Constraint-Generating Dependencies *

Marianne Baudinet,[1] Jan Chomicki,[2] and Pierre Wolper[3]

[1] Université Libre de Bruxelles, Informatique,
50 Avenue F.D. Roosevelt, C.P. 165, 1050 Brussels, Belgium
Email: mb@cs.ulb.ac.be
[2] Kansas State University, Dept of Computing and Information Sciences,
234 Nichols Hall, Manhattan, KS 66506-2302, U.S.A.
Email: chomicki@cis.ksu.edu
[3] Université de Liège, Institut Montefiore, B28
4000 Liège Sart-Tilman, Belgium
Email: pw@montefiore.ulg.ac.be

Abstract. Traditionally, dependency theory has been developed for uninterpreted data. Specifically, the only assumption that is made about the data domains is that data values can be compared for equality. However, data is often interpreted and there can be advantages in considering it as such, for instance obtaining more compact representations as done in constraint databases. This paper considers dependency theory in the context of interpreted data. Specifically, it studies *constraint-generating dependencies*. These are a generalization of equality-generating dependencies where equality requirements are replaced by constraints on an interpreted domain. The main technical results in the paper are a general decision procedure for the implication and consistency problems for constraint-generating dependencies, and complexity results for specific classes of such dependencies over given domains. The decision procedure proceeds by reducing the dependency problem to a decision problem for the constraint theory of interest, and is applicable as soon as the underlying constraint theory is decidable. The complexity results are, in some cases, directly lifted from the constraint theory; in other cases, optimal complexity bounds are obtained by taking into account the specific form of the constraint decision problem obtained by reducing the dependency implication problem.

1 Introduction

Relational database theory is largely built upon the assumption of uninterpreted data. While this has advantages, mostly generality, it foregoes the possibility of exploiting the structure of specific data domains. The introduction of constraint databases [21] was a break with this uninterpreted-data trend. Rather than defining the extension of relations by an explicit enumeration of tuples, a

* This work was supported by NATO Collaborative Research Grant CRG 940110 and by NSF Grant IRI-9110581.

constraint database uses constraint expressions to implicitly specify sets of tuples. Of course, for this to be possible in a meaningful way, one needs to consider interpreted data, that is, data from a specific domain on which a basic set of predicates and functions is defined. A typical example of constraint expressions and domain are linear inequalities interpreted on the reals. The potential gains from this approach are in the compactness of the representation (a single constraint expression can represent many, even an infinite number of, explicit tuples) and in the efficiency of query evaluation (computing with constraint expressions amounts to manipulating many tuples simultaneously).

Related developments have concurrently been taking place in temporal databases. Indeed, time values are intrinsically interpreted and this can be exploited for finitely representing potentially infinite temporal extensions. For instance, in [19] infinite temporal extensions are represented with the help of periodicity and inequality constraints, whereas in [10, 11] and [3] deductive rules over the integers are used for the same purpose. Constraints have also been used recently for representing incomplete temporal information [31, 23].

If one surveys the existing work on databases with interpreted data and implicit representations, one finds contributions on the expressiveness of the various representation formalisms [2, 5, 4], on the complexity of query evaluation [9, 12, 25, 31], and on data structures and algorithms to be used in the representation of constraint expressions and in query evaluation [28, 7, 8, 22]. However, much less has been done on extending other parts of traditional database theory, for instance schema design and dependency theory. It should be clear that dependency theory is of interest in this context. For instance, in [18], one finds a taxonomy of dependencies that are useful for temporal databases. Moreover, many *integrity constraints* over interpreted data can be represented as generalized dependencies. For instance, the integrity constraints over databases with ordered domains studied in [17, 33] can be represented as generalized dependencies. Also, some versions of the constraint checking problem studied in [16] can be viewed as generalized dependency implication problems.

One might think that the study of dependency theory has been close to exhaustive. While this is largely so for dependencies over uninterpreted data (that is, the context in which data values can only be compared for equality) [29], the situation is quite different for dependencies over data domains with a richer structure. The subject of this paper is the theory of these interpreted dependencies.

Specifically, we study the class of *constraint-generating dependencies*. These are the generalization of equality-generating dependencies [6], allowing arbitrary constraints on the data domain to appear wherever the latter only allow equalities. For instance, a constraint-generating dependency over an ordered domain can specify that if the value of an attribute A in a tuple t_1 is less than the value of the same attribute in a tuple t_2, then an identical relation holds for the values of an attribute B. This type of dependency can express a wide variety of constraints on the data. For instance, most of the temporal dependencies appearing in the taxonomy of [18] are constraint-generating dependencies.

Our technical contributions address the implication and the consistency[4] problems for constraint-generating dependencies. The natural approach to these problems is to write the dependencies as logical formulas. Unfortunately, the resulting formulas are not just formulas in the theory of the data domain. Indeed, they also contain uninterpreted predicate symbols representing the relations and thus are not a priori decidable, even if the data domain theory is decidable.

To obtain decision procedures, we show that the predicate symbols can be eliminated. Since the predicate symbols are implicitly universally quantified, this can be viewed as a form of second-order quantifier elimination. It is based on the fact that it is sufficient to consider relations with a small finite number of tuples. This then allows quantifier elimination by explicit representation of the possible tuples. The fact that one only needs to consider a small finite number of tuples is analogous to the fact that the implication problem for functional dependencies can be decided over 2-tuple relations [24]. Furthermore, for pure functional dependencies, our quantifier elimination procedures yields exactly the usual reduction to propositional logic. For more general constraint dependencies, it yields a formula in the theory of the data domain. Thus, if this theory is decidable, the implication and the consistency problems for constraint-dependencies are also decidable. Our approach is based on simple general logical arguments and provides a clear and straightforward justification for the type of procedure based on containment mappings used for instance in [16].

