5-1-2015

Dominion: A Game of Information Exploitation

Jacob Hobbs

Jacob Hobbs

Candidate

Computer Science

Department

This thesis is approved, and it is acceptable in quality and form for publication:

Approved by the Thesis Committee:

Trilce Estrada, Chairperson

Stephanie Forrest

Stephen Verzi

Dominion: A Game of Information Exploitation

by

Jacob Hobbs

B.S., University of New Mexico, 2012

THESIS

Submitted in Partial Fulfillment of the
Requirements for the Degree of

Master of Science
Computer Science

The University of New Mexico

Albuquerque, New Mexico

May, 2015

DEDICATION

To the Sumerian goddess of knowledge. Nisaba be praised.

ACKNOWLEDGEMENTS

I thank Trilce Estrada, Steve Verzi, and Stephanie Forrest for their patience and support. I thank my wife for the hours of babysitting duties she let me avoid for this.

Dominion: A Game of Information Exploitation

by

Jacob Hobbs

B.S., University of New Mexico, 2012

M.S., Computer Science, University of New Mexico, 2015

ABSTRACT

FlipIt is an abstract cyber-security game published in 2012[23] to investigate optimal strategies for managing security resources in response to Advanced Persistent Threats. In this thesis, we place FlipIt within a more general category of "stealthy move" games, and provide an approach towards solving such games. We produce a new stealthy move game, "Dominion", and derive Nash equilibria for it. We establish bounds for the optimal rates of play and benefits for FlipIt, and show that the best strategy to apply to real cyber security threats includes presenting a credible threat to potential players. We also explore the effects of initial game information asymmetry in Dominion.

Contents

List of Figures

List of Tables

Glossary

β_0	Benefit rate for the defender in FlipIt
β_1	Benefit rate for the attacker in FlipIt
δ	Dirac delta function
φ	Likelihood of defender starting Dominion in control
ψ	Likelihood of attacker starting Dominion in control Equal to $1 - \varphi$
E_0	Expected defender score in Dominion
E_1	Expected attacker score in Dominion
k	Integration limit for Dominion move distributions
k_0	Cost of moving for defender in FlipIt
k_1	Cost of moving for attacker in FlipIt
p_0	Move probability distribution function for defender in Dominion
p_1	Move probability distribution function for attacker in Dominion
r_0	Optimal defender average move rate in FlipIt
r_1	Optimal attacker average move rate in FlipIt

Chapter 1

Introduction and Overview

1.1 Motivation

Because game theory is based on the study of competition, conflict, and cooperation, it is heavily tied to the fields of sociology, economics, finance, politics, and evolution. When the FlipIt game was developed[23], it supplied a game-theoretic framework to investigate the use of renewal defenses against persistent, targeted attacks in cyber security domains. What made FlipIt stand out in contrast to regular game theoretic constructs was its employment of the idea of *stealthy moves.* That is, players are allowed to move at arbitrary points in time rather than taking turns or moving simultaneously, and the timing of their moves may be kept hidden from the other players. Although originally the phrase ”stealthy takeover” was used, we believe ”stealthy move” is a more appropriate, general term for describing a new class of games, since not all stealthy moves necessarily involve capturing resources.

Stealthy move games may be viewed as games of timing[18], having silent or noisy actions with arbitrary delay. The authors of FlipIt[23] and of numerous papers that succeeded it such as [12, 21, 25], have employed some game theoretic tools to investigate FlipIt and its extensions, but only under very simplified conditions. In this thesis we begin to develop the necessary math to more fully analyze stealthy move games.

1.2 Dominion

Dominion is a multiple player, unit interval, constant sum game, where the utility curves have a linear gradient and may be discontinuous. It has the following rules.

The goal of the game is to have control of a ”resource” as long as possible. This resource is simply designated as ”the resource”, and is an abstraction for something of value to all the players. We say that the player who controls the resource is *in control*. The game starts at time 0 and ends at time 1. In an n-player game, the players are numbered from 0 to $n - 1$. Player 0, whom we will frequently refer to as the defender, is in control at time 0. The other players may be referred to as attackers.

Players have a single kind of move that they can use at any time, called a *take move*, which we also refer to as *taking control*. A take move simultaneously gives a player control of the resource and reveals to that player information about the current and possibly past states of the game. If more than one player uses a take move simultaneously, control is given to the player whose ordinal is below that of the other players (that is, player 0 is always successful, and player 1 is only successful if player 0 doesn't move at the same time, etc).

All players know the starting conditions and goals of the game, but players take control of the resource in secret. The information a player receives upon taking control is not given to other players. Dominion is thus a stealthy move game. In this thesis we only consider Dominion where $n = 2$, so there is one attacker and one defender.

1.3 Relation to FlipIt

Dominion was designed such that if a complete solution of it can be given, that solution can be extended to the original FlipIt game by a reduction from FlipIt to Dominion. Dominion is easier to analyze than FlipIt because it is constant sum and has closed form Nash equilibria. Despite being simpler to analyze, many variations of FlipIt can be mapped to Dominion without loss of essential details.

In a basic configuration of FlipIt, there are two players competing to control one resource. A player is uncertain at almost any given time whether they control the resource. This is because action by either player to take control can happen at any time; the player who acts incurs a movement cost and takes over the resource if they don't own it already. If both players act at the same time, the player already in control of the resource remains in control.

A player's goal is to maximize her own utility, which for each player is defined as the overall time spent controlling the resource minus the cost of all the moves she has made. The game is assumed to be played between two players just once, but to last indefinitely, beginning at time zero. It is assumed that the players do not cooperate (given the imbalance of information in the game, cooperation probably requires more trust than is warranted). Player 0, known as the "defender", has control of the resource at time zero, and the "attacker" is the player who does not have control at time zero. Given universal knowledge of each player's individual control move costs and resource valuations, a number of strategies have been evaluated. Under assumptions of certain play strategies, *a priori* knowledge of opponent strategies, and varying levels of information being revealed about an opponent's moves at the time a player moves, optimal results and Nash equilibria have been found. However, general optimal strategies, where play is not limited to certain strategy classes, have not been successfully investigated and left for future research.

Dominion simplifies FlipIt in the following ways:

- Dominion is played within a finite time horizon. Only the period from 0 to 1 is considered.

- There is no explicit cost of moving. Instead, each player has a predetermined number of moves they may play.

- The value of controlling the resource for each player is unimportant to the analysis. The goal of the game is simply to control the resource for the maximum possible fraction of the time, by playing a mixed strategy fully utilizing all the moves available to a player.

This simplified setup is sufficient to capture most of the complexity of the original FlipIt game. Beginning with a FlipIt setup, relative move costs and resource values can be normalized and rolled up into the relative number of moves each player is given. That is, a more valuable resource and less expensive moves both translate into relatively more moves a player is willing make per unit time relative to the other player, which translate into having more moves in Dominion. A longer, finite time horizon can be incorporated by multiplying the number of moves each player is given. An infinite time horizon can be incorporated by considering the optimal payoffs as the number of moves for each player approaches ∞, while maintaining a constant ratio of defender to attacker moves. Situations like discounted the value of future gains could be employed by adding a utility function that gives different values to different portions of the unit interval, but we are not going to concern ourselves with this. We also do not consider asymmetric benefits for moving (such as upon moving, the attacker receiving precise information about the last defender move, but not vice versa), but leave that to future work.

1.4 A Game Theoretic Framework

The concept of a stealthy move means that, aside from possibly sharing complete knowledge of starting conditions, players only ever receive partial information about the state of the game, and they usually receive it at different, self-selected moments in time. We might imagine a game being played sequentially over a very large, possibly uncountably infinite number of rounds (in the case where moves can be made at any positive real-valued time). In any round, a player may choose to act or not to act. Only upon acting do players receive any information as to the current state of the game, including estimates of their current score. However, acting incurs a cost, and so optimizing players have a cost-limited number of moves they may feasibly make in a given period of time. Under this limited-move condition, the framework for treating a stealthy move-based game as a series of simultaneous move subgames is unwieldy and unable to help us go very far in our analysis. Other authors have attempted to get around the problems raised by stealthy moves by either simplifying the strategy search space (e.g., [14, 24]) or by discretization and simulation (e.g., [25]). While this is helpful, it would be preferable to find closed form solutions where they exist.

It is possible to treat a stealthy move game as a type of Hidden Markov Model game, where players are privy to partial information about the state of the model, and their moves are part of the transition function[3]. However, the ability to move at any time and in any order causes this to suffer from the same unwieldiness as earlier.

We may consider instead differential games. In differential games, players attempt to optimize the outcome of a system of continuous differential equations by supplying input functions which influence a part of them[9]. We can try to fit stealthy move games into this framework. Moves are discrete, but when considering mixed strategies, they may be supplied as probability distributions instead.

There are a few problems that have to be sorted out for a differential game framework. One problem is that players don't have a fixed number of moves to make, so while a single move may be represented as a single "control" probability distribution, it is not clear how to represent a full game with regular control functions. Furthermore, each time a player moves they receive information about the state of the game, and so they may adapt their control function according to this additional knowledge. Thus, a normal differential game framework has a hard time fitting the concept of a "stealthy move".

It is possible to treat a stealthy move game as a set of simultaneous differential games. Given an upper bound on the number of moves a player may choose to make, we could construct a separate differential game for every combination of possible move counts each player may make. Depending on the parameters of the original game, the games can then be solved together in interesting ways. For instance, consider a game where a player i has m moves to make and makes the first move. Once player i has moved, the rest of the game can be considered as an instance of the original game, only where player i has $m - 1$ moves to make, and the initial parameters of the game are updated for the new situation.

From this point of view, it seems useful to approach a stealthy move game by first solving instances where each player has one or fewer moves. From there we can build solutions for higher-move count games. When the game has only two players, we can put the expected scores for the best strategies of each game into a "rates" matrix M, indexed by the number of moves each player makes. That is, one player can be called the "row" player and the other player the "column" player, and if the row player has i moves in a given game, and the column player has j moves, then their expected scores can be recorded in row i and column j of M. Then the optimal strategy to follow for the original game may be determined by first finding an optimal strategy for the "rates" game on M, where each player picks an index of M (the former row player picking the row index, and the column player picking the column index). Then the strategy to follow for the original game is the strategy corresponding to the selected row or column a player chose in the ”rates” game. This strategy can be extended in a natural way to games with additional players.

Future work should highlight under what circumstances the above strategy corresponds to an optimal strategy in the original game. For instance, what are the consequences of having mixed strategy equilibria in the rates game? In this work we will forego the direct matrix approach, and instead extend it to consider rates of play as the matrix approaches infinite size. We begin by looking at a simple class of stealthy move games which we call Dominion. After defining the game and its general characteristics, we will solve it and then show how we can use the rates matrix framework to answer questions about optimal play in FlipIt. We will then look at extensions of Dominion that help us gain insights into optimal play under conditions of asymmetric information. We hope our approach will be expanded on in the future to begin solving broader classes of stealthy move games.

1.5 Contributions

This thesis contributes to the advancement of game theory by defining stealthy move games as a new extension to games of timing in Section 1.1 and outlining an approach to solving them in Section 1.4. We also define a simple stealthy move game called Dominion. Dominion may be extended to incorporate additional resources to control, types and numbers of moves, and rules about the starting conditions and goals of the game. Chapter 2 solves a probabilistic extension of the basic two player form of Dominion, where the defender begins in control of the unit interval each player has one move to make. This solution may be seen as an extension to the results of [22], dealing with resource control rather than duels. We analyze FlipIt to provide bounds on optimal play in Chapter 3. We look at how asymmetric information relating to starting conditions affects decision making in Dominion in Chapter 4.

Chapter 5 concludes with additional ways our current framework may be extended in the future.

