

Queries with Guarded Negation

Vince Bárány
TU Darmstadt, Germany

barany@mathematik.tu-darmstadt.de

Balder ten Cate
UC Santa Cruz, CA, USA

btencate@ucsc.edu

Martin Otto
TU Darmstadt, Germany

otto@mathematik.tu-darmstadt.de

ABSTRACT

A well-established and fundamental insight in database theory is that negation (also known as complementation) tends to make queries difficult to process and difficult to reason about. Many basic problems are decidable and admit practical algorithms in the case of unions of conjunctive queries, but become difficult or even undecidable when queries are allowed to contain negation. Inspired by recent results in finite model theory, we consider a restricted form of negation, *guarded negation*. We introduce a fragment of SQL, called GN-SQL, as well as a fragment of Datalog with stratified negation, called GN-Datalog, that allow only guarded negation, and we show that these query languages are computationally well behaved, in terms of testing query containment, query evaluation, open-world query answering, and boundedness. GN-SQL and GN-Datalog subsume a number of well known query languages and constraint languages, such as unions of conjunctive queries, monadic Datalog, and frontier-guarded tgds. In addition, an analysis of standard benchmark workloads shows that many uses of negation in SQL in practice are guarded.

1. INTRODUCTION

A well-established and fundamental insight of database theory is that *negation* (also called *complementation* or *difference*) tends to make queries difficult to reason about. Recall that the unions of conjunctive queries are the first-order queries that can be expressed without using negation. Many basic problems are decidable and admit practical algorithms in the case of unions of conjunctive queries, but are undecidable in the case of arbitrary first-order queries. Examples include *query containment* and *open world query answering*.

We argue that most queries in practice use only a restricted form of negation, which is called *guarded negation* and was first considered in [7] (in the study of decidable fragments of first-order logic). By guarded negation we mean that queries may involve negative conditions only if these conditions, intuitively, pertain to

a single record in the database. For instance, if a database schema contains relations $\text{Author}(\text{AuthID}, \text{Name})$ and $\text{Book}(\text{AuthID}, \text{Title}, \text{Year}, \text{Publisher})$, the query that asks for *authors that did not publish any book with Elsevier* since “not publishing a book with Elsevier” is a property of an author. The query that asks for *pairs of authors and book titles where the author did not publish the book*, on the other hand, is not allowed, since it involves a negative condition (in this case, an inequality) pertaining to two values that do not necessarily co-occur in a record in the database. The requirement of guarded negation can be formally stated most easily in terms of the Relational Algebra: we allow the use of the difference operator $E_1 - E_2$ provided that E_1 is a projection of a relation from the database.

Based on an analysis of standard SQL benchmark workloads, we argue that guarded negation covers many uses of negation in SQL in practice. Furthermore, building on recent results in logic and finite model theory [7, 9], we show that queries with guarded negation are computationally very well behaved. For instance, query containment and open world query answering are decidable for first-order queries with guarded negation, and boundedness is decidable for the guarded-negation fragment of Datalog with stratified negation. We also determine the complexity of query evaluation for queries with guarded negation, which (under reasonable complexity theoretic assumptions) is easier than the same problem for queries with unguarded negation.

Our results show that guarded negation is a fruitful concept for databases, in the sense that it enables solving central decision problems in database theory more efficiently. We also believe that guarded negation is a fruitful concept from a more practical point of view, allowing for efficient query plans and query optimization strategies. This is something we are exploring in a separate line of investigation.

Outline and Main Results. In Section 2 we review the definition of GNFO, guarded-negation first-order logic, and GNFP, guarded-negation fixed-point logic, as well as the main known decidability and complexity results for these logics [7]. We also provide an equivalent characterization of GNFO in terms of the Relational Algebra.

In Section 3, we investigate what it means for an SQL query to be negation-guarded. Specifically, we identify syntactic restrictions on the use of negation in SQL queries, and we show that the first-order queries satisfying these restrictions can be translated to GNFO, and, in fact, are expressively complete for GNFO, in the sense of Codd’s completeness theorem. Furthermore, by means of an analysis of standard SQL benchmark workloads, we show that many

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Articles from this volume were invited to present their results at The 38th International Conference on Very Large Data Bases, August 27th - 31st 2012, Istanbul, Turkey.

Proceedings of the VLDB Endowment, Vol. 5, No. 11
Copyright 2012 VLDB Endowment 2150-8097/12/07... \$ 10.00.

SQL queries in practice satisfy the syntactic restrictions.

In Section 4, similarly, we introduce a syntactic fragment of Datalog with stratified negation, called GN-Datalog, which admits a translation into GNFP.

In Section 5, we show that GN-SQL and GN-Datalog subsume a number of important existing query languages and constraint languages. In particular, GN-Datalog subsumes both monadic Datalog (which it extends by allowing IDBs of arbitrary arity, and negation, subject to guardedness conditions) and unions of conjunctive queries.

In Section 6, we show that query containment is 2ExpTime-complete for GN-SQL queries as well as for GN-Datalog queries (note that the decidability of these problems follows via translations into GNFO and GNFP).

In Section 7, we determine the complexity of query evaluation and open-world query answering for GN-SQL and for GN-Datalog. While the data complexity of query evaluation is in PTime, both for GN-SQL and for GN-Datalog, in terms of combined complexity, the problem is complete for the complexity class $P^{NP[\log^2]}$ (for GN-SQL) and P^{NP} (for GN-Datalog). The data complexity of *open world query answering* for GN-SQL with respect to incomplete databases is coNP-complete. The problem can be solved in PTime for a considerable fragment of GN-SQL.

In Section 8, we prove decidability of the *boundedness* problem for GN-Datalog. Boundedness is a classical decision problem in the study of query optimization for recursive queries. It is known to be undecidable for Datalog, but decidable for monadic Datalog. Our result can be viewed as a powerful generalization of the decidability of boundedness for monadic Datalog queries [15].

We conclude in Section 9 by discussing possibilities for further extending GN-SQL and GN-Datalog.

2. PRELIMINARIES

In this section, we review definitions and results concerning the guarded-negation logics GNFO and GNFP from [7]. These results will be put to extensive use in the rest of this paper. We assume familiarity with the basic syntax and semantics of first-order logic.

For clarity, we will maintain a distinction between instances and structures: a structure has an associated domain, which may be a superset of its active domain, and which may depend on the structure in question. Furthermore, structures may interpret not only relation symbols but also constant symbols (which denote, not necessarily distinct, domain elements). Thus, instances may be viewed as a special case of structures, where the domain is the active domain and there are no constant symbols. Unless explicitly stated otherwise (by means of the adjective “unrestricted”), we always assume structures and instances to be finite.

GNFO. Guarded Negation First-Order Logic (GNFO) is the fragment of first-order logic consisting of all formulas built up from atomic formulas (including equalities) using conjunction, disjunction, existential quantification, and guarded negation, that is, negation in the specific form $\alpha \wedge \neg\phi$ where α is an atomic formula (possibly an equality statement) and all free variables of ϕ occur in α . Note that, since the guard α is allowed to be an equality statement, we are essentially able to negate any formula with at most one free variable (by writing $x = x \wedge \neg\phi(x)$). Formally, the

formulas of GNFO are generated by the recursive definition

$$\phi ::= R(t_1, \dots, t_n) \mid t_1 = t_2 \mid \phi_1 \wedge \phi_2 \mid \phi_1 \vee \phi_2 \mid \exists x\phi \mid \alpha \wedge \neg\phi$$

where each t_i is either a variable or a constant symbol, and, in the last clause, α is an atomic formula containing all free variables of ϕ . Note that function symbols (of arity greater than zero) are not considered.

In the above definition, we required α to be an atomic formula containing all free variables of the negated formula ϕ . Occasionally, it is convenient to allow a slightly more liberal syntax. Let us say that α is a *generalized guard* for ϕ if α is a disjunction of existentially quantified atomic formulas such that the free variables of ϕ are included in the free variables of each disjunct. One could extend GNFO by allowing generalized guards in the definition of guarded negation, thus admitting formulas such as $(\exists uv R(x, y, u, v) \vee \exists uv R(y, x, u, v)) \wedge \neg Sxy$. This would not increase the expressive power of GNFO: if a negation is guarded by a generalized guard, we can “pull out” the disjunction and the existential quantification to obtain an equivalent formula without generalized guards (at the cost of a possibly exponential blow-up in formula size). In particular, the above example can be equivalently expressed by $\exists uv(R(x, y, u, v) \wedge \neg Sxy) \vee \exists uv(R(y, x, u, v) \wedge \neg Sxy)$. Therefore, for simplicity, our definition of GNFO does not allow for generalized guards.

GNFP. Guarded Negation Fixed Point Logic (GNFP) further extends GNFO with an operator for least fixed points of positively definable monotone operations on relations. That is, we introduce second-order variables (also called fixed-point variables) of arbitrary arity, which may be used to form atomic formulas in the same way as ordinary relation symbols, and if ϕ is any GNFP formula, X an n -ary second-order variable ($n \geq 1$) occurring only positively in ϕ (i.e., under an even number of negations), $\mathbf{x} = x_1, \dots, x_n$ a sequence of first-order variables, and $\mathbf{t} = t_1, \dots, t_n$ a sequence of terms (first-order variables or constant symbols), and the free first-order variables of ϕ are included in \mathbf{x} , then

$$[\text{LFP}_{X, \mathbf{x}} \alpha \wedge \phi](\mathbf{t})$$

is also a formula of GNFP, where α is a generalized guard for ϕ , i.e., a disjunction of existentially quantified atomic formulas (involving only atomic relations, no second-order variables), such that all free first-order variables of ϕ are also free variables of each disjunct of α .

