Computer Science Department Technical Report
Cognitive Systems Laboratory
University of California
Los Angeles, CA 90024-1596

AXIOMS AND ALGORITHMS FOR INFERENCES INVOLVING PROBABILISTIC INDEPENDENCE

Dan Geiger
Azaria Paz
Judea Pearl

June 1989
CSD-890031

Submitted to [redacted] Information and Computation, 1989.

TECHNICAL REPORT
R-119
January 1989

AXIOMS AND ALGORITHMS FOR INFERENCES INVOLVING PROBABILISTIC INDEPENDENCE*

Dan Geiger, Azaria Paz† and Judea Pearl

Cognitive Systems Laboratory
Computer Science Department
Los Angeles, CA 90024-1596

* This work was supported in part by the National Science Foundation Grant # IRI-8610155.

† Permanent affiliation: Technion IIT, Haifa, Israel.

AXIOMS AND ALGORITHMS FOR INFERENCES

Dan Geiger
4731 Boelter Hall
UCLA Computer Science Department
Los Angeles, California, 90024-1596.

ABSTRACT

This paper offers an axiomatic characterization of the probabilistic relation "X is independent of Y (written $I(X, Y)$)," where X and Y are two disjoint sets of variables. Three axioms for $I(X, Y)$ are presented and shown to be *complete*. Based on these axioms, a polynomial *membership* algorithm is developed to decide whether any given *statement* $I(X, Y)$ *logically follows* from a set Σ of such statements i.e., whether $I(X, Y)$ holds in every probability distribution that obeys Σ. The complexity of the algorithm is $O(|\Sigma| \cdot k^2 + |\Sigma| \cdot n)$ where $|\Sigma|$ is the number of given statements, n is the number of variables in $\Sigma \cup \{I(X, Y)\}$ and k is the number of variables in $I(X, Y)$.

SYMBOLS

Symbol	Meaning
P_{σ}	P sub sigma
$P_{\sigma'}$	P sub sigma prime
$\nvdash_{\mathbf{A}}$	not derive
$\models_{\mathbf{P}}$	entails
$\not\models_{\mathbf{P}}$	not entails
$\vdash_{\mathbf{A}}$	derives
$\otimes$	citimes
σ	sigma
$\mathbf{\Sigma}$	bold capital sigma
γ	gamma
$\mathbf{\Gamma}$	bold capital gamma
α	alpha
$\emptyset$	empty set
$\in$	member of
$\notin$	not member of
$\cup$	union
$\supseteq$	supset (or equal)
$\subseteq$	subset (or equal)
$\rightarrow$	then

1. INTRODUCTION

Consider a collection of information sources, each reflecting a different aspect of some underlying probabilistic phenomenon. These sources can be regarded as a set of random variables, governed by an unknown probability distribution, some of which are dependent and some independent. In this paper, we are concerned with the following problem: Assume we know that some groups of variables are mutually independent, either by statistical analysis or by conceptual understanding of the underlying phenomenon; we need to infer new independencies without resorting to additional measurements or expensive numerical analysis.

We formalize this question as follows: Let a (independence) *statement* $\sigma = I(X, Y)_P$ stand for the probabilistic relation "the variables in X are independent of those in Y," i.e., $P(X\ Y) = P(X) \cdot P(Y)$ where X and Y are disjoint sets of variables and P is a distribution over these variables. We say that a statement σ is *logically implied* by a set of statements Σ iff every distribution that obeys Σ obeys σ as well. We ask: Is the statement $\sigma = I(X, Y)_P$ logically implied by a given set Σ of such statements, each characterized by a different pair of subsets X' and Y'?

The answer is given in two steps. First, (in Section 2), we provide the following inference rules, considered as axioms:

Trivial Independence	$I(X, \varnothing)$	(1.a)
Symmetry	$I(X, Y) \rightarrow I(Y, X)$	(1.b)
Decomposition	$I(X, Y \cup Z) \rightarrow I(X, Y)$	(1.c)

Mixing $$I(X, Y) \,\&\, (X \cup Y, Z) \rightarrow I(X, Y \cup Z) \tag{1.d}$$

These axioms clearly hold for all distributions and therefore are *sound.* For example, to prove (1.d), we observe that $P(X\ Y) = P(X) \cdot P(Y)$ and $P(X\ Y\ Z) = P(X\ Y) \cdot P(Z)$ imply that $P(X\ Y\ Z) = P(Y)P(X)P(Z)$. Moreover, summing over X yields $P(Y\ Z) = P(Y)P(Z)$, hence $P(X\ Y\ Z) = P(X)P(Y\ Z)$ which establishes the right hand side of (1.d). We show that these axioms are also *complete, i.e.,* capable of deriving by repeated applications all independencies that are logically implied by the input set of independencies.

