ELSEVIER

Theoretical Computer Science 288 (2002) 101–128

Theoretical Computer Science

www.elsevier.com/locate/tcs

On the complexity of data disjunctions ☆

Thomas Eiter[a,*], Helmut Veith[b]

[a] *Institut für Informationssysteme, Abteilung für Wissensbasierte Systeme, Technische Universität Wien, Favoritenstraße 9-11, A-1040 Wien, Austria*

[b] *Institut für Informationssysteme, Abteilung für Datenbanken und AI, Technische Universität Wien, Favoritenstraße 9-11, A-1040 Wien, Austria*

Abstract

We study the complexity of data disjunctions in disjunctive deductive databases (DDDBs). A data disjunction is a disjunctive ground clause $R(\bar{c}_{-1\,pt1})\cdots R(\bar{c}_k)$, $k \geqslant 2$, which is derived from the database such that all atoms in the clause involve the same predicate R. We consider the complexity of deciding existence and uniqueness of a minimal data disjunction, as well as actually computing one, both for propositional (data) and nonground (program) complexity of the database. Our results extend and complement previous results on the complexity of disjunctive databases, and provide newly developed tools for the analysis of the complexity of function computation.

Keywords: Data disjunction; Deductive databases; Computational complexity; Complexity upgrading; Conversion lemma

1. Introduction

During the past decades, a lot of research has been spent to overcome the limitations of conventional relational database systems. The field of deductive databases, which has emerged from logic programming [30], uses logic as a tool for representing and querying information from databases. Numerous logical query languages, which extend Horn clause programming for dealing with aspects such as incomplete or indefinite information, have been proposed to date, cf. [1, 34].

☆ The results of this paper have been presented at the international workshop "Colloquium Logicum: Complexity", Vienna, October 9–10, 1998. This work was partially supported by the Austrian Science Fund Project N Z29-INF.

* Corresponding author.

E-mail addresses: eiter@kr.tuwien.ac.at (T. Eiter); veith@dbai.tuwien.ac.at (H. Veith).

0304-3975/02/$ - see front matter
PII: S0304-3975(01)00147-5

In particular, the use of disjunction in rule heads for expressing indefinite information was proposed in Minker's seminal paper [33] which started interest in disjunctive logic programming [31, 11]. For example, the rule

$$lives\_in(x, us) \vee lives\_in(x, can) \vee lives\_in(x, mex) \leftarrow lives\_in(x, n\_america) \qquad (1)$$

informally states that a person living in North America lives in one of the three countries there. The semantical and computational aspects of disjunctive logic programming, and in particular, disjunctive deductive databases, have been investigated in many papers (see [34] for an overview).

In this paper, we are interested in a restricted type of disjunction which has been previously considered e.g. in [6, 5, 13, 20, 16]. A *data disjunction* [20] is a ground clause $R(\bar{c}_1) \vee \cdots \vee R(\bar{c}_k)$, $k \geqslant 2$, in which all atoms are different and involve the same predicate R. For example, the head of the rule (1) for $x = Joe$, is a data disjunction, as well as the disjunctive fact

$$loves(bill, monica) \vee loves(bill, hillary).$$

A data disjunction expresses indefinite information about the truth of a predicate on a set of arguments; in database terminology, it expresses a *null value* on this predicate, whose range is given by the arguments $\bar{c}_1, \ldots, \bar{c}_k$ of its atoms. In the context of deductive databases, null values of this form in the extensional database and their complexity have been considered e.g. in [21], and in many other papers.

If, in the above example, the fact $lives\_in(joe, n\_america)$ is known, then the data disjunction

$$\gamma = lives\_in(joe, us) \vee lives\_in(joe, can) \vee lives\_in(joe, mex)$$

can be derived from rule (1). If a clause γ is entailed from a database, then also any clause γ' subsumed by γ is entailed. For example, the clause $\gamma \vee lives\_in(joe, uzbekistan)$ is entailed by virtue of γ as well. We thus adopt the natural condition that a data disjunction γ must be minimal, i.e., no proper subclause of γ is entailed.

The question we address in this paper is the complexity of data disjunctions in a disjunctive deductive database (DDDB). Table 1 summarizes the problems studied (see Section 3 for precise definitions), and the main complexity results obtained. They complement previous results on reasoning from DDDBs. Deciding whether an arbitrary disjunction, rather than a data disjunction, follows from a DDDB has Π_2^P data and propositional complexity, and exponentially higher expression and combined complexity [15]; various syntactic restrictions lower the complexity to coNP or even polynomial time [10]. On the other hand, evaluating a conjunctive query over a disjunctive extensional database is coNP-complete [21], and hence deciding entailment of a single ground atom a has coNP data and propositional complexity. Thus, data disjunctions have intermediate complexity between arbitrary clauses and single atoms.

Observe that Table 1 contains also results on actually computing a data disjunction for a predicate (assuming at most one exists). While all the results in this table could

Table 1
Complexity of data disjunctions

Instance: A disjunctive deductive database $\mathsf{DB} = (\pi, E)$, where E is a collection of (possibly disjunctive) ground facts, and π are the inference rules, plus a distinguished relation symbol R.

Problem:

$\exists$DD: does DB have a data disjunction on R?
$\exists!$DD: does DB have a unique data disjunction on R?
λDD: Computation of the unique data disjunction on R.
k-λDD: Computation of the unique data disjunction on R, if it has at most k disjuncts (k constant).

Complexity:	*propositional c.* (π ground)	*data c.* (π fixed)	*expression c.* (E fixed)	*combined c.* (π, E vary)
$\exists$DD:	Θ_2^{P}	Θ_2^{P}	$\mathrm{PSpace}^{\mathrm{NP}}$	$\mathrm{PSpace}^{\mathrm{NP}}$
$\exists!$DD:	Θ_2^{P}	Θ_2^{P}	$\mathrm{PSpace}^{\mathrm{NP}}$	$\mathrm{PSpace}^{\mathrm{NP}}$
λDD:	$\mathrm{FP}_{\parallel}^{\mathrm{NP}}$	$\mathrm{FP}_{\parallel}^{\mathrm{NP}}$	$\mathrm{FPSpace}^{\mathrm{NP}}$	$\mathrm{FPSpace}^{\mathrm{NP}}$
k-λDD:	$\mathrm{FL}_{\log}^{\mathrm{NP}}[\log]$	$\mathrm{FL}_{\log}^{\mathrm{NP}}[\log]$	$\mathrm{FPSpace}^{\mathrm{NP}}[\mathrm{pol}]$	$\mathrm{FPSpace}^{\mathrm{NP}}[\mathrm{pol}]$

in principle be derived in the standard way, i.e., by proving membership in class C and reducing a chosen C-hard problem to the problem in question, we pursue here an "engineering" perspective of complexity analysis in databases, proposed e.g. in [19], which utilizes tools from descriptive and succinct complexity theory and exploits properties of the deductive database semantics. By means of these tools, hardness results can be derived at an abstracted level of consideration, without the need for choosing a fixed C-hard problem. Such tools (in particular, complexity upgrading) have been developed for decision problems, but are not available for function problems. We overcome this by a suitable generalization of the tools to the case of function problems.

The main contributions of this paper can be thus summarized as follows:

- Firstly, we determine the complexity of data disjunctions. We obtain natural and simple logical inference problems complete for the class Θ_2^{P} of the refined polynomial hierarchy [46], and, in their computational variants, complete problems for the function classes $\mathrm{FP}_{\parallel}^{\mathrm{NP}}$ and $\mathrm{FL}_{\log}^{\mathrm{NP}}[\log]$ and their exponential analogs.
- Secondly, we provide upgrading techniques for determining the complexity of function computations. They generalize available tools for decision problems and may be fruitfully applied in other contexts as well.

The rest of this paper is organized as follows. Section 2 states preliminaries, and Section 3 formalizes the problems. In Section 4, the decision problems are considered, while Section 5 is devoted to computing data disjunctions. Logical characterizations of function computations are given through a generalization of the *Stewart normal form* (SNF) [39, 40, 18], which has been introduced to characterize the class Θ_2^{P}. For deriving the expression and combined complexity of function computations, upgrading results are developed in Section 6. Section 7 considers restricted data disjunctions and applies the results to closed-world reasoning in databases. The final Section 8 concludes the paper.

2. Preliminaries

2.1. Deductive databases

For a background on disjunctive deductive databases, we refer to [31].

Syntax. A finite relational language is a tuple $\tau = (R_1, \ldots, R_n, c_1, \ldots, c_m)$ where the R_i are relation symbols (also called predicate symbols) with associated arities $a_1, \ldots, a_n$, and the c_i are constant symbols. An atom is a formula of the form $R_i(\bar{v})$, where $\bar{v}$ is a tuple of first-order variables and constant symbols.

A *disjunctive datalog rule* is a clause of the form

$$a_1 \vee \cdots \vee a_n \leftarrow b_1, \ldots, b_m, \quad 1 \leqslant n, \ 0 \leqslant m,$$

over a finite relational language, where the a_i's are atoms forming the *head* of the clause, and the b_j's are *atoms* or inequalities of the form $u \neq v$ (where u and v are variables or constants) forming the *body* of the clause.

A *disjunctive deductive program* (short *program*) is a finite collection of disjunctive datalog rules; it is *ground*, if no variables occur in the rules.

If a predicate symbol occurs only in rule bodies, it is called an *input predicate*, otherwise it is called a *derived predicate*.

A *disjunctive deductive program with input negation* is a program where input predicates are allowed to appear negated.

A *ground fact* is a clause of the form

$$a_1$$

where a_1 is a variable-free atom; a *disjunctive ground fact* is a clause of the form

$$a_1 \vee \cdots \vee a_n, \quad n \geqslant 1,$$

where the a_i's are variable-free atoms. Note that, for our concerns, any ground fact a_1 is semantically equivalent to the disjunction $a_1 \vee a_1$. We thus syntactically subsume ground facts by disjunctive ground facts, and distinguish them from proper disjunctions ($n \geqslant 2$) by referring to them as non-disjunctive ground facts.

A *disjunctive deductive database* (DDDB) is a tuple $\mathsf{DB} = (\pi, E)$ where π is a program and E is a finite set of disjunctive ground facts. Here, E represents the input database, also called the *extensional part*, and π are inference rules, called the *intensional part* of the database DB.

Remark 2.1. Note that π, E, and $\pi \cup E$ are all disjunctive deductive programs, i.e., ground facts can be included into the programs, and in fact we shall do this for defining the semantics. However, for methodological and complexity issues, it is important to distinguish the input data from the inference rules. For example, the complexity of evaluating DB is exponentially lower when π is fixed. In Section 3, we shall define data and expression complexity to give a formal meaning to this intuition.

Semantics. The semantics of DDDBs has been defined in terms of their minimal models [33, 31]. For a DDDB $\mathsf{DB} = (\pi, E)$, we denote by $\mathsf{HU}_{\mathsf{DB}}$ its Herbrand universe, i.e., the set of all constants occurring in DB.[1] The Herbrand base $\mathsf{HB}_{\mathsf{DB}}$ (resp., *disjunctive Herbrand base* $\mathsf{DHB}_{\mathsf{DB}}$) is the set of all ground atoms (resp., disjunctive ground facts) of predicates in DB over $\mathsf{HU}_{\mathsf{DB}}$. The ground instantiation of a program π over a set of constants C is denoted by ground(π, C); the ground instance of DB, denoted ground(DB), is ground$(\pi, \mathsf{HB}_{\mathsf{DB}}) \cup E$.

