A Polynomial Time Algorithm for Testing Implications of a Join Dependency and Embodied Functional Dependencies

J. Leuchner L. Miller G. Slutzki

Iowa State University, Ames, Iowa

Abstract

The problem of deciding whether a full join dependency (JD) $\bowtie[\mathcal{R}]$ and a set of functional dependencies (FDs) $\mathcal{F}$ imply an embedded join dependency (EJD) $\bowtie[\mathcal{S}]$ is known to be NP-complete. We show that the problem can be decided in polynomial time if $\mathcal{S} \subseteq \mathcal{R}$ and $\mathcal{F}$ is embedded in $\mathcal{R}$. Our work uses arguments based on an extension of complete intersection graphs rather than tableaus. This approach has facilitated our results and should prove useful for future research.

1 Introduction

The problem of deciding whether a full join dependency (JD) $\bowtie[\mathcal{R}]$ and a set of functional dependencies (FDs) $\mathcal{F}$ imply an embedded join dependency (EJD) $\bowtie[\mathcal{S}]$ is known to be NP-complete [MSY]. In this paper we examine the complexity of the implication problem when $\mathcal{S} \subseteq \mathcal{R}$ and $\mathcal{F}$ is required to be embedded in $\mathcal{R}$. We show that under these restrictions the problem can be decided in polynomial time.

The problem is of interest because of its implications for query optimization. Suppose, for example, that a query involves a relation r which is stored implicitly as its projections onto a collection of relation schemes $\mathcal{R}$ such that r satisfies the join dependency $\bowtie[\mathcal{R}]$. (I.e., we have a database satisfying a universal relation assumption [FMU,U2].) Answering this query would normally require taking the join of all the relations in $\{\pi_R(r) \mid R \in \mathcal{R}\}$. If, however, the query uses only the attributes appearing in some subset $\mathcal{S}$ of $\mathcal{R}$ and r satisfies the embedded join dependency $\bowtie[\mathcal{S}]$, it suffices to take only the join of the relations in $\{\pi_S(r) \mid S \in \mathcal{S}\}$. It would be useful in this situation to have an efficient algorithm for determining whether the JD $\bowtie[\mathcal{R}]$ and $\mathcal{F}$, a set of FDs satisfied by r, implies the EJD $\bowtie[\mathcal{S}]$.

To show that a polynomial time algorithm exists for the restricted problem, we define *join agreement graphs* (JAGs), an extension of complete intersection graphs (CIGs). The JAG corresponding to an instance of the implication problem exposes the combinatorial structure of the problem and offers a valuable extension of the use of graphs in database theory. (See [GS] and [FMUY] for examples of how graphs have been used by other researchers in relational database theory.) For this reason our work will not make use of tableaus, even though tableau arguments could easily be used in some places.

In Section 2 we will present some basic definitions and consider the case in which the set of FDs is empty. Section 3 will give the definition of JAG and agreement mapping. Section 4 contains the main result of the paper. Because of space limitations, we state some of our results without proof. We assume that the reader is familiar with basic relational database theory as presented in, for example, [U1,M,TL].

2 Hypergraphs and Join Dependencies

A *hypergraph* is a pair $(\mathcal{N}, \mathcal{E})$, where $\mathcal{N}$ is a finite set whose elements are called *nodes* and $\mathcal{E}$ is a set of nonempty subsets of $\mathcal{N}$ such that $\cup\mathcal{E} = \mathcal{N}$. The elements of $\mathcal{E}$ are called *hyperedges* (or simply *edges*). A standard reference for the theory of hypergraphs is [B]. It is immediately apparent that a database scheme can be represented by a hypergraph and any hypergraph can be interpreted as representing some database scheme. Thus, when we say "hypergraphs" and "edges", the reader may safely think of database schemes and relation schemes. We will usually represent a hypergraph by its edge set. I.e., we will use "hypergraph $\mathcal{E}$" to mean "hypergraph $(\mathcal{N}, \mathcal{E})$". Also, since our hypergraphs are representations of database schemes, we will usually refer to nodes as attributes.

If $\mathcal{H} = (\mathcal{N}, \mathcal{E})$ is a hypergraph and $\mathcal{E}'$ is a subset of $\mathcal{E}$, then $(\cup\mathcal{E}', \mathcal{E}')$ is the *subhypergraph* of $\mathcal{H}$ generated by $\mathcal{E}'$. The question of whether a JD $\bowtie[\mathcal{R}]$ implies an EJD $\bowtie[\mathcal{S}]$, where $\mathcal{S} \subseteq \mathcal{R}$ and no FDs are involved, can be answered by examining the topology of the hypergraph $\mathcal{R}$ [MLKL,GP]. In particular, the implication depends on how the subhypergraph $\mathcal{S}$ is connected to the rest of $\mathcal{R}$. To understand why this is so, we need to define "connected" precisely.

For a set of attributes V , we say that the edges R_1 and R_2 of $\mathcal{R}$ are *connected* in V if $R_1 \cap R_2 \cap V \neq \emptyset$. The sequence of edges $(S_1, \ldots, S_n)$ is an *edge path* in V from S_1 to S_n if S_i and S_{i+1} are connected in V for $1 \leq i \leq n-1$. The subhypergraph $\mathcal{S}$ is *connected in* V if for every $S_1, S_2 \in \mathcal{S}$, there is an edge path in V from S_1 to S_2 consisting of edges of $\mathcal{S}$. If V is $\alpha(\mathcal{R})$ (where $\alpha(\mathcal{R})$ denotes the set of attributes appearing in $\mathcal{R}$), the "in V" can be dropped from the above terms.

DEFINITION 2.1 The subhypergraph $\mathcal{S}$ of $\mathcal{R}$ is *externally connected* if there is some subset $\mathcal{A}$ of $\alpha(\mathcal{R})$ and some subset $\mathcal{G}$ of $\mathcal{R} - \mathcal{S}$ such that

1. $\mathcal{A} \subseteq \cup_{S \in \mathcal{S}} S$,
2. $\mathcal{A}$ is not contained in any single edge in $\mathcal{S}$,
3. $\mathcal{A} \subseteq \cup_{G \in \mathcal{G}} G$,
4. $\mathcal{G}$ is connected in $\overline{\alpha(\mathcal{S})}$ (the complement of $\alpha(\mathcal{S})$ with respect to $\alpha(\mathcal{R})$).