The second step of our solution is a membership algorithm that efficiently answers whether a statement σ is a member of the closure $\mathbf{cl}(\Sigma)$ of Σ under axioms (1). In light of the completeness results, $\sigma \in \mathbf{cl}(\Sigma)$ iff σ holds in every distribution that obeys Σ. This step is covered in Section 3. Section 4 extends the results to the problem of deciding consistency: Given a set of independence statements mixed with dependence statements, to decide if the set is consistent, i.e., if there exists a probability distribution that satisfies all the statements simultaneously. We show that the axioms provided in section 2 are *strongly complete* and, hence, the consistency problem can be translated into a sequence of membership problems.

Similar problems of membership and axiomatic characterization are treated in the literature on database dependencies, for example (Beeri et al, 1977; Fagin, 1977; Beeri, 1980; Fagin, 1982; Sagiv and Walecka, 1982). Our notations and definitions were particularly influenced by (Beeri et al, 1977; Fagin, 1977). A survey on database dependency theory can be found in (Rissanen, 1978; Fagin and Vardi, 1986; Vardi, 1988). An extension of this work to conditional independence has been obtained in (Geiger and Pearl, 1987; Pearl et al, 1988; Geiger et al, 1989;

Pearl, 1988).

2. AXIOMATIC CHARACTERIZATION

The following symbols are used: σ for a statement, $\boldsymbol{\Sigma}$ for a set of statements, $\mathbf{P}$ for a class of distributions, such as the class of all probability distributions (PD), the class of normal distributions (PN), and the class of distributions over bi-valued variables (PB). The set union symbol is often dropped and XY is written instead of $X \cup Y$. A statement $I(X, Y)$ is written as a pair (X, Y).

Definition: σ is *logically implied* by Σ denoted $\boldsymbol{\Sigma} \models_{\mathbf{P}} \sigma$, iff every distribution in $\mathbf{P}$ that obeys Σ also obeys σ. $\Sigma \vdash_{\mathbf{A}} \sigma$ iff $\sigma \in \mathbf{cl}_{\mathbf{A}}(\Sigma)$, i.e., there exists a *derivation chain* $\sigma_1 \ldots, \sigma_n = \sigma$ such that for each σ_j, either $\sigma_j \in \Sigma$, or σ_j is derived by an axiom in $\mathbf{A}$ from the previous statements.

Definition: A set of axioms $\mathbf{A}$ is *sound in* $\mathbf{P}$ iff for every statement σ and every set of statements $\boldsymbol{\Sigma}$

$$\boldsymbol{\Sigma} \vdash_{\mathbf{A}} \sigma \quad \textit{only} \text{ if } \boldsymbol{\Sigma} \models_{\mathbf{P}} \sigma$$

The set $\mathbf{A}$ is *complete* for $\mathbf{P}$ iff

$$\Sigma \vdash_{\mathbf{A}} \sigma \text{ if } \Sigma \models_{\mathbf{P}} \sigma .$$

Proposition 1: Axioms (1) are sound for PD (i.e., holds for all distributions).

The proof is achieved by induction on the length of a derivation.

Proposition 2 (After (Fagin, 1977)): A set of axioms A is complete iff for every set of statements $\boldsymbol{\Sigma}$ and every statement $\sigma \notin \mathbf{cl}_{\mathbf{A}}(\boldsymbol{\Sigma})$ there exists a distribution P_σ in $\mathbf{P}$ that satisfies

Σ and does not satisfy σ.