An (Herbrand) interpretation of DB is a subset $H \subseteq \mathsf{HB}_{\mathsf{DB}}$. An interpretation H of DB is a *model* of DB, if it satisfies each rule in ground(DB) in the standard sense. A model H of DB is minimal, if it does not contain any other model of DB properly; by MM(DB) we denote the set of all minimal models of DB. We write $\mathsf{DB} \models_{\mathrm{MM}} \varphi$ if a formula φ is true in every $M \in \mathrm{MM}(\mathsf{DB})$, and say that φ is entailed from DB.

Example 2.2. Let $\mathsf{DB} = (\pi, E)$, where π is the rule $q(x) \leftarrow p(x)$ and E contains the single disjunctive fact $p(a) \vee q(b)$. Then $M = \{p(a), q(a)\}$, $M' = \{p(a), q(a), q(b)\}$ are among the models of DB; M is minimal, while M' is not. The minimal models of DB are $\mathrm{MM}(\mathsf{DB}) = \{\{p(a), q(a)\}, \{q(b)\}\}$.

Remark 2.3. It is easy to see [33] that for each positive clause C, $\mathsf{DB} \models_{\mathrm{MM}} C$ if and only iff $\mathsf{DB} \models C$, where $\models$ denotes satisfaction in all models of DB. We will repeatedly use this fact.

The set of minimal models of DB has been characterized in terms of a unique least model-state *MS* (see [31]), i.e., a subset of $\mathsf{DHB}_{\mathsf{DB}}$, which can be computed by least fixpoint iteration of an operator T_P^S generalizing the standard T_P operator of logic programming [30]. In general, the computation of *MS* takes exponential space and time, even if the program π of DB is fixed.

2.1.1. Negation

Introducing negation in disjunctive deductive databases is not straightforward, and gave rise to different semantics, cf. [34]. In this paper, we restrict negation to *input negation*, i.e., the use of negated atoms $\neg R(\bar{t})$ in rule bodies where R is an extensional predicate, and adopt a *closed-world assumption* (*CWA*) *on models* by imposing the following condition: any accepted model M of $\mathsf{DB} = (\pi, E)$, restricted to the extensional part, must be a minimal model of E. Unless stated otherwise, a model of a DDDB must satisfy this kind of closed-world assumption.

Observe that this condition is satisfied by each $M \in \mathrm{MM}(\mathsf{DB})$ if π is negation-free; furthermore, if E contains no disjunctive facts, then $\neg R(\bar{c})$ is true in every $M \in \mathrm{MM}(\mathsf{DB})$ iff $R(\bar{c}) \notin E$.

[1] As usual, if no constant occurs in DB, we set $\mathsf{HU}_{\mathsf{DB}} := \{c\}$ for an arbitrary constant c.

As for complexity, it is easy to see that checking whether the restriction of M to its extensional part is a minimal model of E is possible in polynomial time. Hence, the complexity of model checking and of deciding $\mathsf{DB} \models_{\mathrm{MM}} \varphi$ is not increased by imposing the CWA on models. Furthermore, if E is restricted to disjunction-free ground facts, input negation can be eliminated in computation as follows.

Definition 2.4. Let τ be a finite relational language, and let $\tau' = \tau \cup \{R' \mid R \in \tau\}$. Then $\mathrm{NEG}_{\tau'}$ denotes the class of all finite τ' structures $\mathscr{A}$ where for all relations R in τ, $R'^{\mathscr{A}}$ is the complement of $R^{\mathscr{A}}$.

Proposition 2.5. *Extending a given τ-structure to its corresponding $\mathrm{NEG}_{\tau'}$ structure and replacing literals $\neg R(\bar{t})$ in a program π by $R'(\bar{t})$ is possible in LOGSPACE.*

In the derivation of hardness results, we shall consider DDDBs $\mathsf{DB} = (\pi, E)$ using input negation but where E is disjunction-free. Hence, all hardness results in this paper hold for DDDBs without negation and non-disjunctive (i.e., relational) facts as well.

2.2. Complexity

In this section, we introduce some of the more specific complexity classes and notions employed in the paper; we assume however some familiarity with basic notions of complexity theory such as oracle computations, NP, PSpace, L etc.

The class Θ_2^{p} contains the languages which are polynomial-time truth-table reducible to sets in NP. It has a wide range of different characterizations [46, 22]. In particular, the following classes coincide with Θ_2^{p}:

- $\mathrm{P}_{\|[k]}^{\mathrm{NP}}$: polynomial time computation with k rounds of parallel queries to an NP oracle [28, 9].
- $\mathrm{P}_{\log}^{\mathrm{NP}}$: polynomial time computation where the number of queries to an NP oracle is at most logarithmic in the input size [26].
- $\mathrm{L}_{\log}^{\mathrm{NP}}$: logarithmic space computation where the number of queries to an NP oracle is at most logarithmic in the input size [27]. [2]

For an overview of different characterizations and their history, consult [46, 22]. It is shown in [22, 4, 38, 41] that this picture changes when we turn to function computation. The above-mentioned list gives rise to at most three presumably different complexity classes $\mathrm{FP}_{\|}^{\mathrm{NP}}, \mathrm{FP}_{\log}^{\mathrm{NP}}$, and $\mathrm{FL}_{\log}^{\mathrm{NP}}$, which are shown in Fig. 1. Here, for any function class FC, we denote by FC[log] the restriction of FC to functions with logarithmic output size. Moreover, $\|[k]$ denotes k rounds of parallel queries, where k is a constant.

The relationships between the complexity classes in Fig. 1 have been attracting quite some research efforts which led to a number of interesting results.

- $\mathrm{II} = \mathrm{III}$ is equivalent to $\mathrm{P} = \mathrm{L}$ [22].

[2] Observe that the space for the oracle tape is not bounded. Unbounded oracle space is also assumed for all other classes using an oracle in this paper.

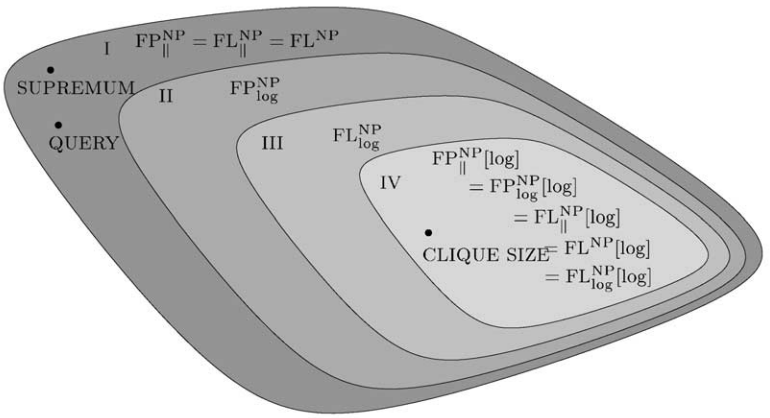

Fig. 1. Function classes corresponding to Θ_2^{P}.

- $\mathrm{I} = \mathrm{II}$ is equivalent to the property that SAT is $\mathrm{O}(\log n)$ approximable. This was shown in [2], after $\mathrm{I} \Rightarrow \mathrm{II}$ was proved in [8]. Here *f-approximability* of a set A means that there is a function g such that for all $x_1, \ldots, x_m$ where $m \geqslant f(\max_i |x_i|)$ it holds that $g(x_1, \ldots, x_m) \in \Sigma^m$ and $g(x_1, \ldots, x_m) \neq \chi_A(x_1, \ldots, x_m)$, where by abuse of notation $\chi_A(x_1, \ldots, x_m)$ denotes the function which maps the list of strings $x_1, \ldots, x_m$ to the m-bit vector whose ith bit is $\chi_A(x_i)$.
- Furthermore, if $\mathrm{I} = \mathrm{II}$, then (1SAT,SAT), i.e., promise SAT, is in P [4, 41]; $\mathrm{FewP} = \mathrm{P}$; $\mathrm{NP} = \mathrm{R}$ [38]; $\mathrm{coNP} = \mathrm{US}$; $\mathrm{SAT} \in \mathrm{NP}(n/\log^k n)$; and, furthermore, $\mathrm{NP} \subseteq \mathrm{DTIME}(2^{n^{\mathrm{O}(1/\log\log n)}})$ [22].

To compare the complexity of functions, and to obtain a notion of completeness in function classes, we use Krentel's notion of metric reducibility [26]:

Definition 2.6. A function f is *metric reducible* ($\leqslant^{\mathrm{mr}}$-reducible) to a function g (in symbols, $f \leqslant^{\mathrm{mr}} g$), if there is a pair (h_1, h_2) of polynomial-time computable functions h_1 and h_2 such that for every x, $f(x) = h_2(x, g(h_1(x)))$.

Proviso 1. Let C be a complexity class. Unless stated otherwise, we use the following convention: C-completeness is defined with respect to LOGSPACE reductions, if C is a class of decision problems, and with respect to metric reductions, if C is a class of function problems.

Some complete problems for function classes are shown in Fig. 1. The canonical $\mathrm{FP}_{\|}^{\mathrm{NP}}$-complete problem is QUERY, i.e., computing the string $\chi(I_1)\chi(I_2)\cdots\chi(I_n)$ of given SAT instances $I_1, \ldots, I_n$; SUPREMUM is computing, given a Boolean formula $F(x_1, \ldots, x_n)$, the string $s_1 \ldots s_n$ where $s_i = 1$ if there is a satisfying assignment to the variables of F such that $x_i = 1$, and $s_i = 0$ otherwise; CLIQUE SIZE is computing the size of a maximum clique in a given graph. Note that this problem is also complete for $\mathrm{FP}_{\log}^{\mathrm{NP}}$. All these problems, turned into proper decision problems, are Θ_2^{P}-complete. In

particular, deciding whether the maximum clique size in a graph is even and deciding whether the answer string to QUERY contains an even number of 1's are Θ_2^{P}-complete, cf. [46].

2.3. Queries and descriptive complexity

Definition 2.7. Let τ be a finite relational language, and let $\delta = \{R\}$ be a language containing a single relational symbol R. A query Q is a function which maps τ-structures to δ-structures over the same domain, such that $Q(\mathscr{A})$ and $Q(\mathscr{B})$ are isomorphic, if $\mathscr{A}$ and $\mathscr{B}$ are isomorphic. If R is nullary, then Q is a Boolean query.

A Boolean query Q is regarded as a mapping from τ-structures to $\{0,1\}$ such that for isomorphic $\mathscr{A},\mathscr{B}$, $Q(\mathscr{A}) = Q(\mathscr{B})$. If $Q(\mathscr{A}) = 1$, we also write $\mathscr{A} \models Q$.

Remark 2.8. (1) If we disregard queries of non-elementary complexity, we can identify queries with higher order definable relations. (2) Note that "query" is also used for oracle calls. (3) Since queries are functions, we shall also write them as sets of pairs $(\mathscr{A}, Q(\mathscr{A}))$.

Definition 2.9. Let τ be a finite relational language with a distinguished binary relation succ and two constant symbols min, max. Then SUCC_τ is the set of all finite structures $\mathscr{A}$ with at least two distinct elements where $\mathrm{succ}^{\mathscr{A}}$ is a successor relation on $|\mathscr{A}|$, and $\min^{\mathscr{A}}, \max^{\mathscr{A}}$ are the first and last element with respect to the successor relation, respectively.

