SIAM J. COMPUT.
Vol. 12, No. 1, February 1983

0097-5397/83/1201-0003 $01.25/0

TOOLS FOR TEMPLATE DEPENDENCIES*

RONALD FAGIN†, DAVID MAIER‡, JEFFREY D. ULLMAN§
AND MIHALIS YANNAKAKIS¶

Abstract. Template dependencies (TD's) are a class of data dependencies that include multivalued and join dependencies and embedded versions of these. A collection of techniques, examples and results about TD's are presented. The principal results are:

1) Finite implication (implication over relations with a finite number of tuples) is distinct from unrestricted implication for TD's.

2) There are, for TD's over three or more attributes, infinite chains of increasingly weaker and increasingly stronger full TD's.

3) However, there are weakest (nontrivial) and strongest full TD's over any given set of attributes.

4) Over two attributes, there are only three distinct TD's.

5) There is no weakest (not necessarily full) TD over any set of three or more attributes.

6) There is a finite relation that obeys every strictly partial TD but no full TD.

7) The conjunction of each finite set of full TD's is equivalent to a single full TD. However, the conjunction of a finite set of (not necessarily full) TD's is not necessarily equivalent to a single TD and the disjunction of a finite set of full TD's is not necessarily equivalent to a single TD.

8) There is a finite set of TD's with an infinite Armstrong relation but no finite Armstrong relation.

9) A necessary and sufficient condition for the existence of finite Armstrong relations for sets of TD's can be formulated in terms of the implication structure of TD's.

Key words. relational database, template dependency, finite implication, multivalued dependency, join dependency

1. Introduction. Template dependencies (TD's) were introduced by Sadri and Ullman [SU] and, independently, by Beeri and Vardi [BV2]. Both sets of authors introduced TD's to provide a class of dependencies (sentences about relations) that include join dependencies [Ri] and embedded multivalued dependencies [Fa2] and that also has a complete axiomatization (no complete axiomatization is known for either join dependencies or embedded multivalued dependencies). TD's are examples of the "tuple-generating dependencies" of Beeri and Vardi [BV2]. Tuple-generating dependencies, along with "equality-generating dependencies" (which include functional dependencies [Co]) together comprise Fagin's [Fa3] class of embedded implicational dependencies (which is equivalent to Yannakakis and Papadimitriou's [YP] class of algebraic dependencies). This paper is a compendium of techniques, examples and counterexamples for TD's.

In § 2, we present definitions. In § 3, we demonstrate the existence of a strongest TD and a weakest nontrivial full TD. (*Note.* Unless stated otherwise, TD's are not assumed to be full.) We show that there is no weakest TD. In § 4, we show that there are only three distinct TD's on two attributes. In § 5, we demonstrate a useful correspondence between TD's and graphs and introduce the notion of an lp-homomorphism (label-preserving homomorphism). In § 6, we utilize this correspondence to help prove the existence of infinite chains of progressively weaker and progressively stronger full TD's. In § 7, we show that for TD's, implication is distinct

* Received by the editors May 21, 1981, and in revised form March 15, 1982.

† IBM Research Laboratory, San Jose, California 95193.

‡ State University of New York, Center at Stony Brook, Stony Brook, New York, 11794. The research of this author was supported in part by the National Science Foundation under grant IST-79-18264.

§ Stanford University, Stanford, California 94305. The research of this author was supported in part by the National Science Foundation under grant MCS-79-04528.

¶ Bell Laboratories, Murray Hill, New Jersey 07974.

from implication restricted to finite relations. In § 8, we show that the conjunction of a finite set of full TD's is equivalent to a single full TD. However, we show that the conjunction of a finite set of TD's is not necessarily equivalent to a single TD, and the disjunction of a finite set of full TD's is not necessarily equivalent to a single TD. In § 9, we show that there is a finite relation that obeys every strictly partial TD but no nontrivial full TD. In § 10, we demonstrate a finite set of TD's with no finite Armstrong relation [Fa3] (although we know [Fa3] that there is an infinite Armstrong relation). We also give a necessary and sufficient condition for the existence of finite Armstrong relations for sets of TD's.

2. Definitions. A relational database scheme consists of a universal set of attributes U and a set of "dependencies". The attributes in U are names for the components (columns) of relations in the database. The most common forms of dependencies are functional dependencies, or FD's [Co], and multivalued dependencies, or MVD's [Fa2]. We shall not discuss FD's in this paper.

In database theory, a tuple is formally regarded as a mapping from attributes to values, rather than as a list of component values, although the latter viewpoint is handy when the order of the attributes in the list is understood. We often use $t[Z]$, where t is a tuple and Z is a set of attributes, to stand for t restricted to domain Z, that is, the components of t for the attributes in Z. If A is an attribute, then we call $t[A]$ the A *entry* or A *value* of t.

Multivalued dependencies are denoted syntactically by $X \twoheadrightarrow Y$. The meaning of this dependency is that if relation R obeys the dependency, and if t_1 and t_2 are tuples of R with $t_1[X]=t_2[X]$, then there exists t_3 in R such that:

1. $t_3[X]=t_1[X]=t_2[X]$,
2. $t_3[Y]=t_1[Y]$ and
3. $t_3[U-XY]=t_2[U-XY]$.

Intuitively, the set of Y-values associated with each given X-value is independent of the values in all other attributes. By XY in 3 above, we mean $X \cup Y$.

Example. Consider the relation R, in Fig. 1.1, where $U=\{A, B, C, D\}$.

A	B	C	D
0	1	2	3
0	2	1	4
0	1	1	4
0	2	2	3
5	1	3	2

FIG. 1.1. *The relation R.*

The MVD $A \twoheadrightarrow B$ holds in R. For example, if t_1 and t_2 are the first two tuples in Fig. 1.1, then we may check that the tuple t_3, where $t_3[A]=t_1[A]=t_2[A]=0$, $t_3[B]=t_1[B]=1$, and $t_3[CD]=t_2[CD]=14$, is present; it is row three. (By 14, we mean the tuple with first entry 1 and second entry 4; we shall sometimes find this type of abbreviation convenient.)

Let Σ be a set of dependencies, and let σ be a single dependency. When we say that Σ *logically implies* σ or that σ *is a logical consequence of* Σ, we mean that whenever every dependency in Σ holds for a relation R, then σ also holds for R. That is, there is no "counterexample relation" such that every dependency in Σ holds for R, but such that σ fails in R. We write $\Sigma \models \sigma$ to mean that Σ logically implies σ. For example, if A, B, and C are attributes, then $\{A \twoheadrightarrow B, B \twoheadrightarrow C\} \models A \twoheadrightarrow C$.

It appears that FD's and MVD's are almost sufficient to describe the "real world," and thus could be used for a database design theory. However, there is at least one, more general form of dependency that appears naturally, and this form causes severe difficulties when we try to infer dependencies. This type of dependency, called an embedded multivalued dependency (EMVD), was first studied by Fagin [Fa2] and Delobel [De]. For disjoint X, Y and Z, we say $X \twoheadrightarrow Y|Z$ holds if, when any "legal" relation over the set of attributes is projected onto the set of attributes XYZ (we project by restricting tuples to these attributes), then the MVD $X \twoheadrightarrow Y$ holds. (Note that $X \twoheadrightarrow Y$ holds in XYZ if and only if $X \twoheadrightarrow Z$ holds [Fa2]).

Another way of looking at the EMVD $X \twoheadrightarrow Y|Z$ is that if the relation R over attributes U obeys the dependency, then whenever we have two tuples t_1 and t_2 in R, and $t_1[X]=t_2[X]$, it follows that there is some t_3 in R, where

1. $t_3[X]=t_1[X]=t_2[X]$,
2. $t_3[Y]=t_1[Y]$ and
3. $t_3[Z]=t_2[Z]$.

Note that $t_3[U-XYZ]$ can be arbitrary; we can assert nothing about the values t_3 has in these components.

Unfortunately, when we try to make inferences about EMVD's we appear to run into a stone wall. It is not known whether the decision problem for EMVD's is decidable (the *decision problem* for EMVD's is the problem of deciding whether $\Sigma \models \sigma$, when Σ is a set of EMVD's and σ is a single EMVD). Neither is a complete axiomatization for EMVD's known. It is known [SW], [CFP] that there is no k-ary complete axiomatization for EMVD's for any fixed k, and, in particular, no finite complete axiomatization.

To tackle these problems for EMVD's, some more general types of dependencies have been studied recently, with the hope that the more general class would have a complete axiomatization or would provide insights on the EMVD decision problem. In particular, Sadri and Ullman [SU] and, independently, Beeri and Vardi [BV2] introduced *template dependencies*, or TD's, and provided a complete axiomatization. TD's include as special cases (a) MVD's, (b) EMVD's, (c) subset dependencies [SW], (d) mutual dependencies [Ni], (e) generalized mutual dependencies [MM] and (f) join dependencies [Ri]. The class of TD's was studied independently by Beeri and Vardi [BV2] and by Paradaens and Jannsens [PJ], and still more general classes were considered by Fagin [Fa3] and Yannakakis and Papadimitriou [YP]. Vardi [Va1] and, independently, Gurevich and Lewis [GL] have recently shown that the decision problem for TD's is undecidable.

A template dependency is an assertion about a relation R, that if we find tuples $r_1, \cdots, r_k$ in R with certain specific equalities among the entries of these tuples, then we can find in R a tuple r that has certain of its entries equal to certain of the entries in $r_1, \cdots, r_k$. Other entries of r may be arbitrary. Formally, we write a template dependency as $r_1, \cdots, r_k/r$, or as

$$\begin{array}{c} r_1 \\ \vdots \\ \underline{r_k} \\ r \end{array}$$

where the r_i's and r are strings of abstract symbols (sometimes called *variables*). The length of the r_i's and r equals the number of attributes in the universal set, and positions in these strings are assumed to correspond to attributes in a fixed order. No symbol may appear in two distinct components among the r_i's and r. It is, of course, permissible that one symbol appear in the same component of several of the r_i's or r.

Let R be a relation and let T be a TD. Let h be a homomorphism that maps symbols in T into entries of R. By saying that h is a homomorphism, we mean that $h(a_1 \cdots a_n)$ is defined to be $h(a_1) \cdots h(a_n)$. We call h a *valuation.* Relation R is said to obey TD T if whenever there is a valuation h on the symbols appearing in the r_i's such that $h(r_i)$ is a tuple in R for all i, then we can extend h to those symbols that appear in r but do not appear among the r_i's, in such a way that $h(r)$ is also in R.