In the above formula, the LFP operator is a generalized quantifier binding the variables X and \mathbf{x} . The formula expresses that the tuple \mathbf{t} belongs to the least fixed-point of the monotone operation on relations defined by $\alpha \wedge \phi$. Incidentally, here, unlike in the case of GNFO, it is important that α is allowed to be a *generalized guard*.

In what follows, whenever we consider LFP formulas, we will always assume that they do not have any free second-order variables. The formal semantics of $[\text{LFP}_{X, \mathbf{x}} \alpha \wedge \phi](\mathbf{y})$ is the familiar one from least fixed-point logic, cf. [1]. If the formula ϕ has at most one free variable x , we may omit the guard α , which can be assumed to be the equality statement $x = x$. For example, the GNFP formula

$$[\text{LFP}_{X, x} P(x) \vee \exists y R(x, y) \wedge X(y)](z)$$

says that there is an R -path from z to some element in P .

Definability of Greatest Fixed Points. Besides the above *least fixed-point* operator, we can consider an analogous operator GFP for taking the *greatest fixed point* of a definable monotone operation on relations. However, as it turns out, it is possible to define the GFP operator in terms of the LFP operator (and vice versa) using a dualization via guarded negation. Specifically, $[\text{GFP}_{X,\mathbf{x}} \alpha(\mathbf{x}) \wedge \phi(\mathbf{x})](\mathbf{t})$ can be equivalently expressed as $\alpha(\mathbf{t}) \wedge \neg[\text{LFP}_{X,\mathbf{x}} \alpha(\mathbf{x}) \wedge \neg\phi'(\mathbf{x})](\mathbf{t})$, where ϕ' is obtained from ϕ by replacing all subformulas of the form $X(\mathbf{t}')$ by $\alpha(\mathbf{t}') \wedge \neg X(\mathbf{t}')$. For this reason, the above definition of GNFP does not include GFP as a primitive operator.

Definability of Simultaneous Fixed Points. It is common, in the literature on fixed point logics, to consider also a *simultaneous* least fixed point operator, that takes as arguments not a single formula but a tuple of formulas. More precisely, in the context of GNFP it is natural to consider also formulas of the form $[\text{LFP}_{X_i, S}](\mathbf{t})$ where

$$S = \begin{cases} X_1(\mathbf{x}_1) \leftarrow \alpha_1(\mathbf{x}_1) \wedge \phi_1(X_1, \dots, X_n, \mathbf{x}_1) \\ \vdots \\ X_n(\mathbf{x}_n) \leftarrow \alpha_n(\mathbf{x}_n) \wedge \phi_n(X_1, \dots, X_n, \mathbf{x}_n) \end{cases}$$

is a system of GNFP formulas, with each X_k a distinct second-order variable, whose arity matches the length of the tuple \mathbf{x}_k , and which occurs only positively in ϕ_1, \dots, ϕ_n , and where \mathbf{t} is a tuple of terms of the same length as \mathbf{x}_i . Here, the system S can be viewed as defining a monotone operation on tuples of relations, and the above formula expresses that \mathbf{t} belongs to the i -th component of the least fixed point of this operation. It is well known that simultaneous fixed point expressions of this form can be expressed equivalently using a nesting of ordinary, single-variable, fixed point operators, possibly at the cost of an exponential blow-up in formula size (cf. for example [2]). The transformation preserves guardedness. Hence, extending GNFP with such a simultaneous least fixed point operator does not increase its expressive power.

Disjunctive Normal Form for GNFO and Width. We say that a GNFO formula is in *Disjunctive Normal Form* (DNF) if it is a disjunction of disjunction-free GNFO formulas, no existential quantifier occurs directly below a conjunction sign, and no conjunction sign occurs directly below a negation sign. Equivalently, a GNFO formula is in DNF if it is a disjunction of GNFO formulas ϕ generated by the following recursive definition:

$$\begin{aligned} \phi & ::= \exists x_1, \dots, x_n (\zeta_1 \wedge \dots \wedge \zeta_m) \\ \zeta & ::= R(t_1, \dots, t_n) \mid (t_1 = t_2) \mid \alpha \wedge \neg\phi \end{aligned} \quad (1)$$

where, in the last clause, α is an atomic formula containing all free variables of ϕ . Every GNFO formula is equivalent to one in DNF, of possibly exponential size, that can be obtained by repeatedly applying the following equivalences.

$$\begin{aligned} (\exists x\phi) \wedge \psi & \simeq \exists x'(\phi[x'/x] \wedge \psi), & \phi \wedge (\psi \vee \chi) & \simeq (\phi \wedge \psi) \vee (\phi \wedge \chi) \\ \exists x(\phi \vee \psi) & \simeq \exists x\phi \vee \exists x\psi, & \alpha \wedge \neg(\phi \wedge \psi) & \simeq (\alpha \wedge \neg\phi) \vee (\alpha \wedge \neg\psi) \end{aligned}$$

The *width* of a GNFO formula ϕ is the number of variables occurring (free or bound) in the DNF formula obtained from ϕ by applying the above rules.

A *union of conjunctive queries* (UCQ) is a GNFO formula in DNF without negation. Thus, GNFO can be naturally viewed as an extension of UCQs with guarded negation.

Known Decidability and Complexity Results. The following theorem summarizes what is known about GNFO and GNFP that is relevant for present purposes. Recall that the satisfiability problem has as input a formula $\phi(\mathbf{x})$, and asks whether there exists a structure M and a tuple of elements \mathbf{a} such that $M \models \phi(\mathbf{a})$. The entailment problem takes as input two formulas $\phi(\mathbf{x}), \psi(\mathbf{x})$, and asks whether it is the case that, for every structure M and for every tuple of elements \mathbf{a} , $M \models \phi(\mathbf{a})$ implies $M \models \psi(\mathbf{a})$. The model checking problem has as input a formula $\phi(\mathbf{x})$, a structure M , and a tuple of elements \mathbf{a} , and asks whether $M \models \phi(\mathbf{a})$.

- Theorem 1 ([7])** 1. *The satisfiability problem and the entailment problem for GNFO and for GNFP are decidable and 2ExpTime-complete. This holds both for finite structures and for unrestricted structures.*
2. *For GNFO formulas, satisfiability over finite structures coincides with satisfiability over unrestricted structures, and similarly for entailment. The same does not hold for GNFP.*
3. *The model checking problem for GNFO is $\text{P}^{\text{NP}[\log^2]}$ -complete (combined complexity). For GNFP, the problem is P^{NP} -hard and is contained in $\text{NP}^{\text{NP}} \cap \text{coNP}^{\text{NP}}$.*

In the above theorem, $\text{P}^{\text{NP}[\log^2]}$ refers to those problems that can be solved by a polynomial time deterministic algorithm that is allowed to ask $O(\log^2(n))$ queries to an NP-oracle, cf. Section 7.1. A close analysis of the 2ExpTime upper bound argument for the satisfiability and entailment problems of GNFP shows that these results extend to the case with simultaneous fixed-point operators (both on finite structures and on unrestricted structures).¹

2.1 Guarded Negation in Relational Algebra

The concept of guarded negation can be equivalently cast in terms of the Relational Algebra, where negation is expressed by means of the difference operator. Consider the Relational Algebra (RA) defined over a schema consisting of relation symbols of specified arity using the following primitive operators (cf. [1] for their semantics).

Atomic Relations: every relation symbol belongs to RA.

Selection: if $E \in \text{RA}$ has arity k and $1 \leq i, j \leq k$, then $\sigma_{i=j}(E)$ belongs to RA and has arity k .

Projection: if $E \in \text{RA}$ has arity k and $1 \leq i_1, \dots, i_n \leq k$, then $\pi_{i_1 \dots i_n}(E)$ belongs to RA and has arity n .

Crossproduct: if $E_1, E_2 \in \text{RA}$ have arity k and n , respectively, then $E_1 \times E_2$ belongs to RA and has arity $k+n$.

¹Specifically, the proof of the 2ExpTime upper bound for GNFP is based on a satisfiability preserving translation from GNFP to guarded fixed-point logic (GFP). The translation may give rise to an exponential blow-up in the size of the formula, but it preserves the width (following a suitable definition of width, analogous to the definition of width for GNFO formulas). The satisfiability problem for GFP formulas, in turn, is decidable in time $2^{\text{poly}(|\phi|) \cdot \text{exp}(\text{width}(\phi))}$ (where $\text{poly}(n)$ is short for $n^{O(1)}$ and $\text{exp}(n)$ is short for $2^{\text{poly}(n)}$) by a reduction to the emptiness problem for a suitable type of automata [20, 6]. The translation from GFP formulas to automata extends immediately to the case for formulas containing simultaneous fixed-point operators. Furthermore, the polynomial-time inductive satisfiability-preserving translation from GNFP to GFP given in [7] (which in fact simply commutes with the fixed-point operators) extends in a straightforward manner to the case where the input and output formulas may contain simultaneous fixed-point operators.

$query$:= $select (t_1 \text{ as } ATTR_1, \dots, t_n \text{ as } ATTR_n) \text{ from } (REL_1 R_1, \dots, REL_m R_m) \text{ where } condition$
| $query \text{ union } query$ | $query \text{ intersect } query$ | $query \text{ except } query$
 $condition$:= $true$ | $t_1 = t_2$ | $t \text{ in } query$ | $exists(query)$
| $condition \text{ and } condition$ | $condition \text{ or } condition$ | $not(condition)$

Figure 1: Grammar for FO-SQL queries

Union, Intersection, and Difference: if $E_1, E_2 \in \text{RA}$ both have arity k , then $E_1 \cup E_2$, $E_1 \cap E_2$ and $E_1 - E_2$ belong to RA and have arity k .