Chapter 2

Creating and Solving the Equations for Dominion

2.1 Terminology

Play is defined over the time interval $[0,1)$. There are n players. We let p_i be a probability distribution function defined over this interval, describing the move strategy of player i. A distribution may be generalized a function. Specifically, we will later use the Dirac delta distribution, δ, in our equilibria.

We let E_i be the expected score for player i, given p_j for all $j \in \{0...n-1\}$. In a mixed strategy, players will not necessarily have a nonzero probability of moving over the entire range. In fact, in a Nash Equilibrium all players will only play up to time $k = 1 - \min_i(E_i)$ We can therefore generally describe the Nash distributions p_i as being over the range $[0,k]$.

It's possible to consider game play wherein no one controls the resource in the beginning. We will call this a "neutral advantage" game. Games where a player starts out in control of the resource are the default. If a player moves when she is already in control, she essentially updates her knowledge of the state of the game but doesn't gain or lose control because of it. To indicate how many moves each player is allowed to use in a game, we will write $\{m_0, m_1, ..., m_{n-1}\}$, where player i has m_i moves. We therefore may refer the 2 player version of Dominion where each player has 1 move and player 0 starts out in control as the $\{1, 1\}$ version of Dominion.

2.2 On the Existence of Nash equilibria

In game theory, a Nash equilibrium is a solution wherein no player can improve her expected payouts by individually changing her strategy. In a constant sum game such as Dominion, equilibria are guaranteed to exist so long as the payoff function is upper or lower semicontinuous[7, 10]. Thus, we can guarantee equilibria in the 2 player versions of Dominion if we consistently allow either the defender or the attacker to successfully have control each time both players choose to move simultaneously, which we have done.

Since the resource is always controlled by some player in Dominion and the total score for the game (since it is played over the unit interval) is 1, it follows that sum of the expected scores for all players will be 1.

2.3 Equation definitions

We will be solving the $\{1, 1\}$ version of Dominion. Our basic equation for the expected score of the defender can be written as

$$E_0 = \int_0^k p_1(x) \left(x + \int_x^k (1 - y)\, p_0(y)\, \mathrm{d}y \right) \mathrm{d}x \tag{1}$$

Here, k is the currently unknown upper bound of the range over which players are willing to play. The first part of the equation, $p_1(x) * x$, describes the score the defender expects to get from time 0 up to the time the attacker moves. The second part of the equation, $p_1(x)\int_x^k (1-y)p_0(y)\,\mathrm{d}y$, describes the additional score the defender expects to get whenever she moves after the attacker does, up to end of the game.

Our basic equation for the expected score of the attacker can be written as $E_1 = 1 - E_0$ (since the game is zero sum), but it can also be written

$$E_1 = \int_0^k p_0(y)\left(\int_0^y (y-x)\,p_1(x)\,\mathrm{d}x + \int_y^k (1-x)p_1(x)\,\mathrm{d}x\right)\,\mathrm{d}y \tag{2}$$

Here, the expected score for the attacker is the sum of the expected score for moving before the defender moves and the expected score for moving after the defender moves. Because we are allowing generalized functions, care should be taken to note that the last integral range is open on the low end, since if the attacker moves at precisely the same time the defender does, she loses to the defender.

We can subsume and generalize these equations together by treating the likelihood of being the defender, or in other words being in control at the beginning of the game, as a variable φ. The main equation, then, describing the expected score for the defender in the $\{1,1\}$ version of Dominion, is

$$\begin{aligned} E_0(\varphi) &= \varphi E_0 + (1-\varphi)E_1 \\ &= \int_0^k p_1(x)\left(\varphi x + (1-\varphi)\int_0^x (x-y)\,p_0(y)\,\mathrm{d}y + \int_x^k (1-y)p_0(y)\,\mathrm{d}y\right)\,\mathrm{d}x \end{aligned} \tag{3}$$

Note that we treated E_1 as if the defender was the attacker, so $p_1(x)$ and $p_0(y)$ switched places. Also, since we are dealing with generalized functions, in our analysis the final integral might need to be split apart, with fraction φ of it covering the interval $[x, k]$, and the fraction $1-\varphi$ of it covering the interval $(x, k]$, as in our discussion following Equation 2.

Since we are computing a Nash equilibrium, we can assume that the optimal p_0 forces the payoffs for each pure strategy component of the attacker equilibrium strategy to all be equal. This is sometimes called the *condition of indifference* for mixed strategy Nash equilibria. Thus, for the purposes of finding p_0 we can set $p_1(x) = \delta(x - T)$, for some T among the set of valid values x takes in the Nash. Thus we can simplify Equation 3, for the purposes of finding the Nash:

$$\begin{aligned} E_0(\varphi) &= \int_0^k \delta(x-T)\left(\varphi x + (1-\varphi)\int_0^x (x-y)\, p_0(y)\,\mathrm{d}y + \int_x^k (1-y)p_0(y)\,\mathrm{d}y\right)\,\mathrm{d}x \\ &= \varphi T + (1-\varphi)\int_0^T (T-y)p_0(y)\,\mathrm{d}y + \int_T^k (1-y)p_0(y)\,\mathrm{d}y \end{aligned} \tag{4}$$

We now expand this equation and treat it as a function of T:

$$E_0(\varphi, T) = \varphi T + (1-\varphi)T\int_0^T p_0(y)\,\mathrm{d}y + \varphi\int_0^T y p_0(y)\,\mathrm{d}y + \int_T^k p_0(y)\,\mathrm{d}y - \int_0^k y p_0(y)\,\mathrm{d}y \tag{5}$$

2.4 Solving the Equations

Notice that we can employ a trick to Equation 5. Since in the Nash, the choice of T shouldn't affect the expected score, we can set the derivative of $E_0(\varphi, T)$ with respect to T to 0.

$$\frac{\partial E_0(\varphi, T)}{\partial T} = \varphi + (1-\varphi)\int_0^T p_0(y)\,\mathrm{d}y - (1-T)p_0(T) = 0 \tag{6}$$

Equation 6 leads to the following constraint when solving for p_0:

$$(1-T)p_0(T) = \varphi + (1-\varphi)\int_0^T p_0(y)\,\mathrm{d}y \tag{7}$$

When $\varphi \neq 1$, we have

$$\int_0^T p_0(y)\,\mathrm{d}y = \frac{(1-T)p_0(T) - \varphi}{1-\varphi} \tag{8}$$

Since p_0 is a probability distribution function, we can use our knowledge that

$$\int_0^T p_0(y)\,\mathrm{d}y + \int_T^k p_0(y)\,\mathrm{d}y = 1 \tag{9}$$

(minding that one of the T integral endpoints is open, in the case of generalized functions) to get the additional constraint:

$$\int_T^k p_0(y)\,\mathrm{d}y = 1 - \frac{(1-T)p_0(T) - \varphi}{1-\varphi} = \frac{1-(1-T)p_0(T)}{1-\varphi} \tag{10}$$

We can substitute the constraints of Equation 8 and Equation 10 into Equation 5 to get

$$E_0(\varphi, T) = T(1-T)p_0(T) + \varphi \int_0^T y p_0(y)\,\mathrm{d}y + \frac{1-(1-T)p_0(T)}{1-\varphi} - \int_0^k y p_0(y)\,\mathrm{d}y \tag{11}$$

We can again exploit the Nash condition of indifference as we did in Equation 4 by setting the derivative of this new equation with respect to T to 0.

$$\begin{aligned} \frac{\partial E_0(\varphi, T)}{\partial T} &= (1-2T)p_0(T) + T(1-T)\frac{\mathrm{d}p_0(T)}{\mathrm{d}T} \\ &\quad + \varphi T p_0(T) + \frac{p_0(T) - (1-T)\frac{\mathrm{d}p_0(T)}{\mathrm{d}T}}{1-\varphi} \\ &= \frac{(1-T+\varphi T)\left((2-\varphi)\,p_0(T) - (1-T)\frac{\mathrm{d}p_0(T)}{\mathrm{d}T}\right)}{1-\varphi} \end{aligned} \tag{12}$$

Setting Equation 12 to 0 we find

$$p_0(T) = \frac{1-T}{2-\varphi}\frac{\mathrm{d}p_0(T)}{\mathrm{d}T} \tag{13}$$

Thus, for some α,

$$p_0(T) = \frac{\alpha}{(1-T)^{2-\varphi}} \tag{14}$$

Applying Equation 14 to Equation 6, we get

$$\begin{aligned}\frac{\partial E_0(\varphi,T)}{\partial T} &= \varphi + (1-\varphi)\int_0^T \frac{\alpha}{(1-y)^{2-\varphi}}\,\mathrm{d}y - \frac{\alpha}{(1-T)^{1-\varphi}} \\ &= \varphi + \alpha\left((1-T)^{\varphi-1} - 1\right) - \frac{\alpha}{(1-T)^{1-\varphi}} \\ &= \varphi - \alpha\end{aligned} \tag{15}$$

Setting this to 0 leads us to the result that $\varphi = \alpha$. Thus,

$$p_0(x) = \frac{\varphi}{(1-x)^{2-\varphi}} \tag{16}$$

Since p_0 is a probability distribution function, we need p_0 from 0 to k to integrate to 1. It follows that

$$\int_0^k p_0(x)\,\mathrm{d}x = \frac{\varphi}{1-\varphi}\left((1-k)^{\varphi-1} - 1\right) = 1 \tag{17}$$

Thus,

$$k = 1 - \varphi^{\frac{1}{1-\varphi}} \tag{18}$$

Using Equation 3, we can compute

$$E_0(\varphi) = 1 - \varphi^{\frac{\varphi}{1-\varphi}} \tag{19}$$

Does Equation 3 work for both attacker and defender distributions? Checking our work, for a pair of solutions to be Nash, we have the additional constraint that the expected scores sum to 1. If we assume the attacker and defender draw use the same strategy form, then the attacker likelihood of starting the game in control is $1-\varphi$.

$$E_0(\varphi) + E_0(1-\varphi) = 1 \tag{20}$$

Unfortunately, in our case we get

$$E_0(\varphi) + E_0(1-\varphi) = 2 - \left(\varphi^{\frac{\varphi}{1-\varphi}} + (1-\varphi)^{\frac{1-\varphi}{\varphi}}\right) \tag{21}$$

This only matches our constraint at $\varphi = \frac{1}{2}$. We also could have noted that the bound $k(\varphi) = 1 - \varphi^{\frac{1}{1-\varphi}}$ for a pair of strategies (φ and $(1-\varphi)$ should be the same, since in a Nash equilibrium if one player's strategy stops at k, the other player shouldn't move beyond k either. Unfortunately this means our solution is incorrect, or actually incomplete.

What we are missing is the fact that a strategy in this game could include nonzero point distributions. Thus we must allow generalized functions into our solution space. Given the nature of the equations, it only makes sense to place a point distribution, the Dirac delta function, at time 0. Placing such a function at any other time introduces discontinuities in the solution space, which can't be reconciled with the need maintain indifference among solution choices in a Nash equilibrium. Furthermore, it only makes sense to use a delta function if $\varphi \le \frac{1}{2}$ (there is no reason to have a finite probability of moving at time 0 if you are fairly certain you are in control at that time anyway).