□

The relationship between join dependency implication and the topology of the hypergraph representations of the join dependencies has been explored in [GP] where a methodology is presented for decomposing a join dependency into an equivalent set of smaller join dependencies. The equivalence between *hinges*, as defined in [GP], and subhypergraphs which are not externally connected is shown in [MLKL]. Proposition 2.2 (below) gives the condition, stated in terms of external connections, under which a full join dependency $\bowtie[\mathcal{R}]$ implies the embedded join dependency $\bowtie[\mathcal{S}]$ for $\mathcal{S} \subseteq \mathcal{R}$. A proof is given in [MLKL]. Equivalent conditions have been given in [G] and [V]. The *if* direction of the proposition can also be derived directly from the results of [GP] using Sciore's axioms for join dependencies [S].

PROPOSITION 2.2 *Let $\mathcal{R}$ be a database scheme with $\mathcal{S} \subseteq \mathcal{R}$ and $\mathcal{S} \neq \emptyset$. Then $\bowtie[\mathcal{R}] \models \bowtie[\mathcal{S}]$ if and only if $\mathcal{S}$ is not externally connected.* □

Example 1: (See Figure 1.) Let $\mathcal{R} = \{ABD, BCE, ACF, AFG, BHI, CJ, GH, IJ\}$ and $\mathcal{S} = \{ABD, BCE, AFC\}$. The subhypergraph $\mathcal{S}$ is externally connected, so by Proposition 2.2, $\bowtie[\mathcal{R}]$ does not imply $\bowtie[\mathcal{S}]$. □

Since determining whether a subhypergraph is externally connected can be done in polynomial time, the implication problem can be decided in polynomial time for the special case in which the set of functional dependencies $\mathcal{F}$ is empty.

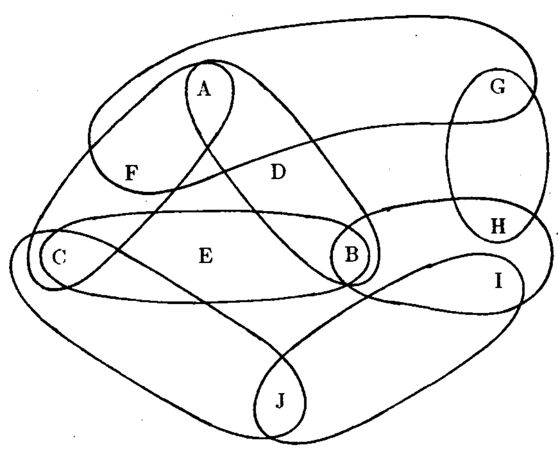

Figure 1. Hypergraph used in examples.

3 Join Agreement Graphs and Agreement Mappings

We now introduce *join agreement graphs* (JAGs) and show how these can be used to determine the effect of functional dependencies when the JD implication does not follow from the hypergraph topology alone.

It should be noted that we could define JAGs using tableaus rather than by starting with CIGs. This would allow us to avoid several proofs by appealing to known properties of tableaus. We have chosen not to do so for two reasons. First, the main result of this paper makes use of arguments that are naturally viewed in terms of graphs. By not using tableaus we achieve a consistency of approach that would be lost by the introduction of tableau arguments. Second, the formalism we introduce often allows proofs that are more algebraic, and perhaps more readily comprehensible, than the corresponding tableau arguments. It has been our aim to test the usefulness and power of this formalism. We believe it has shown its worth and offers opportunities for future research.

We will informally point out correspondences between tableaus and the various structures we define. A detailed exposition of tableaus can be found in [M].

The *complete intersection graph* (CIG) for the subhypergraph $\mathcal{S}$ of hypergraph $\mathcal{R}$ is an edge-labeled graph

$$\mathcal{I}_{\mathcal{S}} = (\mathcal{S}, \mathcal{E}_{\mathcal{S}}, \mathcal{L})$$

where $\mathcal{E}_{\mathcal{S}} = \{(S, S') \mid S, S' \in \mathcal{S}, S \neq S'\}$ is the edge set and $\mathcal{L} : \mathcal{E}_{\mathcal{S}} \to 2^{\alpha(\mathcal{S})}$ with $\mathcal{L}(S, S') = S \cap S'$. In other words, a CIG is an undirected graph whose vertices are hyperedges and whose edge set consists of all (unordered) pairs of distinct vertices. The edge labeling function $\mathcal{L}$ assigns to each edge the intersection of the hyperedges incident on the edge. Note that we use (S, S') to stand for $\{S, S'\}$, the edge adjacent to the vertices (hyperedges) S and S', i.e., the pair (S, S') is taken to be unordered. Also, we write $\mathcal{L}(S, S')$ for $\mathcal{L}((S, S'))$.

Clearly a database scheme can be represented as either a hypergraph or the corresponding CIG. Both contain the same information. As we will see, the effect of a set of FDs can be incorporated by modifying the edge labels of the CIG. For the following, let $\mathcal{R}$ be a database scheme with $\mathcal{S} \subseteq \mathcal{R}$. Also, let $\mathcal{F}$ be a set of functional dependencies over $\alpha(\mathcal{R})$. For any $X \subseteq \alpha(\mathcal{R})$, X^+ will denote the closure of X with respect to $\mathcal{F}$.

DEFINITION 3.1 Let $\mathcal{I}_{\mathcal{S}} = (\mathcal{S}, \mathcal{E}_{\mathcal{S}}, \mathcal{L})$ be the CIG for the subhypergraph $\mathcal{S}$. Define $\mathcal{M}(\mathcal{S}, \mathcal{F})$ to be the set of functions from $\mathcal{E}_{\mathcal{S}}$ to $2^{\alpha(\mathcal{R})}$ satisfying

(P1) $f(E) \supseteq \mathcal{L}(E)$, for all $E \in \mathcal{E}_{\mathcal{S}}$,

(P2) $f(E) = (f(E))^+$, for all $E \in \mathcal{E}_{\mathcal{S}}$, and

(P3) for any distinct $R_1, R_2, R_3 \in \mathcal{S}$, $f(R_1, R_2) \cap f(R_2, R_3) \subseteq f(R_1, R_3)$.