Codd’s completeness theorem states that RA has the same expressive power as the domain-independent fragment of first-order logic, cf. [1]. Let us briefly recall here the definition of domain independence for first-order formulas without constant symbols [1]. The *active domain* of a structure M is the set $adom(M)$ of all elements that occur in a tuple belonging to one of the relations. For any structure M , let M' be a copy of the same structure but where all elements outside $adom(M)$ are removed. A first-order formula $\phi(\mathbf{x})$ without constant symbols is *domain-independent* if (i) whenever $M \models \phi(\mathbf{a})$, then the tuple \mathbf{a} consists of elements of $adom(M)$, and (ii) for all tuples \mathbf{a} consisting of elements of $adom(M)$, $M \models \phi(\mathbf{a})$ if and only if $M' \models \phi(\mathbf{a})$. The same definition applies to formulas with fixed-point operators. Examples of first-order formulas that are *not* domain-independent are $P(x) \vee Q(y)$, $x = x$, and $\neg P(x)$.

We say that a relation algebra expression is *negation-guarded* if every occurrence of the difference operator is of the form $\pi_{i_1 \dots i_m}(R) - E$ where R is a relation symbol. We denote by GN-RA the negation-guarded fragment of RA. It can be shown by straightforward inductive translations that GN-RA captures GNFO in the following sense.

Theorem 2 *Every k -ary GN-RA expression is equivalent to a domain-independent GNFO formula $\phi(x_1, \dots, x_k)$, and vice versa, via a linear translation from GN-RA to GNFO and an exponential translation backwards.*

Let R, S be relation symbols of arity 2 and 1, respectively. The following RA expressions are *not* negation-guarded.

$(\pi_1(R) \times S) - \pi_{1,1}(R)$ (distinct pairs from $\pi_1(R) \times S$)
 $\pi_{1,4}(\sigma_{2=3}(R \times R)) - R$ (reachability in two steps, not one)
 $\pi_1(R) - \pi_1((\pi_1(R) \times S) - R)$ (the quotient $R \div S$)

In fact, it follows from results in [7] that none of these expressions is equivalent to a GN-RA expression.

It is easy to show that testing if a given RA expression is equivalent to a GN-RA expression is undecidable.

Observe that in the above definition of GN-RA we did not allow for the use of constant values in selections and projections. This was only to simplify presentation. All complexity results that we will present go through in the presence of constant values, cf. Section 9.

3. GUARDED NEGATION IN SQL

In this section, we discuss what it means for an SQL query to have guarded negation. More precisely, we consider a simple, first-order expressively complete, fragment of SQL with a set-based semantics, that we call FO-SQL, and we characterize the queries in this fragment that can be expressed in GNFO.

FO-SQL: a Simple First-Order Fragment of SQL. In this section, unlike in the rest of the paper, we work with named schemas. A *named schema* is a collection of relation names, each with an associated list of attribute names. For the discussion below, assume we have a fixed schema, say, consisting of `BOOK(ISBN,AUTHOR,TITLE)` and `LOCATION(ISBN,SHELF,NUMBER)`. We also fix an infinite supply of “tuple variables” (also known as *aliases*, and denoted by R_1, R_2, \dots). By a *term* t we will mean an expression of the form $R_i.ATTR$ where R_i is a tuple variable and `ATTR` is an attribute name.

We consider SQL expressions that are generated by the simple grammar given in Figure 1, where each t_i is a term, each REL_i is a relation name, $ATTR_1, \dots, ATTR_n$ are distinct attribute names, and R_1, \dots, R_m are distinct tuple variables. This grammar generates queries that may have free tuple variables, i.e., there may be occurrences of tuple-variables R_i that are not in the scope of a `select-from-where` clause where they are declared. We will refer to queries with free tuple variables as *open queries* (or *correlated subqueries*), and we refer to queries without free tuple variables as *closed queries* (or *uncorrelated subqueries*). We will denote by $FV(q)$ the set of free tuple variables of q . We will be mainly interested in closed queries. Note, however, that closed queries *are* allowed to contain subexpressions of the form $exists(q)$ or of the form $t \text{ in } q$ where q is an open query.

We only consider queries that are well-typed in the sense that each (open or closed) query can be consistently assigned a (unique) *type*, which is a list of attribute names, where

1. the type of a `select-from-where` query is the set of attribute names specified in its `select` clause;
2. the `union`, `intersect`, and `except` operators take as arguments two queries of equal type, yielding a query of the same type.

Furthermore, terms $R_i.ATTR$ are only allowed to occur when `ATTR` belongs to the schema of the relation to which the occurrence of R_i in question is bound, and conditions of the form $t \text{ in } q$ are allowed only when q is a unary query, i.e., when the type of q consists of a single attribute.

By an FO-SQL query, we will mean a closed query satisfying the above requirements. Two examples are given in Figure 2. We assume that the reader is familiar with the semantics of SQL, and hence omit the formal semantics of the fragment FO-SQL. We just mention that we disregard order and duplicates, treating relations as *sets* of tuples. It is known that, under this set-based semantics, FO-SQL is expressively complete for first-order logic, in the sense of Codd’s expressive completeness theorem [1, 25]. That is, FO-SQL queries have the same expressive power as the domain-independent fragment of first-order logic. Since FO-SQL queries are defined in terms of named schemas, while the syntax of first-order logic is based on unnamed schemas in which the attributes of a relation are identified by natural numbers instead of by attribute names, here, we consider a

```
select A.NAME from AUTHOR A where not exists(
  select B.TITLE from BOOK B where B.AUTH = A.NAME)
```

```
select A.NAME from AUTHOR A where not exists(
  select B.TITLE from BOOK B where not B.AUTH = A.NAME)
```

Figure 2: Two examples of FO-SQL queries, the first negation-guarded and the second not

FO-SQL query q of type $\{A_1, \dots, A_n\}$ to be “equivalent” to a first-order formula $\phi(x_1, \dots, x_n)$, containing the relation names REL_i occurring in q as relation symbols of appropriate arity, if for every instance I and for every n -tuple \mathbf{a} , the tuple \mathbf{a} is an answer to q in I if and only if $I \models \phi(\mathbf{a})$.

The most important features of (full) SQL that are excluded in the above definition of FO-SQL are *constants*, *arithmetical comparison*, and *aggregation*. We will discuss the importance of these restrictions later in Section 9.

GN-SQL: the Guarded-Negation Fragment of FO-SQL.

We say that an SQL query is *negation-guarded* if the following two conditions hold:

1. each **except** operator has as its first argument a simple projection and as its second argument an uncorrelated subquery.
2. each **not** operator has as its argument a condition with at most one free tuple variable.

Here, by a *simple projection*, we mean a **select-from-where** query, where the **where**-clause is ‘true’. GN-SQL is the fragment of FO-SQL consisting of all negation-guarded queries.

To illustrate this definition, consider the two queries given in Figure 2. The first query involves a single occurrence of negation, which is guarded, since the negated condition has only one free tuple variable, namely A . The second query, on the other hand, is *not* a GN-SQL query, since the second occurrence of negation is not guarded. Indeed, the condition $B.AUTH = A.NAME$ has two free tuple variables.

The next theorem states that GN-SQL captures GNFO, in the same way that FO-SQL captures full first-order logic, as we discussed above (the same conventions apply, concerning what it means for a FO-SQL query to be equivalent to a first-order formula).

Theorem 3 (GN-SQL is Codd-complete for GNFO)

Each GN-SQL query can be translated in linear time into an equivalent domain-independent GNFO formula. Conversely, each domain-independent GNFO formula can be translated in exponential time into an equivalent GN-SQL query.

It can be shown that the exponential complexity of the translation from GNFO to GN-SQL is in general unavoidable for formulas of the form $(R(x_1) \vee S(x_1)) \wedge \dots \wedge (R(x_n) \vee S(x_n))$. On the other hand, the proof of Theorem 3 shows that if the schema includes a unary relation $ADOM$ that is guaranteed to denote the active domain of the instance, then there is a polynomial translation.

3.1 Negation in Practice: A Benchmark Study

In order to assess the usage of negation in SQL queries in practice, we have studied the workloads of two standard SQL benchmarks, namely TPC-H [32] and TPC-DS [31]. These

benchmark	queries	queries with negation ¹	queries with unguarded negation	queries with inequalities ²	queries with unguarded inequalities
TPC-H	22	4	0	3	1
TPC-DS	99	8	1	8	7
SkyServer	48	2	0	8	1

¹ By negation, we mean any occurrence of **not** or **except**.

² An inequality is any occurrence of $<>$ or $!=$. An inequality is *guarded* if the corresponding negation $\text{not}(\dots = \dots)$ is guarded.

Figure 3: Usage of negation in SQL benchmarks

benchmarks were designed to evaluate and compare the performance of relational database management systems. In addition, we studied the sample queries published on the Sloan Digital Sky Survey (SDSS) SkyServer website [30], a selection of actual queries submitted by SDSS users. For each query, we investigated whether the query uses negation, and, if so, whether the query is negation-guarded. We also studied the use of inequalities, and investigated which of these inequalities can be expressed using guarded negation. The results, given in Figure 3, shows that most queries using negation use only guarded negation. We should note here that most queries contain SQL constructs, such as aggregation, that do not belong to FO-SQL. Therefore, the queries are not necessarily expressible in GN-SQL. The statistics in Figure 3 are only concerned specifically with the explicit use of *negation*. We also did not investigate the use of other SQL constructs such as *outer joins*, that can, in some sense, be viewed as involving an implicit form of negation.