Because $\varphi \le \frac{1}{2}$, we will refer to the distribution that includes a Dirac delta function as p_1, and replace φ with ψ in all of the defender equations, to make equivalent attacker equations. Henceforth we may assume $\varphi \ge \frac{1}{2}$, and $\psi = 1 - \varphi$. Now we have, for some α and β,

$$p_1(T) = \frac{\delta(T)}{\beta} + \frac{\alpha}{(1-T)^{2-\psi}} \tag{22}$$

This new equation is basically orthogonal to Equation 16, with respect to being an additional solution to the constraining Equation 8 and Equation 13, if parameterized correctly. Solving Equation 7 with this new equation, where $T \ne 0$, we get

$$(1-T)p_1(T) = \psi + (1-\psi)\int_0^T p_1(x)\,\mathrm{d}x \tag{23}$$

Assuming for the moment $T \ne 0$, and noting that the lower limit in the integral is included, we have

$$\frac{\alpha}{(1-T)^{1-\psi}} = \psi + \alpha\left((1-T)^{\psi-1} - 1\right) + \frac{1-\psi}{\beta} \tag{24}$$

Solving for α,

$$\alpha = \frac{1-\psi+\psi\beta}{\beta} \tag{25}$$

We can use the above to rewrite Equation 22 as

$$p_1(T) = \frac{1}{\beta}\left(\delta\left(T\right) + \frac{1-\psi+\psi\beta}{(1-T)^{2-\psi}}\right) \tag{26}$$

Lastly, we solve for β by exploiting our knowledge that p_1 is a probability distribution.

$$\int_0^k p_1(T)\,\mathrm{d}T = \frac{1}{\beta}\left(1 + \frac{1-\psi+\psi\beta}{1-\psi}\left((1-k)^{\psi-1} - 1\right)\right) = 1 \tag{27}$$

Solving for β and factoring,

$$\beta = \frac{(1-\psi)(1-k)^{\psi}}{1-k-\psi(1-k)^{\psi}} \tag{28}$$

Applying this to Equation 3, we find

$$\begin{aligned} E_1(\psi) &= \frac{(1-\psi+\psi\beta)(1-(1-k)^{\psi})}{\psi\beta} \\ &= \frac{1}{\psi}\left(\frac{1-\psi+\psi\beta}{\beta} - \left(\frac{1-\psi+\psi\beta}{\beta}\right)^{\frac{1}{1-\psi}}\right) \end{aligned} \tag{29}$$

We wish to maximize this expected value, and set β accordingly.

$$\frac{\partial E_1(\psi)}{\partial\beta} = \frac{(\frac{1-\psi+\psi\beta}{\beta})^{\frac{\psi}{1-\psi}} - (1-\psi)}{\psi\beta^2} \tag{30}$$

Setting this to 0 implies

$$\beta = \frac{(1-\psi)^2}{\psi^2 + (1-\psi)^{\frac{1}{\psi}} - \psi} \tag{31}$$

Combining this with Equation 29 we have

$$E_1(\psi) = (1-\psi)^{\frac{1-\psi}{\psi}} \tag{32}$$

Finally, Equation 28 and Equation 31 combine to give us

$$k = 1 - (1-\psi)^{\frac{1}{\psi}} \tag{33}$$

When $\psi = 1 - \varphi$, we have

$$k = 1 - (1-\psi)^{\frac{1}{\psi}} = 1 - \varphi^{\frac{1}{1-\varphi}} \tag{34}$$

and also,

$$E_0(\phi) + E_1(\psi) = 1 - \varphi^{\frac{\varphi}{1-\varphi}} + (1-\psi)^{\frac{1-\psi}{\psi}} = 1 - \varphi^{\frac{\varphi}{1-\varphi}} + \varphi^{\frac{\varphi}{1-\varphi}} = 1 \tag{35}$$

Thus, the conditions for these equation pairs to constitute a Nash equilibrium are satisfied. The holes from discontinuities that we ignored earlier, such as when $\varphi = 1$, disappear in the limit.

In conclusion, the optimal strategies and payoffs for the $\{1, 1\}$ version of Dominion where the defender has probability $\varphi \geq \frac{1}{2}$ of being in control at the beginning of the game, and the attacker has probability $\psi = 1 - \varphi$ of being in control, are

$$p_0(x) = \frac{\varphi}{(1-x)^{2-\varphi}} \tag{36}$$

$$p_1(x) = \frac{\varphi^{\frac{\varphi}{1-\varphi}} - (1-\varphi)}{\varphi}\delta(x) + \frac{\varphi^{\frac{\varphi}{1-\varphi}}}{(1-x)^{1+\varphi}} \tag{37}$$

$$k = 1 - \varphi^{\frac{1}{1-\varphi}} \tag{38}$$

$$E_0(\varphi) = 1 - \varphi^{\frac{\varphi}{1-\varphi}} \tag{39}$$

$$E_1(\psi) = E_1(1-\varphi) = \varphi^{\frac{\varphi}{1-\varphi}} \tag{40}$$

2.5 Notes on the Solution

What follows are some visualizations and remarks on the solutions that were produced in the last section.

Figure 2.1 shows the surface of equilibrium strategy distributions based on φ. Note from Equation 37 that when $\varphi = 0$, the slope for the distribution is proportional to $\frac{1}{(1-x)^2}$, and from Equation 36 that when $\varphi = 1$, the slope is proportional to $\frac{1}{1-x}$. Furthermore, at the endpoint k of the distribution, $p_0(k) = p_1(k)$. That is, the likelihood of moving at any point besides $x = 0$ is higher for the defender than it is for the attacker, up until the highest x-values of the distributions, where the likelihoods overlap.

Figure 2.2 shows the equilibrium distributions for a few select values of φ, disregarding the initial δ functions where $\varphi < 0.5$. Where $\varphi \geq 0.5$, the value at $x = 0$ is φ. The gap between the distributions shrinks as φ approaches 0.5, where they become identical. A player's likelihood of moving over a fixed-width open interval increases as time increases.

Figure 2.3 shows the expected payout for a player given φ, the likelihood that player starts out as the defender. This figure shows that the payoff doesn't quite increase linearly as φ varies.

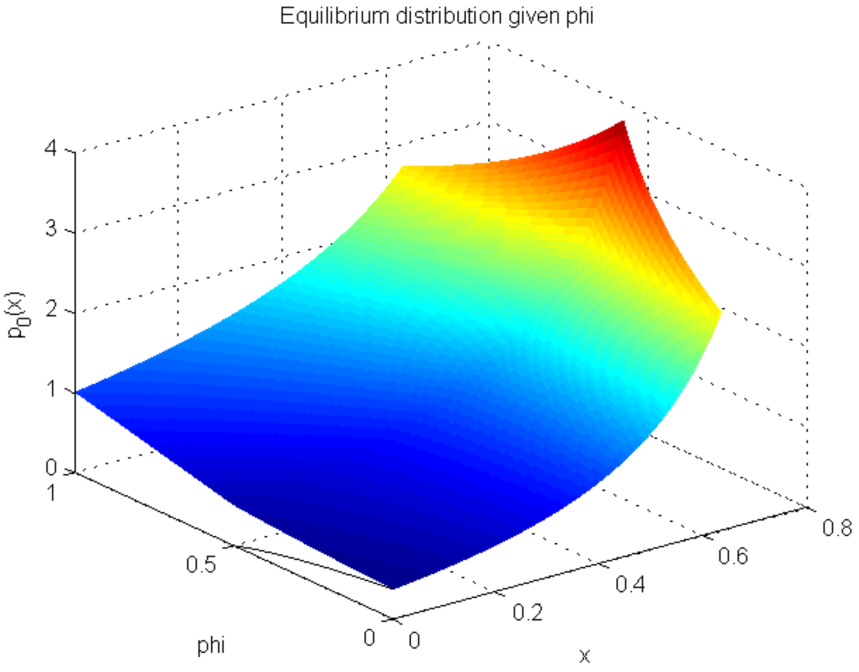

Figure 2.1: The equilibrium strategies, where the black line at $x = 0$ and $phi \in [0, 0.5)$ represents δ function strength. "phi" is the likelihood of starting the game in control, and "x" is the time axis.

We can use the equations we solved in this chapter to make observations about the roles different components of Dominion play in the expected payouts of the game. For instance, so long as $\varphi = 0.5$, the expected score for the $\{k, k\}$ version of Dominion for all $k \in \mathbb{N}$ should remain as 0.5. It is the point at which disparate strategies of "usually" attacking versus "usually" defending players converge.

In the $\{1, 1\}$ version of Dominion that we covered in this chapter, we can think

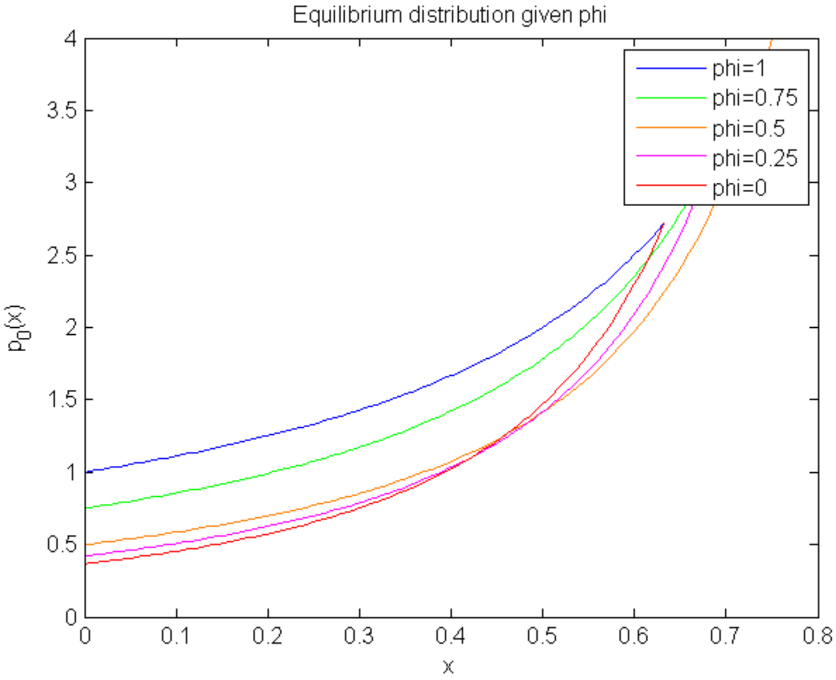

Figure 2.2: Equilibrium distributions for select values of φ.

of the defender's initial control as a first "bonus" move the defender automatically has, only this move is revealed to all players. Thus, it is almost as if the defender has twice as many "moves" as the attacker in this game. We can ask how expected scores will change as the number of moves each player has increases in proportion with each other. That is, what is the defender's expected score for the $\{2k-1, k\}$ version of Dominion as $k \to \infty$?

The defender's expected score in the $\{2k-1, k\}$ version of Dominion will always fall somewhere between $\frac{1}{2}$ and $\frac{3}{4}$, depending on how much information moving grants a player. If upon moving a player receives no additional information about the state of the game, then the defender can hide all information about the timing of her moves. If we assume her moves partition the unit interval into roughly $2k$ parts, and the attacker can only move within k of these intervals, then with no additional knowledge she should expect to gain control of the resource for half the length of each

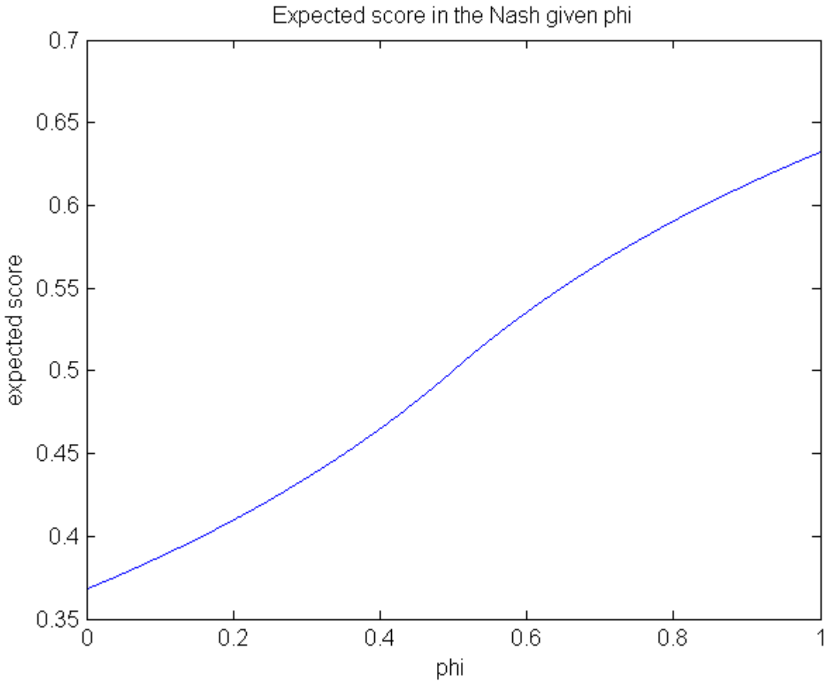

Figure 2.3: The equilibrium payouts in Dominion.

interval she moves in. Thus, $\frac{3}{4}$ is an upper bound to the defender's expected score. Size bias[1] suggests some strong limitations on the optimal shape of the defender distribution in this case, which we will not further explore here.