The complexity of the decision procedure depends on the specific data domain being considered and on the exact form of the constraint dependencies. We consider three typical constraint languages: equalities/inequalities, ordering constraints, and linear arithmetic constraints. We give a detailed picture of the complexity of the implication problem for dependencies over these theories and show the impact of the form of the dependencies on tractability.

2 Constraint-Generating Dependencies

Consider a relational database where some attributes take their values in specific domains, such as the integers or the reals, on which a set of predicates and functions are defined. We call such attributes *interpreted*. For the simplicity of the presentation, let us assume that the database only contains one (universal) relation r and let us ignore the noninterpreted attributes. In this context, it is natural to generalize the notion of equality-generating dependency [6]. Rather than specifying the propagation of equality constraints, we write similar statements involving arbitrary constraints (i.e., arbitrary formulas in the theory of the data domain). Specifically, we define *constraint-generating k-dependencies* as follows (the constant k specifies the number of tuples the dependency refers to).

[4] Though consistency is always satisfied for equality-generating dependencies, more general constraints turn it into a nontrivial problem.

Definition 1. Given a relation r, a *constraint-generating k-dependency* over r (with $k \geq 1$) is a first-order formula of the form

$$(\forall t_1)\cdots(\forall t_k)\Big[\big[r(t_1) \wedge \cdots \wedge r(t_k) \wedge C[t_1,\ldots,t_k]\big] \Rightarrow C'[t_1,\ldots,t_k]\Big]$$

where $C[t_1,\ldots,t_k]$ and $C'[t_1,\ldots,t_k]$ denote arbitrary constraint formulas relating the values of various attributes in the tuples $t_1,\ldots,t_k$. There are no restrictions on these formulas, they can include all constructs of the constraint theory under consideration, including quantification on the constraint domain. For instance, a constraint $C[t_1,t_2]$ could be $\exists z(t_1[A] < z \wedge z < t_2[A])$.

Note that we have defined constraint-generating dependencies in the context of a single relation, but the generalization to several relations is immediate.

Constraint-generating 1-dependencies as well as constraint-generating 2-dependencies are the most common. Notice that functional dependencies are a special form of constraint-generating 2-dependencies. Constraint-generating dependencies can naturally express a variety of arithmetic integrity constraints. The following examples illustrate their definition and show some of their potential applications.

Example 1. In [18], an exhaustive taxonomy of dependencies that can be imposed on a temporal relation is given. Of the more than 30 types of dependencies that are defined there, all but 4 can be written as constraint-generating dependencies. These last 4 require a generalization of tuple-generating dependencies [6] (see Section 5).

For instance, let us consider a relation $r(tt, vt)$ with two temporal attributes: transaction time (tt) and valid time (vt). The property of r being "strongly retroactively bounded" with bound $c \geq 0$ is expressed as the constraint-generating 1-dependency

$$(\forall t_1)\big[r(t_1) \Rightarrow [(t_1[tt] \leq t_1[vt] + c) \wedge (t_1[vt] \leq t_1[tt])]\big].$$

The property of r being "globally nondecreasing" is expressed as the constraint generating 2-dependency

$$(\forall t_1)(\forall t_2)\big[[r(t_1) \wedge r(t_2) \wedge (t_1[tt] < t_2[tt])] \Rightarrow (t_1[vt] \leq t_2[vt])\big].$$

Example 2. Let us consider a relation $emp(name, boss, salary)$. Then the fact that an employee cannot make more than her boss is expressed as

$$(\forall t_1)(\forall t_2)\big[[emp(t_1) \wedge emp(t_2) \wedge (t_1[boss] = t_2[name])] \Rightarrow (t_1[salary] \leq t_2[salary])\big].$$

3 Decision Problems for Constraint-Generating Dependencies

There are two basic decision problems for constraint-generating dependencies.

- *Implication*: Does a finite set of dependencies D imply a dependency d_0?
- *Consistency*: Does a finite set of dependencies D have a non-trivial model, that is, is D true in a nonempty relation?

The implication problem is a classical problem of database theory. Its practical motivation comes from the need to detect redundant dependencies, that is, those that are implied by a given set of dependencies. It is also the basis for proving the equivalence of dependency sets, and consequently for finding covers with desirable properties, such as minimality. The consistency problem has a trivial answer for uninterpreted dependencies: every set of equality- and tuple-generating dependencies has a 1-element model. However, even a single constraint-generating dependency may be inconsistent, as illustrated by $(\forall t)[r(t) \Rightarrow t[1] < t[1]]$. We only study the *implication* problem since the consistency problem is its dual: a set of dependencies D is inconsistent if and only if D implies a dependency of the form $(\forall t)[r(t) \Rightarrow C]$, where C is any unsatisfiable constraint (we assume the existence of at least one such unsatisfiable constraint formula).

The result we prove in this section is that the implication problem for constraint-generating dependencies reduces to the validity problem for a formula in the underlying constraint theory. Specific dependencies and theories will be considered in Section 4, and the corresponding complexity results provided. The reduction proceeds in three steps. First, we prove that the implication problem is equivalent to the implication problem restricted to finite relations of bounded size. Second, we eliminate from the implication to be decided the second-order quantification (over relations). Third, we eliminate the first-order quantification (over tuples) from the dependencies themselves and replace it by quantification over the domain – a process that we call *symmetrization*. This gives us the desired result.