4. GUARDED NEGATION IN DATALOG

In this section, we present a powerful variant of Datalog with stratified negation, which we call GN-Datalog and which, in terms of its expressive power, is contained in GNFP. We first briefly recall the syntax and semantics of Datalog, with and without stratified negation.

Definition 4 (Datalog) A *Datalog program* is specified by a triple $\Pi = (\text{EDB}^\Pi, \text{IDB}^\Pi, \text{Rules}^\Pi)$, where EDB^Π and IDB^Π are disjoint sets of relation names, each with an associated arity, and Rules^Π is a finite set of rules of the form

$$\phi \leftarrow \psi_1, \dots, \psi_n$$

where $\phi, \psi_1, \dots, \psi_n$ are atomic formulas of the form $R(x_1, \dots, x_n)$ with $R \in \text{EDB}^\Pi \cup \text{IDB}^\Pi$ and x_1, \dots, x_n a sequence of first-order variables of appropriate length. We refer to ϕ as the head of the rule, and ψ_1, \dots, ψ_n as the body of the rule. In addition, we require that (i) every first-order variable occurring in the head of a rule must occur in the body, and (ii) the relation in the head of each rule must be an IDB relation.

A Datalog query is a pair (Π, Ans) , where Π is a Datalog program and Ans is a union of conjunctive queries over the schema $\text{EDB}^\Pi \cup \text{IDB}^\Pi$. The semantics of a Datalog query is defined as follows: first, if Π is a Datalog program, I an instance over the schema EDB^Π , and k a natural number, then we denote by $\Pi^k(I)$ the instance over the schema $\text{EDB}^\Pi \cup \text{IDB}^\Pi$ containing all facts that can be derived from the facts in I using at most k rounds of applications of rules

of Π . In addition, we denote by $\Pi^\infty(I)$ the union $\bigcup_k \Pi^k(I)$. If $q = (\Pi, Ans)$ is a Datalog query and I an instance over the schema EDB^Π , then we denote by $q(I)$ the set of all tuples that are an answer to the query Ans in $\Pi^\infty(I)$.

We remark that the above definition differs slightly from the standard presentation of Datalog. Usually, Ans is required to be a designated relation from IDB^Π instead of a union of conjunctive queries. The presentation we use here is convenient as it helps simplify the definitions below. On the other hand, note that this is not essential: a Datalog program can always be extended with an additional IDB relation and with additional rules computing the Ans query.

Definition 5 (Datalog with Stratified Negation) A *Datalog⁻ program* is a Datalog program Π where the body of each rule may, in addition, contain atomic formulas of the form $\neg R(x_1, \dots, x_n)$ provided that $R \in EDB^\Pi$, and provided that each first-order variable occurring in the head or body of the rule occurs positively in the body. A *Datalog program with stratified negation* is a sequence $\tilde{\Pi} = (\Pi_1, \dots, \Pi_n)$ of *Datalog⁻ programs*, called *strata*, with $n \geq 1$, where for each $i = 2 \dots n$, $EDB^{\Pi_i} = EDB^{\Pi_{i-1}} \cup IDB^{\Pi_{i-1}}$. We use $EDB^{\tilde{\Pi}}$ and $IDB^{\tilde{\Pi}}$ to denote EDB^{Π_1} and $\bigcup_{i=1 \dots n} IDB^{\Pi_i}$, respectively.

A *Datalog query with stratified negation* is a pair $(\tilde{\Pi}, Ans)$, where $\tilde{\Pi} = (\Pi_1, \dots, \Pi_n)$ is a Datalog program with stratified negation and Ans is a union of conjunctive queries over the schema $EDB^{\tilde{\Pi}} \cup IDB^{\tilde{\Pi}}$. The semantics of Datalog programs and of Datalog queries extends naturally to Datalog with stratified negation, by defining $\tilde{\Pi}^\infty(I)$ as $\Pi_n^\infty(\Pi_{n-1}^\infty(\dots \Pi_1^\infty(I) \dots))$ for $\tilde{\Pi} = (\Pi_1, \dots, \Pi_n)$.

We say that a Datalog program Π is *non-recursive* if no IDB occurs in the body of any of its rules, and hence, in particular, for all instances I we have that $\Pi^\infty(I) = \Pi^1(I)$. We say that a Datalog program with stratified negation is non-recursive if it consists entirely of non-recursive strata.

Definition 6 (GN-Datalog) A GN-Datalog program is a Datalog program with stratified negation $\tilde{\Pi} = (\Pi_1, \dots, \Pi_n)$, where each rule

$$\phi_0 \leftarrow (\neg)\phi_1, \dots, (\neg)\phi_n \quad \in \text{Rules}^{\Pi_k} \quad (1 \leq k \leq n)$$

is *negation guarded*, meaning that the following holds:

For each atom ϕ_i that either occurs negated in the body or is the head, the body includes a positive atom ϕ_j containing all first-order variables occurring in ϕ_i , and ϕ_j uses a relation from EDB^{Π_k} .²

A *GN-Datalog query* is a Datalog query with stratified negation, where each rule is negation guarded. Note that this requirement concerns only the rules; the answer query Ans can be any union of conjunctive queries.

Theorem 7 (Non-recursive GN-Datalog is Codd-complete for GNFO) *Each non-recursive GN-Datalog query is equivalent to a domain-independent GNFO formula, and vice versa, via exponential translations.*

²To understand why this is the appropriate definition of negation guardedness, observe that a rule of the form $\phi \leftarrow \psi_1, \dots, \psi_n$ expresses that $\neg \exists \mathbf{x}(\psi_1 \wedge \dots \wedge \psi_n \wedge \neg \phi)$, i.e., the head of the rule plays the same role as a negated atom in the body.

The translation from non-recursive GN-Datalog to GNFO given in the proof of Theorem 7 can be extended in a straightforward manner to a translation from GN-Datalog to the extension of GNFP with simultaneous fixed-point operators. Since simultaneous fixed-point operators can be eliminated (at the cost of an additional exponential blow-up, cf. Section 2), we obtain the following:

Theorem 8 *Each GN-Datalog query is equivalent to a domain-independent GNFP formula.*

The translation from non-recursive GN-Datalog to GNFO provided by Theorem 7 involves an exponential blow-up, due to an elimination of subformula sharing. The translation from GN-Datalog to GNFP provided by Theorem 8 involves another exponential blow-up, due to the elimination of simultaneous fixed-point operators. These sources of exponential complexity can be avoided (i) if we transcribe GN-Datalog queries into GNFP formulas over a larger schema (containing a relation symbol not only for each EDB of the GN-Datalog query, but also for each IDB), and (ii) freely use simultaneous fixed-point operators in the GNFP formula. More precisely, the proof of Theorem 8 can be adapted in a straightforward manner to show the following result, which will be useful later on (where, for two schemas $S \subseteq \hat{S}$, an \hat{S} -*expansion* of an instance I over \hat{S} is an instance over \hat{S} that agrees with I on all facts over S).

Theorem 9 *For every k -ary GN-Datalog query q over a schema S one can compute in polynomial time a GNFP sentence ϕ_q and a GNFP formula $\psi_q(x_1, \dots, x_k)$, both with simultaneous fixed point operators, and over a possibly larger schema \hat{S} , such that*

1. *each instance I has a unique \hat{S} -expansion \hat{I} satisfying ϕ_q .*
2. *for all instances I and k -tuples \mathbf{a} , $\mathbf{a} \in q(I)$ iff $\hat{I} \models \psi_q(\mathbf{a})$.*

5. RELATIONSHIPS WITH EXISTING LANGUAGES

Monadic Datalog is a well-known Datalog fragment that combines an interesting level of expressiveness with good algorithmic behavior thanks to a tight connection with tree automata, which also make monadic Datalog suitable for a number of applications, e.g. [19]. It also stands out as a fragment for which the boundedness problem is decidable [15], Theorem 21 below. Monadic Datalog does not allow any form of negation and since all IDB predicates are unary, guardedness of rule heads is guaranteed, so that monadic datalog rules are trivially negation guarded. We will show that boundedness remains decidable for GN-Datalog.

Datalog LIT is a fragment of stratified Datalog whose model checking has linear-time data complexity [18]. Each Datalog LIT rule must either contain in its body as ‘guard’ a positive literal containing all variables occurring in the rule, or must solely be comprised of unary literals (including its head). While the ‘guard’ of guarded rules need not be an EDB atom, [18] shows that every Datalog LIT program can (in exponential time) be transformed into an equivalent one having only EDB atoms as guards. The latter are trivially negation guarded. Every Datalog LIT program is thus equivalent to a GN-Datalog program.