If the attacker is able to know when the defender moves, but not vice versa, then evenly spacing her moves guarantees the defender a score of $\frac{1}{2}$. It is interesting that in the $\{1, 1\}$ version of Dominion which we covered in this chapter, the defender's best expected score was about 16% worse than $\frac{3}{4}$ and 26% better than $\frac{1}{2}$. This can be seen as an information advantage the attacker gets for knowing the timing of the defender's first move. Solving the $\{3, 2\}$ version of Dominion with differing information assumptions would help us see more clearly the role information players in formulating optimal strategies and their payoffs.

Chapter 3

Applying Dominion to FlipIt

3.1 Application to FlipIt

The preceding chapter demonstrated the optimal strategies for the $\{1, 1\}$ version of Dominion where each player had shared, correct beliefs about her probability of being in control in the beginning of the game. Furthermore, we showed that we can loosely bound the expected utilities for the different move count Dominion games based on assumptions we make about information a player receives upon moving, and the relative number of moves each player makes.

Much of the following analysis can be fruitfully compared to the proof of Theorem 5 in Appendix A of [23]. Suppose the defender has $a - 1$ moves, and the attacker has b moves. If a is considerably greater than b, and the attacker is unable to make use of additional information about the timing of her moves, then at best the defender can achieve an expected score of $1 - \frac{b}{2a}$. If instead a is relatively close in value to b, and each player receives symmetrically the same kind of information upon moving as the other player, then we may suppose that each player's moves are about as powerful as the other's. Then the defender's best expected score should instead be close to $\frac{a}{a+b}$. Note that if we replace a with ka and b with kb, this best expected score doesn't change.

Let us now map this expected score to FlipIt. Let us assume the equivalent to this Dominion game is played at a rate λ in a FlipIt game. That is, the defender moves periodically once every $\frac{1}{\lambda}$, and then plays an additional $a-1$ moves within each period, while the attacker plays b moves within each period. Let $r_0 = \frac{a}{\lambda}$ be the average move rate for the defender, and $r_1 = \frac{b}{\lambda}$ the average move rate for the defender. Then the expected benefit rate for the defender is roughly

$$\beta_0 = E_0 - \frac{a * k_0}{\lambda} = E_0 - r_0 k_0 \tag{41}$$

The expected benefit rate for the attacker is

$$\beta_1 = (1 - E_0) - \frac{b * k_1}{\lambda} = (1 - E_0) - r_1 k_1 \tag{42}$$

Now consider what happens when we replace a with ka, b with kb, and λ with $k\lambda$ in this mapping. The expected benefits for the players remain the same, and the period over which they play the Dominion-based strategy is lengthened by k times. We can extend this process as $k \to \infty$, and therefore can simply consider FlipIt games where the average rates of play for each player is similar to the other.

Consider having the defender and attacker play a new game before playing FlipIt, in which, for fixed costs of moving k_0 and k_1, the defender and attacker each simultaneously choose average rates of play r_0 and r_1 to maximize their expected payoffs when using those rates of play in FlipIt. We can compute the expected benefit rates β_0^* and β_1^* of these optimal strategies, even though we don't yet know what the strategies are. Noting that $\lim\limits_{k\to\infty} E_0 = \lim\limits_{k\to\infty} \frac{ka}{ka+kb} = \frac{r_0}{r_0+r_1}$, and $\lim\limits_{k\to\infty} \frac{kak_0}{k\lambda} = \frac{\alpha k_0}{\lambda} = r_0 k_0$,

$$\beta_0^* = \frac{r_0}{r_0 + r_1} - r_0 k_0 \tag{43}$$

$$\beta_1^* = \frac{r_1}{r_0 + r_1} - r_1 k_1 \tag{44}$$

We can set the derivative of Equation 43 with respect to r_0 to 0 to find the best defender move rate versus a given attacker rate.

$$\frac{\partial \beta_0}{\partial r_0} = \frac{r_1}{(r_0 + r_1)^2} - k_0 = 0 \implies r_0 k_0 = \sqrt{r_1 k_0} - r_1 k_0 \tag{45}$$

This is a local maximum. Thus, given a fixed attacker average move rate, we have the best corresponding defender rate. By symmetry, given a fixed average defender move rate, the corresponding best attacker move rate is $r_1 k_1 = \sqrt{r_0 k_1} - r_0 k_1 \implies r_1 k_0 = \sqrt{r_0 k_1} \frac{k_0}{k_1} - r_0 k_0$. We draw these best response curves in Figure 3.1. Note that each curve is scaled differently. The best response curve for the attacker is scaled by the attacker move cost, and the best response curve for the defender is scaled by the defender move cost.

Combining the best response curves, we find the equilibrium at

$$r_0 = \frac{k_1}{(k_0 + k_1)^2} \tag{46}$$

$$r_1 = \frac{k_0}{(k_0 + k_1)^2} \tag{47}$$

Applying these to Equations 43 and 44, we have

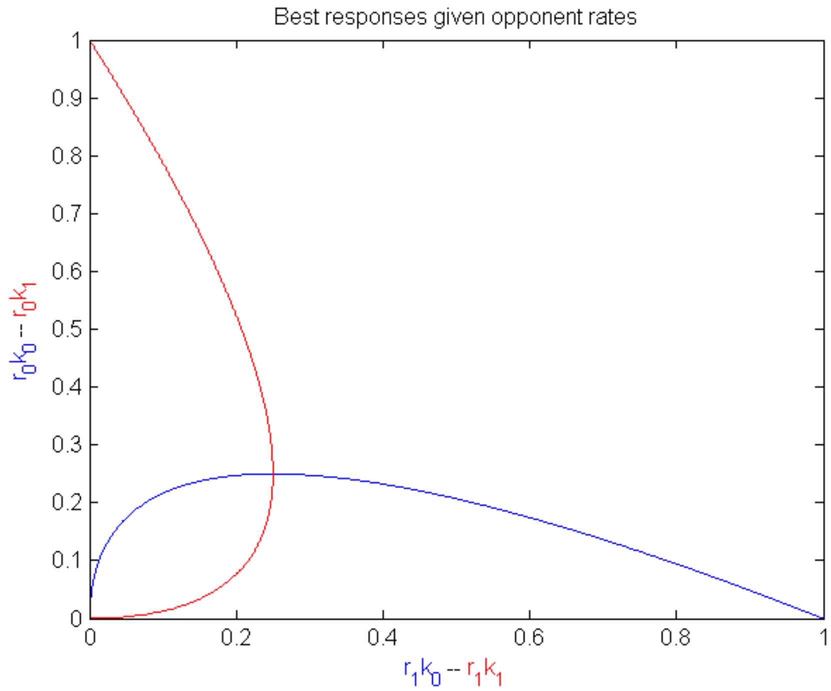

Figure 3.1: Best scaled move rate given an opponents' scaled rate.

$$\beta_0^* = \frac{k_1^2}{(k_0 + k_1)^2} \tag{48}$$

$$\beta_1^* = \frac{k_0^2}{(k_0 + k_1)^2} \tag{49}$$

Suppose instead of this equilibrium, the defender wished to drive the attacker out of the game. From Equation 45, we can see that $r_1 k_0 = 1$ is the point at which the defender drops out of the game, and by symmetry, $r_0 k_1 = 1$ is the point at which the attacker drops out of the game. Thus, the defender will play at an average rate of $r_0 = \frac{1}{k_1}$. Notice that the attacker will have no incentive to play at a nonzero rate, unlike what happens under the assumptions of periodic or exponential play in [23],

$\frac{k_0}{k_1}$	r_0	r_1	β_0^*	β_1^*
$(0, \frac{\sqrt{5}-1}{2}]$	$\frac{1}{k_1}$	0	$1-\frac{k_0}{k_1}$	0
$[\frac{\sqrt{5}-1}{2}, \frac{1+\sqrt{5}}{2}]$	$\frac{k_1}{(k_0+k_1)^2}$	$\frac{k_0}{(k_0+k_1)^2}$	$\frac{k_1^2}{(k_0+k_1)^2}$	$\frac{k_0^2}{(k_0+k_1)^2}$
$[\frac{1+\sqrt{5}}{2}, \infty)$	0	$\frac{1}{k_0}$	0	$1-\frac{k_1}{k_0}$

Table 3.1: Bounds on optimal strategy payoffs in FlipIt

Sections 5.2 and 5.3. The defender's expected benefit in this case will be $1-\frac{k_0}{k_1}$. It is more beneficial for the defender to drive the attacker out when

$$1-\frac{k_0}{k_1} > \frac{k_1^2}{(k_0+k_1)^2} \implies k_1^2 > k_0^2 + k_0k_1 \implies \frac{k_1}{k_0} > \frac{1+\sqrt{5}}{2} \tag{50}$$

At the point $\frac{k_1}{k_0} = \frac{1+\sqrt{5}}{2}$, the defender is indifferent to driving the attacker out or playing her equilibrium strategy, and in either case receives an expected benefit of $1-\frac{2}{1+\sqrt{5}} \approx 0.382$. Thus, depending on the defender strategy, the attacker may have an expected score of 0 or $\frac{(1-\sqrt{5})^2}{(1+\sqrt{5})^2} \approx 0.146$. Symmetric results can be derived for when $\frac{k_0}{k_1} = \frac{1+\sqrt{5}}{2}$.

Table 3.1 summarizes the results of this section. Under the assumption that each player derives symmetric information benefits upon moving, and that their move costs, and therefore average rates of play, are within $\frac{1+\sqrt{5}}{2}$ of each other, we are additionally assuming that each defender move gives her roughly the same benefit as each attacker move gives the attacker. This is necessarily an upper bound where $\frac{k_0}{k_1} < 1$ and a lower bound where $\frac{k_0}{k_1} > 1$. We can usefully compare these results to Theorem 8 of [23], to see the most the defender can lose by playing exponentially rather than optimally.