3.1 Statement of the Problem and Notation

Let r denote a relation with n interpreted attributes. Let $d_0, d_1, \ldots, d_m$ denote constraint-generating k-dependencies over the attributes of r. The value of k need not be the same for all d_i's. We denote by k_0 the value of k for d_0.

The *dependency implication problem* consists in deciding whether d_0 is implied by the set of dependencies $D = \{d_1, \ldots, d_m\}$. In other words, it consists in deciding whether d_0 is satisfied by every interpretation that satisfies D, which can be formulated as

$$(\forall r)\big[r \models D \Rightarrow r \models d_0\big], \tag{1}$$

where D stands for $d_1 \wedge \cdots \wedge d_m$. We equivalently write (1) as $(\forall r)\big[D(r) \Rightarrow d_0(r)\big]$ when we wish to emphasize the fact that the dependencies apply to the tuples of r.

3.2 Towards a Decision Procedure

Reduction to k-tuple Relations. The following three lemmas establish that when dealing with constraint-generating k-dependencies, it is sufficient to consider relations of size[5] k. Their proofs are straightforward.

Lemma 2. *Let d denote any constraint-generating k-dependency. If a relation r does not satisfy d, then there is a relation r' of size k that does not satisfy d. Furthermore, r' is obtained from r by removing and/or duplicating tuples.*

Lemma 3. *If a relation r satisfies a set of constraint-generating k-dependencies $D = \{d_1, \ldots d_m\}$ and does not satisfy a constraint-generating k_0-dependency d_0, then there is a relation r' of size k_0 that satisfies D but does not satisfy d_0.*

Lemma 4. *Consider an instance (D, d_0) of the dependency implication problem where d_0 is a constraint-generating k_0-dependency. The dependency d_0 is implied by D over all relations if and only if it is implied by D over relations of size k_0; i.e., $(\forall r)\big[r \models D \Rightarrow r \models d_0\big]$ iff $(\forall r')\big[|r'| = k_0 \Rightarrow \big[r' \models D \Rightarrow r' \models d_0\big]\big]$.*

The above lemmas generalize properties of uninterpreted dependencies.

Second-order Quantifier Elimination. By Lemma 4, in order to decide the implication problem, we just need to be able to decide this problem over relations of size k for a given k. Deciding the implication (1) thus reduces to deciding

$$(\forall r')\big[[|r'| = k \wedge D(r')] \Rightarrow d_0(r')\big]. \tag{2}$$

Let $r' = \{t_{x_1}, \ldots, t_{x_k}\}$ denote an arbitrary relation of size k where $t_{x_1}, \ldots, t_{x_k}$ are arbitrary tuples. We can eliminate the (second-order) quantification over relations from the implication (2) and replace it with a quantification over tuples (that is, over vectors of elements of the domain). We get

$$(\forall t_{x_1}) \cdots (\forall t_{x_k})\big[D(\{t_{x_1}, \ldots, t_{x_k}\}) \Rightarrow d_0(\{t_{x_1}, \ldots, t_{x_k}\})\big]. \tag{3}$$

Symmetrization. Next, we simplify the formula (3), whose validity is equivalent to the constraint dependency implication problem, by eliminating the quantification over tuples that appears within the dependencies of $D \cup \{d_0\}$. We refer to this quantifier elimination procedure for dependencies as *symmetrization*. For the sake of clarity, we present the details of the symmetrization process for the case where $k = 2$. The process can be extended directly to the more general case.

For the case where $k = 2$, the formula (3) to be decided is the following.

$$(\forall t_x)(\forall t_y)\big[D(\{t_x, t_y\}) \Rightarrow d_0(\{t_x, t_y\})\big].$$

We can simplify this formula further by eliminating the quantification over tuples that appears in the dependencies $d(\{t_x, t_y\})$ in $D \cup \{d_0\}$. Every such dependency $d(\{t_x, t_y\})$ can indeed be rewritten as a constraint formula $cf(d)$ in the following manner.

---

[5] In what follows, we consider relations as multisets rather than sets. This has no impact on the implication problem, but simplifies our procedure.

1. Let d be a 1-dependency, that is, d is of the form $(\forall t)\big[[r'(t) \wedge C[t]] \Rightarrow C'[t]\big]$. This dependency considered over $r' = \{t_x, t_y\}$ is equivalent to the constraint formula
$$cf(d) : \big[C[t_x] \Rightarrow C'[t_x]\big] \wedge \big[C[t_y] \Rightarrow C'[t_y]\big],$$
which is a conjunction of $k = 2$ constraint implications. Notice that the t_x and t_y appearing in this formula are just tuples of variables ranging over the domain of the constraint theory of interest.
2. Let d be a 2-dependency, that is, d is of the form
$$(\forall t_1)(\forall t_2)\big[[r'(t_1) \wedge r'(t_2) \wedge C[t_1, t_2]] \Rightarrow C'[t_1, t_2]\big].$$
This dependency considered over $r' = \{t_x, t_y\}$ is equivalent to the constraint formula
$$\begin{aligned} cf(d) : &\big[C[t_x, t_y] \Rightarrow C'[t_x, t_y]\big] \wedge \big[C[t_y, t_x] \Rightarrow C'[t_y, t_x]\big] \wedge \\ &\big[C[t_x, t_x] \Rightarrow C'[t_x, t_x]\big] \wedge \big[C[t_y, t_y] \Rightarrow C'[t_y, t_y]\big], \end{aligned}$$
which is a conjunction of $k^k = 4$ constraint implications.