A labeling function that satisfies (P2) is said to be *closed* (with respect to $\mathcal{F}$), and one that satisfies (P3) is said to be *transitive*.

Let $\mathcal{L}^* : \mathcal{E}_S \to 2^{\alpha(\mathcal{R})}$ be defined by $\mathcal{L}^*(E) = \bigcap_{h \in \mathcal{M}(S,\mathcal{F})} h(E)$. The *join agreement graph* (JAG) for S with respect to $\mathcal{F}$ is $\mathcal{I}^*_{S,\mathcal{F}} = (S, \mathcal{E}_S, \mathcal{L}^*)$. □

We will usually write $\mathcal{I}^*$ for $\mathcal{I}^*_{S,\mathcal{F}}$ where S and $\mathcal{F}$ can be determined from context. It is easy to prove that $\mathcal{L}^*$ is transitive and closed and that for each $E \in \mathcal{E}_S$, $f(E) \supseteq \mathcal{L}(E)$, i.e., $\mathcal{L}^* \in \mathcal{M}(S,\mathcal{F})$. Furthermore, $\mathcal{L}^*$ can be computed in polynomial time by the following algorithm:

Algorithm $\mathcal{L}^*$:

begin
for (each $S, S' \in S$ with $S \neq S'$) *do*
    $\mathcal{L}^*(S,S') := S \cap S'$;
repeat
    for (each $E \in \mathcal{E}_S$) *do*
        $\mathcal{L}^*(E) := (\mathcal{L}^*(E))^+$;
    for (each triple S_1, S_2, S_3 of distinct elements of S) *do*
        $\mathcal{L}^*(S_1,S_3) := \mathcal{L}^*(S_1,S_3) \cup (\mathcal{L}^*(S_1,S_2) \cap \mathcal{L}^*(S_2,S_3))$
until ($\mathcal{L}^*$ is not changed)
end

The JAG $\mathcal{I}^*(S,\mathcal{F})$ corresponds to the tableau chase$_{\mathcal{F}}(T_S)$, where T_S is the extended tableau (i.e., extended to include all attributes in $\alpha(\mathcal{R})$) for $\bowtie[S]$. The edge-labeling function $\mathcal{L}^*$ can be interpreted as giving the sets of attributes on which pairs of distinct rows in chase$_{\mathcal{F}}(T_S)$ agree. The time bound on Algorithm $\mathcal{L}^*$ is not surprising since it is well known that chase$_{\mathcal{F}}(T_S)$ can be computed in polynomial time.

Example 2: Consider the hypergraph of Example 1. Let $\mathcal{F} = \{A \to B, B \to H\}$. The corresponding CIG is shown in Figure 2. Edge labels given in parentheses are values of $\mathcal{L}^*$. The JAG $\mathcal{I}^*$ consists of the three nodes ABD, BCE, ACF, the edges between them, and the edge labels in parentheses. □

Our study of the relationship between the implication problem $\{\bowtie[\mathcal{R}]\} \cup \mathcal{F} \models \bowtie[S]$ and the structures of the JAG $\mathcal{I}^*_{S,\mathcal{F}}$ depends on the following definition.

DEFINITION 3.2 A mapping $f : \mathcal{R} \to S$ is an *agreement mapping* (AM) with respect to $\mathcal{F}$ if

1. $f(S) = S$ for any $S \in S$, and
2. for any $R_1, R_2 \in \mathcal{R}$, if $f(R_1) \neq f(R_2)$, then $\mathcal{L}^*(f(R_1), f(R_2)) \supseteq R_1 \cap R_2$.

□

An agreement mapping $f : \mathcal{R} \to S$ corresponds to an application of the tableau $T_{\mathcal{R}}$ to chase$_{\mathcal{F}}(T_S)$ with the restriction that each row (in $\mathcal{R}$) labeled by a scheme in S is mapped to the corresponding row in chase$_{\mathcal{F}}(T_S)$. It is clear that such an application produces a row with distinguished values for all attributes in $\alpha(S)$, and this justifies (see [M]) the following proposition which can also be proved directly using JAGs.

PROPOSITION 3.3 *If there exists an agreement mapping $f : \mathcal{R} \to S$, then $\{\bowtie[\mathcal{R}]\} \cup \mathcal{F} \models \bowtie[S]$.* □

The converse of Proposition 3.3 is not true in general, but it is true when the set of FDs $\mathcal{F}$ is embedded in $\mathcal{R}$ as proved in the next proposition. Let $\mathcal{Q}$ be a subset of $\mathcal{R}$. We say that $\mathcal{F}$ is *embedded in* $\mathcal{Q}$ if for every $X \to Y \in \mathcal{F}$ there is some $R \in \mathcal{Q}$ such that $XY \subseteq R$.

PROPOSITION 3.4 *If $\{\bowtie[\mathcal{R}]\} \cup \mathcal{F} \models \bowtie[S]$ and $\mathcal{F}$ is embedded in $\mathcal{R}$, then there exists an agreement mapping $f : \mathcal{R} \to S$.*

Proof:

Assume that $\{\bowtie[\mathcal{R}]\} \cup \mathcal{F} \models \bowtie[S]$ and $\mathcal{F}$ is embedded in $\mathcal{R}$. We will define a relation which satisfies both $\bowtie[\mathcal{R}]$ and $\mathcal{F}$ (and therefore $\bowtie[S]$) and use this relation to define an agreement mapping $f : \mathcal{R} \to S$.

For convenience, assume that all attributes in $\alpha(\mathcal{R})$ have the same domain, and let a be some element of that domain. Let r be any relation which satisfies

1. $r = \{t_S \mid S \in S\}$ where t_S is a tuple over $\alpha(\mathcal{R})$ for each $S \in S$,
2. for any $S \in S$ and any $A \in S$, $t_S(A) = a$, and
3. for any $S_1, S_2 \in S$, with $S_1 \neq S_2$, and any $A \in \alpha(\mathcal{R})$,

$$t_{S_1}(A) = t_{S_2}(A) \text{ if and only if } A \in \mathcal{L}^*(S_1, S_2).$$

It can be shown that such a relation exists. We will now show that $r \in SAT(\mathcal{F})$.