Example. Let $U = \{A, B, C, D\}$ and let R be the relation previously given in Fig. 1.1. Let T be the TD

$$\begin{array}{cccc} a_1 & b_1 & c_1 & d_1 \\ a_2 & b_1 & c_2 & d_2 \\ a_1 & b_2 & c_2 & d_3 \\ \hline a_2 & b_3 & c_2 & d_1 . \end{array}$$

Define h by: $h(a_1) = h(a_2) = 0$; $h(b_1) = h(c_1) = 1$; $h(b_2) = h(c_2) = 2$; $h(d_2) = h(d_3) = 3$, and $h(d_1) = 4$. Then $h(a_1b_1c_1d_1) = 0114$, $h(a_2b_1c_2d_2) = 0123$, and $h(a_1b_2c_2d_3) = 0223$, which are rows three, one, and four of Fig. 1.1. Thus, we must exhibit a value b for $h(b_3)$ such that $h(a_2b_3c_2d_1)$ is in the relation of Fig. 1.1, if that relation is to obey the TD T. However, for no value of b is $0b24$ a row of Fig. 1.1, so we may conclude without further ado that R does not obey T. Of course, if a value of b had been found, we would then have to check all other possible valuations that mapped the first three rows of T into rows of Fig. 1.1.

When we say that a relation is *finite* (respectively, *infinite*), we mean that it has a finite (respectively, infinite) set of tuples. Database theory is most concerned with finite relations; however, sometimes it is convenient to consider infinite relations. If Σ is a set of dependencies, such as TD's, then by SAT (Σ), we mean the collection of relations (finite or infinite) that obey all of Σ. Note that $\Sigma \models \sigma$ if and only if SAT $(\Sigma) \subseteq$ SAT (σ). If we wish to consider only finite relations, then we can write $\mathrm{SAT_{fin}}(\Sigma)$ to mean the collection of finite relations that obey Σ. Similarly, we can define $\Sigma \models_{\mathrm{fin}} \sigma$ to mean that every finite relation that obeys Σ also obeys σ. As above, $\Sigma \models_{\mathrm{fin}} \sigma$ if and only if $\mathrm{SAT_{fin}}(\Sigma) \subseteq \mathrm{SAT_{fin}}(\sigma)$. Note that if $\Sigma \models \sigma$, then $\Sigma \models_{\mathrm{fin}} \sigma$. As we shall show in § 7, the converse fails for TD's.

When we speak of two dependencies σ and τ being *equivalent,* we mean that SAT (σ) = SAT (τ), or equivalently, that $\sigma \models \tau$ and $\tau \models \sigma$. Similarly, we can define equivalent *sets* of dependencies. We shall sometimes speak of conjunctions or disjunctions of TD's. A relation obeys the conjunction (respectively, disjunction) of a set of TD's precisely if it obeys all (respectively, at least one) of them. Thus,

$$\mathrm{SAT}\,(\wedge\{\sigma : \sigma \in S\}) = \bigcap\{\mathrm{SAT}\,(\sigma) : \sigma \in S\},$$

$$\mathrm{SAT}\,(\vee\{\sigma : \sigma \in S\}) = \bigcup\{\mathrm{SAT}\,(\sigma) : \sigma \in S\}.$$

The following terminology will prove helpful. If $r_1, \cdots, r_k/r$ is a TD, then $r_1, \cdots, r_k$ are called the *hypothesis rows,* or *hypotheses,* and r is the *conclusion row,* or simply the *conclusion.* Each symbol that appears in the conclusion is said to be *distinguished.* A TD is said to be *full* if each of its distinguished symbols also appears in the hypotheses; otherwise, it is said to be *strictly partial.* If T is a TD, and if V is exactly the set of attributes for which the hypothesis rows of T contain distinguished variables, then we may call T a *V-partial* TD (we allow the possibility that $V = U$, the set of all attributes). A TD is *trivial* if it always holds (in relations over the appropriate attributes).

Remark. A V-partial TD is trivial precisely if some hypothesis row of T contains distinguished variables for every one of its V entries. For if no hypothesis row of T contains distinguished variables for every one of its V entries, then the relation that consists of all of the hypothesis rows of T but not the conclusion is a relation not in SAT (T); hence, T is nontrivial.

Example. Let $U=\{A, B, C, D\}$. Then the MVD $A \twoheadrightarrow B$ is synonymous with the TD:

$$\begin{array}{cccc} a_1 & b_1 & c_1 & d_1 \\ a_1 & b_2 & c_2 & d_2 \\ \hline a_1 & b_1 & c_2 & d_2 . \end{array}$$

The EMVD $A \twoheadrightarrow B \mid C$ is written:

$$\begin{array}{cccc} a_1 & b_1 & c_1 & d_1 \\ a_1 & b_2 & c_2 & d_2 \\ \hline a_1 & b_1 & c_2 & d_3 . \end{array}$$

Note that this EMVD is a strictly partial TD. However, MVD's are full TD's.

3. Strongest and weakest TD's. An important tool in the study of dependencies is the chase process [ABU], [MMS], [SU]. When TD's alone are involved, could the chase go on forever in a nontrivial way? The question of the existence of infinite chases where "things keep happening" can be related to the existence of certain infinite sequences of TD's as follows. The set of rows in the tableau at any time during a chase may be taken to be the hypothesis rows of a TD whose conclusion row is the goal row for the chase. It is easy to show that as the chase proceeds, these TD's get progressively weaker. If the chase is successful, then we eventually arrive at a TD so weak that it is trivial.

If the chase is unsuccessful, then we might obtain an infinite sequence of TD's that, although some could be equivalent to the previous TD, would include an infinite subsequence of strictly weaker TD's. Or, we might necessarily reach a point where all successive TD's were equivalent but not trivial, and if we knew that we had reached that point, then we could deduce that the chase was unsuccessful.

These observations lead to the consideration of the structure of the space of TD's. Are there infinite sequences of strictly weaker TD's? Can we construct such a sequence by showing that for every nontrivial TD there is a weaker nontrivial TD? The answers to these (yes and no, respectively) and related questions are contained in later sections.

THEOREM 3.1. *For each set of attributes, there is a strongest* TD. *That is, there is a* TD T *such that* $T \models T'$ *for each* TD T' *over the same set of attributes as* T.

Proof. The TD that states a relation is a Cartesian product is the strongest TD. For example, the Cartesian product TD over three attributes is

$$\begin{array}{ccc} a_1 & b_1 & b_2 \\ b_3 & a_2 & b_4 \\ b_5 & b_6 & a_3 \\ \hline a_1 & a_2 & a_3 . \end{array}$$

The Cartesian product TD is strongest because each relation that is a Cartesian product is easily seen to obey every TD (over the same attributes). □

Recall that a TD is said to be V-partial if V is the set of attributes for which the hypothesis rows of T contain distinguished variables.

COROLLARY 3.2. *There is a strongest V-partial* TD. *That is, there is a V-partial* TD *T such that $T \models T'$ for every V-partial* TD *T' over the same attributes.*

Proof. The V-partial TD that says of a relation that its projection onto V is a Cartesian product is the strongest V-partial TD. Thus, if U is ABC and V is AB, then this TD is

$$\begin{array}{ccc} a_1 & b_1 & b_2 \\ b_3 & a_2 & b_4 \\ \hline a_1 & a_2 & a_3 . \end{array}$$

□

THEOREM 3.3. *Assume that V contains at least two attributes. Then there is a weakest nontrivial V-partial* TD. *That is, there is a nontrivial V-partial* TD *T such that $T' \models T$ for every nontrivial V-partial* TD *T' over the same attributes. In particular (when $V = U$) there is a weakest nontrivial full* TD.

Note. The assumption that V contains at least two attributes is necessary, since it is easy to see that if V contains 0 or 1 attribute, then every V-partial TD is trivial.

Proof. Assume that the attributes in V are $A_1, \cdots, A_m$. Denote by W the attributes not in V. (Possibly, W is empty.) Assume that the attributes in W are $A_{m+1}, \cdots, A_n$. The variables of T that appear in the column A_i $(1 \leqq i \leqq m)$ of T are a_i and b_i. The only variable that appears in the hypothesis rows of A_j, for $j > m$, is c_j. The projection of the hypothesis of T into V contains all possible rows $e_1 \cdots e_m$, where e_i is either a_i or b_i, *except* that the row of all a's does not appear. The conclusion row contains all a's. For example, if $V = A_1A_2A_3$ and $W = A_4A_5$, then T is

$$\begin{array}{ccccc} a_1 & a_2 & b_3 & c_4 & c_5 \\ a_1 & b_2 & a_3 & c_4 & c_5 \\ a_1 & b_2 & b_3 & c_4 & c_5 \\ b_1 & a_2 & a_3 & c_4 & c_5 \\ b_1 & a_2 & b_3 & c_4 & c_5 \\ b_1 & b_2 & a_3 & c_4 & c_5 \\ b_1 & b_2 & b_3 & c_4 & c_5 \\ \hline a_1 & a_2 & a_3 & a_4 & a_5 . \end{array}$$

Clearly, T is nontrivial (see the remark near the end of § 2). We now show that if T' is a nontrivial, V-partial TD, then SAT $(T') \subseteq$ SAT (T), that is, that $T' \models T$. Let r be a relation (over set of attributes U) that is not in SAT (T); we shall show that r is not in SAT (T'). Let g be a valuation that maps every hypothesis row of T to a tuple in r, but such that $g(a_1 \cdots a_m)$ does not appear in the projection $r[V]$ of r onto V. We know that g exists since r is not in SAT (T). We define a valuation h on T' as follows. We assume for convenience that T' and T have the same distinguished variables $a_1, \cdots, a_n$. For each distinguished variable a, let $h(a) = g(a)$. For each nondistinguished variable d in T', if d is in the A_i column, for some A_i in V, then let $h(d) = g(b_i)$; if d is in the A_j column for A_j in W, then let $h(d) = g(c_j)$.

Since T' is nontrivial, no hypothesis row of T' contains $a_1 \cdots a_m$ as its V entries. Let w' be an arbitrary hypothesis row of T' and let w be the row in T that has a's in its V entries exactly where w' does. Since those entries are not all a's, we know that w exists. By definition of h, we know that $h(w')=g(w)$, and so $h(w')$ is a tuple in r. However, $h(a_1 \cdots a_m)=g(a_1 \cdots a_m)$ is not in $r[V]$, so r violates T', as was to be shown. □

We shall conclude this section by showing that there is no weakest nontrivial TD (including full and strictly partial TD's) if the number of attributes is at least 3. We first need a preliminary result.