The rewriting of d as $cf(d)$ is what we call the *symmetrization* of d, for rather obvious reasons. It extends directly to any value of k. Notice that for a given k, any j-dependency d is rewritten as a constraint formula $cf(d)$, which is a conjunction of k^j constraint implications. Interestingly, in the case of functional dependencies, symmetrization is not needed. This is due to the fact that the underlying constraints are equalities, which are already symmetric. Hence, in that special case, symmetrization would produce several instances of the same constraint formulas.

Applying the symmetrization process to all the dependencies appearing in the formula (3), we get
$$(\forall t_{x_1}) \cdots (\forall t_{x_k})\big[cf(d_1) \wedge \cdots \wedge cf(d_m) \Rightarrow cf(d_0)\big]. \tag{4}$$
Notice that in formula (4), each tuple variable can be replaced by n domain variables, and thus the quantification over tuples can be replaced by a quantification over elements of the domain. For the sake of clarity, we simply denote by $(\forall *)$ the adequate quantification over elements of the domain (the *universal closure*). Formula (4) thus becomes
$$(\forall *)\big[cf(d_1) \wedge \cdots \wedge cf(d_m) \Rightarrow cf(d_0)\big], \tag{5}$$
where each $cf(d)$ is a conjunction of k^j constraint implications if d is a j-dependency and d_0 is a k-dependency. Thus, we have proved the following theorem.

Theorem 5. *For constraint-generating k-dependencies, with bounded k, the implication problem is linearly reduced to the validity of a universally quantified formula of the constraint theory.*

Example 3. Let us consider the following constraint-generating 2-dependencies over a relation r with a single attribute.

$$d_1 : (\forall x)(\forall y)\big[r(x) \wedge r(y) \Rightarrow x \leq y\big]$$
$$d_2 : (\forall x)(\forall y)\big[r(x) \wedge r(y) \Rightarrow x = y\big]$$

Symmetrizing them produces the following constraint formulas.

$$cf(d_1) : x \leq y \wedge y \leq x \wedge x \leq x \wedge y \leq y$$
$$cf(d_2) : x = y \wedge y = x \wedge x = x \wedge y = y$$

It is clear that these two constraint formulas are equivalent, as they should be.

4 Complexity Results

4.1 Clausal dependencies

In this section, we study the complexity of the implication problem for some classes of constraint-generating dependencies occurring in practice, in particular dependencies with equality, order, and arithmetic constraints. We restrict our attention to atomic constraints and clausal dependencies as defined below.

Definition 6. An *atomic constraint* is a formula consisting of an interpreted predicate symbol applied to terms. A *clausal* constraint-generating dependency is a constraint-generating dependency such that the constraint in the antecedent is a conjunction of atomic constraints and the constraint in the consequent is an atomic constraint.

Notice that a constraint-generating dependency in which the constraint in the antecedent and the constraint in the consequent are both conjunctions of atomic constraints can be rewritten as a set of clausal constraint-generating dependencies (by decomposing the conjunction in the consequent). Essentially all the dependencies mentioned in [18] can be written in clausal form.

Moreover, we assume that the constraint language is *closed under negation*.[6] This is again satisfied by many examples of interest, the most notable exception being the class of functional dependencies. Finally, we study classes of k-dependencies for fixed values of k (mainly $k = 2$). This makes it possible to contrast our results with the results about functional dependencies which are 2-dependencies and for which the implication problem can be solved in $O(n)$.

We proceed by reducing clausal dependency implication to unsatisfiability. More precisely, we negate the result of the symmetrization (formula 5) to obtain

$$(\exists *)\big[cf(d_1) \wedge \cdots \wedge cf(d_m) \wedge \neg cf(d_0)\big], \tag{6}$$

and then move the negation inwards and put the result in conjunctive normal form. Because the constraint language is closed under negation, the implication

[6] Note that in this context, the distinction between positive and negative atomic constraints is meaningless.

problem for clausal dependencies can thus be reduced to the unsatisfiability of a formula of the form

$$\Psi = (\exists *)\Big[\bigwedge_i \Big(\bigvee_j (c_{ij})\Big)\Big], \tag{7}$$

where each c_{ij} is an atomic constraint. When $|D| = m$ and d_0 is a k-dependency, the number of clauses in the formula Ψ above is at most equal to $m \cdot k^k$ plus the number of constraints in d_0. The number of literals in each clause is equal to the number of atomic constraints in the dependencies of D, or to 1 for the clauses obtained from the decomposition of d_0. Thus deciding the validity of the implication problem for k-dependencies (k fixed) can be done by checking the unsatisfiability of a conjunction of clauses of length that is linear in the size of $D \cup \{d_0\}$. We can replace the variables in the constraint formulas by the corresponding Skolem constants and view Ψ as a ground formula.

The opposite LOGSPACE reduction, from unsatisfiability to implication, also exists and requires only 1-dependencies.

4.2 Equality and order constraints

We consider here atomic constraints of the form $x\theta y$ where $\theta \in \{=, \neq, <, \leq\}$ over integers, rationals, or reals.[7] This constraint language has two sublanguages closed under negation which we also study: $\{=, \neq\}$-constraints and $\{<, \leq\}$-constraints. We make the additional assumption that *no domain constants appear in the dependencies.* (If this assumption is not satisfied, the complexity usually shifts up. For example, in Theorem 7 the first case becomes co-NP-complete for the integers by the results of [26].)