Note that queries are not defined over SUCC_τ, but over arbitrary τ-structures; this is called "order independence" of queries. Many query languages however seem to require a built-in order for capturing complexity classes, i.e., capturing requires that the τ-structures are extended by an arbitrary contingent ordering to structures from SUCC_τ. Thus, when we talk about *ordered structures/databases*, or SUCC_τ, we mean that the queries are computed on τ-structures which are extended to SUCC_τ structures.

The following lemmata provide examples of this phenomenon.

Definition 2.10. An SNF formula (*Stewart normal form*) is a second-order formula φ of the form

$$\exists \bar{x}.\alpha(\bar{x},\bar{y}) \wedge \neg\beta(\bar{x},\bar{y}) \tag{2}$$

where α and β are Π_1^1 second-order formulas with equality having the free first-order variables $\bar{y}$. An SNF sentence is an SNF formula without free variables.

We call the standard Skolem functions for the variables $\bar{x}$ in formula (2) the *SNF witnesses* of φ.

Lemma 2.11 (Stewart [39, 40]; Gottlob [18]). *Every* Θ_2^{P}*-decidable property on* SUCC_τ *is expressible as*

$$\{\mathscr{A} \in \mathrm{SUCC}_\tau\colon\ \mathscr{A} \models \sigma\}$$

where σ *is an SNF sentence.*

This result, in equivalent terms of first-order logic with NP-computable generalized quantifiers, is contained for particular cases of generalized quantifiers in [39, 40], and was shown for broad classes of generalized quantifiers in [18].

On a structure $\mathscr{A}$, a formula $\varphi(\bar{x})$ with free variables $\bar{x}$ defines the relation $\varphi^{\mathscr{A}}$ by $\varphi^{\mathscr{A}} = \{\bar{c} \mid \mathscr{A} \models \varphi(\bar{c})\}$. A program π defines a relation R on $\mathscr{A}$, if $(\pi, \mathscr{A}) \models_{\mathrm{MM}} R(\bar{c})$ iff $\bar{c} \in R$, for every $\bar{c}$ on $\mathscr{A}$. In particular, if R is nullary, φ (resp., π) defines a property on $\mathscr{A}$.

Lemma 2.12 (Immediate from Eiter et al. [14, 15]). *Every* Π_1^1 *definable property* φ *on* SUCC_τ *is expressible by a disjunctive datalog program* π_φ *using input negation.*

Remark 2.13. Note that Lemma 2.12 does not require inequalities in rule bodies, since inequality is definable in the presence of order, cf. [15].

3. Data disjunctions

Definition 3.1. Given a DDDB DB, a disjunctive ground fact $\delta = R\bar{c}_1 \vee \cdots \vee R\bar{c}_n,\ n \geqslant 2$, is called a *data disjunction*, if
(1) $\mathsf{DB} \models_{\mathrm{MM}} R\bar{c}_1 \vee \cdots \vee R\bar{c}_n$, and
(2) for all $S \subset \{1, \ldots, n\}$, $\mathsf{DB} \not\models_{\mathrm{MM}} \bigvee_{i \in S} R\bar{c}_i$.
We say that DB has a data disjunction on R, if any such ground fact δ exists.

A data disjunction can be seen as a kind of null value in a data base.

Example 3.2. The DDDB $\mathsf{DB} = (\pi, E)$

$$\pi : Pc \vee Qx, \quad Pa \vee Py \leftarrow Qy, Rxy, \quad E : Rab, \quad Sa \vee Rbc$$

has a data disjunction $Pa \vee Pb \vee Pc$.

Definition 3.3. Given a DDDB DB, the *maximal disjunction* on R (in symbols, md(DB, R)) is the disjunctive ground fact

$$\bigvee \{R\bar{c} \in \mathrm{HB}_{\mathrm{DB}}\colon\ \mathrm{D}B \not\models_{\mathrm{MM}} R\bar{c}\}.$$

Lemma 3.4. DB *has a data disjunction on* R *if and only if* $\mathsf{DB} \models_{\mathrm{MM}} \mathrm{md}(\mathsf{DB}, R)$.

Proof. If DB has a data disjunction δ on R, then no atom $R\bar{c}$ of δ is implied by DB. Therefore, δ is a subclause of $\mathrm{md}(\mathsf{DB}, R)$, and thus $\mathsf{DB} \models \mathrm{md}(\mathsf{DB}, R)$. Since δ is

a positive clause, this implies $\mathsf{DB} \models_{\mathrm{MM}} \mathrm{md}(\mathsf{DB}, R)$. Conversely, if $\mathsf{DB} \models_{\mathrm{MM}} \mathrm{md}(\mathsf{DB}, R)$, then clearly $\mathrm{md}(\mathsf{DB}, R)$ is not empty. Either $\mathrm{md}(\mathsf{DB}, R)$ is a data disjunction itself, or atoms of $\mathrm{md}(\mathsf{DB}, R)$ can be removed until a minimal disjunction δ^* is reached such that $\mathsf{DB} \models_{\mathrm{MM}} \delta^*$. Since by definition no atomic subformula of $\mathrm{md}(\mathsf{DB}, R)$ is implied by DB, δ^* must contain at least two different atoms. □

In measuring the complexity of data disjunctions, we distinguish several cases following Vardi's [42] distinction between data complexity, expression complexity (alias program complexity), and combined complexity.

Definition 3.5. The problems ∃DD, ∃!DD, λDD, and k-λDD are defined as follows:
Instance: A DDDB $\mathsf{DB} = (\pi, E)$, and a relation symbol R.
Question: ∃DD: Does DB have a data disjunction on R?
∃!DD: Does DB have a unique data disjunction on R?
λDD: Compute the unique data disjunction on R if it exists, and # otherwise.
k-λDD: Compute the unique data disjunction on R, if it exists and has at most k disjuncts, and # otherwise.

Observe that ∃DD, called *ignorance test* in [5], has been used in [5, 6] to discriminate the expressive power of different query languages based on nonmonotonic logics over sets of disjunctive ground facts. Problem ∃!DD corresponds to the unique satisfiability problem. Note that the complexity of the uniqueness variant of a problem is often different. For example, in the case of the satisfiability problem, it is not known whether deciding unique satisfiability is in NP; this would clearly imply NP = coNP.

Definition 3.6. Let Π be any of the problems ∃DD, ∃!DD, λDD, or k-λDD.
- The *data complexity* of Π is the complexity of Π with parameter π fixed.
- The *expression complexity* of Π is the complexity of Π with parameter E fixed.
- The *propositional complexity* of Π is the complexity of Π where π is ground.
- The (unconstrained) complexity of Π is also called the *combined complexity* of Π.

Problem Π has combined (or propositional) complexity C, if Π is C-complete with respect to combined (resp. propositional) complexity. Π has data (or expression) complexity C, if Π is in C with respect to data (resp. expression) complexity for all choices of the parameter, and Π is C-complete with respect to data (resp. expression) complexity for a particular choice of the parameter.

4. Existence of data disjunctions

Our first result shows that deciding the existence of a data disjunction, considered as a Boolean query, is equivalent to evaluating a second-order sentence in SNF. Furthermore, the same is true for deciding the existence of a unique data disjunction.

Table 2

Algorithm **DDExistence**(DB, R)
Input: Disjunctive deductive database DB, relation symbol R;
Output: ''true'' if DB has a data disjunction on R,
''false'', otherwise.
1: $M := \emptyset$;
2: for all $R\bar{c} \in \mathsf{HB}_{\mathsf{DB}}$ do
3: if not (DB $\models_{\mathrm{MM}} R\bar{c}$) then $M := M \cup \{R\bar{c}\}$;
4: $\varphi := \bigvee_{R\bar{c} \in M} R\bar{c}$
5: if DB $\models_{\mathrm{MM}} \varphi$ then output true else output false;

Theorem 4.1. *Let Q be a fixed Boolean query. Over ordered databases, the following are equivalent*:
(1) *Q is Θ_2^{P}-computable.*
(2) *Q is definable by an SNF sentence.*
(3) *There exist a program π and a relation symbol R such that $(\pi, \mathscr{A})$ has a data disjunction over R iff $\mathscr{A} \models Q$.*
(4) *Q is equivalent to an SNF sentence whose SNF witnesses are* uniquely *defined.*
(5) *There exist a program π and a relation symbol R such that $(\pi, \mathscr{A})$ has a* unique *data disjunction over R iff $\mathscr{A} \models Q$.*

Proof. The equivalence of (1) and (2) is stated in [18]. A close inspection of the proof in [18] shows that in fact (1) is also equivalent to (4).

(3) $\rightarrow$ (1): By Lemma 3.4, algorithm **DDExistence** in Table 2 determines if DB has a data disjunction on R:

Note that in line 4, φ equals $\mathrm{md}(\mathsf{DB}, R)$. The algorithm **DDExistence** works in polynomial time and makes two rounds of parallel queries to an NP oracle, and thus the problem is in $\mathrm{P}_{\|[2]}^{\mathrm{NP}} = \Theta_2^P$.

(5) $\Rightarrow$ (1): Consider the algorithm **DDUniqueness** in Table 3, which is an extension of **DDExistence**. Note that lines 1 to 4 coincide with **DDExistence**.

On line 5, the algorithm terminates if no data disjunction exists. Otherwise, all possible data disjunctions are subclauses of $\varphi = \mathrm{md}(\mathsf{DB}, R)$. Lines 6 to 9 construct a subclause ψ; it contains all those literals $R\bar{c}$ of $\mathrm{md}(\mathsf{DB}, R)$ in N which cannot be removed from $\mathrm{md}(\mathsf{DB}, R)$ without destroying the data disjunction, i.e., it contains those literals which necessarily appear in every data disjunction. Thus, if φ is a data disjunction, it is the unique one. On the other hand, if a unique data disjunction exists, it is by construction equal to φ.

Like the algorithm **DDExistence**, this algorithm also works in polynomial time making a constant number of rounds of parallel queries to an NP oracle. Hence, the problem is in Θ_2^{P}.

(2) $\Rightarrow$ (3): Let φ be a second-order sentence of the form

$$\exists \bar{x}.\alpha(\bar{x}) \wedge \neg\beta(\bar{x}).$$

Table 3

Algorithm **DDUniqueness**(DB,R)
Input: Disjunctive deductive database DB, relation symbol R;
Output: "#" (resp., ''true'', ''false'') if DB has no (resp., a unique, more than one) data disjunction on R.
01: $M := \emptyset$;
02: for all $R\bar{c} \in \mathsf{HB}_{\mathsf{DB}}$
03: if not (DB $\models_{\mathrm{MM}} R\bar{c}$) then $M := M \cup \{R\bar{c}\}$;
04: $\varphi := \bigvee_{R\bar{c} \in M} R\bar{c}$;
05: if not (DB $\models_{\mathrm{MM}} \varphi$) then output # and halt;
06: $N := \emptyset$;
07: for all $R\bar{c} \in M$
08: if not (DB $\models_{\mathrm{MM}} \varphi \wedge \neg R\bar{c}$) then $N := N \cup \{R\bar{c}\}$;
09: $\psi := \bigvee_{R\bar{c} \in N} R\bar{c}$;
10: if DB $\models_{\mathrm{MM}} \psi$ then output true else output false;

By Lemma 2.12 there exist programs π_A and π_B containing predicate symbols A and B, respectively, such that for all $A(\bar{c}) \in \mathsf{HB}_{\pi_A}$ and $B(\bar{c}) \in \mathsf{HB}_{\pi_B}$,

$$(\pi_A, \mathscr{A}) \models_{\mathrm{MM}} A(\bar{c}) \text{ iff } \mathscr{A} \models \alpha(\bar{c}) \quad \text{and} \quad (\pi_B, \mathscr{A}) \models_{\mathrm{MM}} B(\bar{c}) \text{ iff } \mathscr{A} \models \beta(\bar{c}).$$

Let π be the program $\pi_A \cup \pi_B$ augmented with the rules

$$P(\bar{x}, \min) \vee P(\bar{x}, \max) \leftarrow A(\bar{x}),$$
$$P(\bar{x}, \min) \leftarrow B(\bar{x}).$$

Observe that P does not occur in $\pi_A \cup \pi_B$. Thus, by well-known modularity properties [15, Section 5], the minimal models of π on $\mathscr{A}$ are obtained by extending the minimal models of $\pi_A \cup \pi_B$ on $\mathscr{A}$.