Figure 3.2 shows the benefits for both players playing optimally, as well as the benefits for both players when the defender is playing exponentially and the attacker

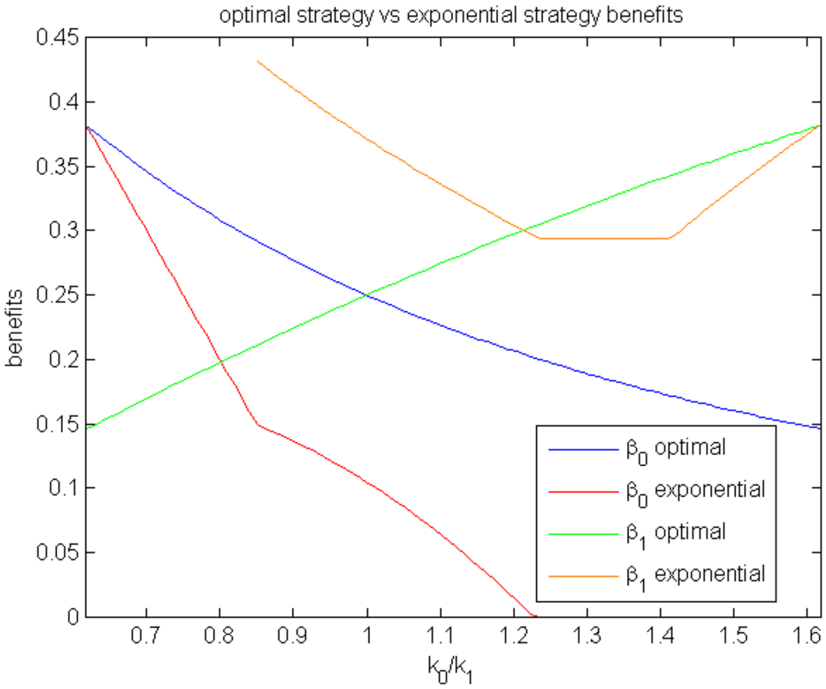

Figure 3.2: Optimal strategy benefits contrasted with exponential strategy benefits.

is playing at her optimal period based on the defender's rate, as in [23]. The window of strategies it covers is $\frac{k_0}{k_1} \in (\frac{\sqrt{5}-1}{2}, \frac{1+\sqrt{5}}{2})$. To the left of the window shown, the optimal defender benefit and the exponential defender benefit are both $\beta_0 = 1 - \frac{k_0}{k_1}$. To the right of the window shown, the optimal attacker and exponential attacker benefits are both $\beta_1 = 1 - \frac{k_1}{k_0}$.

Note that we have the defender drop out of the game if her expected benefit is less than 0, at which point the optimal strategy for the attacker would change. [23] doesn't comment about this, so we deal with it here by supposing the attacker could stick to the same strategy from the point the defender drops out of the game (this is where the orange line remains horizontal) up to the point she would do better by playing periodically at the rate $\frac{1}{k_0}$. Note also that although the defender cost goes up relative to the attacker costs, the attacker's benefit goes down because the defender is expending a lot of effort to reach her exponential strategy equilibrium. This is a large red flag that the players aren't performing optimally.

Solving Dominion should allow us to close the gap and provide the strategy that gives the optimal benefit in FlipIt. Depending again on information assumptions, the exponential distribution seems to be a poor choice, since it assigns a higher probability for the defender to move when the defender has a higher likelihood of being in control, whereas an optimal strategy should call for the likelihood of moving to correspond to the likelihood the defender is *not* in control. [4] also uses the exponential strategy for a real mock application. Given our results in the last chapter, we suggest that distributions shaped more like $f(x) = \frac{1}{1-x}$ be tried on the same problem.

3.2 Optimal Strategies in FlipIt

Assume each player gets the same symmetric informational benefits upon moving. The strategies we found in the last section were based on the optimistic assumption that a player with a lower move rate can still derive the same benefit as the player moving more frequently. If the situation is really at the other extreme, and the player with a lower move rate derives roughly half the utility per move as the player with the higher rate, then assuming $k_0 \leq k_1 \implies r_0 \geq r_1$,

$$\beta_0^* = 1 - \frac{r_1}{2r_0} - r_0 k_0 \tag{51}$$

$$\beta_1^* = \frac{r_1}{2r_0} - r_1 k_1 \tag{52}$$

Here the math simply follows Theorem 1 in [23] and so the optimal player strategies will always be to drive the player with highest cost out of the game.

Let us assume, then, that the player with fewer moves can make better use of them, perhaps because players receive a lot of feedback about the state of the game upon moving. The results of the previous section suggest that the optimal strategies for both players in this case will form a cooperative equilibrium (see [8]), with r_0 and r_1 being rates just strong enough to allow each player the ability to punish the other for deviating from cooperation. This suggests that players would be able to use their move information to punish the other player if she deviates from cooperation, rather than merely using her information to try to exploit her opponent. In a cyber security situation, we may assume that it hurts the defender more to give the attacker additional benefit in exchange for a little more benefit than it does to not receive the extra benefit. Note by Figure 3.2 that the attacker benefits greatly if the defender chooses to play her optimal, cooperative strategy. In this case, cooperating doesn't make sense.

Unfortunately, where players only receive limited information, no amount of cooperation provides more benefits than presenting a credible threat. Figure 3.3 plots the results of Theorem 5 in [23] against the optimal cooperative strategy from above. If the defender credibly sticks to her strategy of guaranteeing the attacker a score of 0 or less, the attacker will not participate in the game and the defender will have a much higher expected score as a result. In a real cyber security situation, the defender presumably has the upper hand in this, being able to set a resource refreshing schedule to discourage potential adversaries from engaging in resource competition in the first place.

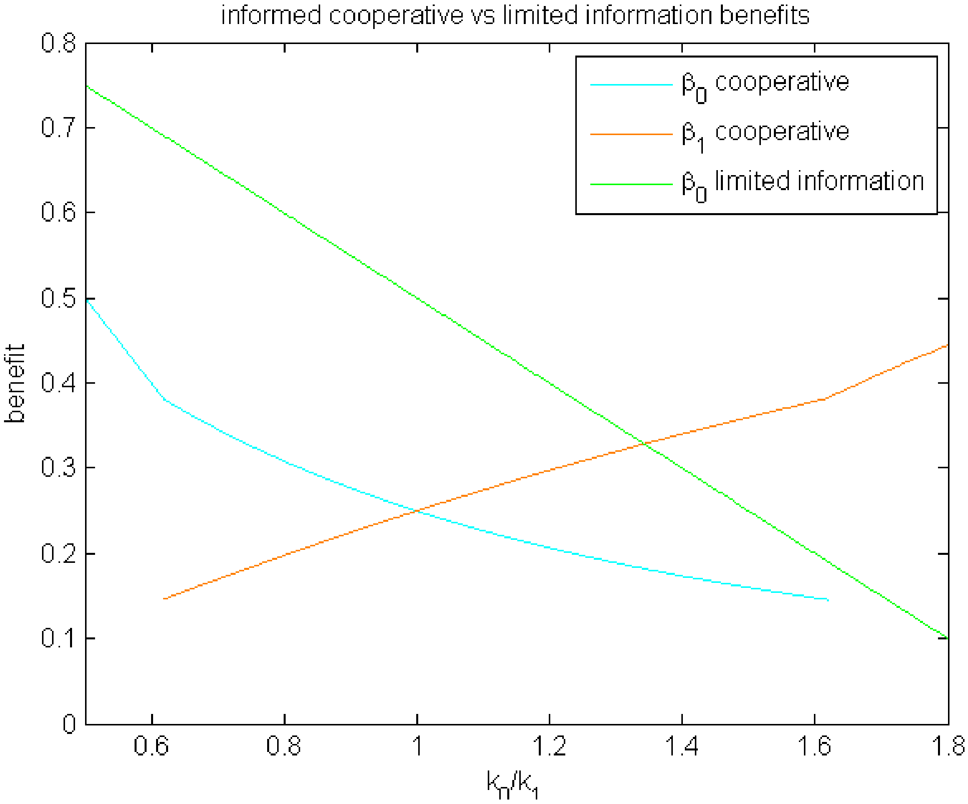

Figure 3.3: The best cooperative strategy does worse than the best limited information strategy

Chapter 4

Extensions of Dominion

4.1 Asymmetric Starting Conditions

Consider the same probabilistic $\{1,1\}$ Dominion game we solved in Chapter 2, only this time let it be a game of asymmetric information. Allow one player, who we will refer to as the defender, to know nature's choice for who starts the game in control of the resource. Let the attacker and defender both know the probability nature used, φ. Then we can imagine a number of scenarios, which we now examine.

4.1.1 Attacker in the Dark

Let the attacker believe the defender only knows φ, so the defender can exploit her additional knowledge, since she will know the attacker's assumed equilibrium strategy.

The attacker's equilibrium strategy tried to equalize the expected benefit of all non-dominated moves the defender could make. Since the defender knows nature's choice and her opponent's strategy, her optimal strategy will be pure, consisting of a single point move.

If nature put the defender in control, and $\varphi < \frac{1}{2}$, then letting $\psi = 1 - \varphi$, we pick $T \in [0, k]$ to maximize the following:

$$(1 - T) \int_0^T \frac{\psi}{(1 - x)^{2-\psi}} \, \mathrm{d}x \tag{53}$$

If nature put the attacker in control, and $\varphi < \frac{1}{2}$, then letting $\psi = 1 - \varphi$, we pick $T \in [0, k]$ to maximize:

$$(1 - T) \int_0^T \frac{\psi}{(1 - x)^{2-\psi}} \, \mathrm{d}x + \int_T^{1-\psi^{\frac{1}{1-\psi}}} (x - T) \frac{\psi}{(1 - x)^{2-\psi}} \, \mathrm{d}x \tag{54}$$

If nature put the defender in control, and $\varphi \geq \frac{1}{2}$, then we pick $T \in (0, k]$ to maximize:

$$(1 - T) \int_0^T \frac{\varphi^{\frac{\varphi}{1-\varphi}} - (1 - \varphi)}{\varphi} \delta(x) + \frac{\varphi^{\frac{\varphi}{1-\varphi}}}{(1 - x)^{1+\varphi}} \, \mathrm{d}x \tag{55}$$

If nature put the attacker in control, and $\varphi \geq \frac{1}{2}$, then we pick $T \in (0, k]$ to maximize:

$$\begin{aligned} &(1 - T) \textstyle\int_0^T \frac{\varphi^{\frac{\varphi}{1-\varphi}} - (1-\varphi)}{\varphi} \delta(x) + \frac{\varphi^{\frac{\varphi}{1-\varphi}}}{(1-x)^{1+\varphi}} \, \mathrm{d}x \\ + &\textstyle\int_T^{1-\varphi^{\frac{1}{1-\varphi}}} (x - T) \left(\frac{\varphi^{\frac{\varphi}{1-\varphi}} - (1-\varphi)}{\varphi} \delta(x) + \frac{\varphi^{\frac{\varphi}{1-\varphi}}}{(1-x)^{1+\varphi}} \right) \mathrm{d}x \end{aligned} \tag{56}$$

In all of these, the defender's best move is to move at an endpoint of the attacker's equilibrium strategy. If $\varphi < \frac{1}{2}$ and the defender starts in control, the defender should move at time $1 - \psi^{\frac{1}{1-\psi}}$, for a total expected score of $\psi^{\frac{\psi}{1-\psi}} + \psi^{\frac{1}{1-\psi}}$ (rather than $1 - \frac{1}{e}$). If $\varphi < \frac{1}{2}$ and the attacker starts in control, the defender should move at time 0 for a total expected score of $\psi^{\frac{\psi}{1-\psi}}$ (rather than $\frac{1}{e}$). If $\varphi \geq \frac{1}{2}$ and the defender starts in control, the defender should move at time $1 - \varphi^{\frac{1}{1-\varphi}}$, for a total expected score of $\frac{\varphi + \varphi^{\frac{1}{1-\varphi}} (\varphi^2 - \varphi - 1)}{\varphi^2}$ (rather than $1 - \frac{1}{e}$). If $\varphi \geq \frac{1}{2}$ and the attacker starts in control, the defender should move at time $\epsilon > 0$, as close to 0 as possible. This is assuming that the player in control will remain in control if both players move at the same time, which we will assume throughout this chapter. Assuming ϵ adds a negligible loss, the defender gets a total expected score of $1 - \varphi^{\frac{\varphi}{1-\varphi}}$ (rather than $\frac{1}{e}$).