Theorem 7. *The implication problem for clausal constraint-generating k-dependencies is:*

1. *in PTIME for dependencies with one atomic $\{=, \neq, <, \leq\}$-constraint (no constraints in the antecedent),*
2. *co-NP-complete for dependencies with two or more atomic $\{=, \neq\}$-constraints,*
3. *co-NP-complete for dependencies with two or more atomic $\{<, \leq\}$-constraints.*

Proof Sketch. The first result follows from [30, page 892]. The membership in co-NP for the two remaining cases follows from the fact that checking the satisfiability of a conjunction of equality and order constraints can be done in polynomial time. To prove the lower bounds, we reduce an NP-complete problem to satisfiability of a set of ground clauses with at most two literals corresponding to the formula Ψ above (formula 7). This reduction is then composed with the reduction from unsatisfiability to dependency implication. We use a reduction from GRAPH-3-COLORABILITY for $\{=, \neq\}$-constraints, and from BETWEENNESS [13, page 279] for $\{<, \leq\}$-constraints. Details are omitted due to space limitation.

[7] In fact, our lower bounds hold for any infinite linearly-ordered set.

Note that for finite domains of size greater than 2, the implication problem is co-NP-complete even for dependencies with one atomic constraint.

The above results are rather negative. To obtain more tractable classes, we propose to further restrict the syntax of dependencies by typing.

Definition 8. A clausal dependency is *typed* if each atomic constraint involves only the values of one given attribute in different tuples.

The second dependency in Example 1 of Section 2 (i.e., the property of r being "globally nondecreasing") is typed, while the first one (the property of r being "strongly retroactively bounded") and the dependency of Example 2 are not. Functional dependencies are also typed.

Notice that for typed dependencies, the reduction from unsatisfiability to dependency implication given above is not useful for obtaining lower bounds. Indeed, it reduces unsatisfiability to implication of 1-dependencies which are not typed. Furthermore, this reduction cannot in general be adapted to yield typed 2-dependencies. Indeed, because of the symmetrization procedure, the constraint problem obtained from typed 2-dependencies has a particular symmetric structure (for 1-dependencies, there is no symmetrization). The question thus is whether this symmetric structure is sufficient for lowering the complexity of the constraint problem that has to be solved. As shown in the following theorem, the answer is fortunately positive.

Theorem 9. *The implication problem for typed clausal constraint-generating* 2-*dependencies with at most two atomic* $\{=, \neq, <, \leq\}$-*constraints is in PTIME* $(O(n))$.

Proof. A typed 2-dependency is of the form

$$(\forall t_x)(\forall t_y)\Big[\big[r(t_x) \wedge r(t_y) \wedge (t_x[i]\ pred_\ell\ t_y[i])\big] \Rightarrow (t_x[j]\ pred_r\ t_y[j])\Big] \qquad (8)$$

where each of $pred_\ell$ and $pred_r$ is one of $\{=, \neq, <, \leq\}$. By Lemma 4, the implication problem for typed 2-dependencies coincides with the implication problem over 2-tuple relations. The remaining steps of the reduction given in Section 3 show how this implication can be reduced to a pure constraint problem. However, since we need to take into account the specific nature of the constraint problem obtained for typed 2-dependencies, our starting point for the proof of this theorem is further upstream. We consider the problem of deciding whether for a typed 2-dependency d_0 and a set D of dependencies of the same kind, $D \models d_0$ over 2-tuple relations, or equivalently whether $D \wedge \neg d_0$ is unsatisfiable over 2-tuple relations. We give a PTIME algorithm for deciding satisfiability (and hence unsatisfiability) over 2-tuple relations of $D \wedge \neg d_0$.

Among the predicates in $\{=, \neq, <, \leq\}$, we distinguish the set *eq-pred* : $\{=, \leq\}$, and the set *diff-pred* : $\{\neq, <\}$. The intuition is that members of *eq-pred* can be satisfied when their arguments are equal, whereas members of *diff-pred* cannot be satisfied in that case. This allows us to define four classes of constraint dependencies:

$$\textit{eq-pred} \Rightarrow \textit{eq-pred} \quad (9) \qquad\qquad \textit{diff-pred} \Rightarrow \textit{eq-pred} \quad (11)$$
$$\textit{eq-pred} \Rightarrow \textit{diff-pred} \quad (10) \qquad\qquad \textit{diff-pred} \Rightarrow \textit{diff-pred} \quad (12)$$

Notice that (10) and (11) are self-contrapositives, whereas (9) and (12) are each other's contrapositives. We thus only need one of the latter two categories and choose to keep (12). Furthermore, all dependencies of the form (10) are unsatisfiable (over nonempty relations). Indeed, if in (8) one chooses $t_x = t_y$, then $(t_x[i]\ \textit{eq-pred}\ t_x[i])$ is true whereas $(t_x[j]\ \textit{diff-pred}\ t_y[j])$ has to be false and the implication is false. Thus, if such a dependency occurs in D, this set is trivially unsatisfiable and we can assume without loss of generality that D only contains dependencies of the forms (11) and (12). Similarly, if d_0 is of the form (10), $\neg d_0$ is valid and, since D is always satisfiable by a one tuple relation if it does not contain dependencies of the form (10), $D \wedge \neg d_0$ is satisfiable. We can thus also assume without loss of generality that d_0 is either of the form (11) or of the form (12).

Since the dependencies are typed, each dependency d involves two attributes of the relation r which we refer to as l_d (the one on the left of the implication) and r_d (the one on the right of the implication). We are looking for a 2-tuple model of $D \wedge \neg d_0$. The first step of the procedure is to classify the attributes of the relation r into the set of those that must have a different value in the two tuples of the relation and those that may have the same value. We call the first *diff*-attributes and the second *eq*-attributes. The set DA of *diff*-attributes is obtained by the following procedure.