THEOREM 3.4. *Let Σ be a set of V_1-partial* TD*'s and let σ be a nontrivial V_2-partial* TD. *If $\Sigma \models \sigma$, then $V_2 \subseteq V_1$.*

Proof. Assume that $\Sigma \models \sigma$ and that it is false that $V_2 \subseteq V_1$; we shall derive a contradiction. Let T_1 be the strongest V_1-partial TD constructed in the proof of Corollary 3.2, and let T_2 be the weakest nontrivial V_2-partial TD constructed in the proof of Theorem 3.3. Since (a) $T_1 \models \Sigma$ (that is, $T_1 \models \tau$ for every τ in Σ), (b) $\Sigma \models \sigma$, and (c) $\sigma \models T_2$, it follows by transitivity of logical implication that $T_1 \models T_2$. Let r be the relation consisting of the hypothesis rows of T_2. Then r violates T_2. We shall show that r obeys T_1, a contradiction.

Since it is false that $V_2 \subseteq V_1$ there is an attribute A in V_2 but not V_1. It is easy to verify that the projection $r[U-A]$ of r onto every attribute except A is the Cartesian product of the projection of r onto each attribute in $U-A$ (see Fig. 3.1). So, r obeys T_1, which was to be shown. □

THEOREM 3.5. *Assume that there are at least three attributes. Then there is no weakest nontrivial* TD. *That is, there is no nontrivial* TD *T such that $T' \models T$ for every nontrivial* TD *T' over the same attributes.*

Note. The assumption that there are at least three attributes is necessary, as we shall see in § 4. Also, observe that unlike Theorem 3.3, which might seem superficially to contradict Theorem 3.5, we are not fixing our attention on V-partial TD's for a given V, but rather considering the whole class of TD's at once.

Proof. Assume that there are at least three attributes, and that a weakest nontrivial TD T exists. Then T is V-partial for some V (possibly $V=U$). Now V is nonempty, since each V-partial TD with $V=\varnothing$ is trivial. So V contains an attribute A. Let $W=U-A$. Then W contains at least two attributes, since U contains at least three attributes. So there is a nontrivial W-partial TD T'. By definition of T, we know that $T' \models T$. This implication contradicts Theorem 3.4, since V is not a subset of W. □

4. TD's over two attributes. In this section, we prove the following result.

THEOREM 4.1. *There are only three distinct* TD*'s (up to equivalence) on two attributes.*

Proof. The three TD's over two attributes are the following:

$$
\begin{array}{cc}
 & \\
 & \\
a_1 & a_2 \\
\hline
a_1 & a_2 \\
\multicolumn{2}{c}{T_1}
\end{array}
\qquad
\begin{array}{cc}
 & \\
a_1 & b_2 \\
b_1 & a_2 \\
\hline
a_1 & a_2 \\
\multicolumn{2}{c}{T_2}
\end{array}
\qquad
\begin{array}{cc}
a_1 & b_2 \\
b_1 & b_2 \\
b_1 & a_2 \\
\hline
a_1 & a_2 \\
\multicolumn{2}{c}{T_3}
\end{array}
$$

TD T_1 is the trivial TD, obeyed by every relation. TD T_2 says that the relation is a Cartesian product; it is the strongest TD. T_3 is the weakest nontrivial TD over two attributes. It is easy to check that none of T_1, T_2, and T_3 are equivalent. We must show that every TD over two attributes, say $T=t_1, t_2, \cdots, t_n/a_1a_2$ is equivalent to one of these.

Case 1. None of $t_1, \cdots, t_n$ has a_1 in the first column, or none of $t_1, \cdots, t_n$ has a_2 in the second column, or some t_i is a_1a_2. It is easy to show that T is trivial. Thus, every strictly partial TD over two attributes is trivial.

Case 2. Case 1 does not hold, but there is no sequence of rows among $t_1, \cdots, t_n$ of the form

$$(*)\qquad \begin{array}{cc} a_1 & b_1 \\ b_2 & b_1 \\ b_2 & b_3 \\ b_4 & b_3 \\ \cdots & \cdots \\ b_k & b_{k-1} \\ b_k & a_2 \end{array}$$

for any $b_1, \cdots, b_k$, with $k \geqq 2$. Then, we can divide $t_1, \cdots, t_n$ into two groups. The first group contains those "reachable" from a_1, in the sense that they appear in some sequence $a_1b_1, b_2b_1, b_2b_3, b_4b_3, \cdots$, and the second contains those that are not. Tuples in the second category may be "reachable" from a_2 or they may be "reachable" from neither a_1 nor a_2.

We now show that T and T_2 are equivalent. We know that $T_2 \vDash T$, since the proof of Theorem 3.1 shows that T_2 implies every TD over two attributes. To show that $T \vDash T_2$, we need only show that when we chase [MMS] the hypothesis rows of T_2, using T, we get the conclusion row of T_2 [SU]. But this chase needs only one step. Map all tuples of T in the first group to a_1b_2 and all others to b_1a_2. This mapping cannot map one symbol of T to two distinct symbols of T_2, or the groups are not defined correctly. That is, we cannot have some tuple $t_i = cd$ mapped to a_1b_2, and then have some tuple $t_j = ed$ or cf mapped to b_1a_2, because ed and cf would be in group 1.

Case 3. A sequence (*) exists, with $k \geqq 2$, and a_1a_2 is not a hypothesis row. Then T is nontrivial, so by the proof of Theorem 3.3, we know that $T \vDash T_3$ (since T_3 is the weakest nontrivial full TD).

To show that $T_3 \vDash T$, we can chase the hypotheses of T with T_3 to infer successively the rows $a_1b_3, a_1b_5, \cdots, a_1b_{k-1}$ and then a_1a_2. □

5. The correspondence between TD's and graphs. For the upcoming examples, it is useful to give a graphical interpretation to TD's and relations. The graph for a TD or relation will have a node for each row or tuple, and edges labeled with attribute symbols, indicating in which components the rows or tuples agree. More precisely:

Definition. Given relation r on relation scheme $R = \{A_1, A_2, \cdots, A_n\}$, the *graph of* r, denoted G_r, is defined as follows. Let $\{t_1, t_2, \cdots, t_m\}$ be the tuples in r; the nodes in G_r will also be $t_1, t_2, \cdots, t_m$. For nodes t_1 and t_2, there is an undirected edge (t_1, t_2) with label A (possibly among others) in R exactly when $t_1(A) = t_2(A)$.

Example. Let r be

	A	B	C
t_1:	0	0	1
t_2:	0	1	0
t_3:	0	1	1
t_4:	1	0	0
t_5:	1	0	1
t_6:	1	1	0.

Then G_r is as in Fig. 5.1. There is always a self-loop from each node to itself, labeled by all the attributes, but we shall omit drawing such edges. We can also omit drawing some of the edges implied by transitivity of equality, to help reduce the clutter. Figure 5.2 represents the same relation as Fig. 5.1 when transitivity of equality is considered.

The graph (denoted G_T) for a template dependency T is defined similarly, except that there is a node denoted $(*)$ that represents the conclusion row.

Example. Let $T=$

$$\begin{array}{llll} w_1: & a & b_1 & c_1 \\ w_2: & a_1 & b & c_1 \\ w_3: & a_1 & b_1 & c \\ \hline & a & b & c\ . \end{array}$$

Then G_T is as in Fig. 5.3.

We can characterize when a relation obeys a TD in terms of certain homomorphisms between their respective graphs.

DEFINITION. An lp-*homomorphism* (label-preserving homomorphism) between labeled, undirected graphs $G_1=(V_1,E_1)$ and $G_2=(V_2,E_2)$ is a mapping $h: V_1 \to V_2$ such that if (v, w) is an edge of E_1 with label A (possibly among others) then $(h(v), h(w))$ is an edge of E_2 with label A.

Example. Let G_r and G_T be the graphs in the last two examples. Define the mappings h_1 and h_2 as follows:

$$\begin{array}{ll} h_1(*)=t_5, & h_2(*)=t_3, \\ h_1(w_1)=t_5, & h_2(w_1)=t_3, \\ h_1(w_2)=t_1, & h_2(w_2)=t_3, \\ h_1(w_3)=t_1, & h_2(w_3)=t_3. \end{array}$$

Then h_1 and h_2 are each lp-homomorphisms from G_T to G_r.

The mapping

$$\begin{array}{l} h_3(*)=t_1, \\ h_3(w_1)=t_3, \\ h_3(w_2)=t_5, \\ h_3(w_3)=t_6 \end{array}$$

is not an lp-homomorphism from G_T to G_r, since $(h(*), h(w_3))=(t_1, t_6)$ does not exist in G_r, and thus certainly does not have label C, as $(*, w_3)$ does.

We can now interpret the criterion for a relation r to obey a TD T in terms of their respective graphs.

THEOREM 5.1. *Relation r obeys T if and only if every* lp-*homomorphism from $G_T-\{*\}$ to G_r can be extended to an* lp-*homomorphism from all of G_T to G_r.*

The straightforward proof of Theorem 5.1 is left to the reader.

Example. Let T and r be the TD and relation used in previous examples. Some lp-homomorphisms from $G_T-\{*\}$ to G_r can be extended, such as h_1 and h_2 below:

$$\begin{array}{ll} h_1(w_1)=t_5, & h_2(w_1)=t_3, \\ h_1(w_2)=t_1, & h_2(w_2)=t_3, \\ h_1(w_3)=t_1, & h_2(w_3)=t_3. \end{array}$$

In fact, any lp-homomorphism that maps $G_T-\{*\}$ to a single node in G_r can be extended to G_T. We shall later use this fact to show that a particular TD T is obeyed

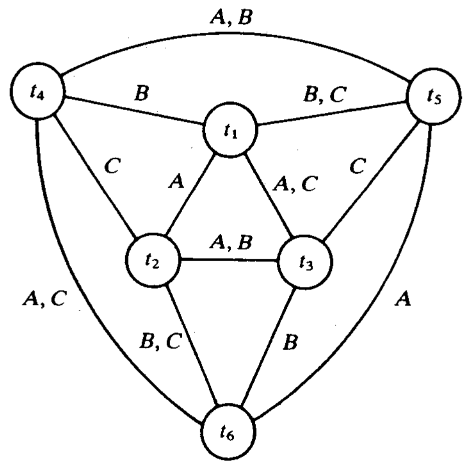

FIG. 5.1

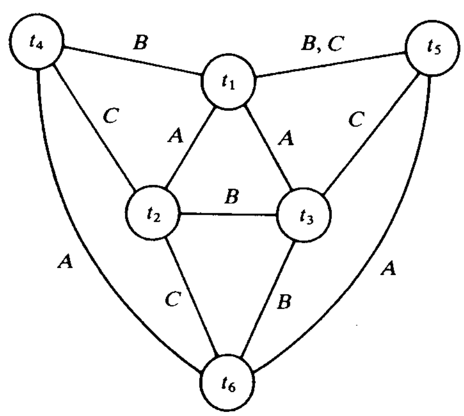

FIG. 5.2

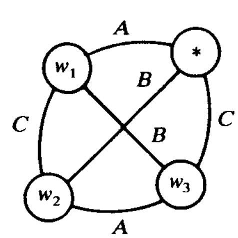

FIG. 5.3

by r, by showing that every lp-homomorphism from $G_T - \{*\}$ to r maps all of $G_T - \{*\}$ to a single node in r.

