34] ——. 2014. Logical Dynamics of Information and Interaction. Cambridge: Cambridge University Press.

[35] van Benthem, J. and Bergstra, J. 1995. Logic of Transition Systems. *Journal of Logic, Language, and Information.* Vol. 3, No. 4 (1995), pp. 247-283

[36] van Benthem, Johan; van Eijck, Jan & Stebletsova, Vera. 1994. Modal Logic, Transition Systems and Processes. *Journal of Logic and Computation.* 4. 10.1093/logcom/4.5.811.

[37] van Benthem Johan; Gerbrandy Jelle; Hoshi Tomohiro, & Pacuit Eric, 2009. Merging frameworks for interaction. *Journal of Philosophical Logic.* 38 (5):491-526.

[38] Berger, Alan. 2011. Kripke On the Incoherency of Adopting a Logic. *Saul Kripke.* Ed. Alan Berger. New York; Cambridge University Press. 177-207.

[39] Bledin, Justin. 2014. Logic Informed. *Mind.* 123: 277–316.

[40] Boghossian, Paul 2008. Epistemic Rules. *Journal of Philosophy.* 105 (9):472-500.

[41] Bolander, Thomas & Andersen, Mikkel Birkegaard. 2011. Epistemic planning for single- and multi-agent systems. *Journal of Applied Non-Classical Logics.* 21 (1):9-34.

[42] Boyle, Deborah. 2012. The Ways of the Wise: Hume's Rules of Causal Reasoning. *Hume Studies.* 38 (2):157-182.

[43] Bratman, Michael. 1998. *Intention, Plans, and Practical Reason.* Center for the Study of Language and Information.

[44] Brink, David O. 1994. Moral conflict and its structure. *Philosophical Review.* 103 (2):215-247.

[45] Brown, B. and Priest, G. 2004. Chunk and Permeate, a Paraconsistent Inference Strategy. Part I: The Infinitesimal Calculus. *Journal of Philosophical Logic.* 33 (4):379-388.

[46] Broome, J. 2000. Normative requirements. In J. Dancy (Ed.). *Normativity.* Oxford: Blackwell.

[47] Caret, Colin R. 2017. The Collapse of Logical Pluralism has been Greatly Exaggerated. *Erkenntnis.* 82 (4):739-760.

[48] Carroll, Lewis. 1895. What the Tortoise Said to Achilles. *Mind.* Volume 4, Issue 14, 278-280.

[49] Cevolani, Gustavo & Schurz, Gerhard. 2017. Probability, Approximate Truth, and Truthlikeness: More Ways out of the Preface Paradox. *Australasian Journal of Philosophy.* 95 (2):209-225.

[50] Cherniak, Christopher. 1986. *Minimal Rationality.* MIT Press.

[51] Choi, Sungho and Fara, Michael. 2016. Dispositions. *The Stanford Encyclopedia of Philosophy* (Spring 2016 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/spr2016/entries/dispositions/>.

[52] Chopra, S. and Parikh, R. 2000. Relevance Sensitive Belief Structures. *Annals of Mathematics and Artificial Intelligence.* pp. 259-285.

[53] Christensen, David. 2004. *Putting Logic in its Place.* Oxford: Oxford University Press.

[54] Clifford, William Kingdon 1999. *The Ethics of Belief and Other Essays.* Prometheus Books

[55] Cohen, L. J. 1992. *An Essay on Belief and Acceptance.* Oxford: Oxford University Press

[56] Colyvan, Mark 2006. Naturalism and the paradox of revisability. *Pacific Philosophical Quarterly.* 87 (1):1–11.

[57] Conee, Earl Brink 2004. *Evidentialism: Essays in Epistemology.* Oxford University Press.

[58] Conee E., & Feldman R.. 1998. The generality problem for reliabilism. *Philosophical Studies.* 89 (1):1-29.

[59] Conte, R., Andrighetto, G., & Campennl, M.(Eds.). 2013. *Minding Norms: Mechanisms and dynamics of social order in agent societies.* : Oxford University Press.

[60] Copeland, B. J. 1979. On When a Semantics is not a Semantics: Some Reasons for Disliking the Routley-Meyer Semantics for Relevance Logic. *Journal of Philosophical Logic.* 8: 399-413.

[61] Copi, Irving M. 1953. *Introduction to Logic.* Pearson/Prentice Hall.

[62] Curry, H. 1942. The Inconsistency of Certain Formal Logics. *Journal of Symbolic Logic.* 7: 115–7.