Proof: This is the counter-positive form of the completeness definition, if $\sigma \notin \mathbf{cl}_A(\Sigma)$ (i.e., $\Sigma \not\vdash_A \sigma$) then $\Sigma \not\models_P \sigma$. $\square$

Theorem 3 (Completeness): Let Σ be a set of statements, and let $\mathbf{cl}(\Sigma)$ be the closure of Σ under the following axioms:

Trivial Independence $\quad (X, \varnothing)$ (1.a)

Symmetry $\quad (X, Y) \rightarrow (Y, X)$ (1.b)

Decomposition $\quad (X, YW) \rightarrow (X, Y)$ (1.c)

Mixing $\quad (X, Y)\ \&\ (XY, W) \rightarrow (X, YW)$ (1.d)

Then for every statement $\sigma = (X, Y) \notin \mathbf{cl}(\Sigma)$ there exists a probability distribution P_σ that obeys all statements in $\mathbf{cl}(\Sigma)$ but does not obey σ.

Proof: Let $\sigma = (X, Y)$ be an arbitrary statement not in $\mathbf{cl}(\Sigma)$. Without loss of generality we assume that for all non-empty sets X' and Y' obeying $X' \subseteq X$, $Y' \subseteq Y$ and $X'Y' \neq XY$ we have $(X', Y') \in \Sigma$. A statement obeying this property, is called a *minimal dependency*. If $\sigma = (X, Y)$ is not a minimal dependency then we can always find a minimal dependency $\sigma' = (X', Y')$ not in $\mathbf{cl}(\Sigma)$, where $X' \subseteq X$ and $Y' \subseteq Y$, by deleting elements of X and Y until we obtain the desired property or until both X' and Y' become singletons, in which case, due to axiom (1.a), it is a minimal dependency. For each such σ', we can construct $P_{\sigma'}$ that satisfies $\mathbf{cl}(\Sigma)$ and violates σ'. Due the decomposition axiom (1.c), which holds for all distributions, we know that any distribution that violates σ', violates σ as well. In particular, $P_{\sigma'}$ violates σ (while obeying $\mathbf{cl}(\Sigma)$), and therefore satisfying the conditions of the theorem.

Let $\sigma = (X, Y)$ be a minimal dependency where X is over the variables $\{x_1, x_2 \cdots x_l\}$, Y is over the variables $\{y_1, y_2 \cdots y_m\}$, and all other variables appearing in statements of Σ are denoted by $Z = \{z_1, z_2 \cdots z_k\}$. Construct P_σ as follows: Let all variables except x_1, be independent identically distributed binary variables (i.e., fair coins) and let

$$x_1 = \sum_{i=2}^{l} x_i + \sum_{i=1}^{m} y_j \quad (mod 2).$$

Clearly, P_σ has the product form:

$$P_\sigma(X\ Y\ Z) = P_\sigma(X\ Y) \cdot \prod_{z_i \in Z} P_\sigma(z_i). \tag{3}$$

We first show that $\sigma = (X, Y)$ does not hold in P_σ. Instantiate x_1 to one and all other variables in XY to zero. For this assignment of values we have

$$P_\sigma(x_1 \cdots x_l, y_1 \cdots y_m) \neq P_\sigma(x_1 \cdots x_l) \cdot P_\sigma(y_1 \cdots y_m) \tag{4}$$

because the LHS of equation (4) is equal to 0 whereas the RHS consists of a product of two non-zero quantities.

It is left to show that every statement in **cl** (Σ) holds in P_σ, or equivalently, that for an arbitrary statement (U, V) we have:

$$(U, V) \in \mathbf{cl}(\Sigma) \Rightarrow P_\sigma(U, V) = P_\sigma(U) \cdot P_\sigma(V).$$

This is done by examining the statement (U, V) for every possible assignment of variables to the sets U and V and showing that either $P_\sigma(U, V) = P_\sigma(U) \cdot P_\sigma(V)$ or that $(U, V) \notin \Sigma$.

Case 1: Either U or V contain only elements of Z.

By equation (4), we get $P_\sigma(U, V) = P_\sigma(U) \cdot P_\sigma(V)$.

Case 2: Both U and V include an element of $X \cup Y$.

Case 2.1: $U \cup V$ does not include all the variables of $X \cup Y$.