Relation r in our previous examples does not obey T, because there are lp-homomorphisms from $G_T - \{*\}$ to G_r that cannot be extended, such as

$$h_3(w_1) = t_3,$$
$$h_3(w_2) = t_5,$$
$$h_3(w_3) = t_6.$$

For, if $h_3(*) = t$, then t would have to agree with t_3 on A, with t_5 on B, and with t_6 on C. Then t would be $(0, 0, 0)$, which is not in the relation r.

6. Chains of full TD's. We now use the correspondence between TD's and graphs to help prove the existence of infinite chains of progressively weaker and stronger full TD's.

LEMMA 6.1. *Let T' be a* TD *derived from* TD *T by the addition of hypothesis rows that use no distinguished symbols not already used in some hypothesis row. Then T is at least as strong as T'. That is, $T \models T'$.*

Proof. This result is easily verified by noting that any lp-homomorphism h' from $G_{T'} - \{*\}$ to a relation r can be restricted to an lp-homomorphism h from $G_T - \{*\}$ to r. Furthermore, if h cannot be extended to $G_{T'}$, then h cannot be extended to G_T. □

THEOREM 6.2 (progressively weaker chain). *There exists an infinite sequence of full* TD's *$T_1, T_2, T_3, \cdots$ such that* SAT $(T_i) \subset$ SAT (T_{i+1}) *for $i \geqq 1$. Thus, $T_i \models T_{i+1}$ for each i, and no T_i's are equivalent.*

Proof. Consider the *infinite* graph G (Fig. 6.1). Let T_i be the TD corresponding to the subgraph of G on nodes $*, 1, 2, \cdots, i+1$. By Lemma 6.1, SAT $(T_i) \subseteq$ SAT (T_{i+1}).

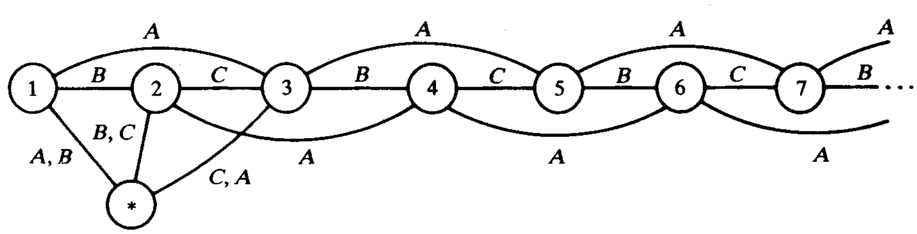

FIG. 6.1

To show proper containment, we need only exhibit a relation r in SAT (T_{i+1}) that does not obey T_i.

Relation r is simply the hypothesis rows of T_i considered as a relation. That is, r is any relation such that G_r is G restricted to nodes $1, 2, \cdots, i+1$. We see that r violates T_i, since the lp-homomorphism h from $G_{T_i} - \{*\}$ to G_r defined by $h(j) = j$, $1 \leqq j \leqq i+1$, cannot be extended to G_{T_i}.

We now show that r obeys T_{i+1}, that is, that each lp-homomorphism h from $G_{T_{i+1}} - \{*\}$ to G_r can always be extended to an lp-homomorphism from $G_{T_{i+1}}$ to G_r.

Case 1. For some nodes j and $j+1$ in $G_{T_{i+1}} - \{*\}$, we have $h(j) = h(j+1)$. Since in G, all odd nodes agree on A, and likewise all even nodes, if $h(j) = h(j+1)$ it follows that $h(p)$ and $h(q)$ agree on A for all p and q. In particular, $h(1)$, $h(2)$ and $h(3)$ agree on A, so we can extend h by letting $h(*) = h(2)$.

Case 2. No nodes j and $j+1$ are mapped to the same node in G_r by h. Let $h(1) = j$. There are 2 subcases, depending on whether j is even or odd.

Case 2a. j is odd. We shall show inductively that $h(k)=j+k-1$ for $1 \leqq k \leqq i+2$.

Assume $h(k-1)=j+k-2$. Suppose k is odd. Since $k-1$ and k are connected by a C-labeled edge, $h(k-1)$ and $h(k)$ must be connected by a C-labeled edge. Since $j+k-2$ is even, the only candidates for $h(k)$ are $j+k-2$ and $j+k-1$. The $j+k-2$ choice is ruled out, since we are not in Case 1. Hence, $h(k)=j+k-1$. A similar argument holds if k is even.

Now look at $h(i+2)$. By our inductive argument, $h(i+2)=j+i+1 \geqq i+2$, which is nonsense, since G_r contains only nodes $1, \cdots, i+1$. Thus, Case 2a cannot occur.

Case 2b. j is even. This case is very similar to Case 2a, except that we show inductively that $h(k)=j+1-k$, for $1 \leqq k \leqq i+2$. Then $h(i+2)=j-i-1 \leqq 0$, which is nonsense, since G_r contains only nodes $1, \cdots, i+1$. Thus, Case 2b cannot occur.

We have shown that Case 2 cannot occur. Thus, r obeys T_{i+1}, and the proof is complete. □

THEOREM 6.3 (progressively stronger chain). *There exists an infinite sequence of full* TD*'s* $T_1, T_2, T_3, \cdots$ *such that* SAT $(T_{i+1}) \subset$ SAT (T_i). *That is,* $T_{i+1} \vDash T_i$ *for each* i, *and no two* T_i*'s are equivalent.*

Proof. Let T_i be the TD corresponding to the finite graph of Fig. 6.2, which we shall call G_i. G_i is just the graph for TD T_{2^i} in the last proof wrapped around with nodes 1 and 2^i+1 overlaid.

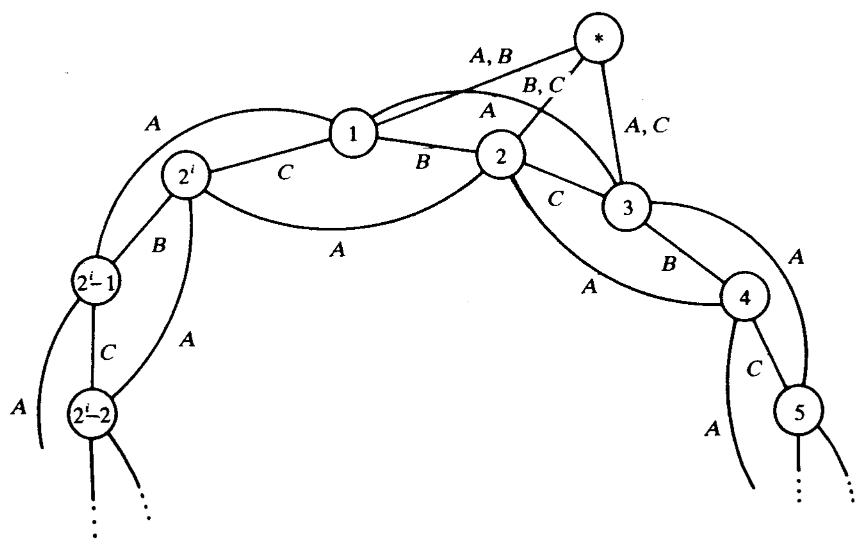

FIG. 6.2

The hard part of this proof is showing that SAT $(T_{i+1}) \subseteq$ SAT (T_i).

Let r be any relation in SAT (T_{i+1}); we shall show that r is in SAT (T_i). To prove this, let h be any lp-homomorphism from $G_i-\{*\}$ to G_r; we must show that h can be extended to an lp-homomorphism from G_i to G_r. We define an lp-homomorphism h' from $G_{i+1}-\{*\}$ to G_r in terms of h, by letting $h'(j)$ be $h(j)$, if $1 \leqq j \leqq 2^i$, and $h(j-2^i)$ if $2^i < j \leqq 2^{i+1}$. Essentially, h' wraps G_{i+1} twice around the image of G_i in G_r under h. Since r is in SAT (T_{i+1}), we know that h' can be extended to G_{i+1}. The reader may check that h can be extended to G_i by letting $h(*)=h'(*)$.

The proof that SAT (T_{i+1}) is a *proper* subset of SAT (T_i) is by a counting argument similar to that used in the proof of Theorem 6.2. The relation r to use is one corresponding to $G_{i+1}-\{*\}$. This relation is not in SAT (T_{i+1}). However, it is in SAT (T_i). For, any lp-homomorphism h from $G_i-\{*\}$ to G_r must map two nodes j and $j+1$ to the same node in G_r, which means the extension of h by $h(*)=h(2)$ will always work. □

7. Finite implication versus implication. In this section we show that finite implication (implication where we restrict our attention to finite relations) and unrestricted implication are distinct for TD's. Thus, the inference rules of Sadri and Ullman [SU] and of Beeri and Vardi [BV2] for TD's, which are complete for unrestricted implication, are incomplete when implication over finite relations only is considered. To state the result another way, let $\mathrm{SAT}_{\mathrm{fin}}(T)$ be the set of all *finite* relations that obey a TD T. We shall exhibit TD's $T_0, T_1, T_2, \cdots, T_k$ such that

$$\mathrm{SAT}_{\mathrm{fin}}(T_1, \cdots, T_k) \subseteq \mathrm{SAT}_{\mathrm{fin}}(T_0),$$

but

$$\mathrm{SAT}(T_1, \cdots, T_k) \not\subseteq \mathrm{SAT}(T_0).$$

Thus, $\{T_1, \cdots, T_k\} \models_{\mathrm{fin}} T_0$, but it is false that $\{T_1, \cdots, T_k\} \models T_0$. Further, we show that there can be no such example with $k=1$. That is, we show that if T_0 and T_1 are TD's, then $T_1 \models_{\mathrm{fin}} T_0$ if and only if $T_1 \models T_0$.