[63] Demey, Lorenz. 2016. The Dynamic Turn in Epistemic Logic. *Tijdschrift Voor Filosofie.* 78 (2):365-391.

[64] Descartes, RenÃ©. 1964. *Philosophical Essays: Discourse on Method; Meditations; Rules for the Direction of the Mind.* Indianapolis: Bobbs-Merrill.

[65] van Ditmarsch, H., Halpern, J., Van Der Hoek, W., and Kooi, B.. 2007. *Dynamic Epistemic Logic.* Springer.

[66] van Ditmarsch, H., Halpern, J., Van Der Hoek, W., and Kooi, B. 2015. *Handbook of Epistemic Logic.* College Publications

[67] Doyle, J. 1987. “A Truth-Maintenance System.” *Readings in Nonmonotonic Reasoning.* Ed. Matthew L. Ginsberg. San Francisco: Morgan Kaufmann. 259-279.

[68] van Eijck J., Wang Y. 2008. Propositional Dynamic Logic as a Logic of Belief Revision. In: Hodges W., de Queiroz R. (eds) *Logic, Language, Information and Computation.* WoLLIC 2008. Lecture Notes in Computer Science, vol 5110. Springer, Berlin, Heidelberg

[69] Ellis, Bryan. 1979. *Rational Belief Systems.* London: Blackwell.

[70] Etchemendy, John 1990. *The Concept of Logical Consequence.* Harvard University Press.

[71] Evans, J. St. B. T., Newstead, S., Allen, J., & Pollard, P. 1994. Debiasing by instruction: The case of belief bias. *European Journal of Cognitive Psychology.* 6, 263–285.

[72] Evans, J. St. B. T., & Over, D. E. (1996). Rationality and reasoning. Hove, England: Psychology Press.

[73] Evans, Jonathan & Stanovich, Keith E. (2013). Dual-Process Theories of Higher Cognition Advancing the Debate. _Perspectives on Psychological Science_ 8 (3):223-241.

[74] Fagin, Ronald, Halpern, Joseph, Moses, Yoram, and Vardi, Moshe. 2004. *Reasoning About Knowledge.* MIT Press.

[75] Field, Hartry. 2001. *Truth and the Absence of Fact.* Oxford: Oxford University.

[76] ——. 2008. *Saving Truth from Paradox.* Oxford: Oxford University Press.

[77] ——. 2009. Epistemology Without Metaphysics. *Philosophical Studies.* Volume 143, Issue 2, pp 249-290.

[78] ——. 2009. Pluralism in Logic. *Review of Symbolic Logic.* Vol. 2 Issue 2, p342-359.

[79] ——.2009. What is the Normative Role of Logic?. *Aristotelian Society Supplementary Volume.* Volume 83, Issue 1, pages 251–268.

[80] ——. 2015. What Is Logical Validity?. In Colin Caret & Ole Hjortland (eds.), *Foundations of Logical Consequence*. Oxford: Oxford University Press.

[81] Fink, Julian. 2011. Are there process-requirements of rationality? *Organon F: Medz-inÃ¡rodnÃœ Časopis Pre AnalytickÃº Filozofiu.* 18 (4):475-88.

[82] Fitting, Melvin ; Mendelsohn, R. & Girle, Roderic A. (2002). First-order modal logic. _Bulletin of Symbolic Logic_ 8 (3):429-430.

[83] Foley, R. 1993. *Working Without a Net.* Oxford University Press.

[84] Forrest, P. 1986. *The Dynamics of Belief.* London: Blackwell.

[85] Forster, M. and Sober, E. 1994. How to Tell when Simpler, More Unified, or Less Ad Hoc Theories will Provide More Accurate Predictions. *British Journal for the Philosophy of Science*. 45, pp. 1-35.

[86] van Fraassen, B. 1980. Rational Belief and Probability Kinematics. *Philosophy of Science*. Vol. 47, No. 2. pp. 165-187

[87] Frigg, Roman and Hartmann, Stephan. 2017. Models in Science. *The Stanford Encyclopedia of Philosophy* (Spring 2017 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/spr2017/entries/models-science/>.

[88] Gabbay, Dov M. (ed.). 1994. *What is a Logical System?*. Oxford University Press.

[89] Gabby, D., Horty, J., and Parent, X. Eds. 2013. Handbook of Deontic Logic and Normative Systems. College Publications

[90] Gabbay, Dov; Rodrigues, Odinaldo & Russo, Alessandra. 2008. Belief revision in non-classical logics. _Review of Symbolic Logic_ 1 (3):267-304.