Suppose $X \to Y \in \mathcal{F}$, $S, R \in S$ (with $S \neq R$), and $t_S[X] = t_R[X]$. By the definition of r, $X \subseteq \mathcal{L}^*(S,R)$. Then, since $\mathcal{L}^*$ is closed with respect to $\mathcal{F}$, $Y \subseteq \mathcal{L}^*(S,R)$. Therefore, $t_S[Y] = t_R[Y]$, and $r \in SAT(\mathcal{F})$.

It is not necessarily the case that r also satisfies the join dependency $\bowtie[\mathcal{R}]$. However, the project join mapping of r, $m_{\mathcal{R}}(r) = \bowtie_{R \in \mathcal{R}} \pi_R(r)$, is in SAT $(\bowtie[\mathcal{R}])$, see [M]. It can be shown that $m_{\mathcal{R}}$ preserves the property that r satisfies $\mathcal{F}$, and so, by hypothesis, $m_{\mathcal{R}}(r) \in$ SAT $(\bowtie[S])$.

Now for each $S \in S$, there is some $u_S \in \pi_S(r)$ such that $u_S(A) = a$ for all $A \in S$. Also, $\pi_S(r) \subseteq \pi_S(m_{\mathcal{R}}(r))$ (since $r \subseteq m_{\mathcal{R}}(r)$), so $u_S \in \pi_S(m_{\mathcal{R}}(r))$. Hence, $m_S(m_{\mathcal{R}}(r))$ contains a tuple

$$w = \bowtie\{u_S \mid S \in S\}$$

and $w(A) = a$ for all $A \in \alpha(S)$. But $m_{\mathcal{R}}(r) \in$ SAT $(\bowtie[S])$ means that

$$m_S(m_{\mathcal{R}}(r)) = \pi_{\alpha(S)}(m_{\mathcal{R}}(r)).$$

So, there is a tuple $v \in m_{\mathcal{R}}(r)$ such that $v[\alpha(s)] = w$. Then $v(A) = a$ for all $A \in \alpha(S)$. Also, $v = \{s_R \mid R \in \mathcal{R}\}$ for some set of tuples $\{s_R \mid R \in \mathcal{R}\}$ such that $s_R \in \pi_R(r)$ for each $R \in \mathcal{R}$, and each s_R is the projection of some tuple in r.

Define $f : \mathcal{R} \to S$ as follows. For $S \in S$, let $f(S) = S$. For $R \in \mathcal{R} - S$, pick some $S \in S$ such that $t_S[R] = s_R$ and let $f(R) = S$. Note that for all $R \in \mathcal{R}$,

$$t_{f(R)}[R] = s_R.$$

To show that f is an agreement mapping, we need to prove that for any $R_1, R_2 \in \mathcal{R}$, $R_1 \cap R_2 \subseteq \mathcal{L}^*(f(R_1), f(R_2))$ if $f(R_1) \neq f(R_2)$. Let $R_1, R_2 \in \mathcal{R}$. Then $t_{f(R_1)}[R_1] = s_{R_1}$ and $t_{f(R_2)}[R_1] = s_{R_2}$. Since $\{s_R \mid R \in \mathcal{R}\}$ is joinable $s_{R_1}[R_1 \cap R_2] = s_{R_2}[R_1 \cap R_2]$. Hence,

$$t_{f(R_1)}\left[R_1 \bigcap R_2\right] = t_{f(R_2)}\left[R_1 \bigcap R_2\right].$$

By the definition of r

$$R_1 \bigcap R_2 \subseteq \mathcal{L}^*(f(R_1), f(R_2)).$$

Therefore, f is an agreement mapping. □

Example 3: (See Figure 2.) Consider the JAG of Example 2. We can define an agreement mapping from $\mathcal{R}$ to $\mathcal{S}$ by mapping the elements of $\mathcal{S}$ to themselves, AG to ACF, BHI to BCE, CJ to BCE, IJ to BCE, and GH to ACF. Therefore $\{\bowtie[\mathcal{R}]\} \cup \mathcal{F} \models \bowtie[\mathcal{S}]$ holds. Note that many other sets of functional dependencies will cause the JD implication to be true. □

For the special case in which $\mathcal{F}$ is embedded in $\mathcal{S}$, the implication problem reduces to the problem of determining whether a particular subhypergraph is externally connected. Let $\mathcal{S}^+$ denote the set of hyperedges consisting of the closures, with respect to $\mathcal{F}$, of the hyperedges of $\mathcal{S}$, i.e., $\mathcal{S}^+ = \{S^+ \mid S \in \mathcal{S}\}$.

PROPOSITION 3.5 *Suppose $\mathcal{F}$ is embedded in $\mathcal{S}$. Then $\{\bowtie[\mathcal{R}]\} \cup \mathcal{F} \models \bowtie[\mathcal{S}]$ iff $\mathcal{S}^+$ is not externally connected in the hypergraph $\mathcal{S}^+ \cup (\mathcal{R} - \mathcal{S})$.* □

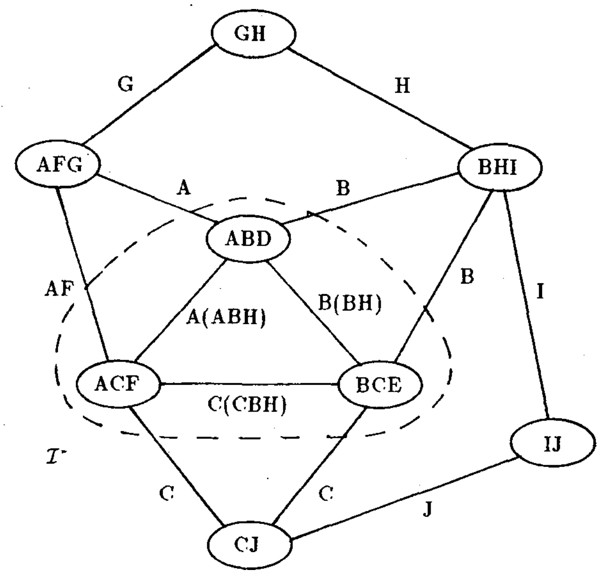

Figure 2.

The CIG for hypergraph $\mathcal{R}$ of Figure 1 and the JAG $\mathcal{I}^*$ for $\mathcal{S} = \{ACF, ABD, BCE\}$ and $\mathcal{F} = \{B \to H, A \to B\}$.