Figure 4.1 summarizes the results of this section, where we may see how much the defender expects to gain from exploiting her knowledge about nature's choice and the attacker's faulty assumptions about what she knows.

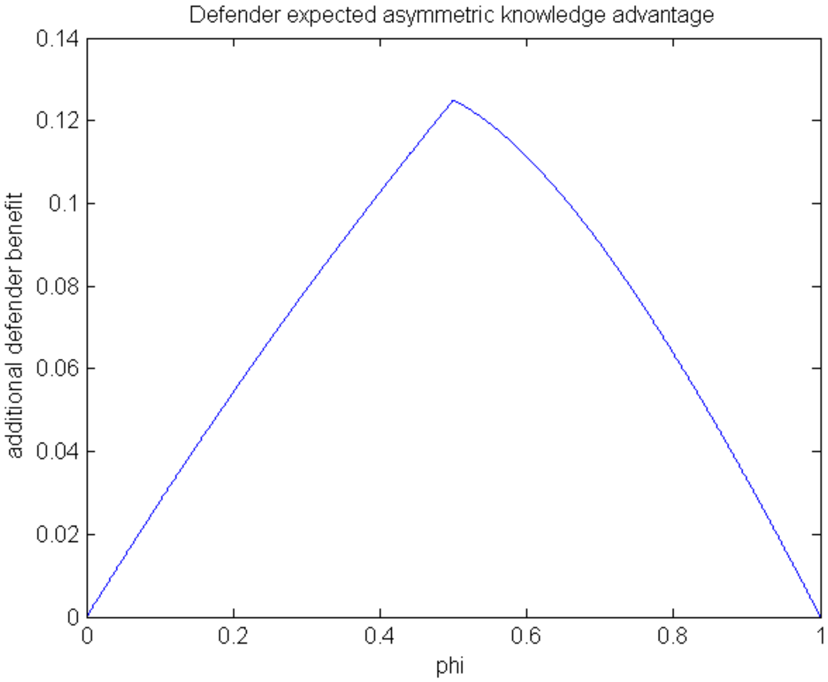

Figure 4.1: The expected advantage the defender gets from secretly knowing nature's choice of who is in control at the start of the game.

4.1.2 Defender is the Dupe

Suppose the attacker only knows the probability the defender has control, but she also knows that the defender knows nature's choice and will try to exploit her, assuming her beliefs and knowledge match what we assumed in Section 4.1.1. Can she exploit the defender?

Section 4.1.1 tells us what the defender's strategies will be, depending on nature's choice. If $\varphi < \frac{1}{2}$, then the attacker knows the defender will move at time $1-(1-\varphi)^{\frac{1}{\varphi}}$ with probability φ, and at time 0 with probability $1-\varphi$. If the attacker moves at time 0, her expected score will be $1-\varphi(1-\varphi)^{\frac{1}{\varphi}}$ (rather than $\varphi^{\frac{\varphi}{1-\varphi}}$).

If $\varphi \geq \frac{1}{2}$, then the attacker knows the defender will move at time $1-\varphi^{\frac{1}{1-\varphi}}$ with probability φ, and at time ϵ with probability $1-\varphi$. If the attacker knows and moves at time ϵ, assuming it is negligibly close to 0, her expected score will be $1-\varphi^{\frac{2-\varphi}{1-\varphi}}$ (rather than $1-\varphi^{\frac{\varphi}{1-\varphi}}$).

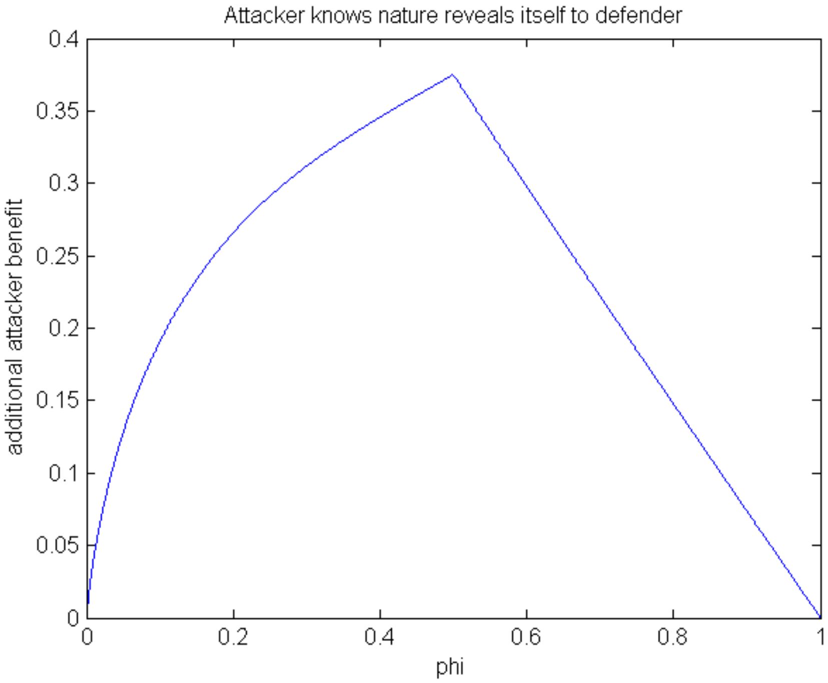

Figure 4.2: How well the attacker can exploit knowing the defender's secret.

As Figure 4.2 demonstrates, the attacker can strongly exploit her knowledge that the defender has a secret, and so the defender might not actually want to exploit her knowledge as previously shown if she has any doubt that her knowledge of nature's choice has leaked.

Since the attacker's strategy here was to uniformly play at time 0 (or ϵ) no matter what, further extensions along these lines, such as the defender knowing the attacker knows the defender knows nature's choice, and also that the attacker doesn't know the defender knows this, are degenerate and trivial.

4.2 Common Knowledge of Asymmetry

Now suppose both players know the defender knows nature's choice of who is in control at the start of the game, the attacker and defender both know φ, and both players know these starting conditions. In this case, the attacker will need to simultaneously optimize against two defender types: the one which knows it is in control at the beginning of the game, and the one which knows it is not in control at the beginning of the game.

We believe that in this case there is no longer any Nash equilibrium in the game. Previous work on games with asymmetric information like this have found equilibria when a game with the exact same initial conditions (including nature's choice) is repeated[2]. However, in our version the game is not repeated, so the defender has no incentive to hide any information, and would rather exploit it immediately. This intuitively seems to mean that if the defender starts out without control, she will probably make a move at around time 0 to maximize her score, and if the defender starts out with control, she will probably move at around time k, again to maximize her score. She has no incentive to act according to a distribution rather than at discrete points, so long as she knows the strategy the attacker has committed to. Given any mixed strategy the attacker may commit to, she would be better off moving as close to time 0 as she can in order to exploit the defender's exploitative strategy. So this situation breaks down and doesn't have a Nash equilibrium.

4.2.1 Defender Exploits Robust Attacker

If the attacker settles on a distribution from which to choose her move, the defender will be able to exploit it. Perhaps the best the attacker can do is choose a strategy which the defender can exploit the least. We could let the attacker assume she is playing the "defender" version of her opponent with some probability p (not necessarily φ), and play her optimal strategy as if against a known defender, $p_1(x) = \frac{\delta(x)}{e} + \frac{1}{e*(1-x)^2}$ the fraction p of the time. The rest of the time, she can play as if against a known attacker, $p_1(x) = \frac{1}{1-x}$. From the defender's point of view, the attacker is drawing from an overall distribution $p_1(x) = \frac{p\delta(x)}{e} + \frac{p}{e*(1-x)^2} + \frac{1-p}{1-x}$ over the range $[0, 1-\frac{1}{e}]$.

The defender may select a single point T at which she will derive her highest benefit, depending on whether nature chooses her or the attacker to be in control at the start of the game. When nature chooses the defender to be in control at the start of the game, we choose $T \in (0, k]$ to maximize

$$\begin{aligned}
E_0 &= \int_0^k \left(\tfrac{p\delta(x)}{e} + \tfrac{p}{e*(1-x)^2} + \tfrac{1-p}{1-x}\right) \left(x + \int_x^k (1-y)\,\delta(y-T)\,\mathrm{d}y\right) \mathrm{d}x \\
&= p + \tfrac{1-3p}{e} + (1-T) \int_0^T \left(\tfrac{p\delta(x)}{e} + \tfrac{p}{e*(1-x)^2} + \tfrac{1-p}{1-x}\right) \mathrm{d}x \\
&= p + \tfrac{1-3p}{e} + (1-T)(\tfrac{p}{e} + \tfrac{pT}{e(1-T)} - (1-p)ln(1-T)) \\
&= p + \tfrac{1-2p}{e} - (1-p)(1-T)ln(1-T)
\end{aligned} \tag{57}$$

We may take the derivative with respect to T to find that this is maximized for $T = k = 1 - \frac{1}{e}$. This matches our results in Section 4.1.1. Playing optimally when nature chooses the defender, $E_0 = \frac{2-(3-e)p}{e}$. When nature chooses the attacker to be in control at the start of the game, we choose $T \in (0, k]$ to maximize

$$\begin{aligned}
E_0 &= \int_0^k \left(\tfrac{p\delta(x)}{e} + \tfrac{p}{e*(1-x)^2} + \tfrac{1-p}{1-x}\right) \\
&\quad \left(\int_0^x (x-y)\,\delta(y-T)\,\mathrm{d}y + \int_x^k (1-y)\delta(y-T)\,\mathrm{d}y\right)\,\mathrm{d}x \\
&= \int_0^T \left(\tfrac{p\delta(x)}{e} + \tfrac{p}{e*(1-x)^2} + \tfrac{1-p}{1-x}\right)(1-T)\,\mathrm{d}x \\
&\quad + \int_T^k \left(\tfrac{p\delta(x)}{e} + \tfrac{p}{e*(1-x)^2} + \tfrac{1-p}{1-x}\right)(x-T)\,\mathrm{d}x \\
&= \tfrac{p}{e} - (1-p)(1-T)ln(1-T) \\
&\quad + \tfrac{p}{e} * (e - 1 - \tfrac{1}{1-T} - ln(1-T)) + (1-p) * (T + ln(1-T) + \tfrac{1}{e}) \\
&\quad + \tfrac{p}{e} * (e - \tfrac{1}{1-T}) + (1-p) * (ln(1-T) + 1) \\
&= p(1-T) + \tfrac{1-2p}{e} - \tfrac{p}{e}ln(1-T)
\end{aligned} \tag{58}$$

We may again take the derivative with respect to T to find that this is minimized for $T = 1 - \frac{1}{e}$. Thus, the best solution is at $T = \epsilon$, matching our results in Section 4.1.1. Playing optimally when nature chooses the attacker, $E_0 \approx \frac{1+(e-2)p}{e}$.

We can now combine these expected scores together to find the optimal p to minimize the defender's expected score, given a φ likelihood of nature choosing the defender to begin the game in control.

$$E_0 = \varphi \frac{2-(3-e)p}{e} + (1-\varphi)\frac{1+(e-2)p}{e} \tag{59}$$

$$\frac{\partial E_0}{\partial p} = \frac{e-2-\varphi}{e} \tag{60}$$

This means the attacker is indifferent to her choice of p when $\varphi = e - 2$. For $\varphi < e - 2$, $p = 0$ is her best choice (the defender can play her Nash equilibrium strategy when nature chooses the attacker to have control at the beginning of the game). For $\varphi > e - 2$, $p = 1$ is her best choice (the defender can play her Nash equilibrium strategy when nature chooses the defender). Playing their best strategies, then, where the defender knows nature's choice, we have

$$E_0 = \frac{1+\varphi}{e}, \varphi \leq (e-2) \tag{61}$$

$$E_0 = 1 - \frac{1}{e}, \varphi > (e-2) \tag{62}$$

Note that the defender can play a robust strategy against the attacker whenever the attacker correctly guessed the starting conditions of the game, and the rest of the time the defender exploits the fact that the attacker was wrong. This also means that part of the time the defender's strategy is fragile, and open to risky exploitation by the attacker.