[91] Gardenfors, Peter. 2008. *Knowledge in Flux.* Reprint. Original: 1988. London: College Publications.

[92] ——. 1992. *Belief revision.* Cambridge: Cambridge University Press.

[93] Garnier, R. and Taylor, J. 2010. *Discrete Mathematics: Proofs, Structures, and Applications.* CRC Press.

[94] Geach. P. 1965. Assertion. *Philosophical Review.* 74: 449–465.

[95] Georgeff, M. Ed. 1987. *Reasoning About Actions and Plans.* Kaufmann.

[96] Gerken, Mikkel. 2014. Same, same but different: the epistemic norms of assertion, action and practical reasoning. *Philosophical Studies.* 168 (3):725-744.

[97] Gibbard, Allen. 1990. *Wise Choices, Apt Feelings.* Cambridge: Harvard University.

[98] ——. 2003. *Thinking How to Live.* Harvard University Press.

[99] Gierasimczuk, Nina. 2009. Bridging learning theory and dynamic epistemic logic. *Synthese.* 169 (2):371-384.

[100] Gigerenzer, G. 2007. *Gut feelings: The intelligence of the unconscious.* London, England: Penguin.

[101] ——. 2011. Personal reflections on theory and psychology. *Theory & Psychology.* 20, 733–743.

[102] Gigerenzer, G., & Regier, T. 1996. How do we tell an association from a rule? Comment on Sloman (1996). *Psychological Bulletin.* 119, 23–26.

[103] Glymour, Clark. 1980. *Theory and Evidence.* Princeton: Princeton University Press.

[104] Goldman, Alvin. 1979. What is Justified Belief?. In George Pappas (ed.). *Justification and Knowledge.* Boston: D. Reidel. pp. 1-25.

[105] ——. (1986). *Epistemology and Cognition.* Harvard University Press.

[106] Goodman, Nelson. 1983. *Fact, Fiction, and Forecast.* Harvard University Press.

[107] Gossett, E. 2009. *Discrete Mathematics with Proof* (2nd ed.). Wiley Publishing.

[108] Grove, Adam. 1988. Two Modellings for Theory Change. *Journal of Philosophical Logic*. 17 (2):157-170.

[109] Guttenplan, Samuel D. (1997). *The Languages of Logic: An Introduction to Formal Logic*. Blackwell.

[110] ——. 1996. *Deviant Logic, Fuzzy Logic: Beyond the Formalism.* University of Chicago Press.

[111] Haack, Susan. 1978. *Philosophy of Logics*. Cambridge University Press.

[112] Hansson, Sven Ove. 2000. Formalization in philosophy. *Bulletin of Symbolic Logic*. 6 (2):162-175.

[113] Harel, David, Kozen, Dexter, and Tiuryn, Jerzy. 2000. *Dynamic Logic*. MIT Press.

[114] David Harel, Dexter Kozen, Rohit Parikh, Process logic: Expressiveness, decidability, completeness. *Journal of Computer and System Sciences*. Volume 25, Issue 2, 1982, Pages 144-170

[115] Haddawy, P.F. 1994. *Representing Plans Under Uncertainty: a Logic of Time, Chance, and Action*. Springer-Verlag.

[116] Hakli, Raul & Negri, Sara. 2012. Does the deduction theorem fail for modal logic? *Synthese*. 187 (3):849-867.

[117] HalldÃ©n, SÃ¶ren. 1966. Preference logic and theory choice. *Synthese*. 16 (3-4):307 - 320.

[118] Hansson, S.O.. 2017. Logic of Belief Revision. *The Stanford Encyclopedia of Philosophy* (Winter 2017 Edition), Edward N. Zalta (ed.), forthcoming URL = <https://plato.stanford.edu/archives/win2017/entries/logic-belief-revision/>.

[119] Harman, Gilbert. 1983. "Logic and Probability theory Versus Canons of Rationality." *Behavioral and Brain Sciences* 6 (2):251.

[120] ——. (1973). *Thought.* Princeton University Press.

[121] ——. 1986. *Change in View.* Cambridge; MIT Press.

[122] ——. 1995. Rationality. In E. E. Smith & D. N. Osherson (eds.). *Invitation to Cognitive Science.* Mit Press.

[123] Harman, G. and S. Kulkarni. 2007. *Reliable Reasoning: Induction and Statistical Learning Theory.* Cambridge: MIT Press.