To verify whether (U, V) holds in P_σ, amounts to checking this statement in the projection of P_σ on the set $U \cup V$. Since $U \cup V$ do not include all the variables of $X \cup Y$ this projection results a simple product of the form $\prod_{w_i \in U \cup V} P_\sigma(w_i)$.

Hence, again, $P_\sigma(U\ V) = P_\sigma(U) \cdot P_\sigma(V)$.

Case 2.2: $U \cup V$ include all elements of $X \cup Y$.

This is the only case for which (U, V) is definitely not in **Σ**.

Let $U = X'Y'Y'$, $V = X''Y''Z''$ where $X = X'X''$, $Y = Y'Y''$ and $Z' \cup Z'' \subseteq Z$. We continue by contradiction. Assume $(U, V) = (X'Y'Z', X''Y''Z'')$ belongs to **cl** (Σ). **cl** (Σ) is closed under decomposition. Therefore, $(X'Y', X''Y'') \in$ **Σ**. To reach a contradiction we show that this statement implies that σ must have been in **cl** (Σ), contradicting our selection of σ. The proof uses the mixing and symmetry axioms to infer $(X'X'', Y'Y'')$ (i.e σ) from $(X'Y', X''Y'')$ by "pushing" all the X's to one side and all Y's to the other side. We further assume that X', X'', Y', Y'' are non-empty sets. If some of these sets is empty, not all the derivations that follow need to be performed to reach the contradicting conclusion that $\sigma \in$ **cl** (Σ). The following is a derivation of σ.

First, (X', Y') belongs to $\mathbf{\Sigma}$ because (X, Y) is a minimal dependency. Due to the mixing axiom

$$(X', Y') \ \& \ (X'Y', X''Y'') \rightarrow (X', Y'X''Y'').$$

We conclude that $(X', X''Y) \in \mathbf{cl}(\Sigma)$. Due to symmetry $(XY'', X') \in \mathbf{cl}(\Sigma)$ as well. $(X'', Y) \in \mathbf{cl}(\Sigma)$ because σ is a minimal dependency and therefore (by symmetry) also (Y, X'') is a member of $\mathbf{cl}(\Sigma)$. Using the mixing axiom again, we get,

$$(Y, X'') \ \& \ (YX'', X') \rightarrow (Y, X'X'')$$

which leads to the conclusion that $(Y, X) \in \mathbf{cl}(\Sigma)$, and by symmetry that (X, Y) is in $\mathbf{\Sigma}$, contradiction. □

Theorem 3 implies that the problem of verifying whether a given statement σ logically follows from an arbitrary set of statements Σ is *decidable;* by applying axioms (1.a) through (1.d) successively, one could in a finite number of steps, generate all statements that logically follow from Σ. We note that the construction of P_σ uses binary variables, therefore, axioms (1.a) through (1.d) are complete also in PB, namely a statement is derivable iff every distribution in PB that satisfies Σ also satisfies σ.

Theorem 4 (Completeness in PN): (Geiger and Pearl, 1987) Let Σ be a set of statements, and let $\mathbf{cl}(\Sigma)$ be the closure of Σ under the following axioms:

Trivial Independence	$(X, \varnothing)$	(2.a)
Symmetry	$(X, Y) \rightarrow (Y, X)$	(2.b)

Decomposition $(X, YW) \rightarrow (X, Y)$ (2.c)

Composition $(X, Y) \;\&\; (X, W) \rightarrow (X, YW)$ (2.d)

Then there exists a normal distribution that satisfies all statements in **cl** (Σ) and none other.

Axioms (2) are stronger than axioms (1). Indeed, the additional knowledge about the distributions (being normal) usually results in additional independencies implied from a set of statements. For normal distributions the membership algorithm is trivial; (U, V) is logically implied from Σ iff for each $u_i \in U$ and $u_j \in V$ there exists a statement (X, Y) in Σ such that u_i appears in either X (or Y) and u_j appears in Y (or X); its complexity is $0(n^2 \cdot |\Sigma|)$. The next section provides a membership algorithm for axioms (1.a) through (1.d), having similar complexity.