Apart from its inherent interest, we note another reason for studying the issue of whether finite and unrestricted implication are distinct. If finite implication and unrestricted implication were the same, then the decision problem would be decidable. That is, it would be decidable whether or not $\Sigma \models \sigma$, whenever Σ is a finite set of TD's and σ is a single TD. For, $\{(\Sigma, \sigma): \Sigma$ is finite and $\Sigma \models \sigma\}$ is r.e. (recursively enumerable), by Gödel's completeness theorem for first order logic [En] (or, in our special case, by the known [BV2], [SU] complete set of inference rules for TD's). Also, $\{(\Sigma, \sigma): \Sigma$ is finite and it is false that $\Sigma \models_{\mathrm{fin}} \sigma\}$ is r.e., since it is possible to systematically check for finite relations that obey Σ but not σ. Hence, if $\models$ and $\models_{\mathrm{fin}}$ were the same, then $\{(\Sigma, \sigma): \Sigma$ is finite and $\Sigma \models \sigma\}$ would be both r.e. and co-r.e., and hence decidable. As we have noted, Vardi [Va1] and, independently, Gurevich and Lewis [GL] have recently shown that the decision problem for TD's is undecidable.

THEOREM 7.1. *$\models$ and $\models_{\mathrm{fin}}$ are distinct. That is, implication of* TD*'s over the universe of all relations is distinct from implication of* TD*'s over the universe of finite relations.*

Proof. This proof draws its basic outline from a proof by Beeri and Vardi [BV3] of the same result for *untyped* TD's, that is, TD's in which a symbol may appear in more than one column. The construction used here is greatly more complicated than Beeri and Vardi's. We exhibit TD's T_0, T_1, T_2, T_3, T_4 for which there is an infinite relation that obeys $T_1, \cdots, T_4$ and violates T_0, but for which there is no such finite relation. The TD's $T_1, \cdots, T_4$ are given by graphs $G_1, \cdots, G_4$ in Fig. 7.1.

There is an underlying logic to these TD's. The intuition is that if we look at a relation r, we interpret the subgraph of G_r in Fig. 7.2 as representing a directed edge from t_1 to t_3. The relation r can then be interpreted as a directed graph D_r on some subset of its tuples. TD's T_1 and T_2 together say that if D_r has an edge $u \to v$ then for some w it has edge $v \to w$. That is, no node v is a sink. TD T_3 says roughly that D_r is transitively closed. What it actually tells us is that if we have the linked configuration of Fig. 7.3, then for some tuple t' we have Fig. 7.4, where t' is the tuple $*$ of G_3. As we shall see, TD T_4 applies nontrivially when D_r has an edge u such that $u \to u$.

The last TD, T_0, corresponds to graph G_0 in Fig. 7.5.

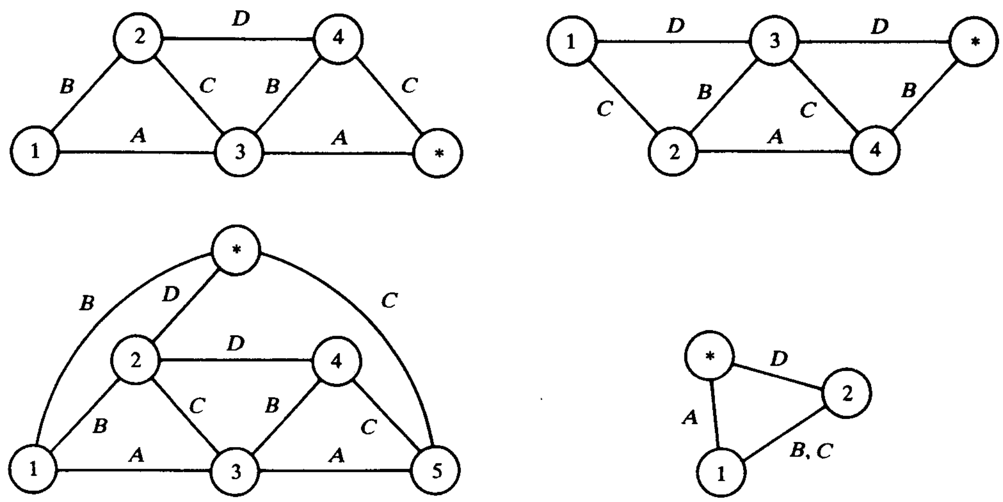

FIG. 7.1

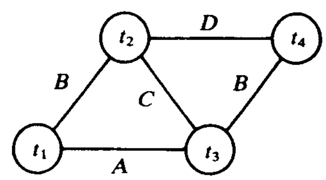

FIG. 7.2

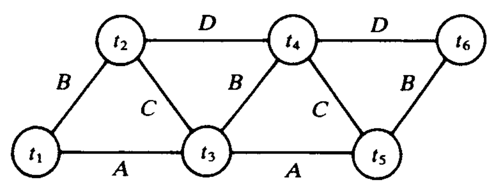

FIG. 7.3

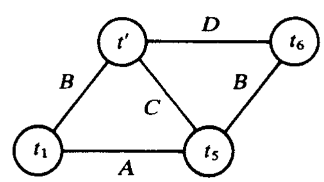

FIG. 7.4

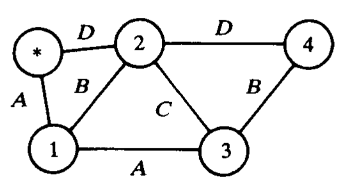

FIG. 7.5

The property of directed graphs we shall exploit is that any finite directed graph D that has no sinks and that is transitively closed has at least one loop edge. This statement is not true for infinite graphs; consider the graph on the natural numbers, where $i \to j$ is an edge if and only if $i < j$.

We now present an infinite relation r_I, and show that r_I obeys T_1, T_2, T_3 and T_4, but violates T_0. Thus, it is false that $\{T_1, T_2, T_3, T_4\} \models T_0$.

Let $r_I = \{(i, i, j, 0): 1 \leqq i < j\} \cup \{(0, i, i, i): 1 \leqq i\}$. We shall refer to tuples of r_I of the form $(i, i, j, 0)$ with $1 \leqq i < j$ as tuples of the first type and tuples $(0, i, i, i)$ with $1 \leqq i$ as tuples of the second type.

1. *r_I obeys T_1.* We shall show that if we chase r_I with T_1, then no new tuples appear. Consider the first time that a new tuple could appear. The only AC combinations not already present in r_I that could be forced by chasing with T_1 are those in which the A entry is i (we write this informally as $A = i$), $C = j$, and $i \geqq j \geqq 1$. To obtain such an AC combination, an application of T_1 must have $t_4 = (\cdot, b, j, \cdot)$ and $t_3 = (i, b, \cdot, \cdot)$. (By this we mean that t_3 and t_4 have the same B entry b, and the $\cdot$'s represent entries we don't care about now.) Since $i \geqq 1$, we know that t_3 is a tuple of the first type, so $b = i$. So t_4 is $(\cdot, i, j, \cdot)$ with $i \geqq j$. Thus, t_4 is a tuple of the second type, so $t_4 = (0, i, i, i)$. Since t_2 agrees with t_4 in D, we know that $t_2 = t_4$. Hence, t_4 agrees with t_3 in C (since t_2 agrees with t_3 in C). So the A and C entries of t_3 are both i, and hence equal. But in no tuple of r_I do the A and C entries agree. This is a contradiction, so chasing r_I with T_1 can produce no new AC entries. Hence, r_I obeys T_1, since T_1 is an AC-partial TD.

2. *r_I obeys T_2.* The only BD combinations that can be generated by chasing r_I with T_2 and that are missing have $B = i$, $D = j$, $i \neq j$ and $j \neq 0$. So $t_3 = (\cdot, \cdot, c, j)$, and $t_4 = (\cdot, i, c, \cdot)$. Since $j \neq 0$, we know that $t_3 = (0, j, j, j)$. Since $j = c \neq i$, we know $t_4 = (i, i, j, 0)$. Now t_2 agrees with t_4 in A, so $t_2 = (i, i, \cdot, 0)$. Thus, t_2 does not agree with t_3 in B, a contradiction.

3. *r_I obeys T_3.* Since t_2 and t_4 agree on D, they are both tuples of the first type or they are both tuples of the second type. If they are both tuples of the second type then they are equal, since they agree on D. In this case, either can serve as $*$ ($*$ must have C from t_4, and BD from t_2). So we can assume that t_2 and t_4 are both of the first type. The only way that no tuple of r_I can serve as $*$ is if the B entry of t_1 (and t_2), say i, is greater than or equal to the C entry of t_5 (and t_4), say j. So assume $i \geqq j$. Let $t_3 = (a, i', j', \cdot)$. Since $t_2 = (i, i, j', 0)$, we know that $i < j'$. Similarly, $t_4 = (i', i', j, 0)$ and $i' < j$. There are now two cases. *Case* 1. $a \neq 0$. Then, t_1, t_3 and t_5 are all of the first type. Since t_3 is of the first type, $a = i'$. Now, the B entry of t_1 is i, so the A entry of t_1 is i. Thus, $a = i$, so $i = i'$. Since $i' < j$, it follows that $i < j$, a contradiction. *Case* 2. $a = 0$. Then $i' = j'$, so $i < j' = i' < j$, a contradiction.

4. *r_I obeys T_4.* Since t_1 and t_2 agree on B and C, it follows easily that $t_1 = t_2$. Thus, $*$ can be taken to be t_1.

5. *r_I violates T_0.* Let $t_1 = (0, 1, 1, 1)$, $t_2 = (1, 1, 2, 0)$, $t_3 = (0, 2, 2, 2)$ and $t_4 = (2, 2, 3, 0)$. Then $*$ must be $(0, \cdot, \cdot, 0)$, and r_I contains no such tuple.