The initial extension of DA is obtained from d_0. If d_0 is of the form $\textit{diff-pred} \Rightarrow \textit{eq-pred}$, then DA initially contains both l_{d_0} and r_{d_0}; whereas if d_0 is of the form $\textit{diff-pred} \Rightarrow \textit{diff-pred}$, then initially $DA = \{l_{d_0}\}$. One then repeatedly applies the following step until saturation: if there is a dependency d of the form $\textit{diff-pred} \Rightarrow \textit{diff-pred}$ such that the attribute l_d is in DA, then the attribute r_d is added to DA. This procedure is similar to the one computing the closure of a set of attributes under a set of functional dependencies and hence can be implemented in linear time. From now on, let DA be the set of attributes obtained by this procedure.

A direct consequence of the way in which DA is constructed is that any 2-tuple model in which both tuples give the same value to attributes in DA cannot satisfy $D \wedge \neg d_0$. Furthermore, we claim that if $D \wedge \neg d_0$ has a 2-tuple model, it has a 2-tuple model in which all attributes in DA have different values in the two tuples, and all attributes not in DA have the same value in both tuples. To prove this, assume there is a model and give an arbitrary identical value in both tuples to the attributes not in DA. Since this can only change the truth value of the dependencies in D and of $\neg d_0$ from *false* to *true*, we still have a model.

Thus, in order to find a model for $D \wedge \neg d_0$, it is sufficient to find values for the attributes in DA. We know that these values have to be different and, since we are working in a limited theory, the only relevant property of these values is their order (which one is smaller than the other). Let us call the two possible orders u (up) and d (down). The choice between u and d for each attribute i can be encoded by one boolean proposition $u[i]$ (*true* if the order for i is u, *false* if it is d). The problem thus is to find truth values for the propositions $u[i]$ in such a way that they define a model of $D \wedge \neg d_0$.

To do this, we encode the conditions imposed by the dependencies referring to attributes that are both in DA. Indeed, for dependencies in D (and for $\neg d_0$), if one of the atomic constraints does not refer to an attribute in DA, the dependency ($\neg d_0$) is satisfied whatever the order chosen for the attributes.

We construct the constraints on the propositions u for dependencies in positive form as they appear in D. For $\neg d_0$, one applies the construction to d_0 and negates the result. There are 9 cases of dependencies of the form *diff-pred* $\Rightarrow$ *diff-pred*:

$$\begin{array}{llllll} \neq \Rightarrow \neq & (13) & < \Rightarrow \neq & (16) & > \Rightarrow \neq & (19) \\ \neq \Rightarrow < & (14) & < \Rightarrow < & (17) & > \Rightarrow < & (20) \\ \neq \Rightarrow > & (15) & < \Rightarrow > & (18) & > \Rightarrow > & (21) \end{array}$$

Cases 13, 16, and 19 translate to *true* (we have imposed that attributes in DA have different values in both tuples). Cases 14 and 15 are always unsatisfiable (by symmetry) and thus translate to the constraint *false*. Cases 17 and 21 translate to $(u[l_d] \Rightarrow u[r_d]) \wedge (\neg u[l_d] \Rightarrow \neg u[r_d])$, whereas cases 18 and 20 translate to $(u[l_d] \Rightarrow \neg u[r_d]) \wedge (\neg u[l_d] \Rightarrow u[r_d])$.

There are also 9 cases of dependencies of the form *diff-pred* $\Rightarrow$ *eq-pred*:

$$\begin{array}{llllll} \neq \Rightarrow = & (22) & < \Rightarrow = & (25) & > \Rightarrow = & (28) \\ \neq \Rightarrow \leq & (23) & < \Rightarrow \leq & (26) & > \Rightarrow \leq & (29) \\ \neq \Rightarrow \geq & (24) & < \Rightarrow \geq & (27) & > \Rightarrow \geq & (30) \end{array}$$

Cases 22, 23, 24, 25, and 28 are contradictory and translate to *false*. Cases 26 and 30 translate as 17 and 21 and, similarly, 27 and 29 are translated as 18 and 20.

The result of this encoding is a set of Boolean clauses with at most two literals per clause. Deciding if it is satisfiable can thus be done with the 2-SAT procedure which is in PTIME ($O(n)$) [1].

Theorem 10. *The implication problem for typed clausal constraint-generating 2-dependencies is:*

1. *co-NP-complete for dependencies with three or more atomic* $\{=, \neq\}$*-constraints,*
2. *co-NP-complete for dependencies with three or more atomic* $\{<, \leq\}$*-constraints.*

Proof Sketch. Proving the lower bounds in the typed case is more difficult than in Theorem 7 because the reverse reduction, from unsatisfiability of ground clauses to dependency implication that uses 1-dependencies, is not available. We can continue, however, to work with ground clauses as in the proof of Theorem 7 provided the clauses can be mapped back to typed 2-dependencies. The proofs in both cases involve a reduction from SET SPLITTING [13, page 221]. They proceed in two steps. First, we reduce SET SPLITTING to a collection of ground clauses C. Then we show how to construct an instance of the implication problem for typed 2-dependencies whose clausal formulation Ψ (see formula 7) is equisatisfiable with C. Details are omitted due to space limitation.

Theorem 9 yields a new class of dependencies with a tractable implication problem. This class properly contains that of unary functional dependencies and is incomparable with the class of all functional dependencies. Together, Theorems 7, 9 and 10 give a *complete classification* of tractable and intractable classes of untyped and typed 2-dependencies with $\{=, \neq, <, \leq\}$-constraints. The case of typed k-dependencies ($k > 2$) with two $\{=, \neq, <, \leq\}$-constraints is open. (The implication problem for such dependencies with three constraints is clearly co-NP-complete by Theorems 7 and 10.)