[124] Hartonas, Chrysafis. 2014. On the Dynamic Logic of Agency and Action. *Studia Logica.* 102 (3):441-478.

[125] Hatfield, Gary. 2009. Perception as Unconscious Inference. *Perception and Cognition: Essays in the Philosophy of Psychology*, Oxford: Oxford University Press, pp. 124–152.

[126] Head, M. L., Holman, L., Lanfear, R., Kahn, A. T., & Jennions, M. D. 2015. The Extent and Consequences of P-Hacking in Science. *PLoS Biology.* 13(3), e1002106. http://doi.org/10.1371/journal.pbio.1002106

[127] Hedden, Brian (2015). Time-Slice Rationality. *Mind.* 124 (494):449-491.

[128] Hindriks, Koen V. and Meyer, John-Jules Ch. 2007. Agent Logics as Program Logics: Grounding KARO. *Advances in Artificial Intelligence: 29th Annual German Conference on AI.* Freksa, Christian and Kohlhase, Michael and Schill, Kerstin. Eds. Bremen, Germany, Springer Berlin Heidelberg: Berlin, Heidelberg. pp. 404–418.

[129] Hintikka, Jaakko. 1962. *Knowledge and Belief.* Ithaca: Cornell University Press.

[130] Horty, J. 2001. *Agency and Deontic Logic.* Oxford: Oxford University Press.

[131] Hoshi, T. 2009. *Epistemic Dynamics and Protocol Information.* ILLC Dissertation Series

[132] Hume, David. 1998. *An Enquiry Concerning Human Understanding: A Critical Edition.* Oxford University Press.

[133] Huber, Franz. 2007. The Logic of Theory Assessment. *Journal of Philosophical Logic.* 36 (5):511-538.

[134] Jago, Mark. 2009. Epistemic logic for rule-based agents. *Journal of Logic, Language and Information.* 18 (1):131-158.

[135] Jenkins, Ralph. 2018. The Technical Foundations of PLEN: A Protocol-theoretic Logic of Epistemic Norms. Unpublished technical document.

[136] Kahneman, Daniel; Slovic, Paul & Tversky, Amos (eds.). 1982. *Judgment Under Uncertainty: Heuristics and Biases.* Cambridge University Press.

[137] Kalish, Donald; Montague, Richard & Mar, Gary. 1964. *Logic: Techniques of Formal Reasoning.* Oxford University Press USA.

[138] Katz, Jerrold J. 2000. *Realistic Rationalism.* A Bradford Book.

[139] Kozen, D., Parikh, R. 1981. An elementary proof of the completeness of PDL. *Theoretical Computer Science.* Volume 14, Issue 1, 1981, Pages 113-118

[140] Kelly, Kevin T. 1988. Formal Learning Theory and the Philosophy of Science. *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association.* 1988: 413 - 423.

[141] ——. 1999. Iterated belief revision, reliability, and inductive amnesia. *Erkenntnis.* 50 (1):11-58.

[142] Kelly, Kevin T.; Schulte, Oliver & Hendricks, Vincent. 1995. Reliable Belief Revision. *Logic and Scientific Methods: Volume One of the Tenth International Congress of Logic, Methodology and Philosophy of Science, Florence*. pp 383-398

[143] Kirsh, David & Maglio, P. 1992. Some Epistemic Benefits of Action-Tetris, a Case Study. *Proceedings of the 14th Annual Conference of the Cognitive Science Society.*

[144] Kirsh, David & Maglio, Paul. 1994. On Distinguishing Epistemic from Pragmatic Action. *Cognitive Science*. 18 (4):513-549.

[145] Kleene, Stephen Cole. 1967. *Mathematical Logic*. Dover Publications.

[146] Kolodny, Niko. 2005. Why be rational? *Mind*. 114. (455):509-563.

[147] Kolodny, Niko. 2007. State or process requirements? *Mind*. 116 (462):371-385.

[148] Konolige, Kurt. 1986. *A Deduction Model of Belief*. Morgan Kaufmann Publishers.

[149] Kooi, Barteld P. 2003. Probabilistic dynamic epistemic logic. *Journal of Logic, Language and Information*. 12 (4):381-408.

[150] Korcz, K. 1997. Recent Work on the Basing Relation. *American Philosophical Quarterly*. 34 (2): 171–191.

[151] Korcz, Keith Allen. 2015. The Epistemic Basing Relation. *The Stanford Encyclopedia of Philosophy* (Fall 2015 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/fall2015/entries/basing-epistemic/>.