4.2.2 Attacker Exploits Defender's Privilege

How much does the attacker stand to gain by trying to exploit the defender when the defender's additional information is common knowledge? For $\varphi < e - 2$, the defender believes the attacker is going to play as if she was in control at the start of the game. The fraction φ of the time, the attacker is wrong (the defender has control at the beginning) and the defender will exploit her by playing at time $1 - \frac{1}{e}$. The rest of the time the defender will play by her robust distribution $\frac{\delta(x)}{e} + \frac{1}{e(1-x)^2}$. The attacker can exploit these facts either by playing at time ϵ, or else by playing at time $1 - \frac{1}{e} + \epsilon$.

If the attacker plays at time ϵ, then she gains an expected score of $1 - \frac{1}{e}$ when the defender moves at time $1 - \frac{1}{e}$, and she also gains an expected score of $\frac{1}{e} + \int_0^{1-\frac{1}{e}} \frac{1}{e(1-x)^2}\,\mathrm{d}x = 1 - \frac{1}{e}$ when the defender plays by the distribution $\frac{\delta(x)}{e} + \frac{1}{e(1-x)^2}$. If the attacker plays at time $1 - \frac{1}{e} + \epsilon$, then she gains an expected score of $\frac{1}{e}$ when the defender moves at time $1 - \frac{1}{e}$, and she gains an expected score of $1 - \frac{1}{e}$ when the defender plays by the distribution $\frac{\delta(x)}{e} + \frac{1}{e(1-x)^2}$. Thus, the exploitative attacker will always play at time ϵ in this scenario.

For $\varphi > e - 2$, the defender believes the attacker is going to play as if she is not in control at the start of the game. The fraction $1 - \varphi$ of the time, the attacker is wrong and the defender will exploit her by playing at time ϵ. The rest of the time, the attacker is correct and the defender will play her robust distribution $\frac{1}{1-x}$. The attacker can exploit the defender, then, by always playing at time 2ϵ. Her expected score will now be $(1 - \varphi) + \frac{\varphi}{e}$.

Figure 4.3 shows how much the defender can expect to gain by playing optimally when her asymmetric knowledge advantage is known to the attacker. Compare this to Figure 4.1, which shows how much the defender can expect to gain by playing optimally when her asymmetric knowledge advantage is not known to the attacker, and note that in both cases, the defender's strategy is exactly the same.

We believe this is the best the attacker can do. If so, it shows that she should simply play against the version of her opponent who has the most to gain if she fails to play against that opponent. This relates to optimal strategies with incomplete information in [2], where the player with less information must equalize her payoffs against all potential versions of her adversary, weighted by the likelihood of playing against each version. One difference here is the attacker doesn't have the ability to equalize the payoffs against different versions of her opponent. The one that starts the game in control has a much larger advantage no matter how the attacker plays, and the one that starts the game without control has the most to gain by exploiting the attacker. So unless the attacker is quite certain she'll be playing against the version of her opponent that starts the game in control a large majority of the time, she maximizes her defenses against the other version of her opponent.

Figure 4.4 summarizes the results of this chapter to this point. It shows that additional, hidden knowledge is not necessarily advantageous. The defender can

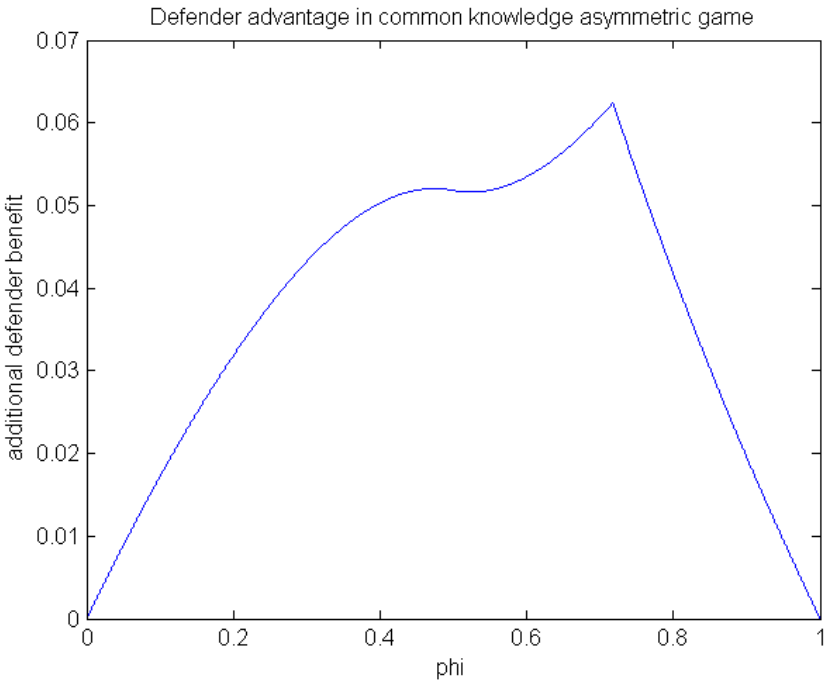

Figure 4.3: The maximum advantage the defender gets from publicly knowing nature's choice of who is in control at the start of the game.

choose to exploit her additional knowledge, but doing so makes her position more fragile, and open to exploitation. It may make strategic sense to announce or prove her additional knowledge.

4.3 Bluffing

We might consider what would happen if the defender decided to bluff, and pretend to knowledge of nature's choice without actually having it. If the attacker believes the defender and wants to play robustly, then she will play as if the attacker starts out in control so long as $\varphi < e - 2$. The defender will then play at time $k = 1 - \frac{1}{e}$ to get an expected score of $\frac{1}{e} * (1 - \varphi) + \frac{2\varphi}{e} = \frac{1+\varphi}{e}$. If the defender had moved at time 0 instead, she would always get an expected score of $\frac{1}{e}$.

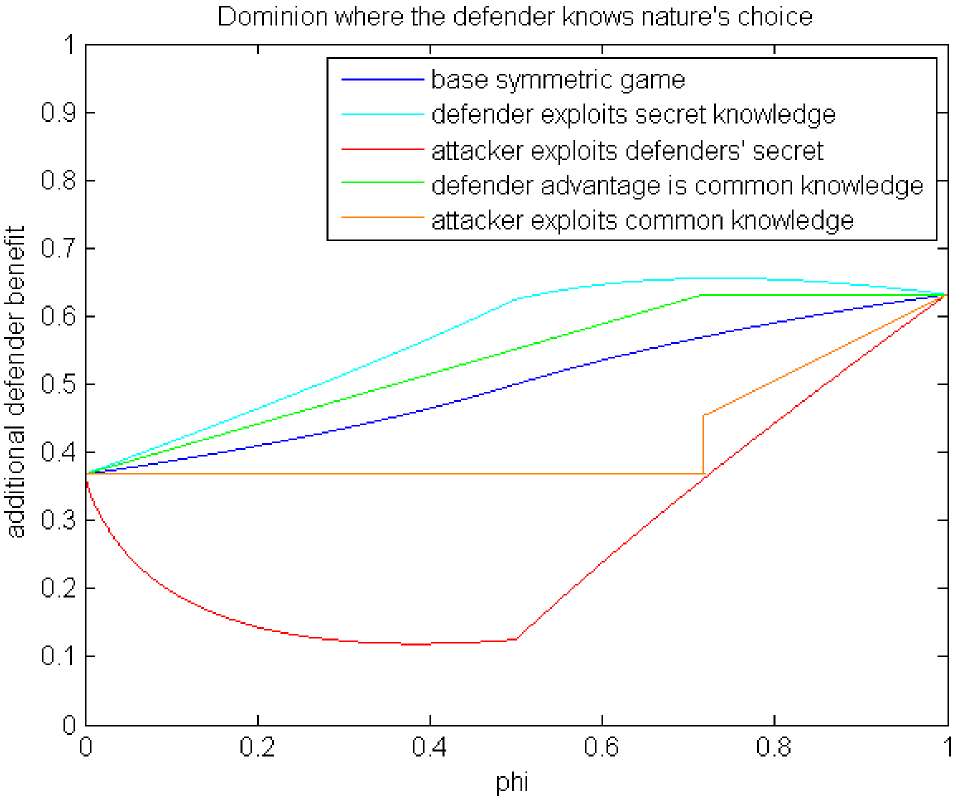

Figure 4.4: Losses and advantages based on who knows and does what.

If $\varphi > e - 2$, the gullible attacker would play as if she starts the game without control of the resource. The defender chooses T to maximize

$$\begin{aligned} E_0(T) &= (1-T)\int_0^T \tfrac{1}{e}(\delta(x) + \tfrac{1}{(1-x)^2})\,\mathrm{d}x \\ &\quad +(1-\varphi)\int_T^{1-\frac{1}{e}}(x-T)(\tfrac{1}{e}\tfrac{1}{(1-x)^2})\,\mathrm{d}x + \varphi\int_0^{1-\frac{1}{e}}\tfrac{1}{e}\tfrac{x}{(1-x)^2}\,\mathrm{d}x \\ &= (\tfrac{1-T}{e} + \tfrac{T}{e}) + (1-\varphi)(1 - T - \tfrac{1}{e}(2 + ln(1-T))) + \varphi(1 - \tfrac{2}{e}) \\ &= 1 + \tfrac{1}{e} - \tfrac{2\varphi}{e} - (1-\varphi)(T + \tfrac{1}{e}(2 + ln(1-T))) \end{aligned} \tag{63}$$

This is the defender's expected score for moving after the attacker plus her expected score for moving before the attacker when the defender doesn't have control in the beginning plus her expected score for being in control in the beginning times the likelihood of being in control in the beginning. It is maximized at $T = \epsilon$ for an expected score of approximately $1 - \frac{1}{e}$.

Comparing these results to those in Section 4.2.1, we find that the defender can derive the same benefits from successfully bluffing as she would had she actually known nature's choice for who controls the resource at the start of the game. Her strategy, however, is much more fragile, making bluffing very risky.

4.4 Other Starting Conditions

What happens when the only the defender knows nature's choice, and only the attacker knows φ? For the sake of brevity, we only concern ourselves with the case where these starting conditions are common knowledge.

The defender is uncertain about the attacker's strategy. However, in Section 4.2.1, we found that the defender doesn't depend upon φ to exploit the attacker, except to the extent of reducing the fragility of her strategy. The defender has the choice of playing robustly, as if the attacker also knew nature's choice, or she can play to exploit the attacker by choosing to move at a point ϵ when nature chooses the attacker to have control at the beginning, and to move at $1 - \frac{1}{e}$ when nature chooses the defender to have control at the beginning. She could note that her exploit is more fragile when nature chooses the attacker to be in control in the beginning and only exploit when she starts the game in control. She could also mix among these options to hedge against her risks. Of course, if the game is repeated and φ remains constant across the repetitions, she could estimate φ to bring the conditions closer to those in Section 4.2.1. In the absence of Nash equilibria, payoffs could be scaled to capture risk taking dispositions.

Finally, consider the case where the defender only probabilistically knows whether she has control in the beginning of the game. That is, some fraction of the time, she absolutely know who is in control, and the rest of the time she has the same knowledge as the attacker. This and related extensions could also be analyzed in a manner similar to how we analyzed other situations in this chapter. We leave this and other possibilities as exercises for ambitious readers, who should be able to construct solutions using this chapter as a guide.