It is easy to see that $\mathsf{DB} = (\mathscr{A}, \pi)$ has a data disjunction on P if and only if there exists a tuple $\bar{c}$ on $\mathscr{A}$ such that $\mathscr{A} \models \alpha(\bar{c}) \wedge \neg\beta(\bar{c})$. Thus, π indeed computes property φ on SUCC_τ.

(4) $\Rightarrow$ (5): From the equivalence of (2) and (4), it follows that the data disjunction of program π in the proof of (2) $\Rightarrow$ (3) is unique.

Corollary 4.2. *The propositional and data complexity of* $\exists$DD *and* $\exists!$DD *are in* Θ_2^{P}. *The expression and combined complexity of* $\exists$DD *and* $\exists!$DD *are in* $\mathrm{PSpace}^{\mathrm{NP}}$.

Proof. It remains to consider expression and combined complexity. When the program is not fixed, the size of $\mathsf{HB}_{\mathsf{DB}}$ is single exponential in the input, and thus the algorithm **DDExistence** takes exponentially many steps. The problem is thus in $\mathrm{EXPTIME}^{\mathrm{NP}}_{\|[2]}$, i.e., exponential time with two rounds of parallel NP oracles. This class coincides with $\mathrm{EXPTIME}^{\mathrm{NP}}_{\|}$, i.e., exponential time with a single round of parallel NP queries, which coincides with $\mathrm{PSpace}^{\mathrm{NP}}$ [19].

Since Θ_2^{P} has complete problems, we obtain from Theorem 4.1 the following.

Corollary 4.3. *There is a program π for which $\exists DD$ and $\exists!DD$ are Θ_2^P-hard.*

Hence, we obtain the announced result on the data and propositional complexity of data disjunctions.

Theorem 4.4. *The data complexity and propositional complexity of $\exists DD$ and $\exists!DD$ is Θ_2^P.*

Note that the propositional complexity of ∃DD has been stated in [13]. The hardness proof there, given by a standard reduction, is far more involved; this indicates the elegance of using the descriptional complexity approach.

Since the data complexity of a query language is uniquely determined by its expressive power, two languages with the same expressive power will always have the same data complexity. Hence, data complexity is a property of semantics. Expression and combined complexity, however, *depend on the syntax of the language*. It is thus, in general, not possible to determine the expression complexity of a query language $\mathscr{L}$ from its expressive power. Indeed, both the syntax and the semantics of $\mathscr{L}$ impact on its expression complexity. In spite of these principal obstacles, the *typical behavior* of expression complexity was often found to respect the following pattern: *If $\mathscr{L}$ captures C, then the expression complexity of $\mathscr{L}$ is hard for a complexity class exponentially harder than C.*

The query language of data disjunctions is an instance of this pattern, too:

Theorem 4.5. *The expression complexity and the combined complexity of $\exists DD$ and $\exists!DD$ is* PSpace^{NP}.

In the rest of this section, we give a proof of this result, which uses a general result linking the expressiveness of a query language to its expression complexity.

The main result of [19] shows that all query languages satisfying simple closure properties match indeed the above observation on expression complexity. Suppose that in a database domain elements are replaced by tuples of domain elements. This operation is natural when a database is redesigned; for instance, entries like “John Smith” in a database $\mathscr{A}$ can be replaced by tuples (“John”,“Smith”) in a database $\mathscr{B}$. It is natural to expect that a query $Q_{\mathscr{A}}$ over $\mathscr{B}$ can be easily rewritten into an equivalent query $Q_{\mathscr{B}}$ over $\mathscr{B}$. We call $Q_{\mathscr{B}}$ a *vectorized variant* of $Q_{\mathscr{A}}$. This is the essence of the following closure property.

Vector closure: A query language is *uniformly vector closed*, if the vectorized variants of query expressions can be computed in LOGSPACE.

Suppose again that a database $\mathscr{A}$ is replaced by a database $\mathscr{B}$ in such a way that all relations of $\mathscr{A}$ can be defined by views which use only unions and intersections of relations in $\mathscr{B}$. Then, it is again natural to expect that a query $Q_{\mathscr{A}}$ over $\mathscr{A}$ can be translated into an equivalent query $Q_{\mathscr{B}}$ over $\mathscr{B}$. In this case, we call $Q_{\mathscr{B}}$ an *interpretational variant* of $Q_{\mathscr{A}}$. We thus can formulate another closure property as follows.

Interpretation closure: A query language is *uniformly interpretation closed*, if the interpretational variants of queries can be computed from the database schemata in LOGSPACE.

In combination, we have the following closure condition (see [19] for a formal definition).

Definition 4.6. A query language is *uniformly closed*, if it is uniformly vector closed and uniformly interpretation closed.

This property, together with expressive capability of Θ_2^{P}, allows one to conclude the following lower bound on expression and combined complexity of a query language.

Lemma 4.7 (Gottlob et al. [19]). *If a query language is uniformly closed and expresses all Θ_2^{P} properties of* SUCC_τ, *then its expression complexity and combined complexity are at least* $\mathrm{PSpace}^{\mathrm{NP}}$.

As for disjunctive deductive databases, the following has been shown.

Lemma 4.8 (Gottlob et al. [19]). *The language of DDDBs is uniformly closed.*

Thus, Theorem 4.5 follows by combining Corollary 4.2, Lemma 4.7, and Lemma 4.8.

5. Computation of data disjunctions

In this section, we consider computing a data disjunction rather than deciding the existence.

5.1. Data complexity and propositional complexity

The following result is a functional analog to Theorem 4.1 for decision problems. It shows that the Stewart normal form also applies to non-Boolean queries over ordered databases.

Theorem 5.1. *Let Q be a fixed query. Then, over ordered databases the following are equivalent.*

(1) *Q is $\mathrm{FP}_{\|}^{\mathrm{NP}}$ computable.*
(2) *Q is definable by an SNF formula.*
(3) *There exist a program π and a relation symbol R such that $Q(\mathscr{A})$ is computable in time polynomial from the unique data disjunction of $(\mathscr{A}, \pi)$ on R.*

Proof. (1) $\Rightarrow$ (2): The problem of deciding whether a given tuple $\bar{c}$ on $\mathscr{A}$ fulfills $\bar{c} \in Q(\mathscr{A})$ is easily seen to be in Θ_2^{P}. Thus, Lemma 2.11 implies that there is an SNF sentence σ such that $\mathscr{A}, \bar{c} \models \sigma$ iff $\bar{c} \in Q(\mathscr{A})$ (provide $\bar{c}$ through a designated singleton

Table 4

DDDB program for $\mathrm{FP}_{\|}^{\mathrm{NP}}$ queries.

$$T(\overline{\min}) \tag{1}$$
$$S(\bar{u}) \vee T(\bar{v}) \leftarrow T(\bar{u}) \wedge \ \mathrm{succ}(\bar{u}, \bar{v}) \tag{2}$$
$$S(\overline{\max}) \leftarrow T(\overline{\max}) \tag{3}$$
$$R(\bar{x}, \bar{y}, \min, \min) \vee R(\bar{x}, \bar{y}, \min, \max) \vee R(\bar{x}, \bar{y}, \max, \max) \leftarrow S(\bar{x}\bar{y}) \tag{4}$$
$$R(\bar{x}, \bar{y}, \min, \min) \vee R(\bar{x}, \bar{y}, \min, \max) \leftarrow S(\bar{x}\bar{y}) \wedge A(\bar{x}\bar{y}) \tag{5}$$
$$R(\bar{x}, \bar{y}, \min, \min) \leftarrow S(\bar{x}\bar{y}) \wedge B(\bar{x}\bar{y}) \tag{6}$$

relation $R_{\bar{c}}$, and use $\exists \bar{y} R_{\bar{c}}(\bar{y})$ to access $\bar{c}$). Hence, there is an open SNF formula $\varphi(\bar{y})$ such that $\mathscr{A}, \bar{c} \models \varphi(\bar{y})$ iff $\bar{c} \in \pi(\mathscr{A})$.

(2) $\Rightarrow$ (3): Similarly as in the proof of Theorem 4.1, let φ be the SNF formula

$$\exists \bar{x}.\alpha(\bar{x}, \bar{y}) \wedge \neg\beta(\bar{x}, \bar{y})$$

having the free variables $\bar{y}$. By Lemma 2.12, there exist programs π_A and π_B containing predicate symbols A and B, respectively, such that for all $A(\bar{c}, \bar{d}) \in \mathsf{HB}_{\pi_A}$ and $B(\bar{c}, \bar{d}) \in \mathsf{HB}_{\pi_B}$ it holds that

$$(\mathscr{A}, \pi_A) \models_{\mathrm{MM}} A(\bar{c}, \bar{d}) \text{ iff } \mathscr{A} \models \alpha(\bar{c}, \bar{d})$$
$$\text{and} \quad (\mathscr{A}, \pi_B) \models_{\mathrm{MM}} B(\bar{c}, \bar{d}) \text{ iff } \mathscr{A} \models \beta(\bar{c}, \bar{d}).$$

We have to construct a program π whose unique data disjunction on input $\mathscr{A}$ over relation R contains the information about all tuples in $\varphi^{\mathscr{A}}$. To this end, consider the program in Table 4. Let a be the arity of A and B there. Then the new relation symbols T and S also have arity a, and R has arity $a+2$.

The program requires that a successor relation over tuples is available. The lexicographical successor relation can be easily defined using datalog rules. Then, lines 1 to 3 enforce that $S(\bar{c})$ holds for at least one $\bar{c}$. Consequently, the unique data disjunction on $\bar{c}$ is the positive clause containing all possible $S(\bar{c})$ on $\mathscr{A}$.