[152] Kruglanski, A. W., & Gigerenzer, G. 2011. Intuitive and deliberative judgements are based on common principles. *Psychological Review*. 118, 97–109.

[153] Kyburg, Henry. 1983. *Epistemology and Inference*. Univ of Minnesota Press.

[154] Lam, Barry. Unpublished. The Dynamic Foundations of Epistemic Rationality. Ph.D Dissertation: Vassar College.

[155] Leitgeb, Hannes. 2017. *The Stability of Belief: How Rational Belief Coheres with Probability.* Oxford University Press.

[156] Leitgeb, Hannes & Pettigrew, Richard. 2010. An Objective Justification of Bayesianism I: Measuring Inaccuracy. *Philosophy of Science.* 77 (2):201-235.

[157] Leitgeb, Hannes & Pettigrew, Richard. 2010. An Objective Justification of Bayesianism II: The Consequences of Minimizing Inaccuracy. *Philosophy of Science* 77 (2):236-272.

[158] Leitgeb, H. and Segerberg, K. 2007. Dynamic doxastic logic: why, how, and where to?. *Synthese.* 155 (2):167-190.

[159] Lepore, Ernest. 2000. *Meaning and Argument: An Introduction to Logic Through Language.* Wiley-Blackwell.

[160] Levi, Isaac. 1996. *For the Sake of the Argument: Ramsey Test Conditionals, Inductive Inference, and Nonmonotonic Reasoning.* Cambridge University Press.

[161] ——. 2004. *Mild Contraction: Evaluating Loss of Information Due to Loss of Belief.* Oxford University Press.

[162] LindstrÃ¶m, Sten & Rabinowicz, Wlodek. 1999. DDL unlimited: Dynamic doxastic logic for introspective agents. *Erkenntnis.* 50 (2-3):353-385.

[163] Lipton, Peter. 2007. Alien abduction: Inference to the best explanation and the management of testimony. *Episteme.* 4 (3):238-251.

[164] Littlejohn, C. and John Turri. *Epistemic Norms: New Essays on Action, Belief, and Assertion.* Oxford: Oxford University Press.

[165] Littlewood, J.E. 1953. *A mathematician's miscellany.* London: Meuthen.

[166] Lomuscio, A. and Sergot, M. 2003. Deontic Interpreted Systems.

[167] MacFarlane, John. 2004. In What Sense (If Any) Is Logic Normative for Thought? Unpublished. (Presented at the 2004 Central Division APA symposium on the normativity of logic.) Accessed online: http://johnmacfarlane.net/normativity_of_logic.pdf

[168] Macpherson, R., & Stanovich, K. E. 2007. Cognitive ability, thinking dispositions, and instructional set as predictors of critical thinking. *Learning and Individual Differences*. 17, 115-127. LAID07.pdf

[169] Maglio, Paul P., Wenger, Michael J., Copeland, Angelina M. 2003. The Benefits of Epistemic Action Outweigh the Costs. *Proceedings of the Annual Meeting of the Cognitive Science Society*. Volume 25 Permalink: https://escholarship.org/uc/item/5bg5p6bv

[170] Makinson, David C. 1965. The Paradox of the Preface. *Analysis*. 25 (6):205-207.

[171] Mandelbaum, Eric & Levy, Neil. 2014. The Powers that bind : doxastic voluntarism and epistemic obligation. In Jonathan Matheson (ed.). *The Ethics of Belief*. New York: Oxford University Press. pp. 12-33.

[172] McGee, Vann. 1985. A counterexample to modus ponens. *Journal of Philosophy*. 82 (9):462-471.

[173] Mill, John Stuart. 1843. *A System of Logic, Ratiocinative and Inductive Being a Connected View of the Principles of Evidence and the Methods of Scientific Investigation*. Longmans, Green, Reader, and Dyer.

[174] Meyer, J. -J.. 1987. A different approach to deontic logic: deontic logic viewed as a variant of dynamic logic. *Notre Dame Journal of Formal Logic*. 29 (1):109-136.

[175] ——. 2000. Dynamic Logic for Reasoning About Actions and Agents" *Logic-based Artificial Intelligence.* Jack Minker (Ed.). Kluwer International Series In Engineering And Computer Science, Vol. 597. Kluwer Academic Publishers, Norwell, MA, USA 281-311.

[176] Meyer, Robert K. 1971. Entailment. *Journal of Philosophy*. 68 (21):808-818.

[177] Montminy, Martin. 2013. Why Assertion and Practical Reasoning Must be Governed By the Same Epistemic Norm. *Pacific Philosophical Quarterly.* 94 (1):57-68.