GNFO subsumes a number of formalisms having currency in ontological reasoning, such as the linear- and guarded tuple generating dependencies (tgds) underlying the recently promoted Datalog[±] [10] framework and the more general frontier-guarded tgds [5] that subsume the description logics DL-Lite_R (which captures RDF Schema), \mathcal{ELI} , and \mathcal{ELH}_1^{dr} [3], which is the core of the proposed OWL-EL profile of the OWL 2.0 ontology language. GNFO can encode query answering and containment assertions involving such specifications of constraints or TBoxes. A tgd is a sentence

$$\forall \mathbf{x}, \mathbf{y} \phi(\mathbf{x}, \mathbf{y}) \rightarrow \exists \mathbf{z} \psi(\mathbf{y}, \mathbf{z}) \quad (2)$$

where both ϕ and ψ , called the body and the head, respectively, of the tgd rule, are conjunctions of positive atoms. A tgd is linear if ϕ consists of a single atom; it is guarded if ϕ contains as conjunct an atom $R(\mathbf{x}, \mathbf{y})$, the ‘guard’, in which all of the variables of the body occur together; and it is frontier-guarded if the body contains an atom $P(\mathbf{y})$ in which all of the variables shared by the body and the head of the rule occur together. Every frontier-guarded tgd naturally translates to a GNFO sentence.

Other query languages that can be viewed as fragments of GNFO include the semi-join algebra [24], as well as Unary Conjunctive View Logic (UCV) and Core XPath, cf. [12].

6. QUERY CONTAINMENT

We now exploit the connection with GNFO and GNFP to show that query containment is decidable for GN-SQL and for GN-Datalog. Recall that a query q is satisfiable if there exists an instance I such that the set of answers $q(I)$ is non-empty, and that a query q_1 is contained in a query q_2 if, for all instances I , $q_1(I) \subseteq q_2(I)$. The satisfiability problem can be viewed as (the complement of) a special case of the query containment problem, where the second query q_2 is any fixed unsatisfiable query.

Theorem 10 *Query containment is 2ExpTime-complete for both GN-SQL queries and GN-Datalog queries. Hardness holds already for satisfiability of non-recursive GN-Datalog, and GN-SQL, over a fixed EDB schema.*

The 2ExpTime upper bounds for GN-SQL follow directly from Theorem 3 and Theorem 1. The 2ExpTime upper bounds for GN-Datalog do not follow directly from Theorem 8 and Theorem 1, due to the exponential complexity of the translation from GN-Datalog to GNFP involved. However, it follows using Theorem 9: let q_1, q_2 be k -ary GN-Datalog queries ($k \geq 0$), and let $\phi_1, \psi_1(x_1, \dots, x_k)$ and $\phi_2, \psi_2(x_1, \dots, x_k)$ be the GNFP-formulas with simultaneous fixed point operators obtained by Theorem 9. We may assume without loss of generality that the only relation symbols that ϕ_1, ψ_1 and ϕ_2, ψ_2 have in common are the relation symbols that appear in q_1 and q_2 . It follows that q_1 is contained in q_2 if and only if $\phi_1 \wedge \psi_1(x_1, \dots, x_k) \models \phi_2 \rightarrow \psi_2(x_1, \dots, x_k)$. This gives us the desired result, since, as we explained in Section 2, the 2ExpTime upper bound for GNFP entailment from Theorem 1 extends to the case with simultaneous least-fixed point operators. The lower bounds are obtained by adapting the proof of an 2ExpTime-hardness result for a fragment of GNFO in [12].

Theorem 10 generalizes the known decidability result for monadic datalog and unions of conjunctive queries [15]. In addition, it easily implies the decidability of containment

of Datalog queries in Unions of Conjunctive Queries [13]. This can be seen as follows. For each Datalog query q , let \hat{q} be the GN-Datalog query obtained from q by guarding each rule using an additional conjunct that is a fresh EDB relation. Then for each UCQ q' over the original schema, we have that q is contained in q' if and only if \hat{q} is contained in q' . One direction follows directly from the fact that \hat{q} is contained in q . For the other direction, note that every counterexample I to the containment of q in q' gives rise to a counterexample I' to the containment of \hat{q} in q' . The instance I' in question extends I by interpreting each new EDB relation as the total relation containing all tuples over the active domain of I .

As a direct consequence of the finite model property of GNFO [7] and of Theorems 3 and 7, respectively, we find that query containment is *finitely controllable* for GN-SQL queries and for non-recursive GN-Datalog queries. By this we mean that one query is contained in another on finite instances if, and only if, the containment holds on unrestricted instances (a finite model property).

Theorem 11 *Satisfiability and containment are finitely controllable for GN-SQL and for non-recursive GN-Datalog.*

7. QUERY ANSWERING

7.1 (Closed-World) Query Evaluation

Since GN-SQL and non-recursive GN-Datalog admit translations into first-order logic, the *data complexity* of query evaluation is in AC^0 for both query languages. Similarly, since GN-Datalog is contained in the fixed-point logic FO(LFP), the data complexity of query evaluation is in PTime. In fact, there is a GN-Datalog query (a monadic Datalog query) for which query evaluation is PTime-hard in terms of data complexity [18].

In what follows we consider the *combined complexity* of query evaluation. Datalog evaluation is known to be ExpTime-complete for combined complexity (implicit in [34]). Monadic Datalog evaluation is known to be NP-complete [19]. The ‘‘guarded fragment of Datalog’’ (every rule contains an EDB atom containing all variables occurring in the rule) evaluation is in PTime [18]. Non-recursive Datalog with stratified negation is PSPACE-complete [16].

Recall that $P^{NP[\log^2]}$ is the class of those problems that can be solved by a polynomial time deterministic algorithm that is allowed to ask $O(\log^2(n))$ queries to an NP-oracle. (It relates to better known complexity classes this way: $NP \subseteq DP \subseteq P^{NP[\log]} \subseteq P^{NP[\log^2]} \subseteq \dots \subseteq P^{NP[\log^i]} \subseteq P^{NP} \subseteq \Sigma_2^P \subseteq PSPACE \subseteq EXPTIME$.)

Theorem 12 *The combined complexity of evaluating GN-SQL queries is $P^{NP[\log^2]}$ -complete.*

PROOF. The upper bound follows directly from Theorem 3 and Theorem 1. For the lower bound, observe that the translation from GNFO to GN-SQL given in the proof of Theorem 3 is polynomial in the presence of a unary relation ADOM containing all elements in the active domain. We may assume without loss of generality that our input instance contains such a relation. Therefore, the lower bound from Theorem 1 extends to GN-SQL as well. \square

We show here that the same problem is P^{NP} -complete for GN-Datalog. Recall that the best known upper bound on the complexity of model checking GNFP is $NP^{NP} \cap coNP^{NP}$.

Theorem 13 *The combined complexity of evaluating GN-Datalog queries is P^{NP} -complete. Hardness holds already for non-recursive GN-Datalog queries with only unary IDB predicates and nullary negation.*

7.2 Open-World Query Answering

Open world (OWA) query answering is the following problem: *given a query q , an instance I , and a tuple of values \mathbf{a} , decide whether it is the case that \mathbf{a} belongs to the answers of q in every instance extending I with additional facts.* An instance of open-world query answering $I \models_{OWA} q(\mathbf{a})$ thus asks for the unsatisfiability of $I \cup \{\neg q(\mathbf{a})\}$ in the usual first-order semantics, when treating I as a set of atomic facts with its elements as constants. Open world semantics is the natural choice when working with incomplete databases, in data exchange settings, and in the context of ontological reasoning. In each of these settings, open world query answering is an extensively researched problem.

In this section, we investigate the data complexity of open-world query answering for queries with guarded negation. Formally, for each query q we denote by OWA_q the problem, given an instance I and a tuple of values \mathbf{a} from $adom(I)$, to decide whether $I \models_{OWA} q(\mathbf{a})$. More generally, for each query q and for each set of constraints Σ , we denote by $OWA_{q,\Sigma}$ the problem, given an instance I , to decide whether $I, \Sigma \models_{OWA} q$.

Note that, in the absence of constraints, for conjunctive queries q , by monotonicity, the problem $I \models_{OWA} q(\mathbf{a})$ coincides with $I \models q(\mathbf{a})$, and therefore OWA_q is in PTime (in fact, in AC^0). For first-order queries q , on the other hand, the problem OWA_q can be undecidable. We will show below that the problem is decidable for first-order queries with guarded negation.

As constraints, we will consider tuple-generating dependencies, cf. (2), and key constraints. As noted above, linear-, guarded- and frontier-guarded tgds [5, 10] are expressible in GNFO. With respect to OWA query answering, conjunctive queries are known to be FO-rewritable relative to linear tgds [10] and possess Datalog rewritings relative to frontier-guarded tgds [4]. Accordingly, the data complexity of open-world query answering for conjunctive queries against linear tgds is in AC_0 , in PTime for frontier-guarded tgds, and PTime-complete already for guarded tgds [10].

We begin by observing that OWA query answering for GNFO queries, as for many description logics [29, 11, 27], has coNP data complexity. For an instance I , we denote by $|I|$ the total number of facts of I , and, for two instances I, J , we write $I \subseteq J$ if every fact of I is a fact of J .