4.3 Linear arithmetic constraints

We consider now *linear arithmetic constraints*, i.e., atomic constraints of the form $a_1 x_1 + \cdots + a_k x_k \leq a$ (domain constants are allowed here). We can use directly the results about the complexity of linear programming [27].

Theorem 11. *For linear arithmetic constraints, the implication problem for clausal constraint-generating k-dependencies with one atomic constraint per dependency is in PTIME for the reals, and co-NP-complete for the integers.*

The case of more than one linear arithmetic constraint per dependency remains to be investigated.

5 Conclusions and Related Work

A brief summary of this paper is that constraint-generating dependencies are an interesting concept, and that deciding implication of such dependencies is basically no harder than deciding the underlying constraint theory, which, a priori, was not obvious. The obvious applications of constraint-generating dependencies are constraint database design theory and consistency checking. Apart from the constraint languages considered in this paper, other languages may be relevant as well, for instance the *congruence constraints* that appear in [18]. Also, the impact that the presence of domain constants in equality and order constraints has on the complexity of implication should be fully studied.

Other forms of constraint dependencies can also be of interest. An obvious candidate is the concept of *tuple-generating* constraint dependency. Unfortunately, the implication problem for these dependencies is harder to decide and more closely linked to the underlying theory. Indeed, tuple-generating constraint dependencies can, for example, specify a dense domain.

As far as related work, we should first mention that Jensen and Snodgrass [18] induced us to think about constraint dependencies. We should note that the integrity constraints over temporal databases postulated there involve both typed and untyped constraint-generating dependencies, as well as tuple-generating ones.

Two recent papers on *implication constraints* by Ishakbeyoğlu, Ozsoyoğlu and Zhang [17, 33], as well as a paper on efficient integrity checking by Gupta, Sagiv, Ullman, and Widom [16] contain work fairly close to ours. However, there are several important differences. Foremost, all three papers discuss a fixed language of constraint formulas, namely equality ($=$), inequality ($\neq$), and order ($<, \leq$) constraints, while our results are applicable to any decidable constraint theory thanks to our general reduction strategy. In particular, the papers [33, 16], which were written independently of the first version of this paper, both present results equivalent to our Theorem 5, but formulated in the context of a fixed constraint language. Also, the proof techniques in those papers, based on the theory of conjunctive queries, are quite different from ours. Moreover, the complexity results of [33] are obtained in a slightly different model. Both the number of database literals and the arity of relations in a dependency are considered as parts of the input, while we consider only the latter. We think that our model is more intuitive because it is difficult to come up with a meaningful dependency that references more than a few tuples in a relation. Our intractability results are stronger than those of [33] while our positive characterizations of polynomial-time decidable problems do not necessarily carry over to the framework of [33]. Also, in [17, 33], the tractable classes of dependencies are not defined syntactically but rather by the presence or absence of certain types of refutations.

A clausal constraint-generating dependency (quantifiers omitted)

$$r(t_1) \wedge \cdots \wedge r(t_k) \wedge C_1 \wedge \cdots \wedge C_n \Rightarrow C_0$$

can be viewed as an integrity constraint (in the notation of [16])

$$\textbf{panic} \; :- \; r(t_1) \,\&\, \cdots \,\&\, r(t_k) \,\&\, C_1 \,\&\, \cdots \,\&\, C_n \,\&\, \neg C_0.$$

Thus the implication of a dependency by a set of dependencies is equivalent to the subsumption of an integrity constraint by a set of integrity constraints. Therefore the results about the complexity of implication from Section 4 transfer directly to the context of constraint subsumption. The paper [16] applies the results about constraint subsumption to develop techniques for efficient integrity checking. Unfortunately, this application requires introducing constants into constraints, so our complexity results, developed under the assumption that constants do not appear in dependencies, are not applicable here, though our general reduction is.

Order dependencies, proposed by Ginsburg and Hull [14, 15], are typed clausal 2-dependencies over the theory of equality and order (without $\neq$). The order is not required to be total. Ginsburg and Hull provided an axiomatization of such dependencies and proved that the implication problem is co-NP-complete for dependencies with at least three constraints. To prove the lower bound they used, however, dependencies with equality and order constraints, while we proved the lower bounds for both theories separately (Theorem 10). Ginsburg and Hull also supplied a number of tractable dependency classes which are, again, different from ours and involve mainly partial orders.

Acknowledgements. We wish to acknowledge anonymous referees for several helpful comments.