[178] Morgan, Mary. 1999. Learning from Models. Morgan, Mary S. and Morrison, Margaret, (eds.) *Models as Mediators: Perspectives on Natural and Social Science. Ideas in context.* Cambridge University Press, Cambridge, UK, pp. 347-388

[179] Moss, Sarah. 2015. Time-Slice Epistemology and Action Under Indeterminacy. In Tamar Szabo Gendler & John Hawthorne (eds.), *Oxford Studies in Epistemology.* Oxford University Press. pp. 172–94.

[180] Moutier, S., Angeard, N., & Houde, O. 2002. Deductive reasoning and matching-bias inhibition training: Evidence from a debiasing paradigm. *Thinking and Reasoning.* 8, 205–224.

[181] Neth, H., & MÃŒller, T. 2008. Thinking by doing and doing by thinking: a taxonomy of actions. In V. Sloutsky, B. Love & K. McRae (Eds.). *30th Annual Conference of the Cognitive Science Society* (pp. 993-998). Austin, TX: Cognitive Science Society.

[182] Oaksford, M., and Chater, N. 2007. *Bayesian Rationality: The Probabilistic Approach to Human Reasoning.* Oxford University Press.

[183] Osherson, Daniel. Unpublished. Formal learning theory in context. Manuscript.

[184] Pacuit, Eric, and Simon, Sunil. 2010. Reasoning With Protocols Under Imperfect Information. *Review of Symbolic Logic.* 4 (3):412-444.

[185] Papineau, David. 1987. *Reality and Representation.* Blackwell.

[186] ——. 2013. There Are No Norms of Belief. In T. Chan (ed.). *The Aim of Belief.* Oxford University Press.

[187] Penther, Brigitte. 1994. A dynamic logic of action. *Journal of Logic, Language and Information.* 3 (3):169-210.

[188] Pettigrew, Richard. Forthcoming. Epistemic Utility Theory and the Normativity of Logic. Unpublished.

[189] ——. 2011. An Improper Introduction to Epistemic utility theory. In Henk de Regt, Samir Okasha & Stephan Hartmann (eds.), *Proceedings of EPSA: Amsterdam '09*. Springer. pp. 287–301.

[190] ——. 2013. Epistemic Utility and Norms for Credences. *Philosophy Compass*. 8 (10):897-908.

[191] ——. 2013. A New Epistemic Utility Argument for the Principal Principle. *Episteme*. 10 (1):19-35.

[192] ——. 2016. Epistemic Utility Arguments for Probabilism. *The Stanford Encyclopedia of Philosophy* (Spring 2016 Edition). Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/spr2016/entries/epistemic-utility/>.

[193] Plaza, J. 2007. Logics of Public Communications. *Synthese*. 158 (2):165 - 179

[194] Pollock, Jon L. 1987. Epistemic Norms. *Synthese*. Vol. 71, No. 1. pp.61 - 95.

[195] ——. 1995. *Cognitive Carpentry*. MIT Press.

[196] ——. 1998. Procedural epistemology. In Terrell Ward Bynum & James Moor (eds.), *The Digital Phoenix: How Computers Are Changing Philosophy*. Blackwell. pp. 17.

[197] ——. 2009. A recursive semantics for defeasible reasoning. *Argumentation in Artificial Intelligence.* : Simari G., Rahwan I. (eds). Springer, Boston, MA. pp 173-197

[198] Pratt, Vaughan. R.. 1980. Application of Modal Logic to Programming. *Studia Logica*. 34, 2/3, 257-274, 1980.

[199] Priest, Graham. 2006. *Doubt Truth to Be a Liar*. Oxford: Oxford University.

[200] ——. 2008. *An Introduction to Non-Classical Logics*. Cambridge: Cambridge University.

[201] ——. 2015. ”Is the Ternary R Depraved?” In Colin Caret & Ole Hjortland (eds.). *Foundations of Logical Consequence*. Oxford: Oxford University Press.

[202] ——. 2016. Logical Disputes and the A Priori. *Logique et Analyse*. 236: pp. 347-366.

[203] Quine, W. V. (1950). *Methods of Logic*. Harvard University Press.

[204] ——. 1969. *Ontological Relativity and Other Essays*. Columbia University Press.

[205] ——. 1970. *Philosophy of Logic*. Harvard University Press.

[206] Read, S. 2010. *Relevant Logic: A Philosophical Examination of Inference*. Originally Published by Basil Blackwell 1988. Revised Edition (c) Stephen Read 2010.