Consider lines 1 to 4 now. If the program would contain only these rules, then line 4 would enforce that the unique data disjunction on R would be the clause

$$\gamma = \bigvee_{\bar{c}, \bar{d} \text{ on } \mathscr{A}} R(\bar{c}, \bar{d}, \min, \min) \vee R(\bar{c}, \bar{d}, \min, \max) \vee R(\bar{c}, \bar{d}, \max, \max).$$

Lines 5 and 6, however, remove certain literals from clause γ. In particular, it holds that $\mathscr{A} \models \varphi(\bar{d})$ iff there is a $\bar{c}$ on $\mathscr{A}$ such that $\mathscr{A} \models \alpha(\bar{c}, \bar{d}) \wedge \neg\beta(\bar{c}, \bar{d})$ iff there is a $\bar{c}$ on $\mathscr{A}$ such that the (unique) data disjunction of $(\pi, \mathscr{A})$ on R contains the clause $R(\bar{c}, \bar{d}, \min, \min) \vee R(\bar{c}, \bar{d}, \min, \max)$ but not $R(\bar{c}, \bar{d}, \max, \max)$. Therefore, $\varphi^{\mathscr{A}}$ is polynomial time computable from the unique data disjunction on R.

(3) $\Rightarrow$ (1): Algorithm **DDUniqueness** in the proof of Theorem 4.1 will compute the unique data disjunction of $(\pi, \mathscr{A})$ in its variable ψ, provided it exists, in polynomial time with two rounds of parallel NP oracle access for fixed π, from which the result

of the query $Q(\mathscr{A})$ is, by the hypothesis, computed in polynomial time. Since $\mathrm{FP}_{\|[2]}^{\mathrm{NP}} = \mathrm{FP}_{\|}^{\mathrm{NP}}$, the result follows.

The data and propositional complexity of λDD is an easy corollary to this result.

Corollary 5.2. *Problem λDD has data complexity and propositional complexity* $\mathrm{FP}_{\|}^{\mathrm{NP}}$.

As we have discussed in Section 2.2, the size of the output may matter in the world of function computations, as far as the hardness of solving a problem is concerned. In the context of query languages, functions with small (logarithmic) output size may be related to the following special class of queries.

Definition 5.3. A *domain element query* (DEQ) is a query whose answer relation is a singleton, i.e., a query Q such that for all $\mathscr{A}$, it holds that $|Q(\mathscr{A})| = 1$.

For this class of queries, we obtain an analog to Theorem 5.1 which shows that again a variant of the Stewart normal form as well as computing data disjunctions may be used for its characterization.

Theorem 5.4. *Let Q be a fixed DEQ. Then, over ordered databases the following are equivalent*:

(1) *Q is* $\mathrm{FP}_{\|}^{\mathrm{NP}}[\log]$ *computable.*
(2) *$Q(\mathscr{A})$ is a tuple of SNF witnesses to an SNF sentence.*
(3) *There exist a program π and a relation symbol R such that $Q(\mathscr{A})$ is definable as a projection of the unique data disjunction over R, where the data disjunction contains at most two atoms.*
(4) *There exist a program π and a relation symbol R such that $Q(\mathscr{A})$ is computable in polynomial time from the unique data disjunction over R where the data disjunction contains at most two atoms.*

Proof. (2) $\Rightarrow$ (1): Let $\varphi(\bar{x})$ be the open SNF formula. It follows from Theorem that for constant $\bar{c}, \mathscr{A} \models \varphi(\bar{c})$ can be decided in Θ_2^{P} by a Turing machine M. Thus, we construct an $\mathrm{FP}_{\|}^{\mathrm{NP}}$ algorithm which loops over all polynomially many $\bar{c}$, simulates M on $\bar{c}$, and returns the unique answer. It is easy to see that the queries are non-adaptive.

(1) $\Rightarrow$ (2): For constant $\bar{c}$, the problem of deciding whether $\bar{c} \in \pi(\mathscr{A})$ is easily seen to be in Θ_2^{P}. Thus, there is an SNF formula $\varphi(\bar{x})$ having the free variables $\bar{x}$ for the constants such that $\mathscr{A}, \bar{c} \models \varphi(\bar{x})$ iff $\bar{c} \in \pi(\mathscr{A})$. By Theorem 4.1 we may w.l.o.g. suppose that the SNF witnesses in φ are unique.

(2) $\Rightarrow$ (3): Let φ be an SNF sentence

$$\exists \bar{x}.\alpha(\bar{x}) \wedge \neg\beta(\bar{x})$$

which has unique SNF witnesses. Then, the program for φ in part (2) $\Rightarrow$ (3) of the proof of Theorem 4.1 has a unique data disjunction of the form $P(\bar{c}, \bar{d}, \min) \vee P(\bar{c}, \bar{d}, \max)$. From this data disjunction, the projection on the $\bar{d}$ tuple yields the desired result.

(3) $\Rightarrow$ (4): Trivial.

(4) $\Rightarrow$ (1): Let M be the polynomial time Turing Machine which computes the answer from the unique data disjunction. Then, we use the algorithm **DDUniqueness** in the proof of Theorem 4.1 which computes the unique data disjunction in its variable ψ. It remains to check if the data disjunction is small, and to simulate M. By assumption, the result has logarithmic size.

The data and propositional complexity of k-λDD is an immediate corollary to this result.

Corollary 5.5. *The data complexity and propositional complexity of k-λDD are* $\mathrm{FL}_{\log}^{\mathrm{NP}}[\log]$.

At this point, the question arises whether we could not have surpassed the reduction in the proof of (2) $\Rightarrow$ (3) in Theorem 5.1, by exploiting the completeness result on k-λDD. The next result tells us that this is (presumably) not possible, and that, in general, disjunctions are needed which are not bounded by a constant in order to have hardness for $\mathrm{FP}_{\|}^{\mathrm{NP}}$.

Proposition 5.6. *Problem λDD is metric reducible to k-λDD with respect to data complexity, for some $k \geqslant 2$, if and only if* $\mathrm{FP}_{\|}^{\mathrm{NP}} = \mathrm{FP}_{\log}^{\mathrm{NP}}$.

Proof. In what follows, we denote for any class C of functions by $\leqslant^{\mathrm{mr}}(C)$ the closure of C under metric reductions.

($\Rightarrow$) Suppose λDD is metric reducible to k-λDD. Then, Corollary 5.2 implies that k-λDD is $\leqslant^{\mathrm{mr}}$-complete for $\mathrm{FP}_{\|}^{\mathrm{NP}}$. Since $\mathrm{FL}_{\log}^{\mathrm{NP}}[\log] \subseteq \mathrm{FP}_{\log}^{\mathrm{NP}}$, this implies $\mathrm{FP}_{\|}^{\mathrm{NP}} \subseteq \leqslant^{\mathrm{mr}}(\mathrm{FP}_{\log}^{\mathrm{NP}})$. Clearly, $\leqslant^{\mathrm{mr}}(\mathrm{FP}_{\log}^{\mathrm{NP}}) = \mathrm{FP}_{\log}^{\mathrm{NP}}$, and thus $\mathrm{FP}_{\|}^{\mathrm{NP}} \subseteq \mathrm{FP}_{\log}^{\mathrm{NP}}$. Combined with $\mathrm{FP}_{\log}^{\mathrm{NP}} \subseteq \mathrm{FP}_{\|}^{\mathrm{NP}}$ (cf. Fig. 1), it follows $\mathrm{FP}_{\|}^{\mathrm{NP}} = \mathrm{FP}_{\log}^{\mathrm{NP}}$.

($\Leftarrow$) Suppose that $\mathrm{FP}_{\log}^{\mathrm{NP}} = \mathrm{FP}_{\|}^{\mathrm{NP}}$. Let f be any function complete for $\mathrm{FP}_{\log}^{\mathrm{NP}}$ (such an f exists). Then, $\lambda\mathrm{DD} \leqslant^{\mathrm{mr}} f$ by hypothesis. We use the following fact.

Fact 5.7. *Every function f in* $\mathrm{FP}_{\log}^{\mathrm{NP}}$ *is* $\leqslant^{\mathrm{mr}}$*-reducible to some function g in* $\mathrm{FL}_{\log}^{\mathrm{NP}}[\log]$.

Indeed, Krentel showed that his class OptP$[O(\log n)]$ satisfies $\mathrm{FP}_{\log}^{\mathrm{NP}} \subseteq \leqslant^{\mathrm{mr}}(\mathrm{OptP}[O(\log n)])$, and that CLIQUE SIZE (cf. Section 2.2) is OptP$[O(\log n)]$-complete [26]. Since CLIQUE SIZE is clearly in $\mathrm{FP}_{\|}^{\mathrm{NP}}[\log] = \mathrm{FL}_{\log}^{\mathrm{NP}}[\log]$, Fact 5.7 follows by transitivity of $\leqslant^{\mathrm{mr}}$.

Now Fact 5.7 and Corollary 5.2, together with the hypothesis $\mathrm{FP}_{\log}^{\mathrm{NP}} = \mathrm{FP}_{\|}^{\mathrm{NP}}$ imply that

$$\lambda\mathrm{DD} \leqslant^{\mathrm{mr}} f \leqslant^{\mathrm{mr}} g \leqslant^{\mathrm{mr}} k\text{-}\lambda\mathrm{DD}.$$

By transitivity of $\leqslant^{\mathrm{mr}}$, we obtain $\lambda\mathrm{DD} \leqslant^{\mathrm{mr}} k$-$\lambda$DD.

5.2. Expression and combined complexity

Finally, we determine the expression complexity of computing the unique data disjunction.

Theorem 5.8. *The expression and combined complexity of λDD is* $\mathrm{FPSpace}^{\mathrm{NP}}$, *and the expression and combined complexity of k-λDD is* $\mathrm{FPSpace}^{\mathrm{NP}}[\mathrm{pol}]$.

The proof of the theorem uses succinct upgrade techniques for function problems whose inputs are given in succinct circuit description. These techniques are described in detail in the following section.

6. Problems with succinct inputs

6.1. Previous work and methodology

A problem is succinct, if its input is not given by a string as usual, but by a Boolean circuit which computes the bits of this string. For example, a graph can be represented by a circuit with $2n$ input gates, such that on input of two binary numbers v, w of length n, the circuit outputs true if there is an edge from vertex v to vertex w. In this way, a circuit of size $\mathrm{O}(n)$ can represent a graph with 2^n vertices. Suppose that a graph algorithm runs in time polynomial in the number of vertices. Then the natural algorithm on the succinctly represented graph runs in exponential time. Similarly, upper bounds for other time and space measures can be obtained.

The question of lower bounds for succinct problems has been studied in a series of papers about circuits, including [36, 23, 32, 25, 3, 7, 45], and also about other forms of succinctness such as representation by Boolean formulas or OBDDs [43, 44]. The first crucial step in these results is a so-called *conversion lemma*. It states that reductions between ordinary problems can be lifted to reductions between succinct problems:

Conversion Lemma. *If* $A \leqslant^X B$, *then* $s(A) \leqslant^Y s(B)$.

Here, $s(A)$ denotes the succinct version of A, while X and Y denote suitable notions of reducibility where $\leqslant^Y$ is transitive.

For the second step, an operator 'long' is introduced which is antagonistic to s in the sense that it reduces the complexity of its arguments. For a binary language A, $\mathrm{long}(A)$ can be taken as the set of strings w whose size $|w|$ written in binary is in A. Contrary to s, long contains instances which are exponentially larger than the input to A. For a complexity class C, $\mathrm{long}(C)$ is the set of languages $\mathrm{long}(A)$ for all $A \in C$. It remains to show a second lemma:

Compensation Lemma. $A \leqslant^Y s(\mathrm{long}(A))$.