Since computing closures and determining whether a subhypergraph is externally connected can both be done in polynomial time, the implication problem in this special case is also decidable in polynomial time. In the next section we show how JAGs and agreement mappings can be applied in the more general case in which $\mathcal{F}$ is only embedded in $\mathcal{R}$.

4 The Implication Problem with Embedded FDs

Suppose that $\mathcal{F}$ is embedded in $\mathcal{R}$. By Propositions 3.3 and 3.4, the implication $\{\bowtie[\mathcal{R}]\} \cup \mathcal{F} \models \bowtie[\mathcal{S}]$ holds if and only if there exists an agreement mapping from $\mathcal{R}$ to $\mathcal{S}$. We will show that the existence of an agreement mapping can be determined in polynomial time. To do this, we need to examine the structure of the JAG $\mathcal{I}^* = (\mathcal{S}, \mathcal{E}_{\mathcal{S}}, \mathcal{L}^*)$ in more detail. We begin with another definition.

DEFINITION 4.1 Let $X \subseteq \alpha(\mathcal{R})$. An *X-clique* of $\mathcal{I}^*$ is any $\mathcal{Q} \subseteq \mathcal{S}$ such that

1. $X \subseteq \mathcal{L}^*(Q_1, Q_2)$ for any $Q_1, Q_2 \in \mathcal{Q}$ with $Q_1 \neq Q_2$, and
2. for any $R \in \mathcal{S} - \mathcal{Q}$ and $Q \in \mathcal{Q}$, $X - \mathcal{L}^*(R, Q) \neq \emptyset$.

An X-clique is *trivial* if it contains only one hyperedge and that hyperedge does not contain X and is *nontrivial* otherwise. If $X = \{A\}$, we will usually drop the set brackets and write 'A-clique' rather than '{ A} -clique'. □

In the JAG of Example 3, $\{ABD\}$ is a trivial C-clique, a trivial G-clique, and a nontrivial D-clique. Also, $\{ACE, BCE\}$ is a nontrivial CBH-clique, and $\mathcal{S}$ is a nontrivial BH-clique. An important property of X-cliques is given in the next lemma.

LEMMA 4.2 *For any $X \subseteq \alpha(\mathcal{R})$ and any $Q \in \mathcal{S}$, there is exactly one X-clique of $\mathcal{I}^*$ containing Q, namely,*

$$\mathcal{C} = \{S \in \mathcal{S} \mid S = Q \text{ or } X \subseteq \mathcal{L}^*(S, Q)\}.$$

Proof:

Let $\mathcal{C} = \{S \in \mathcal{S} \mid S = Q \text{ or } X \subseteq \mathcal{L}^*(S, Q)\}$. Suppose $R, R' \in \mathcal{C}$ and $R \neq R'$. Then $X \subseteq \mathcal{L}^*(R, Q)$ and $X \subseteq \mathcal{L}^*(R', Q)$. By the transitivity of $\mathcal{L}^*$, $X \subseteq \mathcal{L}^*(R, R')$. If $R \in \mathcal{R} - \mathcal{C}$, then $X - \mathcal{L}^*(R, Q) \neq \emptyset$. Hence, $\mathcal{C}$ is an X-clique.

Suppose $\mathcal{C}'$ is an X-clique of $\mathcal{I}$ containing Q. Then for any $R \in \mathcal{C}'$, with $R \neq Q$, $X \subseteq \mathcal{L}^*(R, Q)$. Now for any $R' \in \mathcal{C}$, with $R' = Q$ and $R' \neq R$, $X \subseteq \mathcal{L}^*(R', Q)$. So by the transitivity of $\mathcal{L}^*$, $X \subseteq \mathcal{L}^*(R', R)$. Therefore, $\mathcal{C}' \subseteq \mathcal{C}$. Similarly, $\mathcal{C}' \supseteq \mathcal{C}$. Hence, $\mathcal{C}' = \mathcal{C}$. □

For any $S \in \mathcal{S}$ and $X \subseteq \alpha(\mathcal{R})$, the (unique) X-clique of $\mathcal{I}^*$ containing S will be denoted by $\mathcal{C}^*_X(S)$. Also, if $X = \{A\}$, we will write $\mathcal{C}^*_A(R)$ rather than $\mathcal{C}^*_{\{A\}}(R)$.

The following facts are easy to prove and will be used implicitly in subsequent arguments.

1. For $S, S' \in \mathcal{S}$, with $S \neq S'$, $S' \in \mathcal{C}^*_X(S)$ if and only if $X \subseteq \mathcal{L}^*(S, S')$.

2. If $X \subseteq Y$, then $C^{*}_{Y}(S) \subseteq C^{*}_{X}(S)$ for any $S \in \mathcal{S}$.

3. If $X \rightarrow Y \in \mathcal{F}^{+}$, then for any $S \in \mathcal{S}$, $C^{*}_{X}(S) \subseteq C^{*}_{Y}(S)$.

4. Suppose $R, R' \in \mathcal{R}$ and $f : \mathcal{R} \rightarrow \mathcal{S}$ is an agreement mapping with respect to $\mathcal{F}$. If $X \subseteq R \cap R'$, then $C^{*}_{X}(f(R)) = C^{*}_{X}(f(R'))$.

Although checking whether a given mapping of $\mathcal{R}$ to $\mathcal{S}$ is an agreement mapping is easy, polynomial time is not sufficient to check all possible mappings of $\mathcal{R}$ to $\mathcal{S}$ to determine whether an agreement mapping exists. Even the requirement that each scheme of $\mathcal{S}$ be mapped to itself does not reduce the possibilities to a tractable number. Lemma 4.3 provides a way to further restrict the mappings that must be considered, and this will be the key to a polynomial time algorithm.