[207] ——. 2015. Medieval Theories: Properties of Terms. *The Stanford Encyclopedia of Philosophy* (Spring 2015 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/spr2015/entries/medieval-terms/>.

[208] Restall, Greg 1996. Information flow and relevant logics. In Jerry Seligman & Dag WesterstÃ¥hl (eds.). *Logic, Language and Computation*. CSLI Publications, Stanford. pp. 463–477.

[209] ——. 2000. *An Introduction to Substructural Logics*. Routledge.

[210] ——. 2005. “Multiple Conclusions.” *Logic, Methodology and Philosophy of Science: Proceedings of the Twelfth International Congress.*

[211] Restall, G. and Slaney, J. 1995. Realistic Belief Revision. *Proceedings of the Second World Conference on the Fundamentals of Artificial Intelligence*. edited by M. De Glas. (Paris, July 1995), 367–378.

[212] Reutlinger, Alexander; Hangleiter, Dominik & Hartmann, Stephan. 2016 (Preprint). Understanding (With) Toy Models. *British Journal for the Philosophy of Science.*

[213] Rips, Lance J. 1994. *The Psychology of Proof: Deductive Reasoning in Human Thinking*. MIT Press.

[214] Rott, Hans & Pagnucco, Maurice. 1999. Severe withdrawal (and recovery). *Journal of Philosophical Logic*. 28 (5):501-547.

[215] Russell, Bertrand. 1919. *Introduction to Mathematical Philosophy*. Dover Publications.

[216] Ryan, Mark. 1996. "Belief Revision and Ordered Theory Presentations." In A. Fuhrmann, & H. Rott (Eds.), *Logic, Action, Information.* Berlin: De Gruyter. pp. 129–151.

[217] Schafer, Karl. 2014. "Doxastic planning and epistemic internalism." *Synthese*. 191 (12):2571-2591.

[218] Schmechtig, P. & Grajner, M. (eds.). 2016. *Epistemic Reasons, Norms, and Goals*. De Gruyter.

[219] Schmidt, Renate A. & Tishkovsky, Dmitry. 2007. On combinations of propositional dynamic logic and doxastic modal logics. *Journal of Logic, Language and Information.* 17 (1):109-129.

[220] Schulte, Oliver. 1999. Means-ends epistemology. *British Journal for the Philosophy of Science*. 50 (1):1-31.

[221] ——. 2017. Formal Learning Theory. *The Stanford Encyclopedia of Philosophy* (Spring 2017 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/spr2017/entries/learning-formal/>.

[222] Schurz, Gerhard. 2011. Truth-conduciveness as the primary epistemic justification of normative systems of reasoning. *Behavioral and Brain Sciences*. 34 (5):266-267.

[223] Serjeantson, Richard. W. 2005. Hume's general rules and the 'chief business of philosophers'. In Marina Frasca-Spada & P. J. E. Kail (eds.). *Impressions of Hume.* Oxford University Press. pp. 11–187.

[224] Segerberg, Krister. 1995. AGM as a dynamic, doxastic logic. *Bulletin of Symbolic Logic.* 1.

[225] ——. 1999a. Default logic as dynamic doxastic logic. *Erkenntnis.* 50 (2-3):333-352.

[226] ——. 1999b. Two traditions in the logic of belief: Bringing them together. In H. J. Ohlbach & U. Reyle (Eds.). *Logic, language and reasoning: Essays in honour of Dov Gabbay.* (pp. 135–147). Dordrecht: Kluwer Academic Publishers.

[227] ——. 2011. Routines. *Synthese.* 65 (2):185 - 210.

[228] ——. 2012. $D\triangle L$: Dynamic Deontic Logic. *Synthese.* 185 (S1):1-17.

[229] ——.2013. Applying Modal Logic. *Studia Logica.* 39 (2-3):275 - 295.

[230] Segerberg, Krister, Meyer, John-Jules and Kracht, Marcus. The Logic of Action. *The Stanford Encyclopedia of Philosophy* (Winter 2016 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/win2016/entries/logic-action/>.

[231] Sierra, Carlos, Godo, Lluis, Mantaras, Ramon, and Manzano, Mara. 1996. Descriptive Dynamic Logic and Its Application to Reflective Architectures.

[232] Shapiro, Stewart 2014. *Varieties of Logic.* OUP Oxford.

[233] Spohn, Wolfgang. 2012. *The Laws of Belief: Ranking Theory and its Philosophical Applications.* Oxford: Oxford University Press.