We now show that no finite relation r_F in SAT (T_1, T_2, T_3, T_4) violates T_0. Suppose r_F violates T_0. Then, G_{r_F} contains the configuration in Fig. 7.6 (ignoring X and its edges), where no tuple in r_F can serve as the node marked X (and so $t_1 \neq t_2$), even if we allow other edges connecting X to $t_1, \cdots, t_4$. By TD's T_1 and T_2, we know that r_F must also contain tuples t_5 and t_6 such that G_{r_F} contains the subgraph in Fig. 7.7. We do not require that the tuples be distinct. Further applications of T_1 and T_2 give the subgraph in Fig. 7.8, which we shall abbreviate as in Fig. 7.9. We remarked before that the tuples need not be distinct. Actually, if we extend this chain far enough they

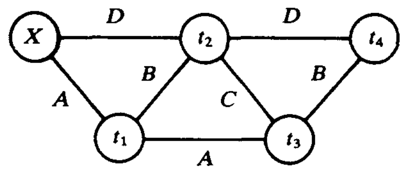

FIG. 7.6

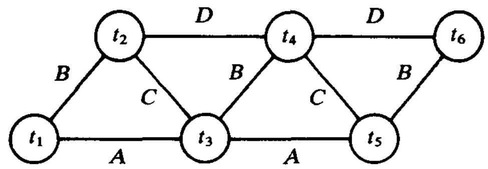

FIG. 7.7

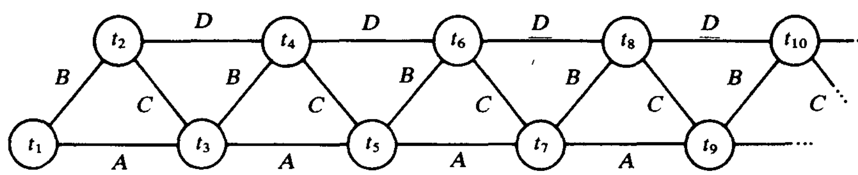

FIG. 7.8

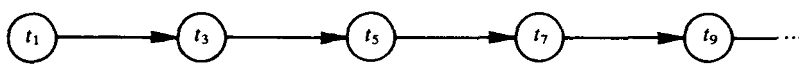

FIG. 7.9

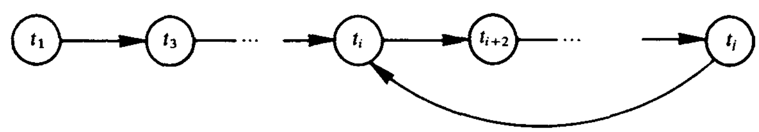

FIG. 7.10

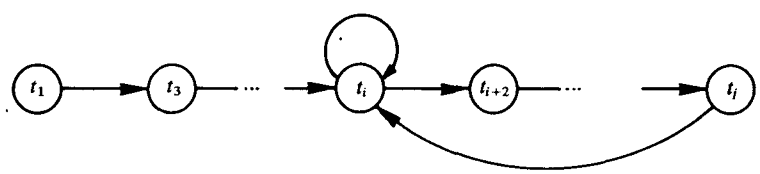

FIG. 7.11

cannot be distinct, since r_F is finite. The chain must eventually loop back on itself (Fig. 7.10). By repeated application of the "transitivity" TD, T_3, we eventually get an edge from t_i to itself (Fig. 7.11). The self-loop from t_i to itself means the same as the configuration shown in Fig. 7.12, where t_i appears twice, and where the exact identities of t' and t'' do not matter (except that $t'[D]=t_{i+1}[D]=t_2[D]$). As we see, t' agrees with t_i on both B and C. T_4 now applies to give us a tuple t where Fig. 7.13 holds. But $t_i[A]=t_1[A]$ and $t'[D]=t_2[D]$, so Fig. 7.14 holds. Hence, t serves as the slot marked by X in the original figure, a contradiction. Relation r_F cannot violate T_0, concluding the proof. □

Although, as we just proved, there are TD's $T_0, T_1, \cdots, T_k$ such that $\{T_1, \cdots, T_k\} \vDash_{\text{fin}} T_0$ but for which $\{T_1, \cdots, T_k\} \vDash T_0$ fails, we now show that this is impossible if $k = 1$.

THEOREM 7.2. *Let T_0 and T_1 be* TD's. *Then $T_1 \vDash_{\text{fin}} T_0$ if and only if $T_1 \vDash T_0$.*

Proof. It is immediate that if $T_1 \vDash T_0$, then $T_1 \vDash_{\text{fin}} T_0$. So assume that $T_1 \vDash_{\text{fin}} T_0$. We must show that $T_1 \vDash T_0$. Assume that T_1 is V_1-partial, and that T_0 is V_0-partial. Now Theorem 3.4 holds when "$\vDash$" is replaced by "$\vDash_{\text{fin}}$", by the same proof. So, since $T_1 \vDash_{\text{fin}} T_0$, it follows that $V_0 \subseteq V_1$. So, when we use T_1 to chase the hypothesis rows of T_0, it is easy to see that we never need to add a new row whose projection onto V_1 is already present. No new variables are added in the V_1 columns during the chase, so the chase terminates after a finite number of steps. Thus, as in the theory of the chase for full TD's [MMS], if there is a "counterexample" relation that obeys T_1 but not T_0, then there is a finite such counterexample. The result follows. □

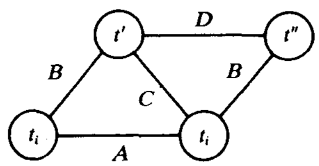

FIG. 7.12

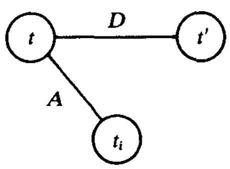

FIG. 7.13

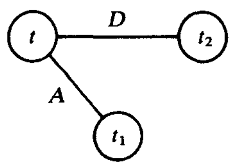

FIG. 7.14

We note that Theorem 7.2 was proven by Sadri [Sa] in the case where T_0 and T_1 are EMVD's. Also, Beeri and Vardi [BV1] showed if Σ is a set of V-partial TD's and σ a TD, then $\Sigma \vDash \sigma$ if and only if $\Sigma \vDash_{\text{fin}} \sigma$. This implies Theorem 7.2.

8. Closure of full TD's under conjunction. In this section, we show that full TD's are closed under finite conjunction. That is, we show that if Σ is a finite set of full TD's, then there is a single full TD T that is equivalent to Σ (in other words, SAT (T) = SAT (Σ)). The same result was obtained independently by Beeri and Vardi [BV2]. However, we show that the conjunction of a finite set of TD's (not necessarily full) is not necessarily equivalent to a single TD, and the disjunction of a finite set of full TD's is not necessarily equivalent to a single TD.

Since every multivalued dependency is equivalent to a full TD, it follows in particular that (the conjunction of) every set of multivalued dependencies is equivalent to a TD. However, sets of multivalued dependencies that are not only equivalent to a TD, but even to a join dependency (which are special cases of TD's), are quite special [BFMMUY], [BFMY], [FMU].

Our main tool is the direct product construction of Fagin [Fa3]. Let r and r' be relations, each with attributes $U = A_1 \cdots A_n$. The direct product $r \otimes r'$ has the same set U of attributes. The possible entries in the A_i column of $r \otimes r'$ are elements $\langle a, a' \rangle$, where a is an entry in the A_i column of r, and a' is an entry in the A_i column of r'. A tuple $(\langle a_1, a_1' \rangle, \cdots, \langle a_n, a_n' \rangle)$ is a tuple of the direct product if and only if $(a_1, \cdots, a_n)$ is a tuple of r and $(a_1', \cdots, a_n')$ is a tuple of r'. Fagin [Fa3] shows that if T is a TD (or even more generally, an embedded implicational dependency), and if r and r' are nonempty relations, then T holds for $r \otimes r'$ if and only if T holds for each of r and r'. This property is called *faithfulness* of T.

THEOREM 8.1. *Full* TD'*s are closed under finite conjunction.*

Proof. It is sufficient to prove that if T_1 and T_2 are full TD's, then there is a TD T that is equivalent to their conjunction; the result then follows by an easy induction. We use the direct product construction on hypothesis rows of the TD's T_1 and T_2. That is, let T_1 be

$$\begin{array}{cccc} c_{11} & c_{12} & \cdots & c_{1n} \\ & & \vdots & \\ c_{r1} & c_{r2} & \cdots & c_{rn} \\ \hline a_1 & a_2 & \cdots & a_n \end{array}$$

and let T_2 be

$$\begin{array}{cccc} d_{11} & d_{12} & \cdots & d_{1n} \\ & & \vdots & \\ d_{s1} & d_{s2} & \cdots & d_{sn} \\ \hline a_1 & a_2 & \cdots & a_n . \end{array}$$

We now define a new TD T, that we shall prove is equivalent to $T_1 \wedge T_2$. The hypothesis rows of T are the direct product of the hypothesis rows of T_1 (treated as a relation) and the hypothesis rows of T_2 (treated as a relation). Thus, let the symbols for the kth column of T be the product symbols $\langle c_{ik}, d_{jk} \rangle$ for $1 \leqq i \leqq r$ and $1 \leqq j \leqq s$, with $\langle a_k, a_k \rangle$ being the distinguished symbol for column k. The rs hypothesis rows of T are all of the rows of the form

$$\langle c_{i1}, d_{j1} \rangle \langle c_{i2}, d_{j2} \rangle \cdots \langle c_{in}, d_{jn} \rangle$$

for all i and j. The conclusion row of T is $\langle a_1, a_1 \rangle \langle a_2, a_2 \rangle \cdots \langle a_n, a_n \rangle$, of course.

$T \vDash T_1$, as we can show in one step of a chase by using the mapping that sends $\langle c_{ij}, d \rangle$ to c_{ij} for each d. Similarly, $T \vDash T_2$.

We shall show, by chasing the hypothesis rows of T, that $\{T_1, T_2\} \vDash T$. First, for each (fixed) j, apply T_1 to the r hypothesis rows of the form $\langle c_{i1}, d_{j1} \rangle \cdots \langle c_{in}, d_{jn} \rangle$ for $1 \leqq i \leqq r$ to infer the rows of the form $\langle a_1, d_{j1} \rangle \cdots \langle a_n, d_{jn} \rangle$ for $1 \leqq j \leqq s$. Then apply T_2 to these rows to infer $\langle a_1, a_1 \rangle \cdots \langle a_n, a_n \rangle$. □

Although the finite conjunction of full TD's is equivalent to a single TD, we now show that the finite conjunction of TD's (not necessarily full) is not necessarily equivalent to a single TD.

THEOREM 8.2. *There is a pair of* TD'*s whose conjunction is not equivalent to a single* TD.

Proof. It is sufficient to show that there is a finite set $T_1, \cdots, T_k$ of TD's such that $T_1 \wedge \cdots \wedge T_k$ is not equivalent to a single TD. For, if the conjunction of a pair of TD's were always equivalent to a single TD, then by induction, the conjunction of a finite set of TD's would be equivalent to a single TD.

Let $T_0, T_1, \cdots, T_4$ be the TD's of § 7 (for which $\{T_1, \cdots, T_4\} \vDash_{\text{fin}} T_0$ but for which $\{T_1, \cdots, T_4\} \vDash T_0$ fails). If $T_1 \wedge \cdots \wedge T_4$ were equivalent to a single TD T, then $T \vDash_{\text{fin}} T_0$, since $\{T_1, \cdots, T_4\} \vDash_{\text{fin}} T_0$. By Theorem 7.2, it follows that $T \vDash T_0$. So, $\{T_1, \cdots, T_4\} \vDash T_0$. This is a contradiction. □

Vardi [Va2] has posed the interesting question as to whether the conjunction of a pair of V-partial TD's (for the same V) is necessarily equivalent to a TD.