References

1. B. Aspvall, M. Plass, and R. Tarjan. A linear-time algorithm for testing the truth of certain quantified boolean formulas. *Inf. Process. Lett.*, 8(3):121–123, 1979.
2. M. Baudinet. On the expressiveness of temporal logic programming. To appear in *Information and Computation*.
3. M. Baudinet. Temporal logic programming is complete and expressive. In *Sixteenth ACM Symposium on Principles of Programming Languages*, pages 267–280, Austin, Texas, Jan. 1989.
4. M. Baudinet, J. Chomicki, and P. Wolper. Temporal deductive databases. In A. Tansel, et al., editors, *Temporal Databases. Theory, Design, and Implementation*, chapter 13, pages 294–320. Benjamin/Cummings, 1993.
5. M. Baudinet, M. Niézette, and P. Wolper. On the representation of infinite temporal data and queries. In *Tenth ACM Symposium on Principles of Database Systems*, pages 280–290, Denver, Colorado, May 1991.
6. C. Beeri and M. Vardi. A proof procedure for data dependencies. *Journal of the ACM*, 31(4):718–741, Oct. 1984.
7. A. Brodsky, J. Jaffar, and M. J. Maher. Toward practical constraint databases. In *19th International Conference on Very Large Data Bases*, Dublin, Aug. 1993.
8. A. Brodsky, C. Lassez, and J.-L. Lassez. Separability of polyhedra and a new approach to spatial storage. In *Proceedings of the First Workhop on Principles and Practice of Constraint Programming*, Newport, Rhode Island, Apr. 1993.
9. J. Chomicki. Polynomial time query processing in temporal deductive databases. In *Ninth ACM Symposium on Principles of Database Systems*, pages 379–391, Nashville, Tennessee, Apr. 1990.
10. J. Chomicki and T. Imieliński. Temporal deductive databases and infinite objects. In *Seventh ACM Symposium on Principles of Database Systems*, pages 61–73, Austin, Texas, Mar. 1988.
11. J. Chomicki and T. Imieliński. Finite Representation of Infinite Query Answers. *ACM Transactions on Database Systems*, 18(2):181–223, June 1993.
12. J. Cox and K. McAloon. Decision procedures for constraint based extensions of Datalog. In F. Benhamou and A. Colmerauer, editors, *Constraint Logic Programming: Selected Research*. MIT Press, 1993.
13. M. R. Garey and D. S. Johnson. *Computers and Intractability : A Guide to the Theory of NP-Completeness*. W.H. Freeman and Company, New York, 1979.
14. S. Ginsburg and R. Hull. Order dependency in the relational model. *Theoretical Computer Science*, 26:149–195, 1983.
15. S. Ginsburg and R. Hull. Sort sets in the relational model. *Journal of the ACM*, 33(3):465–488, July 1986.
16. A. Gupta, Y. Sagiv, J. D. Ullman, and J. Widom. Constraint checking with partial information. In *Thirteenth ACM Symposium on Principles of Database Systems*, pages 45–55, Minneapolis, MN, May 1994.
17. N. S. Ishakbeyoğlu and Z. M. Ozsoyoğlu. On the maintenance of implication integrity constraints. In *Fourth International Conference on Database and Expert Systems Applications*, pages 221–232, Prague, Sept. 1993. LNCS 720, Springer.

18. C. Jensen and R. Snodgrass. Temporal specialization. In *Eighth International Conference on Data Enfineering*, pages 594–603, Tempe, Arizona, Feb. 1992. IEEE.
19. F. Kabanza, J.-M. Stévenne, and P. Wolper. Handling infinite temporal data. In *Ninth ACM Symposium on Principles of Database Systems*, pages 392–403, Nashville, Tennessee, Apr. 1990.
20. P. Kanellakis. Elements of relational database theory. In J. van Leeuwen, editor, *Handbook of Theoretical Computer Science*, volume B, chapter 17, pages 1073–1158. Elsevier/MIT Press, 1990.
21. P. C. Kanellakis, G. M. Kuper, and P. Revesz. Constraint query languages. In *Ninth ACM Symposium on Principles of Database Systems*, pages 299–313, Nashville, Tennessee, Apr. 1990.
22. P. C. Kanellakis, S. Ramaswamy, D. E. Vengroff, and J. S. Vitter. Indexing for data models with constraints and classes. In *Twelfth ACM Symposium on Principles of Database Systems*, pages 233–243, Washington, DC, May 1993.
23. M. Koubarakis. Representation and querying in temporal databases : the power of temporal constraints. In *Ninth International Conference on Data Engineering*, Vienna, Austria, Apr. 1993.
24. D. Maier. *The Theory of Relational Databases.* Computer Science Press, 1983.
25. P. Revesz. A closed form for Datalog queries with integer order. In S. Abiteboul and P. Kanellakis, editors, *ICDT '90, Proceedings of the Third International Conference on Database Theory*, pages 187–201, Paris, Dec. 1990. LNCS 470, Springer.
26. D. Rosenkrantz and H. B. I. Hunt. Processing conjunctive predicates and queries. In *International Conference on Very Large Data Bases*, pages 64–72, 1980.
27. A. Schrijver. *Theory of Linear and Integer Programming.* John Wiley & Sons, 1986.
28. D. Srivastava. Subsumption in constraint query languages with linear arithmetic constraints. In *Second International Symposium on Artificial Intelligence and Mathematics*, Fort Lauderdale, Florida, Jan. 1992.
29. B. Thalheim. *Dependencies in Relational Databases.* Teubner-Texte zur Mathematik, Band 126. B.G. Teubner Verlagsgesellschaft, Stuttgart, 1991.
30. J. D. Ullman. *Principles of Database and Knowledge-Base Systems – Volume II: The New Technologies.* Computer Science Press, 1989.
31. R. van der Meyden. The complexity of querying indefinite data about linearly ordered domains. In *Eleventh ACM Symposium on Principles of Database Systems*, pages 331–345, San Diego, California, June 1992.
32. M. Vardi. Fundamentals of dependency theory. In E. Börger, editor, *Trends in Theoretical Computer Science*, pages 171–224. Computer Science Press, 1988.
33. X. Zhang and Z. M. Ozsoyoğlu. On efficient reasoning with implication constraints. In *Third International Conference on Deductive and Object-Oriented Databases*, Phoenix, Arizona, Dec. 1993.

This article was processed using the LaTeX macro package with LLNCS style