3. THE MEMBERSHIP ALGORITHM

The following notation is employed. σ and γ denote single statements, Σ and Γ sets of statements, and s a set of elements (variables). $\gamma = (X, Y)$ is *trivial* if either X or Y is empty. span(γ) stands for the set of elements represented in a statement γ, and similarly, span(Γ) denotes the set of elements represented in all the statements of Γ e.g., span($\{(1, 2)\ (1, 3)\}$) is $\{1, 2, 3\}$. The *projection* of γ on s, denoted $\gamma(s)$, is the statement derived from γ by removing all elements not in s from γ e.g., if $\gamma = (123, 45)$ then $\gamma(123) = (123, \varnothing)$ and $\gamma(1346) = (13, 4)$. Similarly, the projection of Γ on s, stands for $\{ \gamma(s) \mid \gamma \in \Gamma \}$. The number of elements appearing in γ is denoted by $|\gamma|$ and is called the *size* of γ. The membership algorithm, presented below, uses the procedure Find to answer whether a statement σ is derivable from Σ by axioms (1.a) through (1.d).

Algorithm Membership

Procedure Find (Σ, σ):

1. $\Sigma' := \Sigma(\text{span}(\sigma))$ { Σ' is the projection of Σ on the variables of the target statement σ }

2. If σ is trivial, or σ (or its symmetric image) belongs to Σ' then set Find(Σ, σ) := True and return.

3. Else if for all nontrivial $\sigma' \in \Sigma'$, $\text{span}(\sigma') \neq \text{span}(\sigma)$ then set Find(Σ, σ) := False.

4. Else there exists a statement $\sigma' \in \Sigma'$ such that $\text{span}(\sigma') = \text{span}(\sigma)$, and up to symmetry, $\sigma' = (AP, BQ)$ and $\sigma = (AQ, BP)$ where one of the sets A, B, P, Q may be empty (If several such σ' exist, then choose one arbitrarily).

Set $\sigma_1 := (A, P)$, $\sigma_2 = (B, Q)$,

Find (Σ, σ) := Find $(\Sigma', \sigma_1) \wedge$ Find (Σ', σ_2).

return.

Begin {Membership}

Input(Σ, σ)

Print Find(Σ, σ)

End.

We will show first that the algorithm is correct and then prove its complexity.

Lemma 3: If $s \subseteq \text{span}(\Sigma)$ then $\Sigma(s) \subseteq \mathbf{cl}(\Sigma)$.

Proof: Each statement in $\Sigma(s)$ is derived by the decomposition and symmetry axioms, and therefore belongs to $\mathbf{cl}(\Sigma)$.

Lemma 4: If $\gamma_1 \ldots \gamma_k$ is a derivation chain in Σ (i.e. for $1 \le i \le k, \gamma_i$ is either in Σ or follows from previous γ_js by one of the axioms) and $s \subseteq \text{span}(\Sigma)$ then $\gamma_1(s), \ldots \gamma_k(s)$ is a derivation chain in $\Sigma(s)$ (due to Lemma 3 this derivation can be extended to a derivation in Σ).

Proof: Follows from the fact that the axioms are preserved under the projection operation. □

Lemmas 3 and 4 show that to derive a statement σ from Σ one may start, without loss of generality, by projecting all statements in Σ on the span of σ. This justifies Step 1 of the procedure used by the algorithm. Step 2 is due to lemma 1 and Step 3 stems from the fact that any application of axioms (1.b) through (1.d) does not increase the maximal span, therefore, if all statements in $\Sigma(s)$ have smaller span than σ, σ can not be derived, hence, Find(Σ,σ) is correctly set to False. Step 4 is justified by Lemma 5 and Theorem 6.

Lemma 5: Let $\sigma = (AQ, BP)$, $\sigma' = (AP, BQ)$, $\sigma_1 = (A, P)$, $\sigma_2 = (B, Q)$ be statements. If $\sigma' \in \mathbf{cl}(\Sigma)$ then $\sigma \in \mathbf{cl}(\Sigma)$ iff $\sigma_1 \in \mathbf{cl}(\Sigma_1)$ and $\sigma_2 \in \mathbf{cl}(\Sigma_2)$ where $\Sigma_i = \mathbf{cl}(\Sigma(\text{span}(\sigma_i)))$ (notice that σ, σ', σ_1 and σ_2 are defined as in Step 4 of procedure Find).