We now prove a result that implies (by Corollary 8.4 below) that TD's are not closed under finite disjunction.

THEOREM 8.3. *Let T_1 and T_2 be incomparable* TD's *(that is, neither $T_1 \vDash T_2$ nor $T_2 \vDash T_1$). Then the disjunction $T_1 \vee T_2$ is not equivalent to a single* TD.

Proof. Let r_1 be a relation that obeys T_1 but not T_2, and let r_2 be a relation that obeys T_2 but not T_1. Let r be the direct product $r_1 \otimes r_2$. Then by faithfulness of T_1, we know that r does not obey T_1, since r_2 does not obey T_1. Similarly, r does not obey T_2, and so r does not obey $T_1 \vee T_2$. However, each of r_1 and r_2 obeys $T_1 \vee T_2$, since r_1 obeys T_1 and r_2 obeys T_2. If $T_1 \vee T_2$ were equivalent to a TD T, then the faithfulness of T would be violated. □

COROLLARY 8.4. *There are full* TD's *T_1 and T_2 such that $T_1 \vee T_2$ is not equivalent to a single* TD.

Proof. Let T_1 and T_2 be incomparable full TD's. For example, over three attributes ABC, let T_1 be the MVD $A \twoheadrightarrow B$ and let T_2 be the MVD $B \twoheadrightarrow A$. By Theorem 8.3, it follows that $T_1 \vee T_2$ is not equivalent to a TD. □

We note that Ginsburg and Zaiddan [GZ] have considered questions similar to those discussed in this section, but for FD's instead of TD's, by studying intersections and unions of "functional dependency databases." Classes SAT (Σ), where Σ is a set of FD's, are called *functional dependency classes* by Fagin [Fa3]. *Functional dependency databases* differ from functional dependency classes by explicitly defining the domains for each attribute.

9. A set of strictly partial TD's cannot imply a full TD. In this section, we prove the following result.

THEOREM 9.1. *There is a finite relation that obeys every strictly partial* TD *but no nontrivial full* TD. *In particular, if Σ is a set of strictly partial* TD's *and σ is a nontrivial full* TD, *then it is false that $\Sigma \vDash \sigma$ (or even that $\Sigma \vDash_{\text{fin}} \sigma$).*

We give two proofs of Theorem 9.1, since both proofs are amusing and both give additional information.

Proof 1. This proof is in the spirit of Sadri's [Sa] proof that there is a finite relation that obeys every EMVD that is not a MVD but violates every MVD. Let R be the relation that contains every tuple consisting only of 0's and 1's *except* the tuple of all 0's. For example, if there are three attributes, then R is

0	0	1
0	1	0
0	1	1
1	0	0
1	0	1
1	1	0
1	1	1.

This relation obeys every strictly partial TD, since the projection onto each proper subset V of the attributes U is the Cartesian product of the projection onto each attribute of V. However, R clearly violates the weakest nontrivial full TD T constructed in the proof of Theorem 3.3. Hence, R violates every nontrivial full TD (if R obeyed a nontrivial full TD T', then R would obey T, since $T' \models T$ by Theorem 3.3). □

Proof 2. Let $\mathcal{A}_n$ be the set of all relations (with attributes U) such that every entry of the relation is a member of $\{1, \cdots, n\}$. Thus, $\mathcal{A}_n$ contains 2^{n^u} members, where u is the number of attributes (that is, the size) of U. If P is a property of relations, then we say that "almost all relations have property P" (or "a random relation has property P") if the fraction of members of $\mathcal{A}_n$ with property P converges to 1 as $n \to \infty$. Fagin [Fa1] showed that if P is a first-order property of relations, then either almost all relations have property P or almost all relations fail to have property P. Using his techniques, it is easy to show that if σ is a strictly partial TD, then almost all relations obey σ, while if σ is a nontrivial full TD, then almost all relations violate σ.

Let T_V be the strongest V-partial TD (which exists by Corollary 3.2), and let $\Sigma = \{T_V\colon V \text{ is a proper subset of } U\}$. Then Σ is a finite set of TD's, since U contains only a finite number of subsets. By the above remarks, for each TD T_V in Σ, almost all relations obey T_V (since T_V is strictly partial). Since Σ is finite, it follows from elementary probability theory that almost all relations simultaneously obey every member of Σ. Furthermore, if σ is the weakest nontrivial full TD, whose existence is guaranteed by Theorem 3.3 (with $V = U$), then it follows by our earlier remarks that almost all relations violate σ (since σ is full). Thus, almost all relations obey Σ and violate σ. If a relation R obeys Σ, then it obeys every strictly partial TD, since if T is a V-partial TD, then T is implied by T_V, which is in Σ, if V is a proper subset of U. Further, if a relation R violates the weakest nontrivial full TD σ, then it violates every nontrivial full TD T (since $T \models \sigma$). Thus, almost all relations simultaneously obey every strictly partial TD and violate every nontrivial full TD. This is even stronger than the statement of Theorem 9.1. □

10. Finite Armstrong relations. Let Σ be a set of TD's. Let Σ^*_{fin} be $\{\sigma\colon \Sigma \models_{\text{fin}} \sigma\}$. Thus, Σ^*_{fin} is the set of all TD's that hold in every finite relation obeying Σ. A *finite Armstrong relation* [Fa3] for Σ is defined to be a finite relation that obeys Σ^*_{fin} but no other TD's. The following facts are easy consequences of results by Fagin [Fa3].

Fact 1. There is an Armstrong relation (not necessarily finite) for Σ. This fact can be interpreted in two distinct ways, both of which are correct. One meaning is that there is a relation (not necessarily finite) that obeys every TD in $\Sigma^* = \{\sigma\colon \Sigma \models \sigma\}$, but no other TD's. The second meaning is that there is a relation (not necessarily finite) that obeys every TD in Σ^*_{fin} but no other TD's; this is true because $(\Sigma^*_{\text{fin}})^* = \Sigma^*_{\text{fin}}$.

Fact 2. Let $\mathcal{S}$ be a fixed finite set of TD's (such as the set of all EMVD's over some fixed set of attributes). Then, there is a finite relation that obeys every TD in Σ^*_{fin} but violates every TD in $\mathcal{S}$ that is not in Σ^*_{fin}.

In this section, we shall show (Theorem 10.1 below) that the second sentence of Fact 2 is not necessarily true if $\mathcal{S}$ is the set of all TD's (this set is infinite by § 6, if there are at least three attributes). Also, we note that Fagin shows [Fa3] that the second sentence of Fact 2 is false if "TD" is replaced by "EID" (embedded implicational dependency) and if $\mathcal{S}$ is the set of all EID's.

By Theorem 10.1 below, there is a finite set Σ of TD's that have no finite Armstrong relation (although Σ has an infinite Armstrong relation, by Fact 1 above).

However, there are certainly some sets Σ of TD's that do have a finite Armstrong relation; for example, if Σ is the set of all TD's, then Σ has a finite Armstrong relation, namely, any one-tuple relation. Also, we show at the end of this section that if Σ is the empty set, then Σ has a finite Armstrong relation. In Theorem 10.2 below, we give several characterizations of those sets Σ of TD's that have a finite Armstrong relation.

THEOREM 10.1. *There is a finite set Σ of* TD's *such that Σ has no finite Armstrong relation* (*with respect to* TD's). *That is, there is no finite relation that obeys Σ^*_{fin} and no other* TD's.

Proof. Let Σ be $\{T_3, T_4\}$, where T_3 and T_4 are as in the proof of Theorem 7.1. We shall show that there is no finite Armstrong relation for Σ. Let T^k be the TD that looks like T_0 of Theorem 7.1, except that the quadrangle is repeated k times; i.e., T^k is the TD shown in Fig. 10.1.

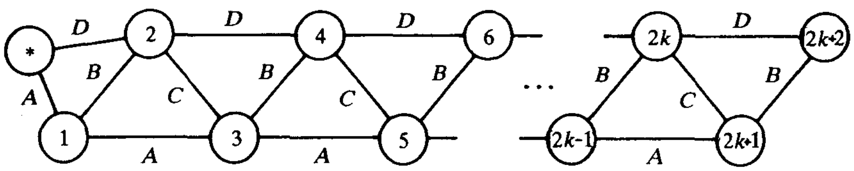

FIG. 10.1

We shall show that 1) for every k, it is false that $\Sigma \models_{\text{fin}} T^k$, and 2) every finite relation obeying Σ also obeys some T^k. It follows easily from 1) and 2) that there is no finite Armstrong relation for Σ.

1) *holds.* Let r^k be the relation $\{(i, i, j, 0): 1 \leqq i < j \leqq k+2\} \cup \{(0, i, i, i): 1 \leqq i \leqq k+1\}$. Then r^k is roughly the truncation of the relation r_I in the proof of Theorem 7.1 to the first $k+2$ positive integers. Now r^k obeys Σ. The proof is exactly the same as the proof in Theorem 7.1 that r_I satisfies T_3 and T_4. However, r^k violates T^k. For, let $t_{2i-1} = (0, i, i, i)$, and $t_{2i} = (i, i, i+1, 0)$ for $i = 1, \cdots, k+1$. Then the role of $*$ in the TD T^k must be filled by $(0, \cdot, \cdot, 0)$, although r^k contains no such tuple. We have shown that r^k obeys Σ but not T^k. This proves 1).

2) *holds.* Let r be a finite relation that obeys Σ and that has exactly k tuples. Consider the TD T^k. Every lp-homomorphism from the graph $G_{T^k} - \{*\}$ to G_r must map two distinct nodes $2i+1$, $2j+1$ to the same node (since there are $k+1$ odd-numbered nodes in $G_{T^k} - \{*\}$ and only k nodes in G_r). Then, as in the proof of Theorem 7.1, we can show that there is a tuple of r that can play the role of $*$. Therefore, r obeys T^k. This completes the proof of 2), and hence the proof of the theorem. □

An alternative proof of Theorem 10.1 can be obtained by using Vardi's result [Va1] that there is a single finite set Σ of TD's such that the set of all TD's σ for which $\Sigma \models_{\text{fin}} \sigma$ is not recursive. This result implies that there is no finite Armstrong relation for Σ, since we could test whether or not $\Sigma \models_{\text{fin}} \sigma$ by simply checking whether or not the finite Armstrong relation obeys σ.

THEOREM 10.2. *Let Σ be a set of* TD's. *The following are equivalent*:

(a) *There is a finite relation that obeys Σ^*_{fin} and no other* TD's ("*Σ has a finite Armstrong relation*").