Proposition 14 *Let $\phi(\mathbf{x})$ be a fixed GNFO formula. For an instance I and a tuple \mathbf{a} of elements from $adom(I)$, if there is an instance $J \models \phi(\mathbf{a})$ with $I \subseteq J$, then there is an instance $J \models \phi(\mathbf{a})$ with $I \subseteq J$ and $|J| = O(|I|)$.*

Proposition 14 tells us that, in solving the open-world query answering problem for GNFO queries, it suffices to consider only instances whose size is linear in the size of the input instance. This gives us the following:

Theorem 15 *For each GNFO query q (in particular, for each GN-SQL query), OWA_q is in coNP. There is a boolean GN-SQL query q for which OWA_q is coNP-hard.*

PROOF. The coNP upper bound is immediate from the above proposition. Given an instance I with distinguished elements \mathbf{a} , Proposition 14 shows that to refute $I \models_{OWA} q(\mathbf{a})$ it suffices to guess a linear size instance J with $I \subseteq J$ and test in polynomial time that J satisfies $\neg q(\mathbf{a})$. The lower bound is established by a reduction from 3-colorability. Let q be the GNFO sentence (for readability, we omit the repeated occurrences of Nx as guard):

$$\exists x(Nx \wedge \neg P_1x \wedge \neg P_2x \wedge \neg P_3x) \vee \bigvee_i \exists xy(Exy \wedge P_ix \wedge P_iy) \quad (3)$$

expressing that P_1, P_2, P_3 do not constitute a valid 3-coloring of the graph (N, E) . It is easy to check that a simple undirected graph G is not 3-colorable iff $G \models_{OWA} q$, and it is straightforward to formulate the domain-independent boolean query (3) in GN-SQL. \square

This is remarkable, given that open-world query answering is in general undecidable for first-order queries, even in the absence of constraints.

Recall that every frontier-guarded tgd can be formulated as a GNFO sentence. This allows us to lift the above result to the open-world query answering problem with constraints that are frontier-guarded tgds. More precisely, if Σ is a set of frontier-guarded tgds, then $OWA_{q,\Sigma}$, by definition, coincides with $OWA_{q \vee \bigvee_{\sigma \in \Sigma} \neg \sigma}$, and therefore we get the following.

Corollary 16 *For each GNFO query q and for each finite set of frontier-guarded tgds Σ , $OWA_{q,\Sigma}$ is in coNP.*

In various contexts, such as data exchange [17], it is useful to consider incomplete databases that contain, besides constant values, also labeled null values. In this case, open world query answering is defined not in terms of extensions of instances, but in terms of homomorphisms that are allowed to map the labeled null values to constant values or to other labeled null values. It is worth observing that the above proofs go through in this more general setting with null values, showing that for GNFO queries q and for finite sets of frontier-guarded tgds Σ , $OWA_{q,\Sigma}$ is in coNP even over instances containing labeled nulls.

Next we identify a subfragment of GNFO that accommodates the earlier mentioned formalisms including conjunctive queries and frontier-guarded tgds and whose queries enjoy PTime data complexity for OWA. Recall that open-world query answering $I \models_{OWA} q$ asks for the unsatisfiability of $I \cup \{\neg q\}$. Under negation, the subformula $\exists x(Nx \wedge \neg P_1x \wedge \neg P_2x \wedge \neg P_3x)$ of the coNP-complete query (3) turns into the disjunctive requirement $\forall x(Nx \rightarrow P_1x \vee P_2x \vee P_3x)$ that is, in an intuitive sense, ultimately responsible for intractability. Indeed, it has been observed in the context of DL-Lite that the introduction of even the weakest form of disjunction renders query answering intractable (see, e.g., [11]). It turns out that the positive occurrence of conjunctions $\neg A(x) \wedge \dots \wedge \neg B(x)$ involving two or more negated conjuncts are the only source of intractability in GNFO queries.

Definition 17 (serial GNFO queries, SGNQ)

A GNFO-formula φ is *serial* if it is in DNF and no conjunction $\neg \chi(x) \wedge \dots \wedge \neg \psi(x)$ with two or more negated conjuncts occurs positively in φ , i.e., in the scope of an even number of negations. Let SGNQ denote the set of serial GNFO queries.

Clearly, every union of conjunctive queries is a serial GNFO query. Furthermore, every frontier-guarded tgds, as well as its negation, is equivalent to a boolean serial GNFO queries. It fact, for every finite set Σ of frontier-guarded tgds and for every serial GNFO query q , we have that $q \vee \bigvee_{\sigma \in \Sigma} \neg \sigma$ is a serial GNFO query. In other words, the reduction from open-world query answering in the presence of frontier-guarded tgds to open-world query answering in the absence of tgds, that we gave earlier, holds also in the case of serial GNFO queries.

Theorem 18 *For each SGNQ q and for each finite set Σ of frontier-guarded tgds, $OWA_{q,\Sigma}$ is in PTime.*

In fact, for every boolean SGNQ q we can effectively compute a boolean Datalog query q' such that for all instances I , we have $I \models_{OWA} q \iff I \models q'$.

There is a boolean SGNQ query q for which OWA_q is PTime-complete.

The proof is based on a reduction from the open-world query answering problem for SGNQs in the presence of frontier-guarded tgds to the open-world query answering problem for conjunctive queries in the presence of frontier-guarded tgds. A PTime solution of the latter problem via Datalog rewritings is due to [4].

Finally, we show that OWA answering GNFO queries under key constraints is undecidable. This holds even for a fixed GNFO query and a fixed key constraint of the form $\forall \mathbf{x}yz(F(\mathbf{x}, y) \wedge F(\mathbf{x}, z) \rightarrow y = z)$ with F a relation symbol.

Theorem 19 (i) *There is a boolean conjunctive query q and a set Σ comprising guarded tgds and a single key constraint, so that $OWA_{q,\Sigma}$ is undecidable.*

(ii) *There is a boolean SGNQ q and a key constraint σ , so that $OWA_{q,\{\sigma\}}$ is undecidable.*

While undecidability of the uniform problem (where the query is part of the input) follows from various similar results for weaker formalisms [29], for a fixed query this seems to be a new result.

8. BOUNDEDNESS AND FIRST-ORDER DEFINABILITY

In this section, we study the boundedness problem for GN-Datalog. Our main result, Corollary 28, states that it is decidable whether a GN-Datalog program is fully bounded, i.e., whether, for every instance, the computation of each stratum of the GN-Datalog program reaches a fixed point in a bounded number of steps.

The semantics of a Datalog program Π can be defined in terms of a *least fixed point* for the IDB predicates. For this we view Π , or rather each of its instantiations Π_I over a given instance I , as a monotone operator. An application of this operator to any instantiation of the IDB predicates produces the result of firing all rules once and in parallel, on these IDB predicates and the static EDB predicates as given in I . This operator Π_I is monotone, and the desired interpretation of the IDB predicates in $\Pi^\infty(I)$ is its unique least fixed point. This view extends to not necessarily finite instances I , where Π_I , due to its monotonicity, still has a unique least fixed point, also referred to as $\Pi^\infty(I)$. As in the case of finite instances, this fixed point is obtained as the limit of the monotone sequence of stages Π_I^α generated by

iterating Π_I as an update operator, starting from the empty instantiation for all IDB predicates in stage 0, and taking unions at limit ordinals, until finally (for cardinality reasons) a stage Π_I^α is reached that is a fixed point, and indeed the unique least fixed point ($\Pi_I^{\alpha+1} = \Pi_I^\alpha$ implies $\Pi_I^\alpha = \Pi^\infty(I)$).

All these considerations hold for any notion of program or recursion scheme that shares the crucial monotonicity with Datalog programs. Monotonicity refers to monotonicity in the IDB arguments, and is guaranteed by syntactic positivity in the IDB predicates in all cases we consider. We are mostly interested in IDB-positive GNFO-programs, which we first investigate in isolation, towards understanding their stratified, and overall no longer monotone, use in GN-Datalog (cf. Definition 23 below).

The notion of *boundedness* captures the semantic and procedural essence of non-recursive behavior (in contrast with syntactic non-recursive behavior as defined in Section 4, which focuses on a trivial reason for boundedness).

Definition 20 A monotone program Π is *c-bounded* (bounded in the classical sense, or over unrestricted instances) if there exists some $n \in \mathbb{N}$ such that $\Pi_I^{n+1} = \Pi_I^n$ for every finite or infinite instance I . It is bounded over a class of instances \mathcal{I} if there is such an n that is good for all $I \in \mathcal{I}$. We call Π *f-bounded* if it is bounded over the class of all finite instances.

$BDD(\mathcal{P}, \mathcal{I})$ stands for the *boundedness problem* for programs from \mathcal{P} over instances from \mathcal{I} : given $\Pi \in \mathcal{P}$, decide whether Π is bounded over \mathcal{I} . We reserve the names $BDD_c(\mathcal{P})$ and $BDD_f(\mathcal{P})$ for $BDD(\mathcal{P}, \text{ALL})$ and $BDD(\mathcal{P}, \text{FIN})$, where ALL and FIN are the classes of unrestricted and of finite instances, respectively.

Despite its basic nature, the boundedness problem is known to be undecidable for even very rudimentary classes of programs – a fact which frustrated all hopes to systematically eliminate bounded, i.e. spurious, recursion in effective tools for query optimization. See for instance [21] for the undecidability of (f-)boundedness for Datalog programs with binary IDB predicates, as well as for Datalog programs with just monadic IDB predicates but with EDB negation or even just with inequalities in the bodies. One of the few major decidability results is the following from [15].

Theorem 21 (Cosmadakis et al. [15]) $BDD_f(\mathcal{P}) = BDD_c(\mathcal{P})$ is decidable for the class \mathcal{P} of all monadic Datalog programs.

The following result from classical model theory is of fundamental importance for links between boundedness and first-order (FO) definability. It speaks about IDB-positive programs Π that are first-order in the sense that the bodies of rules can be expressed in FO, by formulas that are positive in all IDB predicates (which guarantees monotonicity). We use the term *first-order programs* in this sense. We say that the fixed point of Π is FO-definable over the class \mathcal{I} if each IDB predicate in the least fixed point $\Pi^\infty(I)$ is definable in terms of the EDB predicates by some first-order formula, uniformly across all $I \in \mathcal{I}$.