LEMMA 4.3 *Assume that $\mathcal{F}$ is embedded in $\mathcal{R}$, and $f : \mathcal{R} \rightarrow \mathcal{S}$ is an agreement mapping. Suppose that $B \in \alpha(\mathcal{R})$, $S \in \mathcal{S}$, and $R \in \mathcal{R}$, such that $B \in S^{+}$ and $B \in R$. Then $f(R) \in C^{*}_{B}(S)$.*

Proof:

If $R = S$, then $f(R) = S$, and $f(R) \in C^{*}_{B}(S)$. So suppose that $R \neq S$. If $B \in S$, then $B \in S \cap R$, and since f is an agreement mapping,

$$S \cap R \subseteq \mathcal{L}^{*}(f(S), f(R)) = \mathcal{L}^{*}(S, f(R)).$$

Therefore, $f(R) \in C^{*}_{B}(S)$. Suppose that $B \notin S$. Then, since $B \in S^{+}$, there is a sequence of FD's $\{X_i \rightarrow Y_i \mid 1 \leq i \leq n\}$ such that

1. $X_1 \subseteq S$,
2. $X_{i+1} \subseteq SY_1 \ldots Y_i$, $1 \leq i \leq n-1$, and
3. $B \in Y_n$.

Let $R_1, \ldots, R_n \in \mathcal{R}$ such that $X_iY_i \subseteq R_i$, $1 \leq i \leq n$. Such hyperedges exist since $\mathcal{F}$ is embedded in $\mathcal{R}$. Define $\mathcal{C}_0 = C^{*}_{S}(S)$ and, for $1 \leq i \leq n$, let

$$\mathcal{C}_i = \mathcal{C}_{i-1} \cap C^{*}_{Y_i}(f(R_i)).$$

We will show that $S \in \mathcal{C}_n$. Note that $S \in \mathcal{C}_0$. Also,

$$\mathcal{C}_0 = \mathcal{C}_0 \cap C^{*}_{S}(S) = \mathcal{C}_0 \cap C^{*}_{X_1}(S) = \mathcal{C}_0 \cap C^{*}_{X_1}(f(R_1)).$$

The last equality holds since

$$X_1 \subseteq S \cap R_1 \subseteq \mathcal{L}^{*}(f(S), f(R_1)) = \mathcal{L}^{*}(S, f(R_1)),$$

so that S and $f(R_1)$ are in the same X-clique. Since $X_1 \rightarrow Y_1 \in \mathcal{F}$, $C^{*}_{X_1}(f(R_1)) \subseteq C^{*}_{Y_1}(f(R_1))$. So

$$\mathcal{C}_0 \subseteq \mathcal{C}_0 \cap C^{*}_{Y_1}(f(R_1)).$$

Since the reverse inclusion is clearly true,

$$\mathcal{C}_0 = \mathcal{C}_0 \cap C^{*}_{Y_1}(f(R_1)) = \mathcal{C}_1.$$

The same type of argument can be applied to show that $\mathcal{C}_i = \mathcal{C}_{i+1}$ as follows. (Note that $C^{*}_{\emptyset}(R) = \mathcal{S}$ for all $R \in \mathcal{S}$.)

$$\begin{aligned}
\mathcal{C}_i &= C^{*}_{S}(S) \cap C^{*}_{Y_1}(f(R_1)) \cap \ldots \cap C^{*}_{Y_i}(f(R_i)) \\
&= \mathcal{C}_i \cap C^{*}_{S \cap X_{i+1}}(S) \cap C^{*}_{Y_1 \cap X_{i+1}}(f(R_1)) \cap \ldots \\
&\quad \cap C^{*}_{Y_i \cap X_{i+1}}(f(R_i)) \\
&= \mathcal{C}_i \cap C^{*}_{S \cap X_{i+1}}(f(R_{i+1})) \cap C^{*}_{Y_1 \cap X_{i+1}}(f(R_{i+1})) \cap \ldots \\
&\quad \cap C^{*}_{Y_i \cap X_{i+1}}(f(R_{i+1})) \\
&= \mathcal{C}_i \cap C^{*}_{X_{i+1}}(f(R_{i+1})) \subseteq \mathcal{C}_i \cap C^{*}_{Y_{i+1}}(f(R_{i+1})) = \mathcal{C}_{i+1}.
\end{aligned}$$

Since the reverse of the above inclusion is clearly true, we have $\mathcal{C}_0 = \mathcal{C}_1 = \ldots = \mathcal{C}_n$. Hence, $S \in \mathcal{C}_n$. Also,

$$\mathcal{C}_n = \mathcal{C}_{n-1} \cap C^{*}_{Y_n}(f(R_n)) \subseteq C^{*}_{Y_n}(f(R_n)) \subseteq C^{*}_{B}(f(R_n)).$$

The last inclusion is true since $B \in Y_n$. Then $C^{*}_{B}(f(R_n)) = C^{*}_{B}(S)$, and since $B \in R \cap R_n$, $f(R) \in C^{*}_{B}(f(R_n)) = C^{*}_{B}(S)$.
□

The preceding lemma shows that existence of an agreement mapping is closely related to the clique structure of $\mathcal{I}^{*}$. In what follows we develop this relationship further. We will call the JAG $\mathcal{I}^{*}$ *coherent* if it has no more than one nontrivial A-clique for each attribute A. Also, let $\alpha^{*}(\mathcal{S})$ denote the set of attributes that appear in $\mathcal{I}^{*}$, i.e., $\alpha^{*}(\mathcal{S}) = \alpha(\mathcal{S}) \cup [\cup\{\mathcal{L}^{*}(E) \mid E \in \mathcal{E}_{\mathcal{S}}\}]$. The following technical lemma will be used implicitly in subsequent proofs.

LEMMA 4.4 *Assume that $\mathcal{F}$ is embedded in $\mathcal{R}$. Let $R \in \mathcal{R}$, $S \in \mathcal{S}$, and $B \in \alpha(\mathcal{R})$.*

(1) If $C^{*}_{B}(S)$ is nontrivial, then $B \in S^{+}$.

(2) If $B \in \alpha^{*}(\mathcal{S})$, the JAG $\mathcal{I}^{*}$ contains at least one nontrivial B-clique.

Suppose also that $f : \mathcal{R} \rightarrow \mathcal{S}$ is an agreement mapping. Then:

(3) If $B \in R$ and $C^{*}_{B}(S)$ is nontrivial, then $f(R) \in C^{*}_{B}(S)$.

(4) If $R' \in \mathcal{R} - \mathcal{S}$ and $R \cap R' - \alpha^{*}(\mathcal{S}) \neq \emptyset$, then $f(R) = f(R')$.

(5) $\mathcal{I}^{*}$ is coherent.