Then the following theorem can be derived:

Theorem 6.1. *Let C_1, C_2 be complexity classes such that* $\log(C_1) \subseteq C_2$, *and let A be C_2-hard under $\leqslant^X$ reductions. Suppose that the Conversion Lemma and the Compensation Lemma hold. Then $s(A)$ is C_1-hard under $\leqslant^Y$ reductions.*

Proof. To show C_1-hardness, let B be an arbitrary problem in C_1. By assumption, $\log(B) \in C_2$, and therefore, $\log(B) \leqslant^X A$. By the Compensation Lemma, $B \leqslant^Y s(\log(B))$, and by the Conversion Lemma, we obtain $s(\log(B)) \leqslant^Y s(A)$. Since $\leqslant^Y$ is transitive, $s(A)$ is C_1-hard.

6.2. Queries on succinct inputs

For any τ-structure $\mathscr{A}$, let $enc(\mathscr{A})$ denote the encoding of $\mathscr{A}$ by a binary string. The standard way to encode $\mathscr{A}$ is to fix an order on the domain elements, and to concatenate the characteristic sequences of all relations in $\mathscr{A}$.[3] All Turing machine based algorithms (and in particular, all reductions) in fact work on $enc(\mathscr{A})$. Therefore, we shall usually identify $\mathscr{A}$ and $enc(\mathscr{A})$ without further notice. We use the further notation:

- $enc(\tau) \subseteq \{0,1\}^*$ denotes the language of all encodings of finite τ-structures.
- $\mathrm{char}(\mathscr{A})$ is the value of the binary number obtained by concatenating a leading 1 with $enc(\mathscr{A})$.

Given a Boolean circuit C with k input gates, $gen(\mathrm{C})$ denotes the binary string of size 2^k obtained by evaluating the circuit for all possible assignments in lexicographical order.

The idea of succinct representation is to represent $enc(\mathscr{A})$ in the form $gen(\mathrm{C})$. To overcome the mismatch between the fact that the size of $enc(\mathscr{A})$ can be almost arbitrary, while the size of $gen(\mathrm{C})$ has always the form 2^k, we use self-delimiting encodings:

Definition 6.2. Let $w = (x_1, \ldots, x_n) \in \{0,1\}^+$. The *self-delimiting* encoding of w is defined as $sd(w) = (x_1, 0, x_2, 0, \ldots, x_{n-1}, 0, x_n, 1)$. For a number n, $sd(n) = sd(\bar{n})$, where $\bar{n}$ denotes the binary representation of the number n.

Thus, from a string $sd(w)v$, where v is an arbitrary string, the bits of w can be easily retrieved as those on odd positions before the first 1 at an even position in the string.

Definition 6.3 (Veith [43]). For a language $L \subseteq \{0,1\}^*$, let $sd(L)$ be defined as the language

$$sd(L) = \{sd(|w|)wv : w \in L \text{ and } |sd(|w|)wv| = 2^r \text{ for some } r > 1\}.$$

[3] The characteristic sequence of a relation R is the binary string $\chi_R(\bar{c}_1) \cdots \chi_R(\bar{c}_m)$ where $\chi_R(x)$ is the characteristic function of R and $\bar{c}_1, \ldots, \bar{c}_m$ is the lexicographic enumeration of all tuples of the arity of R. For a graph, it is the concatenation of the rows of its adjacency matrix.

That is, $sd(L)$ is the language obtained from L by adding to every word in L the length descriptor and padding the string length to a power of 2.

Definition 6.4. A function f is computable in polylogarithmic time, if there are two polylogarithmic time bounded deterministic Turing machines N and M such that on input x, N computes the size of the output $|f(x)|$, and on input x and i, M computes the ith bit of $f(x)$. The class of all functions computable in polylogarithmic time is denoted by FPLT.

We call any reduction computed by an FPLT function a PLT reduction.

Modulo PLT reductions, self-delimiting encoding is equivalent to standard encoding:

Lemma 6.5 (Veith [43]). *Let $L \subseteq \{0,1\}^*$ be any nonempty language. Then $L \equiv^{\mathrm{PLT}} sd(L)$.*

In particular, this means that there exists an FPLT function 'extract', which extracts a word from its self-delimiting encoding.

Definition 6.6. Let τ be a relational signature. Associate with every Boolean circuit C a τ structure $gen_\tau(\mathrm{C})$ as follows:

$$gen_\tau(\mathrm{C}) = \begin{cases} \mathrm{extract}(gen(\mathrm{C})) & \text{if } gen(\mathrm{C}) \in sd(enc(\tau)), \\ \mathscr{A}_0, & \text{otherwise,} \end{cases}$$

where $\mathscr{A}_0$ is some default τ-structure. Let Q be any query on τ-structures. The succinct version of Q, denoted $s(Q)$, is given by

$$s(Q) = \{(\mathrm{C}, Q(\mathscr{A})) \colon \mathscr{A} = gen_\tau(\mathscr{A})\}.$$

A suitable instance of $\leqslant^Y$ reducibility as in the conversion lemma is given by a restriction of metric reductions which we call *forgetful*. In such reductions, the complexity of the inner function is restricted to FPLT and the outer "forgetful" function may not access the original problem input.

Definition 6.7. A function f is *forgetfully metric reducible* to a function g (in symbols, $f \leqslant_f^{\mathrm{mr}} g$), if there exist a FPLT function h_1 and a polynomial time computable function h_2 such that $f(x) = h_2(g(h_1(x)))$ for every string x.

It is not hard to see that $\leqslant_f^{\mathrm{mr}}$ is transitive. The crucial observation needed to generalize the results about succinct decision problems to succinct function problems is that the succinct representation affects only the inner computation in the metric reduction (i.e., h_1), because the result of the succinct function $s(F)$ is not succinct. Thus, if we are able to lift the inner reductions from ordinary instances to succinct instances, then we can leave the outer computation (i.e., h_2) unchanged. This lifting is achieved by the following lemma:

Lemma 6.8 (Immediate from Veith [45]). *Let f be a FPLT function which maps τ-structures to σ-structures. Then there exists an FPLT function F such that for every circuit* C *it holds that* $f(gen_\tau(\mathrm{C})) = gen_\sigma(F(\mathrm{C}))$.

A conversion lemma for query computations can now be shown as follows.

Lemma 6.9 (Conversion lemma for queries). *Let F be a query over τ-structures, and G be a query over σ-structures. If $F \leqslant_f^{\mathrm{mr}} G$ then $s(F) \leqslant_f^{\mathrm{mr}} s(G)$.*

Proof. By assumption we have $F(\mathscr{A}) = h_2(G(h_1(\mathscr{A})))$. We have to show that there exist an FPLT function H_1 and a polynomial time function H_2 such that

$$[(s(F)](\mathrm{C}) = H_2([s(G)](H_1(\mathrm{C}))).$$

By Lemma 6.8, there is an H_1 such that $h_1(gen_\tau(\mathrm{C})) = gen_\sigma(H_1(\mathrm{C}))$, and thus,

$$[s(G)](H_1(\mathrm{C})) = G(h_1(gen_\tau(\mathrm{C}))).$$

Then we can set $H_2 = h_2$, and the lemma is proved.

It remains to define a suitable 'long' operator. Recall that it has to simplify the complexity of its argument. Following [43], we define long on queries as follows:

Definition 6.10. Let $\delta = (R^1)$ be a signature with a single unary relation symbol, and let Q be a query over signature τ. Then the query $\mathrm{long}(Q)$ is defined as follows:

$$\mathrm{long}(Q) = \{(\mathscr{B}, Q(\mathscr{A})) \colon \mathscr{B} \text{ is a } \delta\text{-structure, } \mathrm{char}(\mathscr{A}) = |R^{\mathscr{B}}|\},$$

where $\mathrm{char}(\mathscr{A})$ is the value of the binary number obtained by concatenating a leading 1 with the characteristic sequence of the tuples in $\mathscr{A}$ in lexicographical order.

We now obtain a compensation lemma for query computations.

Lemma 6.11 (Compensation lemma for queries). *Let F be a query. Then $F \leqslant_f^{\mathrm{mr}} s(\mathrm{long}(F))$.*

Proof. As in Lemma 6.9, it is sufficient to show that every input $\mathscr{A}$ of F can be translated into a circuit C such that $|R^{gen_\delta(\mathrm{C})}| = \mathrm{char}(\mathscr{A})$. This was shown (using somewhat different terminology) in [43, Lemma 6].

We thus obtain the following upgrading theorem for query computations.

Theorem 6.12. *Let $\mathbf{F}_1, \mathbf{F}_2$ be two classes of functions, such that* $\mathrm{long}(\mathbf{F}_1) \subseteq \mathbf{F}_2$. *If a query F is hard for $\mathbf{F}_2$ under $\leqslant_f^{\mathrm{mr}}$-reductions, then $s(F)$ is hard for $\mathbf{F}_1$ under $\leqslant_f^{\mathrm{mr}}$-reductions.*

6.3. Succinctness and expression complexity

Succinct problems and expression complexity are related by the following method, which was used in [24, 15] and generalized in [19]. Suppose that a query language L can express some C_2-complete property A. Then, the data complexity of L is trivially at least C_2. If the language is rich enough to simulate a Boolean circuit by a program of roughly the same size, then it is possible to combine a program for A with a program for circuit simulation, thus obtaining a program for $s(A)$. Consequently, there is a reduction from $s(A)$ to the problem of expression complexity for L.

In [15], it was shown how a negation-free DDDB can simulate a Boolean circuit C, which is represented as collection of its gates. Each gate g is identified by an integer and described by a unique tuple (a, j, l) where $a \in \{in, and, or, not\}$ is the type of the gate and j, l are the identifiers of the gates resp. bits supplying input to g, where for $a = in$, j is the number of the bit in the input string accessed, and for $a \in \{in, not\}$, l is a dummy value. Let $\mathrm{C} = \{g_i = (a_i, j_i, l_l), 1 \leqslant i \leqslant t\}$ be a Boolean circuit that decides a k-ary predicate R over $\{0, 1\}$, i.e., for any tuple $\bar{t} \in \{0, 1\}^k$ supplied to C as input, a designated output gate of C, which we assume is g_t, has value 1 iff $\bar{t} \in R$.

We describe a program $\pi_c ircuit$ that simulates C using the universe $\{0, 1\}$. For each gate g_i, πC uses a k-ary predicate G_i, where $G_i(\bar{x})$ informally states that on input of tuple $\bar{x}$ to C, the circuit computation sets the output of g_i to 1. Moreover, it uses a nullary predicate *False*, which is true in those models in which the G_i do not have the intended interpretation; none of these models will be minimal.