Theorem 22 (Barwise–Moschovakis [8]) *An IDB-positive first-order program Π is bounded in the classical sense if, and only if, the fixed point of Π is FO-definable over the class of all (finite and infinite) instances.*

Analogous equivalences can be derived for many natural fragments $L \subseteq \text{FO}$, where boundedness of IDB-positive L-programs is equated with L-definability of their fixed points. This is true in particular also for the guarded negation fragment $\text{GNFO} \subseteq \text{FO}$.

Moreover, for many well-behaved fragments $L \subseteq \text{FO}$ there are model theoretic transfer results that say that an L-program Π is bounded over \mathcal{I} if, and only if, it is bounded over some subclass $\mathcal{I}_0 \subseteq \mathcal{I}$. A case of particular interest is a *finite model property* for boundedness, which links the classical notion to its finite model theory version. This, too, is available in the case of GNFO.

Definition 23 A GNFO-program is an IDB-positive program Π with rules of the form

$$X\mathbf{x}_s \leftarrow \alpha_s(\mathbf{x}_s) \wedge \phi_s(\mathbf{X}, \mathbf{x}_s)$$

where $\phi_s \in \text{GNFO}$ is positive in the IDB predicates \mathbf{X} and α_s is an EDB atom guarding the variable tuple \mathbf{x}_s in the head.

The following say that for GNFO we are in the ideal situation that f-boundedness and c-boundedness coincide, and that the classical and finite model theory variants of the Barwise–Moschovakis correspondence hold. The finite model theory analogue is the least straightforward of these.³

Proposition 24 For GNFO-programs Π and their least fixed points Π^∞ , t.f.a.e.:

- (i) Π^∞ is FO-definable over all finite instances.
- (ii) Π^∞ is FO-definable over all unrestricted instances.
- (iii) Π^∞ is GNFO-definable over all finite instances.
- (iv) Π^∞ is GNFO-definable over all unrestricted instances.
- (v) Π is bounded over all finite instances.
- (vi) Π is bounded over all unrestricted instances.

Another crucial transfer property for $\text{BDD}(\text{GNFO})$ is based on the notion of *treewidth*. In [7], it was suggested that the key to the good computational behavior of GNFO and GNFP lies in the fact that these logics have a *tree-like model property*: for testing the satisfiability and the entailment of formulas, it suffices to consider structures of bounded treewidth. The same notion provides the key to decidability of boundedness as well.

The *width* $w(\Pi)$ of a GNFO-program Π is the maximum number of element variables used in any of its rules in DNF.

Lemma 25 A GNFO-program Π of width $\leq w$ is bounded over all unrestricted instances if, and only if, it is bounded over the class of all (possibly infinite) instances of treewidth at most w .

PROOF. Each finite stage \mathbf{X}^n of Π can be defined by a sequence of GNFO-formulas whose width is bounded by w . In particular, for each natural number n , boundedness of Π at stage $n \geq 1$, w.r.t. a class of structures, is equivalent to the validity of a certain GNFO-sentence of width w , on that class of structures. Since a GNFO-sentence of width

³It is known, for instance, that the universal fragment of FO, despite its finite model property, does not satisfy this analogue: there is a purely universal program whose limit is uniformly definable in universal FO across all finite instances, although it is unbounded over finite instances.

w is valid on arbitrary structures if and only if it is valid on structures of treewidth at most w (cf. [7, 28]), the claim follows. \square

We turn to decidability of $\text{BDD}_c(\text{GNFO})$ and of *full boundedness* (to be defined below) of GN-Datalog. Given the meager history of decidability results concerning boundedness for database purposes, it is interesting that here is one considerable extension of the early decidability result for monadic Datalog from [15], cf. Theorem 21 above.

We note that GN-Datalog is strictly more expressive than monadic Datalog, but avoids the dangers of negation that render boundedness undecidable, for instance, in the extension of monadic Datalog by just inequalities, or by negative as well as positive access to some binary EDB predicates.

Technically, the following decidability assertion is an easy corollary to the decidability results for monadic second-order logic and guarded second-order logic over tree-like structures in [9]. These results in turn are based on a non-trivial reduction to an automata theoretic decidability result of Colcombet and Löding, which, in the relevant strength needed here, has not been published yet. As in [9] we indicate this caveat formally as an assumption (ILT), which refers to the decidability of *limitedness* for weighted parity automata on infinite trees, as announced in connection with progress on earlier work in [14].

Recall that $\text{BDD}_c(\text{GNFO})$ and $\text{BDD}_f(\text{GNFO})$ coincide.

Theorem 26 (assuming ILT) Boundedness for GNFO-programs is decidable.

PROOF. The GNFO-formulas in an GNFO-program can be translated into explicitly guarded formulas of guarded second-order logic GSO (denoted GSO^* in [9]). The result then follows from the decidability of $\text{BDD}(\text{GSO}^*, \mathcal{W}_k)$, boundedness for GSO^* over the class of structures of treewidth k , where both the GSO^* -formulas and the parameter k are treated as input (Theorem 8.8 in [9]). We apply this to the GSO^* -translation of the input GNFO-program Π over the class \mathcal{W}_k for $k := w(\Pi)$.⁴ By Lemma 25, (ii), this is a valid reduction. \square

Towards our interest in GN-Datalog, with its stratified use of guarded negation as defined in Section 4, we extend the notion of boundedness from Definition 20 as follows.

Definition 27 A GN-Datalog program $\tilde{\Pi} = (\Pi_i)_{i \leq n}$ is called *fully f/c-bounded* over a class of instances \mathcal{I} if each stratum Π_i is f/c-bounded over the class of all instances obtained from instances in \mathcal{I} by evaluating all IDB predicates from lower strata according to $\tilde{\Pi}_{<i}$ and treating them as EDB for Π_i . Equivalently, a GN-Datalog program $\tilde{\Pi} = (\Pi_i)_{i \leq t}$ is fully f/c-bounded if there are natural numbers k_1, \dots, k_t such that for all finite/unrestricted instances I , $\Pi^\infty(I) = \Pi_t^{k_t}(\Pi_{t-1}^{k_{t-1}}(\dots \Pi_1^{k_1}(I) \dots))$.

Corollary 28 (assuming ILT) For GN-Datalog, full f-boundedness is decidable and coincides with full c-boundedness.

⁴NB: since Π really corresponds to a *system* of least fixed points in several IDB predicates X , we need the result for systems of simultaneous fixed points in GSO^* from [9], as discussed in the proof sketch for Theorem 11.5 there.

PROOF. The proof is by induction on the number of strata. Note that by definition of full boundedness, a stratified GN-Datalog program $\bar{\Pi} = (\Pi_i)_{i \leq n}$ fails to be fully bounded if, and only if, there is a least stratum $m \leq n$ such that Π_m is unbounded over the class of instances obtained by evaluating all IDB predicates of lower strata according to $\bar{\Pi}_{< m}$. Since these lower strata are bounded, this partial evaluation is in fact GNFO-definable. It follows that the above arguments concerning the GNFO-variant of the Barwise–Moschovakis theorem and its finite model theory version carry through – stratum by stratum, and up to the first stratum that turns out to be unbounded, if any. This also reduces the decidability claim to that in Theorem 26. \square

We remark that the passage through boundedness for GSO^* over \mathcal{W}_k , which is known to be of non-elementary complexity even for $k = 1$, prevents us from extracting any reasonable complexity bounds. It is conceivable, of course, that alternative methods yield such bounds (as is the case for other special cases of interest, besides that of monadic Datalog, that also follow from the master result of [9]).

9. DISCUSSION

9.1 Further Extensions of GN-SQL

Inequalities. GN-SQL can be viewed as a well-behaved query language extending unions of conjunctive queries with a restricted form of negation. In this sense, it is natural to compare GN-SQL to UCQ(\neq), the language of unions of conjunctive queries with inequalities. Like GN-SQL, UCQ(\neq) is computationally well-behaved: query containment is Π_2^p -complete [23, 33], the combined complexity of query evaluation is NP-complete, and the data complexity of open world query answering is NP-complete w.r.t. a large class of constraints [17], cf. also [26]. In this light, and in the light of Figure 3, the question arises whether we can extend GN-SQL to allow for the use of (unguarded) inequalities.

Let us denote by GN-SQL(\neq) the extension of GN-SQL where conditions may make use of the inequality relation (\neq), but the inequality relation cannot be used to guard negations. It is easy to see that Theorem 12 extends to GN-SQL(\neq) — we may view the inequality as just another relation that is part of the input instance. All the other results we obtained for GN-SQL, however, fail for GN-SQL(\neq). This follows from the fact that it is possible to express functional dependencies in GN-SQL(\neq). Indeed, every functional dependency

$$\forall \mathbf{x} \mathbf{y} \mathbf{z} \mathbf{u} v (F(\mathbf{x}, \mathbf{y}, u) \wedge F(\mathbf{x}, \mathbf{z}, v) \rightarrow u = v)$$

is equivalent to the GNFO sentence with inequality

$$\neg \exists \mathbf{x} \mathbf{y} \mathbf{z}, u, v (F(\mathbf{x}, \mathbf{y}, u) \wedge F(\mathbf{x}, \mathbf{z}, v) \wedge u \neq v),$$

which can easily be expressed in GN-SQL(\neq) as well. Recall that inclusion dependencies too can be expressed in GN-SQL. This, together with classical results in dependency theory (cf. [1]) and Theorem 19(ii)), implies the following:

Theorem 29 (i) GN-SQL(\neq) is not finitely controllable for satisfiability or query containment.