[234] Staffel, Julia. 2013. Can there be reasoning with degrees of belief? *Synthese.* 190 (16):3535-3551.

[235] Stalnaker, Robert. 1984. *Inquiry.* Cambridge University Press.

[236] ——. 2002. Epistemic consequentialism: Robert Stalnaker. *Aristotelian Society Supplementary.* Volume_ 76 (1):153–168.

[237] ——. 2009. Iterated Belief Revision. *Erkenntnis.* 70 (2):189-209.

[238] Stanovich, Keith E. 2005. *The Robot's Rebellion: Finding Meaning in the Age of Darwin.* University of Chicago Press.

[239] ——. 2008. Individual differences in reasoning and the algorithmic/intentional level distinction in cognitive science. In L. Rips & J. Adler (Eds.), *Reasoning: Studies of human inference and its foundations.* Cambridge: Cambridge University Press. (pp. 414-436).

[240] Steinberger, Florian 2017. Consequence and normative guidance. *Philosophy and Phenomenological Research.* 95 (1).

[241] ——. 2017. The normative status of logic. *Stanford Enyclopedia of Philosophy.* (Spring 2017 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/spr2017/entries/logic-normative/>.

[242] Stenning, Keith & van Lambalgen, Michiel. 2008. *Human reasoning and cognitive science.* Boston, USA: MIT Press.

[243] Sylvan, Richard. 1992. Process and action: Relevant theory and logics. *Studia Logica.* 51 (3-4):379 - 437.

[244] Tennant, Neil. 2006. On the Degeneracy of the Full AGM-Theory of Theory-Revision. *Journal of Symbolic Logic.* 71 (2):661 - 676.

[245] ——. 2012. *Changes of Mind: An Essay on Rational Belief Revision.* Oxford: Oxford University Press.

[246] van der Torre, Leon & Tan, Yao-Hua. 1998. An Update Semantics for Deontic Reasoning. *Norms, Logics and Information Systems: New Studies on Deontic Logic and Computer Science*: 73-90.

[247] Troquard, Nicolas, and Balbiani, Phillipe. 2015. Propositional Dynamic Logic. *The Stanford Encyclopedia of Philosophy* (Spring 2015 Edition). Edward N. Zalta (ed.). URL = <https://plato.stanford.edu/archives/spr2015/entries/logic-dynamic/>.

[248] Turri, John. 2011. .Believing for a Reason. *Erkenntnis*. 74(3): 383–397.

[249] Vetter, Barbara. 2013. Multi-track dispositions. *Philosophical Quarterly*. 63: 330–352. doi:10.1111/1467-9213.12024

[250] Vranas, Peter. 2008. New foundations for Imperative Logic I: Logical Connectives, Consistency, and Quantifiers. *NoÃ»s*. 42: 529–572.

[251] ——. 2010. In Defense of Imperative Inference. *Journal of Philosophical Logic*. 39: 59–71.

[252] Walton, Douglas. 1992. Rules for Plausible Reasoning. *Informal Logic*. 14 (1): 33-51.

[253] ——. 2013. *Methods of Argumentation*. Cambridge University Press.

[254] Walton, D., Tindale, C., and Gordon, T. 2014. Applying Recent Argumentation Methods to Some Ancient Examples of Plausible Reasoning. *Argumentation*. 28 (1):85-119.

[255] Wang, Yanjing. 2010. *Epistemic Modeling and Protocol Dynamics*. Ph.D Dissertation.

[256] Wassermann, Renata (2011). On AGM for Non-Classical Logics. _Journal of Philosophical Logic_ 40 (2):271 - 294.

[257] Wetzel, Linda. Types and Tokens. *The Stanford Encyclopedia of Philosophy* (Spring 2014 Edition). Edward N. Zalta (ed.). URL = <https://plato.stanford.edu/archives/spr2014/entries/types-tokens/>.

[258] Willer, Malte. 2015. “An Update on Epistemic Modals.” *Journal of Philosophical Logic.* 44: 835–49.

[259] Williamson, Timothy. 2000. *Knowledge and its Limits.* Oxford University Press.

[260] Wimsatt, William. 2007. *Re-Engineering Philosophy for Limited Beings: Piecewise Approximations to Reality.* Harvard University Press.

[261] von Wright, George H.. 1960. *Norm and Action.* Routledge.

[262] Xu, Ming. 2010. Combinations of Stit and Actions. *Journal of Logic, Language and Information.* 19 (4):485-503.

Index