Proof: If $\sigma' \in \mathbf{cl}(\Sigma)$, $\sigma_i \in \mathbf{cl}(\Sigma_i)$ $i = 1, 2$, then σ can be derived as follows:

(i) $(AP, BQ) = \sigma'$; $(A, P) = \sigma_1$; $(B, Q) = \sigma_2$

(ii) (A, PBQ) :	Apply axiom (2.d) on σ_1 and σ'
(iii) (APB, Q) :	Apply axioms (1.b) and (1.d) on σ_2 and σ'
(iv) (PB, Q) :	Apply axioms (1.b) and (1.c) on (iii)
(v) $(AQ, BP) = \sigma$:	Apply axioms (1.b) and (1.d) on (ii) and (iv)

If $\sigma \in \mathbf{cl}(\Sigma)$ then let $\gamma_1, \gamma_2 ..., \gamma_k = \sigma = (AQ, BP)$ be a derivation chain for σ in Σ. Let $s = \text{span}((\sigma_1))$. Then $\gamma_1(s), \gamma_2(s), ... \gamma_k(s) = \sigma_1 = (A, P)$ is a derivation chain for σ_1 in Σ_1. Thus, $\sigma_1 \in \mathbf{cl}(\Sigma_1)$. Similarly, a derivation chain for σ_2 can be constructed. □

Lemma 5 shows that the selection of σ' in Step 4 can be made arbitrarily because any selection provides a necessary and sufficient means to check whether σ belongs to $\mathbf{cl}(\Sigma)$.

Theorem 6: The procedure in the algorithm halts and when it halts

$$\text{Find}(\Sigma, \sigma) = \text{true} \ \text{ iff } \ \sigma \in \mathbf{cl}(\Sigma)\,.$$

Proof: Every time the algorithm passes through Step 4 the size of the statements involved strictly decrease. If it did not halt before, it will halt when the size of the two statements have reached the value 2 (at Step 2 or 3). We show correctness by induction on the size of σ. If $|\sigma| = 1$ then σ is trivial, $\sigma \in \mathbf{cl}(\Sigma)$ and Find $(\Sigma, \sigma) =$ true. If $|\sigma| = 2$ then $\sigma \in \mathbf{cl}(\Sigma)$ iff Find $(\Sigma, \sigma) =$ true as follows from Steps 2 and 3 of the algorithm.

Assume that the Lemma holds for all $|\gamma| < k$ and let σ be a statement such that $|\sigma| = k$ and $\sigma = (AQ, BP)$. Then Find $(\Sigma, \sigma) =$ true iff (by the definition of Step 4) Find $(\Sigma', \sigma_1) =$ true and Find $(\Sigma', \sigma_2) =$ true iff (by the definition of Step 1) Find $(\Sigma_1, \sigma_1) =$ true and Find $(\Sigma_2, \sigma_2) =$ true, where $\Sigma_i = \Sigma(\text{span}(\sigma_i))$ respectively, iff (by induction) $\sigma_i \in \mathbf{cl}(\Sigma_i)$ $i = 1,2$ iff (by Lemma 5) $\sigma \in \mathbf{cl}(\Sigma)$. □

Next we analyze the time complexity. We measure the complexity in terms of basic operations of two types: comparison of two statements and a projection of a statement. Both operations are bounded by n, the number of distinct variables in $\mathbf{\Sigma} \cup \{\sigma\}$. Let $Cost(k)$ be the number of basic operations needed to solve a size k problem where $k = |\sigma|$ and assume (initially) that span(σ) = span(Σ). By Step 4, $Cost(k)$ must satisfy the following equation:

$$Cost(k) \le Cost(k_1) + Cost(k_2) + |\Sigma|$$

where $k_1 + k_2 = k$, $k_1 = |\sigma_1|$ and $k_2 = |\sigma_2|$

The solution to this equation is $O(|\Sigma| \cdot k)$ measured in basic operations. Adding the cost of projecting $\mathbf{\Sigma}$ over the variables of σ, $O(|\Sigma| \cdot n)$, yields the theorem below.

Theorem 8: The complexity of the Membership algorithm is $O(|\Sigma| \cdot k^2 + |\Sigma| \cdot n)$ (Which is $O(|\Sigma| \cdot n^2)$ since $k \le n$).