(ii) The satisfiability and query containment problems for GN-SQL(\neq) are undecidable (both on finite instances and on unrestricted instances).

(iii) There is a GN-SQL(\neq) query for which open world query answering is undecidable.

Known results for various description logics contained in GNFO imply that OWA answering for GNFO(\neq) queries is undecidable when the query is part of the input [29]. Theorem 29 strengthens this by showing undecidability already for a fixed GNFO(\neq) query. Naturally, similar results can be obtained for the extension of GN-Datalog with inequalities.

Constants and Comparisons. GN-SQL queries, as we defined them, cannot contain constant values, nor arithmetical comparisons (i.e., conditions of the form $t_1 < t_2$). Indeed, over linearly ordered domains, inequalities can be expressed using arithmetical comparisons ($x \neq y$ is equivalent to $x < y \vee y < x$), and hence, by Theorem 29, most problems immediately become undecidable when arithmetical comparisons are allowed. However, as we will show, our results *do* generalize to the extension of GN-SQL where (i) queries may contain constant values, and (ii) arithmetical comparisons of the form $t_1 < t_2$ are allowed provided that at least one of t_1, t_2 is a constant value.

In what follows, let $\text{LIN} = (D, <)$ be any ordered domain (where D is a countable set and $<$ is a total order on D) that is “reasonable” in the sense that the following problems are all solvable in polynomial time (for some appropriate representation of the elements of D):

1. given $d_1, d_2 \in D$, is it the case that $d_1 < d_2$?
2. given $d \in D$, does there exist $d' \in D$ with $d' < d$?
3. given $d \in D$, does there exist $d' \in D$ with $d < d'$?
4. given $d_1, d_2 \in D$, is there a $d' \in D$ with $d_1 < d' < d_2$?

Essentially, all the usual ordered domains, such as the natural numbers $(\mathbb{N}, <)$, the rational numbers $(\mathbb{Q}, <)$, and the strings $(A^*, <_{lex})$ over a finite ordered alphabet A , are reasonable in this sense.

Let GN-SQL(LIN) be the extension of the GN-SQL syntax where (i) all terms t are allowed to be either of the form $R.ATTR$ (as before) or to be an element of LIN (in which case we call t a *constant*); and (ii) for all terms t_1, t_2 of which at least one is a constant value, $t_1 < t_2$ is allowed as an atomic condition. The semantics of GN-SQL(LIN) queries is only well-defined for instances whose active domain is a subset of LIN. Therefore, we restrict attention to such instances.

All results for GN-SQL that we have presented can be extended to GN-SQL(LIN). For simplicity, we sketch the relevant construction here only for the query containment problem.

Theorem 30 Let LIN be any reasonable ordered domain. GN-SQL(LIN) query containment is 2ExpTime-complete.

Aggregation. Recall that GN-SQL does not allow for any form of aggregation that is available in SQL. This is for good reason: allowing even simple forms of aggregation such as counting would quickly lead to undecidability, since query containment for unions of conjunctive queries under the bag semantics is undecidable [22].

9.2 Further Extensions of GN-Datalog

Allowing IDBs As Guards. If in the definition of negation-guarded Datalog rules one permits also the use of IDB atoms from the same or lower stratum as guards, this can result

in an exponential gain in succinctness but does not increase the expressive power. (A simple induction on strata and on stages of inductive definitions of IDB predicates confirms that all tuples added to the interpretation of IDB predicates are guarded by some EDB atom.) Query evaluation complexity, however, suffers an exponential blow-up as a consequence of this relaxation.

Proposition 31 (GN-Datalog with IDB guards)

Answering GN-Datalog queries with IDB atoms allowed as guards is ExpTime-complete in combined complexity.

Capturing the Alternation-Free Fragment of GNFP.

In [18], an extension of Datalog-LITE was presented, called Datalog-LITE, which includes “generalized literals” and was shown to capture the alternation-free fragment of guarded fixed point logic (GFP). We expect that GN-Datalog can be similarly extended, in order to subsume Datalog-LITE and capture the alternation-free fragment of GNFP.

Acknowledgements: This research was partially supported by NSF Grant IIS-0905276. We thank Alkis Polyzotis for helpful comments. Detailed proofs of the results can be found at <http://tiny.cc/guardednegation>.

10. REFERENCES

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

[2] A. Arnold and D. Niwinski. *Rudiments of μ -calculus*. Elsevier, 2001.

[3] J.-F. Baget, M. Leclère, M.-L. Mugnier, and E. Salvat. On rules with existential variables: Walking the decidability line. *Art. Intell.*, 175(10):1620–1654, 2011.

[4] J.-F. Baget, M.-L. Mugnier, S. Rudolph, and M. Thomazo. Complexity Boundaries for Generalized Guarded Existential Rules. Technical Report lirmm-00568935, LIRMM, 2011.

[5] J.-F. Baget, M.-L. Mugnier, S. Rudolph, and M. Thomazo. Walking the complexity lines for generalized guarded existential rules. In *Proc. IJCAI*, pages 712–717, 2011.

[6] V. Bárány and M. Bojanczyk. Finite satisfiability for guarded fixpoint logic. *IPL*, 112(10):371–375, 2012.

[7] V. Bárány, B. ten Cate, and L. Segoufin. Guarded negation. In *Proc. ICALP*, pages 356–367, 2011.

[8] J. Barwise and Y. N. Moschovakis. Global inductive definability. *J. Symb. Log.*, 43(3):521–534, 1978.

[9] A. Blumensath, M. Otto, and M. Weyer. Decidability results for the boundedness problem. Preprint at www.mathematik.tu-darmstadt.de/~otto/, 2011.

[10] A. Cali, G. Gottlob, and T. Lukasiewicz. Tractable query answering over ontologies with Datalog[±]. In *Proc. Description Logics Workshop*, 2009.

[11] D. Calvanese, G. D. Giacomo, D. Lembo, M. Lenzerini, and R. Rosati. Data complexity of query answering in description logics. In *Proc. KR*, pages 260–270, 2006.

[12] B. ten Cate and L. Segoufin. Unary negation. In *Proc. STACS*, pages 344–355, 2011.

[13] S. Chaudhuri and M. Y. Vardi. On the equivalence of recursive and nonrecursive datalog programs. *J. Comput. Syst. Sci.*, 54(1):61–78, 1997.

[14] T. Colcombet and C. Löding. The nesting-depth of disjunctive μ -calculus for tree languages and the limitedness problem. In *Proc. CSL*, pages 416–430, 2008.

[15] S. Cosmadakis, H. Gaifman, P. Kanellakis, and M. Vardi. Decidable optimization problems for database logic programs. In *Proc. STOC 1988*, pages 477–490, 1988.

[16] E. Dantsin, T. Eiter, G. Gottlob, and A. Voronkov. Complexity and expressive power of logic programming. *ACM Comput. Surv.*, 33(3):374–425, 2001.

[17] R. Fagin, P. G. Kolaitis, R. J. Miller, and L. Popa. Data exchange: semantics and query answering. *Theoretical Computer Science*, 336(1):89–124, 2005.

[18] G. Gottlob, E. Grädel, and H. Veith. Datalog LITE: a deductive query language with linear time model checking. *ACM Trans. Comput. Log.*, 3(1):42–79, 2002.

[19] G. Gottlob and C. Koch. Monadic datalog and the expressive power of languages for Web information extraction. *J. ACM*, 51(1):74–113, 2004.

[20] E. Grädel and I. Walukiewicz. Guarded fixed point logic. In *In Proc. LICS*, pages 45–54, 1999.

[21] G. Hillebrand, P. Kanellakis, H. Mairson, and M. Vardi. Undecidable boundedness problems for datalog programs. *J. of Logic Prog.*, 25(2):163–190, 1995.

[22] Y. E. Ioannidis and R. Ramakrishnan. Containment of conjunctive queries: beyond relations as sets. *ACM Trans. Database Syst.*, 20(3):288–324, 1995.

[23] A. Klug. On conjunctive queries containing inequalities. *J. ACM*, 35(1):146–160, 1988.

[24] D. Leinders, M. Marx, J. Tyszkiewicz, and J. Bussche. The semijoin algebra and the guarded fragment. *J. of Logic, Lang. and Inf.*, 14(3):331–343, 2005.

[25] L. Libkin. Expressive power of SQL. *Theor. Comput. Sci.*, 296(3):379–404, 2003.

[26] A. Madry. Data exchange: On the complexity of answering queries with inequalities. *Inf. Process. Lett.*, 94(6):253–257, 2005.

[27] M. Ortiz, D. Calvanese, and T. Eiter. Data complexity of Answering Unions of Conjunctive Queries in SHIQ. In *Proc. Description Logics Workshop*, 2006.

[28] M. Otto. Expressive completeness through logically tractable models. Submitted, 2012.

[29] R. Rosati. The limits of querying ontologies. In *Proc. ICDT*, pages 164–178, 2006.

[30] Sloan Digital Sky Survey. Sky server sample SQL queries. <http://skyserver.sdss.org/public/en/help/docs/realquery.asp>. Accessed Sept. 25, 2011.

[31] TPC. TPC-DS benchmark. <http://www.tpc.org/tpcds/>. Release 2008-01-21.

[32] TPC. TPC-H benchmark. <http://www.tpc.org/tpch/>. Release 2.14.3.

[33] R. van der Meyden. The complexity of querying indefinite data about linearly ordered domains. *J. Comput. Syst. Sci.*, 54(1):113–135, 1997.

[34] M. Y. Vardi. The complexity of relational query languages. In *Proc. STOC 1982*, pages 137–146, 1982.