The clauses of π_{C} are the following ones. For each gate $g_i = (a_i, j_i, l_i)$ of C, it contains the clause

$$(00)\ G_i(\bar{x}) \leftarrow False.$$

Depending on the type a_i, π_{C} contains for g_i additional clauses:

$a_i = in$:
$$(01)\ G_i(x_1, \ldots, x_{j_i-1}, 1, x_{j_i+1}, \ldots, x_k) \leftarrow .$$
$$(02)\quad False \leftarrow G_i(x_1, \ldots, x_{j_i-1}, 0, x_{j_i+1}, \ldots, x_k).$$

$a_i = not$:
$$(03)\ G_i(\bar{x}) \vee G_{j_i}(\bar{x}) \leftarrow .$$
$$(04)\quad False \leftarrow G_{j_i}(\bar{x}), G_i(\bar{x}).$$

$a_i = and$:
$$(05)\quad G_i(\bar{x}) \leftarrow G_{j_i}(\bar{x}), G_{l_i}(\bar{x}).$$
$$(06)\quad G_{j_i}(\bar{x}) \leftarrow G_i(\bar{x}).$$
$$(07)\quad G_{l_i}(\bar{x}) \leftarrow G_i(\bar{x}).$$

$a_i = or$:
$$(08)\quad G_i(\bar{x}) \leftarrow G_{j_i}(\bar{x}).$$
$$(09)\quad G_i(\bar{x}) \leftarrow G_{l_i}(\bar{x}).$$
$$(10)\ G_{j_i}(\bar{x}) \vee G_{l_i}(\bar{x}) \leftarrow G_i(\bar{x}).$$

The clauses (00) ensure that if a model of $ground(\pi_{\mathrm{C}}, \{0, 1\})$ contains *False*, then it is the maximal interpretation possible in which all facts are true (which is

clearly a model of π_C). In fact, this is the only model of π_C that contains *False*. Let M_C denote the interpretation given by $M_C = \{G_i(\bar{t}) \mid g_i \in C,\ \bar{t} \in \{0,1\}^k$, and g_i takes value 1 on input $\bar{t}$ to C$\}$.

Lemma 6.13 (Eiter et al. [15]). *For any Boolean circuit* C, M_C *is the unique minimal model of* π_C.

Armed with these results, we now prove the result stated in Section 5.2.

Proof of Theorem 5.8. Membership of λDD in FPSpace^{NP} and k-λDD in FPSpace^{NP} [pol], respectively, follows from the usual increase of the data complexity by a single exponential. We thus concentrate on the hardness parts.

λDD: circuit solved. It is not hard to see that $FP_{\|}^{NP}$ contains some query Q which is complete under $\leqslant_f^{mr}$-reductions. For example, QUERY from Section 2.2 is such a problem: the function h_1 in any metric reduction (h_1, h_2) of a function g to QUERY can be shifted inside the oracle queries in polylogarithmic time, and the bits of the input string x can be provided to h_2 through dummy oracle queries. Furthermore, long(FLinSpace^{NP}) $\subseteq FP_{\|}^{NP}$ holds. Thus, by Theorem 6.12, $s(Q)$ is complete for FLinSpace^{NP}. By standard padding arguments, completeness for FLinSpace^{NP} implies completeness for FPSpace^{NP}. Thus, it remains to reduce $s(Q)$ to λDD.

By Lemma 6.13, a circuit C with k input gates can be converted into a disjunctive program π_C whose k-ary output relation R describes the string *gen*(C). Consider the query $Q'(\mathscr{A}) = Q(\text{extract}(\mathscr{A}))$, where $\mathscr{A}$ is an ordered input structure which describes a string by a unary relation. Clearly, $Q'(\mathscr{A})$ is in $FP_{\|}^{NP}$. Theorem 5.1 thus implies that there exist a program π and a relation symbol S such that the unique data disjunction of $(\mathscr{A}, \pi)$ on S describes the result of $Q'(\mathscr{A})$. Since DDDBs are uniformly closed (Lemma 4.8), π can be rewritten into a program π' whose input relation R has arity k. As in the proof of Theorem 5.1, we can assume that there is a lexicographical successor relation on k-tuples computed by π'. By well-known modularity properties of DDDBs [15, Section 5], the program $\pi' \cup \pi_C$ indeed computes Q on $s(\mathscr{A})$. Since $\pi' \cup \pi_C$ can be constructed in polynomial time, it follows $s(Q) \leqslant^{mr} \lambda$DD.

k-λDD: Like FL^{NP}, also $FL_{\log}^{NP}$ contains $\leqslant_f^{mr}$-complete queries. For example, a variant of CLIQUE SIZE in which circuits C_1, C_2 computing the functions h_1, h_2 in a metric reduction of a function g to CLIQUE SIZE are part of the problem instance, is $\leqslant_f^{mr}$-complete for $FL_{\log}^{NP}$. The proof is then similar to the one for λDD.

7. Further results

In this section, we consider a restricted form of data disjunction and apply our results to closed-world reasoning in databases.

7.1. Restricted data disjunctions

In [16], a stronger notion of data disjunction $R(\bar{c}_1) \vee \cdots \vee R(\bar{c}_m)$ is considered, which requests in addition that all disjuncts $R(\bar{c}_i)$ are identical up to one argument of the list of constants $\bar{c}_i$; that is, if $\bar{c}_i = c_{i,1}, \ldots, c_{i,k}$ for $i \in \{1, \ldots, m\}$, then $c_{i,j} = c_{i',j}$ holds for all $i, i' \in \{1, \ldots, m\}$ and $j \in \{1, \ldots, k\} \setminus \{j'\}$, where j' is the argument on which the tuples $\bar{c}_i$ may disagree. We call any such data disjunction *restricted*. Note that all data disjunctions considered in Section 1 are restricted.

For the problems reformulated to restricted data disjunctions, Table 1 in Section 1 is the same except that the expression and combined complexity of λDD is $\mathrm{FPSpace}^{\mathrm{NP}}$[pol]. Indeed, a restricted data disjunction C has at most m disjuncts where $m = |\mathsf{HU}_{\mathsf{DB}}|$ is the number of constants, and thus λDD has $O(n^2 \log n)$ many output bits in the combined complexity case, where n is the size of DB. The number of maximal disjunctions $\mathrm{md}(\mathsf{DB}, R)$, adapted to restricted data disjunctions, is polynomial in the data size, and thus the same upper bounds can be easily derived as for unrestricted data disjunctions. All hardness results are immediate from the proofs except for propositional and data complexity of λDD; here, mapping tuples $\bar{c}$ of elements to newly introduced (polynomially many) domain elements is a suitable technique for adapting the construction in Table 4 in the proof of Theorem 5.1.

Finally, we remark that Lemma 2.12 remains true even if all disjunctions in the program π_φ describe restricted data disjunctions. Thus, by a slight adaptation of the programs in proofs and exploiting the fact that disjunction-free datalog with input negation is sufficient for upgrading purposes [15], the complexity results for restricted data disjunction remain true even if all disjunctions in DB must be restricted data disjunctions.

7.2. Closed world reasoning

The results on data disjunctions that we have derived above have an immediate application to related problems in closed world reasoning in databases.

Reiter [37] has introduced the closed-world assumption (CWA) as a principle for inferring negative information from a logical database. Formally,

$$CWA(\mathsf{DB}) = \mathrm{ground}(\mathsf{DB}) \cup \{\neg A \mid A \in \mathsf{HB}_{\mathsf{DB}} \text{ and } \mathsf{DB} \not\models A\}.$$

For example, $CWA(\{P(a), Q(b)\}) = \{P(a), Q(b), \neg P(b), \neg Q(a)\}$. It follows from results in [12] that computing $CWA(\mathsf{DB})$ has propositional complexity $\mathrm{FP}_{\|}^{\mathrm{NP}}$.

Observe that $CWA(\mathsf{DB})$ may not be classically consistent (under Herbrand interpretations); for example, $CWA(\{P \vee Q\}) = \{P \vee Q, \neg P, \neg Q\}$ which has no model. As shown in [12], deciding whether $CWA(\mathsf{DB})$ is consistent is in Θ_2^{P} and coNP-hard in the propositional case; the precise complexity of this problem is open.

In a refined notion of partial CWA (cf. [17]), which is in the spirit of protected circumscription [35], only atoms A involving a particular predicate P or, more general, predicates P_i from a list of predicates $\mathbf{P} = P_1, \ldots, P_n$ may be negatively concluded

from DB:

$$PCWA(\mathsf{DB};\mathbf{P}) = \text{ground}(\mathsf{DB}) \cup \{\neg P_i(\bar{c}) \mid P_i \in \mathbf{P},\ P_i(\bar{c}) \in \mathsf{HB}_{\mathsf{DB}},\ \mathsf{DB} \not\models P_i(\bar{c})\}.$$

E.g.,

$$CWA(\{P(a), Q(b)\}; P) = \{P(a), Q(b), \neg P(b)\}.$$

Semantically, the partial closed world assumption can be characterized in terms of minimal models.

Definition 7.1 (**P**-*minimal model*). Let $\mathbf{P} = P_1, \ldots, P_n$ be a list of predicates and DB a DDDB. For any model M of DB, let $M[\mathbf{P}] = \{P_i(\bar{c}) \in M\colon\ P_i \in \mathbf{P}\}$ be the restriction of M to **P**. Then, define the preorder $\preccurlyeq_{\mathbf{P}}$ on the models of a DB by $M \preccurlyeq_{\mathbf{P}} M' \Leftrightarrow M[\mathbf{P}] \subseteq M'[\mathbf{P}]$; as usual, $M' \prec M$ denotes $M' \preccurlyeq_{\mathbf{P}} M \wedge M \not\preccurlyeq_{\mathbf{P}} M'$. A model M is **P**-minimal for DB, if there exists no model M' such that $M' \prec M$.

We remark that a **P**-minimal model is a special case of the notion of **P**; **Z** minimal model [29], given by an empty list of fixed predicates in a circumscription.

The following proposition characterizes consistency of *PCWA* in terms of data disjunctions and **P**-minimal models, respectively.

Proposition 7.2. *Let* DB *be a DDDB and P a predicate. Then, the following statements are equivalent*:
(1) DB *has a data disjunction on P.*
(2) $PCWA(\mathsf{DB}; P)$ *is not consistent.*
(3) DB *does not have a single P-minimal model M.*

Proof. (1) $\Rightarrow$ (2): Suppose $\gamma = P(\bar{c}_1) \vee \cdots \vee P(\bar{c}_n)$, $n \geqslant 2$, is a data disjunction of DB. Then, $\mathsf{DB} \not\models P(\bar{c}_i)$, which means $\neg P(\bar{c}_i) \in PCWA(\mathsf{DB}; P)$, for every $i = 1, \ldots, n$. Since $\text{ground}(\mathsf{DB}) \subseteq PCWA(\mathsf{DB}; P)$, $\mathsf{DB} \models \gamma$ implies that $PCWA(\mathsf{DB}; P) \models \gamma$. Consequently, $PCWA(\mathsf{DB}; P)$ is not consistent.

$\neg(3) \Rightarrow \neg(2)$: Suppose DB has a single P-minimal model M. Then, for each atom $P(\bar{c}) \in \mathsf{HB}_{\mathsf{DB}}$ it holds that $\mathsf{DB} \not\models P(\bar{c})$ iff $M \not\models P(\bar{c})$, since M has the unique smallest P-part over all models of DB. Hence, M is model of $PCWA(\mathsf{DB}; P)$, and thus $PCWA(\mathsf{DB}; P)$ is consistent.