Remarks:

1. It is reasonable to assume that the bound is pessimistic at least in its $|\Sigma|$ part, since as the algorithm proceeds the number of statements in $\mathbf{\Sigma}'$ decreases.

2. The algorithm can be slightly modified so as to produce a derivation chain for σ if $\sigma \in$ **cl** (Σ), whose length is $0(k)$.

3. The algorithm can be expanded into a polynomial algorithm (provided that $|\Sigma|$ is polynomial) for the following problems:

   a. Given $\mathbf{\Sigma}$ and $\mathbf{\Gamma}$, is **cl** (Σ) = **cl** (Γ), or is **cl** (Σ) $\subset$ **cl** (Γ)?

   b. Minimize the size of Σ while preserving **cl** (Σ):

To solve problem b start with 2 "independent" maximal size statements in Σ and probe all other statements in some non-increasing (as to size) order. Any statement which is found "dependent" on the set of previously probed statements is deleted.

4. EXTENSIONS

In this section we show that axioms (1.a) through (1.d) characterize the independence relation in the sense that any binary relation obeying these axioms is induced by some distribution P. The theorem below formalizes this assertion.

Theorem 9 (Strong Completeness): For every set of statements Σ closed under axioms (1.a) through (1.d) there exists a distribution P such that

$$\sigma \text{ holds for } P \quad \textit{iff} \quad \sigma \in \Sigma$$

Proof: Assume an operation $\otimes$ that maps finite sets of distributions $\{ P_i \mid i = 1..n \}$ into a single distribution P such that

$$\sigma \text{ holds for } P \quad \textit{iff} \quad \sigma \text{ holds for each } P_i,\ i = 1..n. \tag{4}$$

Using $\otimes$, the proof follows. By Theorem 2, for each $\sigma \notin \Sigma$ there exists a distribution P_σ that obeys Σ and does not obey σ. Let $P = \otimes \{ P_\sigma \mid \sigma \notin \Sigma \}$. Due to equation (4), P satisfies all statements in Σ and none other, hence P satisfies the requirements of the theorem.

We shall construct the operation $\otimes$ using a binary operation $\otimes'$ such that if $P = P_1 \otimes' P_2$ then for every independency statement σ we get

$$\sigma \text{ holds for } \otimes' P_i \quad \textit{iff} \quad \sigma \text{ holds for } P_1 \text{ and for } P_2 . \tag{5}$$

The operation $\otimes$ is recursively defined in terms of $\otimes'$ as follows:

$$\otimes \{ P_i \mid i=1..n \} = ((P_1 \otimes' P_2) \otimes' P_3) \otimes' \cdots P_n).$$

Clearly, if $\otimes'$ satisfies equation (5), then $\otimes$ satisfies equation (4). Therefore, it suffices to show that $\otimes'$ satisfies (5).

Let P_1 and P_2 be two distributions sharing the variables $x_1, \cdots, x_n$. Let $A_1, \cdots, A_n$ be the domains of $x_1, \cdots, x_n$ in P_1 and let an instantiation of these variables be $\alpha_1, \cdots, \alpha_n$. Similarly, let $B_1, \cdots, B_n$ be the domains of $x_1, \cdots, x_n$ in P_2 and $\beta_1, \cdots, \beta_n$ an instantiation of these variables. Let the domain of $P = P_1 \otimes' P_2$ be the product domain $A_1 B_1, \cdots, A_n B_n$ and denote an instantiation of the variables of P by $\alpha_1\beta_1, \cdots, \alpha_n\beta_n$. Define $P_1 \otimes' P_2$ by the following equation:

$$P(\alpha_1\beta_1, \alpha_2\beta_2, \cdots, \alpha_n\beta_n) = P_1(\alpha_1, \alpha_2, \cdots, \alpha_n) \cdot P_2(\beta_1, \beta_2, \cdots, \beta_n).$$

The proof that P satisfies equation (5) uses only the definition of independence and can be found in (Geiger and Pearl, 1987). □

Theorem 9 was called *strong completeness* in (Beeri et al, 1977). Equivalent definitions for strong completeness are given in (Fagin, 1977; Beeri et al, 1977) and the construction of $\otimes$ for database dependencies (called *Armstrong relation) is given* in (Fagin, 1982). The immediate consequence of these theorems is that axioms (1.a) through (1.d) are powerful to derive all disjunctions of independency statements that are logically implied from a given set of disjunctions and not merely single statements as advertized in section 3. (Geiger and Pearl, 1987) extend these results to other probabilistic independence relations, most notably, conditional independence.