(b) *There is a finite set $\mathcal{T}$ of* TD's, *disjoint from Σ^*_{fin}, such that for each* TD *T not in Σ^*_{fin} there is a* TD *T' in $\mathcal{T}$ where $T \models T'$.*

(c) *There is a finite set* $\mathcal{T}$ *of* TD's, *disjoint from* Σ^*_{fin}, *such that* $T \models \vee\{T': T: \in \mathcal{T}\}$ *for each* TD T *not in* Σ^*_{fin}.

(d) *There is a finite set* $\mathcal{T}$ *of* TD's, *disjoint from* Σ^*_{fin}, *such that* $\vee\{T: T \notin \Sigma^*_{\text{fin}}\}$ *is equivalent to* $\vee\{T': T' \in \mathcal{T}\}$.

Note that $\{T': T' \in \mathcal{T}\}$ in (d) is a finite subset of $\{T: T \notin \Sigma^*_{\text{fin}}\}$ in (d). So, (d) is a kind of compactness result, that says that a certain set has a finite subcover (that is, it says that a finite number of disjuncts of $\vee\{T: T \notin \Sigma^*_{\text{fin}}\}$ "covers" all of it).

Proof. (a)$\Rightarrow$(b). Let R be a finite relation that obeys Σ^*_{fin} and no other TD's. We now define a finite set $\mathcal{T}$ of TD's, each of which R violates. For each set P of rows of R and for each set V ($V \subseteq U$) of attributes, let $\mathcal{T}$ contain every V-partial TD with P as its hypothesis rows that is false about R. It is easy to see that $\mathcal{T}$ is a finite set of TD's. The set $\mathcal{T}$ is disjoint from Σ^*_{fin}, since R obeys Σ^*_{fin} and violates every member of $\mathcal{T}$. Now let T be a TD not in Σ^*_{fin}. We must show that there is a TD T' in $\mathcal{T}$ where $T \models T'$. Assume that T is V-partial. Since T is not in Σ^*_{fin}, we know that R violates T. So, there is a valuation h that maps the hypothesis rows of T onto rows of R such that there is no way to extend h to get the conclusion row of T mapped onto a row of R.

Let T' be the V-partial member of $\mathcal{T}$ whose hypothesis rows are the images under h of the hypothesis rows of T, and such that for each attribute A in V, the A entry of the conclusion row of T' is the image under h of the A entry of the conclusion row of T. We now show that $T \models T'$. For, assume that a relation S obeys T; we must show that S obeys T'. To show this, assume that the hypothesis rows of T' can be mapped by a valuation h' onto rows of S. We must show that h' can be extended to a mapping from the conclusion row of T' onto a row of S. Now $h \circ h'$ is a valuation from the hypothesis rows of T onto these same rows of S. Then $h \circ h'$ is already defined on the V entries of the conclusion row of T, and (since T holds for S) can be extended to map all of the conclusion row of T onto a row of S. This gives us an extension of h' to map all of the conclusion row of T' onto the same row of S, by mapping the A entry of the conclusion row of T' (for each A not in V) onto the same entry of S as the extension of $h \circ h'$ maps that entry. This was to be shown. So, $T \models T'$, as desired.

(b)$\Rightarrow$(c). Let the set $\mathcal{T}$ of (c) equal the set $\mathcal{T}$ of (b). Take T not in Σ^*_{fin}. By (b), there is some T' in $\mathcal{T}$ such that $T \models T'$. Hence, $T \models \vee\{T': T' \in \mathcal{T}\}$.

(c)$\Rightarrow$(d). Let the set $\mathcal{T}$ of (c) equal the set $\mathcal{T}$ of (d). It is obvious that $\vee\{T': T' \in \mathcal{T}\} \models \vee\{T: T \notin \Sigma^*_{\text{fin}}\}$, since $\{T': T' \in \mathcal{T}\} \subseteq \{T: T \notin \Sigma^*_{\text{fin}}\}$. Conversely, we must show that $\vee\{T: T \notin \Sigma^*_{\text{fin}}\} \models \vee\{T': T' \in \mathcal{T}\}$. Let R be a relation that obeys $\vee\{T: T \notin \Sigma^*_{\text{fin}}\}$; we must show that R obeys $\vee\{T': T' \in \mathcal{T}\}$. Since R obeys $\vee\{T: T \notin \Sigma^*_{\text{fin}}\}$, this means that R obeys some T not in Σ^*_{fin}. By (c), we know that $T \models \vee\{T': T' \in \mathcal{T}\}$, so R obeys $\vee\{T': T' \in \mathcal{T}\}$, which was to be shown.

(d)$\Rightarrow$(a). Assume that (d) holds. By Fact 2 above, there is a finite relation R that obeys every TD in Σ^*_{fin} but violates every member of $\mathcal{T}$. Since R violates every member of $\mathcal{T}$, we know that R violates $\vee\{T': T' \in \mathcal{T}\}$. By assumption, $\vee\{T': T' \in \mathcal{T}\}$ is equivalent to $\vee\{T: T \notin \Sigma^*_{\text{fin}}\}$. Thus, R violates $\vee\{T: T \notin \Sigma^*_{\text{fin}}\}$. Hence, R obeys Σ^*_{fin} and no other TD's, which was to be shown. □

As a simple application of Theorem 10.2, we now show that there is a finite Armstrong relation for the empty set, that is, that there is a finite relation that violates every nontrivial TD. Let $\mathcal{T}$ be the set of weakest nontrivial V-partial TD's, one for every subset V, with at least two members, of the set U of attributes. These weakest nontrivial V-partial TD's exist by Theorem 3.3. But this set $\mathcal{T}$ can play the role of $\mathcal{T}$ in (b) of Theorem 10.2. Hence, (a) of Theorem 10.2 holds, and so the empty set has a finite Armstrong relation.

11. Acknowledgment. The authors are grateful to Moshe Vardi for several useful suggestions.

REFERENCES

[BFMMUY] C. BEERI, R. FAGIN, D. MAIER, A. O. MENDELZON, J. D. ULLMAN AND M. YANNAKAKIS, *Properties of acyclic database schemes*, Proc. Thirteenth Annual ACM Symposium on the Theory of Computing, 1981, pp. 355–362.

[BFMY] C. BEERI, R. FAGIN, D. MAIER AND M. YANNAKAKIS, *On the desirability of acyclic database schemes*, J. Assoc. Comput. Mach., to appear.

[BV1] C. BEERI AND M. Y. VARDI, *A proof procedure for data dependencies*, Technical Report, Hebrew Univ. of Jerusalem, August 1980.

[BV2] ———, *Formal systems for tuple and equality generating dependencies*, Technical Report, Hebrew Univ. of Jerusalem, April 1981.

[BV3] ———, *The implication problem for data dependencies*, Proc. 8th ICALP, Acre, Israel, July 1981, in Lecture Notes in Computer Science 115, Springer-Verlag, New York, 1981, pp. 73–85.

[CFP] M. A. CASANOVA, R. FAGIN AND C. PAPADIMITRIOU, *Inclusion dependencies and their interaction with functional dependencies*, Proc. First ACM SIGACT-SIGMOD Principles of Database Systems, 1982, pp. 171–176.

[Co] E. F. CODD, *Further normalization of the data base relational model*, in Courant Computer Science Symposia 6: Data Base Systems, May 24–25, 1971, R. Rustin, ed., Prentice-Hall, Englewood Cliffs, NJ, 1971, pp. 33–64.

[De] C. DELOBEL, *Normalization and hierarchical dependencies in the relational data model*, ACM Trans. Database Systems, 3 (1978), pp. 201–222.

[En] H. B. ENDERTON, *A Mathematical Introduction to Logic*, Academic Press, New York, 1972.

[Fa1] R. FAGIN, *Probabilities on finite models*, J. Symbolic Logic, 41 (1976), pp. 50–58.

[Fa2] ———, *Multivalued dependencies and a new normal form for relational databases*, ACM Trans. Database Systems, 2 (1977), pp. 262–278.

[Fa3] ———, *Horn clauses and database dependencies*, Proc. 1980 ACM SIGACT Symposium on Theory of Computing, pp. 123–134. J. Assoc. Comput. Math., to appear.

[FMU] R. FAGIN, R. A. O. MENDELZON AND J. D. ULLMAN, *A simplified universal relation assumption and its properties*, ACM Trans. Database Systems, to appear.

[GZ] S. GINSBURG AND S. M. ZAIDDAN, *Properties of functional dependency families*, J. Assoc. Comput. Math., 29 (1982), pp. 678–698.

[GL] Y. GUREVICH AND H. R. LEWIS, *The inference problem for template dependencies*, Proc. First ACM SIGACT-SIGMOD Principles of Database Systems, 1982, pp. 221–229.

[MMS] D. MAIER, A. MENDELZON AND Y. SAGIV, *Testing implications of data dependencies*, ACM Trans. Database Systems, 4 (1979), pp. 455–469.

[MM] A. MENDELZON AND D. MAIER, *Generalized mutual dependencies*, in Proc. 1979 Very large Data Bases Conference, pp. 75–82.

[Ni] J.-M. NICOLAS, *Mutual dependencies and some results on undecomposable relations*, in Proc. 1978 Very Large Data Bases Conference, pp. 360–367.

[PJ] J. PARADAENS AND D. JANNSENS, *Decomposition of relations: A comprehensive approach*, in Formal Bases for Databases, J. Minker and H. Gallaire, ed., Plenum, New York, 1978.

[Ri] J. RISSANEN, *Theory of relations for databases—A tutorial survey*, Proc. 7th Symposium on Mathematical Foundations of Computer Science, Lecture Notes in Computer Science, 64, Springer-Verlag, New York, pp. 537–551.

[Sa] F. SADRI, Personal communication.

[SU] F. SADRI AND J. D. ULLMAN, *Template dependencies: a large class of dependencies in relational databases and its complete axiomatization*, J. Assoc. Comput. Mach., 29 (1982), pp 363–372.

[SW] Y. SAGIV AND S. WALECKA, *Subset dependencies and a completeness result for a subclass of embedded multivalued dependencies*, J. Assoc. Comput. Mach., 29 (1982), pp. 103–117.

[Va1] M. Y. VARDI, *The implication and finite implication problems for typed template dependencies*, in Proc. First ACM SIGACT-SIGMOD Principles of Database Systems, 1982, pp. 230–238.

[Va2] ———, Private communication, Oct. 1981.

[YP] M. YANNAKAKIS AND C. PAPADIMITRIOU, *Algebraic dependencies*, Proc. 1980 IEEE Symposium on Foundations of Computer Science, pp. 328–332.