□

If the JAG $\mathcal{I}^{*}$ contains only one nontrivial A-clique for some $A \in \alpha(\mathcal{R})$, this A-clique is called the *major A-clique* of $\mathcal{I}^{*}$ and will be denoted by C^{*}_{A}. It follows that if $\mathcal{I}^{*}$ is coherent and $B \in S$ for some $S \in \mathcal{S}$, then $C^{*}_{B}(S) = C^{*}_{B}$.

If $\mathcal{F}$ is embedded in $\mathcal{R}$, then by Lemma 4.4 (Part 4) two schemes of $\mathcal{R}$ containing a common attribute which does not appear in either the vertices or edge labels of $\mathcal{I}^{*}$ must have the same image under any agreement mapping. This motivates the following definition. For $R_1, R_2 \in \mathcal{R}$, we write $R_1 \equiv_0 R_2$ if $(R_1 \cap R_2) - \alpha^{*}(\mathcal{S}) \neq \emptyset$. Let $\equiv^{*}$ be the reflexive, transitive closure of $\equiv_0$. Since $\equiv_0$ is symmetric, $\equiv^{*}$ is an equivalence relation. For $R \in \mathcal{R}$ the *agreement mapping equivalence class* of R is

$$\mathcal{D}^{*}(R) = \{R' \in \mathcal{R} \mid R' \equiv^{*} R\}.$$

It is easy to see that if $\mathcal{F}$ is embedded in $\mathcal{R}$ and $f : \mathcal{R} \rightarrow \mathcal{S}$ is an agreement mapping, then $f(R) = f(R')$ whenever R and R' are in the same agreement mapping equivalence class.

PROPOSITION 4.5 *Assume that $\mathcal{F}$ is embedded in $\mathcal{R}$ and that $\mathcal{I}^*$ is coherent. Then there exists an agreement mapping from $\mathcal{R}$ to $\mathcal{S}$ if and only if for each $R \in \mathcal{R}$,*

$$\bigcap \left\{ C^*_B \mid B \in \alpha(\mathcal{D}^*(R)) \bigcap \alpha^*(\mathcal{S}) \right\} \neq \emptyset.$$

Proof:

(only if): Suppose $f : \mathcal{R} \to \mathcal{S}$ is an agreement mapping. Let $R \in \mathcal{R}$.

Now f maps all the hyperedges of $\mathcal{D}^*(R)$ into the same hyperedge of $\mathcal{S}$, say S_0. Suppose $B \in \alpha(\mathcal{D}^*(R)) \bigcap \alpha^*(\mathcal{S})$. Then $\exists R' \in \mathcal{D}^*(R)$ with $B \in R'$. Then $f(R') \in C^*_B$, and $f(R') = S_0$. So, $C^*_B = C^*_B(S_0)$, and $S_0 \in C^*_B$. Thus,

$$S_0 \in \bigcap \left\{ C^*_B \mid B \in \alpha(\mathcal{D}^*(R)) \bigcap \alpha^*(\mathcal{S}) \right\}.$$

(if): Assume that for each $R \in \mathcal{R}$

$$\bigcap \left\{ C^*_B \mid B \in \alpha(\mathcal{D}^*(R)) \bigcap \alpha^*(\mathcal{S}) \right\} \neq \emptyset.$$

Define a mapping $f : \mathcal{R} \to \mathcal{S}$ as follows. For $S \in \mathcal{S}$, let $f(S) = S$. For each $\mathcal{D} \in \{\mathcal{D}^*(R) \mid R \in \mathcal{R} - \mathcal{S}\}$, pick some hyperedge in $\bigcap \{C^*_B \mid B \in \alpha(\mathcal{D}) \bigcap \alpha^*(\mathcal{S})\}$ to be the image, under f, of all the hyperedges in $\mathcal{D}$. We need to show that f is an agreement mapping. Let $R_1, R_2 \in \mathcal{R}$.

Case(1): $R_1, R_2 \in \mathcal{S}$. Then $\mathcal{L}^*(f(R_1), f(R_2)) = \mathcal{L}^*(R_1, R_2)$, and so, $R_1 \bigcap R_2 \subseteq \mathcal{L}^*(f(R_1), f(R_2))$.

Case(2): $R_1 \in \mathcal{S}$, $R_2 \in \mathcal{R} - \mathcal{S}$. Suppose $B \in R_1 \bigcap R_2$. Now $f(R_2) \in C^*_B$. Also, $C^*_B = C^*_B(R_1) = C^*_B(f(R_1))$. Hence, if $f(R_1) \neq f(R_2)$, then at least $f(R_1)$ and $f(R_2)$ are in the same B-clique and $B \in \mathcal{L}^*(f(R_1), f(R_2))$.

Case(3): $R_1, R_2 \in \mathcal{R} - \mathcal{S}$. If $(R_1 \bigcap R_2) - \alpha^*(\mathcal{S}) \neq \emptyset$, then $R_1 \in \mathcal{D}^*(R_2)$ and $f(R_1) = f(R_2)$. Suppose $R_1 \bigcap R_2 \subseteq \alpha^*(\mathcal{S})$, and assume $B \in R_1 \bigcap R_2$. Then $B \in \alpha(\mathcal{D}^*(R_i)) \bigcap \alpha^*(\mathcal{S})$ for $i = 1, 2$. By the definition of f, both $f(R_1)$ and $f(R_2)$ are in C^*_B. Hence, if $f(R_1) \neq f(R_2)$, then $B \in \mathcal{L}^*(f(R_1), f(R_2))$. Thus, for any $R_1, R_2 \in \mathcal{R}$, if $f(R_1) \neq f(R_2)$, then $R_1 \bigcap R_2 \subseteq \mathcal{L}^*(f(R_1), f(R_2))$. Hence, f is an agreement mapping. □

The proof of Proposition 4.5 shows us that we need only consider those mappings that take each scheme R into

$$\bigcap \left\{ C^*_B \mid B \in \alpha(\mathcal{D}^*(R)) \bigcap \alpha^*(\mathcal{S}) \right\}$$

and that any such mapping is an agreement mapping. This allows us to state our main result.

THEOREM 4.6 *If $\mathcal{S} \subseteq \mathcal{R}$ and $\mathcal{F}$ is embedded in $\mathcal{R}$, then it can be decided in polynomial time whether $\{\bowtie[\mathcal{R}]\} \cup \mathcal{F} \models \bowtie[\mathcal{S}]$.*

Proof: (Sketch)

By Propositions 3.3, 3.4, 4.5 and Lemma 4.4, we can determine whether the JD implication holds as follows.