An immediate application of strong completeness is the reduction of the *consistency problem* to a set of membership problems.

Definition: A set of dependencies Σ^- and a set of independencies Σ^+ are consistent iff there exists a distribution that satisfies $\Sigma^+ \cup \Sigma^-$. The task of deciding whether a set is consistent is called the *consistency problem.*

The following algorithm answers whether $\Sigma^+ \cup \Sigma^-$ is consistent: For each member of Σ^- determine, using the membership algorithm, whether its negation logically follows from Σ^+. If the answer is negative for all members of Σ^-, then the two sets are consistent, otherwise they are inconsistent.

The correctness of the algorithm stems from the fact that if the negation of each member σ of Σ^- does not follow from Σ^+ i.e, each member of Σ^- is individually consistent with Σ^+, then there is a distribution P_σ that realizes Σ^+ and $\neg\sigma$. The distribution $P = \otimes\{P_\sigma \mid \neg\sigma \in \Sigma^-\}$ then realizes both Σ^+ and Σ^-, therefore the algorithm correctly identifies that the sets are consistent. In the other direction, namely when the algorithm detects an inconsistent member of Σ^-, the decision is obviously correct.

ACKNOWLEDGEMENT

We thank Ron Fagin for pointing out the usefulness of the notion of Armstrong relation and to Thomas Verma and Norman Dalkey for many discussions on the properties of independence.

REFERENCES

BEERI, C., FAGIN, R., AND HOWARD, J.H. (1977), A complete axiomatization of functional dependencies and multi valued dependencies in database relations. *in* "Proc. 1977 ACM SIGMOD Int. Conf. on Mgmt of Data, Toronto, Canada," pp. 47-61.

BEERI, C. (1980), On the membership problem for functional and multivalued dependencies in relational database. *ACM Trans. Database Syst.* Vol. 5:3 pp. 241-249.

FAGIN, R. (1977), Functional dependencies in a relational database and propositional logic. *IBM J. Res. Dev.* 21:6 pp. 534-544.

FAGIN, R. (1982), Horn clauses and database dependencies. *JACM,* Vol 29:4, pp. 952-985.

FAGIN, R., AND VARDI, M. (1986), The theory of data dependencies - a survey. *in* "Proc. Symp. in Applied Math.," Vol. 34.

GEIGER, D., VERMA, T., AND PEARL, J. (1989), Identifying independence in Bayesian networks. UCLA Cognitive Systems Laboratory, Technical Report (R-116).

GEIGER, D., AND PEARL, J. (1987), Logical and algorithmic properties of conditional independence. UCLA Cognitive Systems Laboratory, Technical Report 870056 (R-97), August 1987. A shorter version (R-97-I) in "*Proc.* 2nd International Workshop on Artificial Intelligence and Statistics", January 1989.

PEARL, J. (1988), "Probabilistic reasoning in intelligent systems: Networks of plausible inference," Morgan Kaufmann, San Mateo.

PEARL, J., GEIGER, D., AND VERMA, T. (1988), The Logic of Influence Diagrams. *in* "Proc. of the Berkeley Conference on Influence Diagrams, May 9-11, 1988", New York: John Wiley & Sons Ltd., 1989.

RISSANEN, J. (1978), Theory of relations for databases - a tutorial survey. *in* "Proc. 7th Symp. on Math. Found. of Comp. Science." Lecture Notes in Computer Sciences - Vol. 64, Springer-Verlag, 537-551.

SAGIV, Y., AND WALECKA, S. (1982), Subset dependencies and completeness results for a subset of *EMVD*. *JACM,* Vol. 29:1, pp. 103-117.

VARDI, M. (1988), Fundamentals of Dependency Theory. *Trends in Theoretical Computer Science* (E. Borger Ed.), pp. 171-224, Computer Science Press, Rockville, Md.