(3) $\Rightarrow$ (1): Suppose DB has not a single P-minimal model. Let $M_1, \ldots, M_n$ be the collection of all P-minimal models M of DB, where w.l.o.g. $M_1 \not\preccurlyeq_P M_2$ and $M_2 \not\preccurlyeq_P M_1$. Let $X = M_1[P] \cup M_2[P]$, and let $P(\bar{c}_1) \in M_1 \setminus M_2$ and $P(\bar{c}_2) \in M_2 \setminus M_1$ be arbitrary atoms. Then, $M_i \models \gamma$ where $\gamma = \bigvee\{P(\bar{c})\colon\ P(\bar{c}) \in (\mathsf{HB}_{\mathsf{DB}} \setminus X) \cup \{P(\bar{c}_1), P(\bar{c}_2)\}\}$ holds for every $i = 1, \ldots, n$. Indeed, clearly $M_1 \models \gamma$ and $M_2 \models \gamma$; if, for some $i > 2$, $M_i \not\models \gamma$, then $M_i[P] \subseteq X \subset M_1[P]$ would hold, which contradicts the P-minimality of M_1. It follows $\mathsf{DB} \models_{\text{MM}} \gamma$. Obviously, $\mathsf{DB} \not\models_{\text{MM}} \gamma \setminus P(\bar{c}_1)$ and $\mathsf{DB} \not\models_{\text{MM}} \gamma \setminus P(\bar{c}_2)$. It follows that γ contains a data disjunction on predicate P (which contains $P(\bar{c}_1) \vee P(\bar{c}_2)$).

As for **P**-minimality, the predicates in a list **P** of predicates can, by simple coding, be replaced with a single predicate P: replace atoms $P_1(x, y)$ and $P_2(z)$, etc. with $P(1, x, y)$, $P(2, z)$, etc. where the first argument in P codes the predicate. This coding is compatible with **P**-minimality, i.e., **P**-minimal and P-minimal models correspond as obvious. From Proposition 7.2 and the results of the previous sections, we thus obtain the following result.

Theorem 7.3. *Deciding consistency of* $PCWA(\mathsf{DB}; \mathbf{P})$ *and existence of a single* **P***-minimal model of* DB *have both* Θ_2^{P} *propositional and data complexity, and* $\mathrm{PSpace}^{\mathrm{NP}}$ *program and expression complexity.*

By the same coding technique, the result in Theorem 7.3 holds even if the language has only two predicates and **P** contains a single predicate. On the other hand, if the language has a only one predicate P, then the existence of a data disjunction on P is equivalent to the consistency of $CWA(\mathsf{DB})$.

8. Conclusion

In this paper, we have considered the complexity of some problems concerning data disjunctions in deductive databases. In our analysis, we have taken an "engineering perspective" on deriving complexity results using tools from descriptive complexity theory, and combined them with techniques for upgrading existing complexity results on ordinary instances to succinct representations of the problem input. In particular, we have also investigated the complexity of actually computing data disjunctions as a function, rather than only the associated decision problem. This led us to generalize upgrading techniques developed for decision problems to computations of functions. These upgrading results, in particular Theorem 6.12, may be conveniently used in other contexts.

The tools as used and provided in this paper allow for a high-level analysis of the complexity of problems, in the sense that establishing certain properties and schematic reductions are sufficient in order to derive intricate complexity results in a clean and transparent way, without the need to deal with particular problems in reductions. This is exemplified by our analysis of data disjunctions. While this high-level approach relieves us from spelling out detailed technical constructions, the understanding of what makes the problem computationally hard may be blurred. In particular, syntactical restrictions under which the complexity remains the same or is lowered can not be immediately inferred. We leave such considerations for further work. Another interesting issue for future work is the consideration of computing data disjunctions viewed as a multi-valued function, which we have not done here.

Acknowledgements

We are grateful to Iain Stewart and Georg Gottlob for discussions and remarks.

References

[1] S. Abiteboul, R. Hull, V. Vianu, Foundations of Databases, Addison-Wesley, Reading, MA, 1995.

[2] V. Arvind, J. Torán, Sparse sets, approximable sets, and parallel queries to NP, in: C. Meinel, S. Tison (Eds.), Proc. 16th Symp. on Theoretical Aspects of Computing (STACS '99), Lecture Notes in Computer Science, Vol, 1563, Trier, March 1999, Springer, Berlin, pp. 281–290.

[3] J. Balcázar, A. Lozano, J. Torán, The complexity of algorithmic problems on succinct instances, in: R. Baeta-Yates, U. Manber (Eds.), Computer Science, Plenum Press, New York, 1992, pp. 351–377.

[4] R. Beigel, NP-hard sets are P-superterse unless R=NP, Tech. Report TR4, CS Dept., Johns Hopkins University, 1988.

[5] P. Bonatti, Autoepistemic logics as a unifying framework for the semantics of logic programs, J. Logic Program. 22 (1995) 91–149.

[6] P.A. Bonatti, T. Eiter, Querying disjunctive databases through nonmonotonic logics, Theoret. Comput. Sci. 160 (1996) 321–363.

[7] B. Borchert, A. Lozano, Succinct circuit representations and leaf languages are basically the same concept, Inform. Process. Lett. 58 (1996) 211–215.

[8] H. Buhrmann, L. Fortnow, L. Torenvliet, Six hypotheses in search of a theorem, in: Proc. 12th IEEE Internat. Conf. on Computational Complexity (CCC '97), 1997, pp. 2–12.

[9] S. Buss, L. Hay, On truth-table reducibility to SAT, Inform. Comput. 91 (1991) 86–102.

[10] M. Cadoli, M. Lenzerini, The complexity of propositional closed world reasoning and circumscription, J. Comput. System Sci. 43 (1994) 165–211.

[11] J. Dix, Semantics of logic programs: their intuitions and formal properties. An overview, in: Fuhrmann, H. Rott (Ed.), Logic, Action and Information. Proc. Konstanz Coll. in Logic and Information (LogIn'92), DeGruyter, Berlin, 1995, pp. 241–329.

[12] T. Eiter, G. Gottlob, Propositional circumscription and extended closed world reasoning are Π_2^{P}-complete, Theoret. Comput. Sci. 114(2) (1993) 231–245, Addendum 118, 315.

[13] T. Eiter, G. Gottlob, The complexity class $\Theta_{\mathbf{2}}^{\mathbf{P}}$: recent results and applications in AI and modal logic, in: B. Chlebus, L. Czaja (Eds.), Proc. 11th Internat. Symp. on Fundamentals of Computation Theory (FCT '97), Lecture Notes in Computer Science, Vol. 1279, Springer, Berlin, 1997, pp. 1–18.

[14] T. Eiter, G. Gottlob, Y. Gurevich, Normal forms for second-order logic over finite structures, and classification of NP optimization problems, Ann. Pure Appl. Logic 78 (1996) 111–125.

[15] T. Eiter, G. Gottlob, H. Mannila, Disjunctive datalog, ACM Trans. Database Systems 22 (3) (1997) (364–417).

[16] Z.K.J. Fernández, J. Minker, A tractable class of disjunctive deductive databases, in: Proc. Workshop on Deductive Databases, JICSLP-92, Washington, DC, 1992. Tech. Report CITRI/TR-92-65, CS Dept, Univ. Melbourne, pp. 11–20.

[17] M. Genesereth, N. Nilsson, Logical Foundations of Artificial Intelligence, Morgan Kaufmann, Los Altos, CA, 1987.

[18] G. Gottlob, Relativized logspace and generalized quantifiers over ordered finite structures, J. Symbolic Logic 62 (2) (1997) 545–574.

[19] G. Gottlob, N. Leone, H. Veith, Succinctness as a source of complexity in logical formalisms, Ann. Pure Appl. Logic 97 (1999) 231–260.

[20] H. Hwang, A. Fu, H.-F. Leung, Data-disjunctive deductive databases, manuscript, 1997.

[21] T. Imielinski, K. Vadaparty, Complexity of query processing in databases with OR-objects, in: Proc. 8th ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems, 1989, pp. 51–65.

[22] B. Jenner, J. Toran, Computing functions with parallel queries to NP, Theoret. Comput. Sci. 141 (1995) 175–193.

[23] M. Karpinski, K. Wagner, The computational complexity of graph problems with succinct multigraph representation, Z. Oper. Res. (ZOR) 32 (1998) 201–211.

[24] P. Kolaitis, C.H. Papadimitriou, Why not negation by fixpoint? J. Comput. System Sci. 43 (1991) 125–144.

[25] M. Kowaluk, K. Wagner, Vector language: simple descriptions of hard instances, in: Mathematical Foundations of Computer Science (MFCS), Lecture Notes in Computer Science, Vol. 452, Springer, Berlin, 1992, pp. 378–384.

[26] M. Krentel, The complexity of optimization problems, J. Computer System Sci. 36 (1988) 490–509.
[27] R. Ladner, N. Lynch, Relativization questions about logspace computability, Math. Systems Theory 10 (1976) 19–32.
[28] R. Ladner, N. Lynch, A. Selman, Comparison of polynomial-time reducibilities, Theoret. Computer Sci. 1 (1975) 103–123.
[29] V. Lifschitz, Computing circumscription, in: Proc. Internat. Joint Conf. on Artificial Intelligence (IJCAI-85) 1985, pp. 121–127.
[30] J. Lloyd, Foundations of Logic Programming, Springer, Berlin, 1984, 1987.
[31] J. Lobo, J. Minker, A. Rajasekar, Foundations of Disjunctive Logic Programming, MIT Press, Cambridge, MA, 1992.
[32] A. Lozano, J. Balcázar, The complexity of graph problems for succinctly represented graphs, in: Proc. 15th Internat. Workshop on Graph-Theoretic Concepts in Computer Science, Lecture Notes in Computer Science, Vol. 411 Castle Rolduc, The Netherlands, 1989, pp. 277–286.
[33] J. Minker, On indefinite data bases and the closed world assumption, in: D. Loveland (Ed.), Proc. 6th Conf. on Automated Deduction (CADE '82), Lecture Notes in Computer Science, Vol. 138, New York, Springer, Berlin, 1982, pp. 292–308.
[34] J. Minker, Logic and databases: a 20 year retrospective, in: Proc. Internat. Workshop on Logic in Databases (LID '96), Lecture Notes in Computer Science, Vol. 1154 Springer, Berlin, 1996, pp. 3–57.
[35] J. Minker, D. Perlis, Computing protected circumscription, J. Logic Program. 2 (4) (1985) 235–249.
[36] C. Papadimitriou, M. Yannakakis, A note on succinct representations of graphs, Inform. Comput. 71 (1985) 181–185.
[37] R. Reiter, On closed-world databases, in: H. Gallaire, J. Minker (Eds.), Logic and Data Bases, Plenum Press, New York, 1978, pp. 55–76.
[38] A. Selman, A taxonomy of complexity classes of functions, J. Comput. System Sci. 48 (1994) 357–381.
[39] I. Stewart, Logical characterizations of bounded query classes I: logspace oracle machines, Fund. Inform. 18 (1993) 65–92.
[40] I. Stewart, Logical characterizations of bounded query classes II: polynomial-time oracle machines, Fund. Inform. 18 (1993) 93–105.
[41] S. Toda, On polynomial-time truth-reducibilities of intractable sets to P-selective sets, Math. Systems Theory 24 (1991) 69–82.
[42] M. Vardi, Complexity of relational query languages, in: Proc. 14th STOC, San Francisco, 1982, pp. 137–146.
[43] H. Veith, Languages represented by Boolean formulas, Inform. Process. Lett. 63 (1997) 251–256.
[44] H. Veith, How to encode a logical structure by an OBDD, in: 13th Ann. IEEE Conf. on Computational Complexity (CCC), 1998, pp. 122–131.
[45] H. Veith, Succinct representation, leaf languages, and projection reductions, Inform. Comput. 142 (2) (1998) 207–236.
[46] K. Wagner, Bounded query classes, SIAM J. Comput. 19 (5) (1990) 833–846.