1. Compute $\mathcal{L}^*$ for $\mathcal{S}$ and $\mathcal{F}$.

2. For each $A \in \alpha(\mathcal{R})$, check whether $\mathcal{I}^*$ is coherent. If it is not, then the JD implication does not hold.

3. Partition $\mathcal{R} - \mathcal{S}$ into equivalence classes under $\equiv^*$. Each equivalence class is $\mathcal{D}^*(R)$ for some $R \in \mathcal{R} - \mathcal{S}$. For each of these, check whether

$$\bigcap \left\{ C^*_B \mid B \in \alpha(\mathcal{D}^*(R)) \bigcap \alpha^*(\mathcal{S}) \right\} \neq \emptyset.$$

   If so, the JD implication holds. Otherwise, it does not.

Note that each $S \in \mathcal{S}$ forms an equivalence class under $\equiv^*$ by itself and that $\mathcal{I}^*$ always has at least one B-clique (which contains S) for each $B \in S$.

The size of the JD implication problem is proportional to the number of attributes appearing in $\mathcal{R}$, the number of relation schemes in $\mathcal{R}$, and the size of the description of $\mathcal{F}$. The edge labeling function $\mathcal{L}^*$ can be computed in polynomial time using Algorithm $\mathcal{L}^*$. For any $A \in \alpha(\mathcal{R})$, $\mathcal{I}^*$ can be partitioned into A-cliques in polynomial time to determine whether more than one nontrivial A-clique exists. Hence, checking whether $\mathcal{I}^*$ is coherent can be done in polynomial time. Finally, the nontrivial cliques found in step 2 can be used to do the checking in step 3. This involves a polynomial number of set intersections. Therefore, the entire process can be accomplished in polynomial time. □

5 Summary

We have shown that the restricted JD implication problem can be decided in polynomial time. This is an interesting result both because of its implications for query optimization and because the unrestricted problem is NP-complete. In fact, the technique used in [MSY] can be extended to show that the problem remains NP-complete even if $\mathcal{S}$ is required to be a subset of $\mathcal{R}$. Thus, requiring that a set of functional dependencies be embedded has a profound effect (assuming that $P \neq NP$) on the complexity of the decision problem.

The results of this paper were made possible by the use of join agreement graphs (JAGs) and agreement mappings. We have found these structures to be extremely useful for analysing the combinatorial structure of the JD implication problem. They have provided insights into the problem beyond those provided by tableau representations of the JDs, and we believe they will continue to prove their utility in future research. There is, of course, a close correspondence between JAGs and tableaus and between agreement mappings and applications of J-rules to a tableaus. (Tableaus and J-rules are defined in [M].) In fact, we could have defined the JAG using chase$_{\mathcal{F}}(T_S)$. However, since none of our arguments involve tableaus, we have defined JAGs directly, as an extension of hypergraphs and complete intersection graphs, thereby avoiding the extra level of complexity in the definitions of tableaus and the *chase* algorithm.

The definitions of JAG and agreement mapping given in this paper open several avenues for future research. Our current efforts are directed to extending the use of JAGs and agreement mappings to other problems. In particular, if the JD implication is not true, we would like to have a method for finding the smallest $\mathcal{S}' \subseteq \mathcal{R}$ such that $\mathcal{S} \subseteq \mathcal{S}'$ and $\{\bowtie[\mathcal{R}]\} \cup \mathcal{F} \models \bowtie[\mathcal{S}']$ holds. This would be of interest for query optimization. Also, although the restricted JD implication problem has been shown to be decidable in polynomial time, it remains to develop an practical algorithm for doing so. A lossless join test based on Proposition 3.5 has been implemented for a universal relation interface to INGRES, and we intend to extend this implementation to use the techniques presented in this paper.

References

[B] C. Berge. *Graphs and Hypergraphs*, North-Holland, 1976.

[FMU] R. Fagin, A.O. Mendelzon and J. Ullman. "A Simplified Universal Relation Assumption and Its Properties", *ACM TODS*, Vol. 7, No. 3, Sept. 1982, pp.343-360.

[FMUY] R. Fagin, D. Maier, J. Ullman and M. Yannakakis. "Tools for Template Dependencies", *SIAM J. Computing*, Vol. 12, No. 1, February 1983, pp. 36-59.

[G] M. Gyssens. "Embedded Join Dependencies as a Tool for Decomposing Full Join Dependencies", *Proceedings of the ACM Symposium on Principles of Database Systems,* April 1985, pp.205-214.

[GP] M. Gyssens and J. Paradaens. "A Decomposition Methodology for Cyclic Databases", *Advances in Database Theory*, Vol. 2, 1984, pp. 85-122.

[GS] N. Goodman and O. Schmueli. "Syntactic Characterization of Tree Database Schemas", *JACM,* Vol. 30, No. 4, October 1983, pp.767-786.

[M] D. Maier. *The Theory of Relational Databases,* Computer Science Press, 1983.

[MSY] D. Maier, Y. Sagiv and M. Yannakakis. "On the Complexity of Testing Implications of Functional and Join Dependencies," *JACM*, Vol. 28, No.4, 1981, pp. 680-695.

[MLKL] L.L. Miller, J.H. Leuchner, S. Kothari and K.C. Liu. *Testing Arbitrary Subhypergraphs for the Lossless Join Property* (Submitted for publication.)

[S] E. Sciore. "A Complete Axiomatization of Full Join Dependencies," *JACM,* Vol. 29, 1982, pp. 373-393.

[TL] D.C. Tsichritzis and F.H.Lochovsky. *Data Models,* Prentice-Hall, 1982.

[U1] J. Ullman. *Principles of Database Systems,* Second Edition, Computer Science Press, 1982.

[U2] J. Ullman. "The U.R. Strikes Back," *Proceedings of the ACM Symposium on Principles of Database Systems*, Los Angeles, March, 1982, pp. 10-22.

[V] M. Y. Vardi. "Axiomatization of Functional and Join Dependencies in the Relational Model," *M. Sc. thesis,* Weizman Institute, Rehovot,Israel, April 1980.