Chapter 5

Future work

5.1 Additional extensions to Dominion

We hope that future work can consider many of the following questions, as well as come up with additional unique ways to look at and solve variants of Dominion.

How does the game change if the utility for a player is the sum of her expected score plus a function of the expected end state. It is reasonable to suppose that the player who ends the game in control should have some benefit proportional to the relative benefit of being in control at the beginning of the game, if we were to extend our thinking about Dominion into repeated interaction games.

What does equilibrium look like when both players only have approximate guesses for φ? What does optimal play look like when both players know that one player's knowledge of an initial parameter is wrong, but they only probabilistically know who is wrong?

What happens to optimal strategies as we extend Dominion into a multiple move game? Is there a closed form solution for different move counts? What do equilibria look like when we include additional players? How about additional move types, that provide different kinds of information upon moving?

What happens when the utility for control of the resource is a strictly increasing or decreasing function of x, rather than constant across the interval $[0, 1)$? How do optimal strategies change if the utility for controlling the resource is proportional to the longest period of control?

Our assumptions for optimal play included ”common knowledge” of the number of moves each player would make, the number of players, the structure of payoffs in the game, the length of the game, and the likelihood of each player's control at the outset of the game. What happens as we perturb these parameters, and also perturb the nature of the knowledge of game parameters from ”common” to some weaker form?

What are the best options when no one starts with control at the outset of the game, so it is no longer constant sum?

5.2 Dominion's Place Within Differential Games

Dominion makes a valuable contribution to the field of Differential Games[9]. It is different than most other differential games in that it is based on a discrete number of moves players are able to make, rather than continuous partial control of a utility function. In this way it functions as a special case of a game of timing[18].

Dominion in its basic form can be treated as a unit square game. Figure 5.1 shows the payoffs for the defender in the $\{1, 1\}$ Dominion game, where the lines represent gradients, moving from red being a low payoff, and yellow being a high payoff. The defender chooses the x coordinate and the attacker chooses the y coordinate, simultaneously. Since the game is constant sum, this square represents the entire game. The attacker payoffs are $1-$ the defender payoffs. For a fixed attacker choice, it's clear that the defender payoff function has a linear or constant gradient with one discontinuity. The same is true for the attacker.

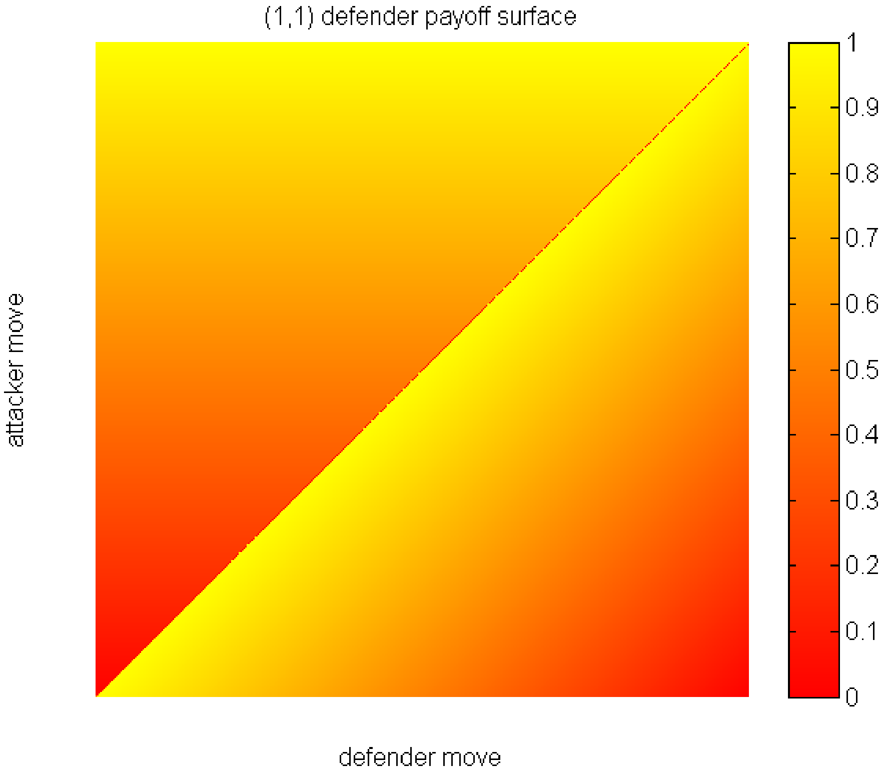

Figure 5.1: A complete visual description of the basic game.

Even though the game is very simple, its analysis has led to some pretty rich results. We could try to visualize extended versions of the game, where players have multiple moves, as a unit hypercube game. Unfortunately, we have to deal with the time-based reveal ordering of the moves, which makes Dominion not quite fit the paradigm. Also, players can't move ”backwards” in time, so if they were picking coordinates along some dimensional sequence, all their choices would be constrained inside an upper triangle of sorts. Given the extremely redundant and linear nature of the payoff function even in high dimensional cases, it may be possible to compress the set of moves and payoffs back into a unit square. This is simple to do if we modified Dominion so that a players' move doesn't give them any information about the state of the game, so that players make all their moves simultaneously. We might call this the blind version of Dominion.

In the blind version of Dominion, we believe a space compressing map could be constructed that walks through the hypercube describing the game in an ordered fashion while simultaneous preserving important gradient features contained in the hypercube. If this is indeed possible, then we could reinterpret the blind, $\{a, b\}$ move Dominion game as a unit square game in which the payoff for one player, holding the move of the other player constant, has a linear gradient with multiple discontinuities. We believe that solving for the equilibria in this, as well as the non-blind version of Dominion, will add considerably to the foundations of Differential Game Theory, and especially its extension into discrete and stealthy move games. It will also enable us to produce a fuller solution to the basic FlipIt game, and provide an alternate avenue of analysis for similar games.

There are many interesting questions and problems to explore in the field of stealthy move games.

References

[1] Arratia, R., Goldstein, L., and Kochman, F. Size bias for one and all, 2013.

[2] Aumann, R. J., Maschler, M. B., and Stearns, R. E. C. *Repeated games with incomplete information.* MIT press, Cambridge (Mass.), 1995.

[3] Benevides, M. R. F., Lima, I., Nader, R., and Rougemont, P. Using HMM in strategic games. In *Proceedings 9th International Workshop on Developments in Computational Models, DCM 2013, Buenos Aires, Argentina, 26 August 2013.* (2014), pp. 73–84.

[4] Bowers, K. D., van Dijk, M., Griffin, R., Juels, A., Oprea, A., Rivest, R. L., and Triandopoulos, N. Defending against the unknown enemy: Applying flipit to system security. *IACR Cryptology ePrint Archive 2012* (2012), 579.

[5] Bowers, K. D., van Dijk, M., Griffin, R., Juels, A., Oprea, A., Rivest, R. L., and Triandopoulos, N. Defending against the unknown enemy: Applying flipit to system security. In *GameSec* (2012), J. Grossklags and J. C. Walrand, Eds., vol. 7638 of *Lecture Notes in Computer Science*, Springer, pp. 248–263.

[6] Gilpin, A., and Sandholm, T. Solving two-person zero-sum repeated games of incomplete information. In *AAMAS (2)* (2008), L. Padgham, D. C. Parkes, J. P. Mller, and S. Parsons, Eds., IFAAMAS, pp. 903–910.

[7] Glicksberg, I. L. A further generalization of the kakutani fixed point theorem with application to nash equilibrium points. *Amer. Math. Soc. 3* (1952), 170–174.

[8] Halpern, J. Y., and Rong, N. Cooperative equilibrium. In *Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems: Volume 1 - Volume 1* (Richland, SC, 2010), AAMAS '10, International Foundation for Autonomous Agents and Multiagent Systems, pp. 1465–1466.

[9] Isaacs, R. *Differential Games: A Mathematical Theory with Applications to Warfare and Pursuit, Control and Optimization.* Dover books on mathematics. Dover Publications, 1999.

[10] Karlin, S. The theory of infinite games. *Annals of Mathematics 58*, 2 (1953), pp. 371–401.

[11] Kurisu, T. On a duel with time lag and arbitrary accuracy functions. *International Journal of Game Theory 19*, 4 (1991), 375–405.

[12] Laszka, A., Horvath, G., Felegyhazi, M., and Buttyn, L. Flipthem: Modeling targeted attacks with flipit for multiple resources. In *Decision and Game Theory for Security*, R. Poovendran and W. Saad, Eds., vol. 8840 of *Lecture Notes in Computer Science*. Springer International Publishing, 2014, pp. 175–194.

[13] Laszka, A., Johnson, B., and Grossklags, J. Mitigating covert compromises. In *Web and Internet Economics*, Y. Chen and N. Immorlica, Eds., vol. 8289 of *Lecture Notes in Computer Science*. Springer Berlin Heidelberg, 2013, pp. 319–332.

[14] Laszka, A., Johnson, B., and Grossklags, J. Mitigation of targeted and non-targeted covert attacks as a timing game. In *Decision and Game Theory for Security*, S. Das, C. Nita-Rotaru, and M. Kantarcioglu, Eds., vol. 8252 of *Lecture Notes in Computer Science*. Springer International Publishing, 2013, pp. 175–191.

[15] Nochenson, A., Grossklags, J., et al. A behavioral investigation of the flipit game. *Proceedings of the 12th Workshop on the Economics of Information Security (WEIS)* (2013).

[16] Nochenson, A., Grossklags, J., and Pennsylvania, T. A behavioral investigation of the flipit game. In *In 12th Workshop on the Economics of Information Security (WEIS* (2013).

[17] Pham, V., and Cid, C. Are we compromised? modelling security assessment games. In *Decision and Game Theory for Security*, J. Grossklags and J. Walrand, Eds., vol. 7638 of *Lecture Notes in Computer Science*. Springer Berlin Heidelberg, 2012, pp. 234–247.

[18] Radzik, T. *Results and problems in games of timing*, vol. Volume 30 of *Lecture Notes–Monograph Series*. Institute of Mathematical Statistics, Hayward, CA, 1996, pp. 269–292.

[19] Rasouli, M., Miehling, E., and Teneketzis, D. A supervisory control approach to dynamic cyber-security. In *Decision and Game Theory for Security*, R. Poovendran and W. Saad, Eds., vol. 8840 of *Lecture Notes in Computer Science*. Springer International Publishing, 2014, pp. 99–117.

[20] Rivest, R. L. Illegitimi non carborundum. Invited keynote talk given at CRYPTO 2011.

[21] Rossides, M., of Technology. Department of Electrical Engineering, M. I., and Science, C. *Extending the Analysis of the FlipIt Game*. 2013.

[22] Sudzute, D. General properties of nash equilibria in duels. *Lithuanian Mathematical Journal 23*, 4 (1983), 398–409. author real name is Sdite.

[23] van Dijk, M., Juels, A., Oprea, A., and Rivest, R. L. Flipit: The game of "stealthy takeover". Cryptology ePrint Archive, Report 2012/103, 2012. `http://eprint.iacr.org/`.

[24] van Dijk, M., Juels, A., Oprea, A., and Rivest, R. L. Flipit: The game of "stealthy takeover". *J. Cryptology 26*, 4 (2013), 655–713.

[25] Wellman, M. P., and Prakash, A. Empirical game-theoretic analysis of an adaptive cyber-defense scenario (preliminary report). In *Decision and Game Theory for Security*, R. Poovendran and W. Saad, Eds., vol. 8840 of *Lecture Notes in Computer Science*. Springer International Publishing, 2014, pp. 43–58.

[26] Wolfe, P., Tucker, A. W., and Dresher, M. Contributions to the theory of games, 3, 1967.