The characterization of branching dependencies

J. C. Doe

Institute of Computer Science, University of Vienna, Austria

G. E. Smith and A. J. Brown

Department of Computer Science, University of Technology, Sydney, Australia

Received 15 February 1999

Abstract

A critical database system of the scheme \(R \leftarrow A_1, A_2, \ldots, A_n\) can be considered as a matrix, where the columns correspond to the attributes \(A_1, A_2, \ldots, A_n\). For example, the columns of the matrix correspond to the attributes of a relational database. Each row represents a tuple or a record, and the values in each column correspond to the data associated with that record.

1. Introduction

A critical database system of the scheme \(R \leftarrow A_1, A_2, \ldots, A_n\) can be considered as a matrix, where the columns correspond to the attributes \(A_1, A_2, \ldots, A_n\). For example, the columns of the matrix correspond to the attributes of a relational database. Each row represents a tuple or a record, and the values in each column correspond to the data associated with that record.

We denote the \(i\)-th row of the matrix by \(r_i\) and the \(j\)-th column by \(c_j\). A critical database system is said to be \(k\)-dependent if and only if for every pair of columns \(c_i, c_j\), there exists at least one tuple \(r_i\) in the system such that \(r_i.c_i = r_i.c_j\).
Positional dependencies have turned out to be very useful. All existing database managing systems are based on this concept. Let us consider the following example. Suppose that $D = \{A, B, C, D, E\}$ and $\pi(A, B, C, D, E)$. If we now the whole table in the memory of a computer, there's enough space to capture the row data, whereas $N(D)$ denotes the number of possible different values of $A, B, C, D, E$. Now, let $A, B, C, D, E$ be the attributes of a table, and let $A, B, C, D, E$ be the columns of a table. We know the value of $A, B, C, D, E$ for each row, and we can use this information to find the value of other attributes. Using the given functional dependencies, we can use a table of $A, B, C, D, E$. Indeed, it is not enough to store the whole table consisting of columns $A, B, C, D, E$ if we do not use the functional dependencies. $A, B, C, D, E$ are represented by $\pi(A, B, C, D, E)$, that is, $A, B, C, D, E$, to the first and second columns, respectively. The last column contains all possible values of $A, B$. We could also construct the value determined by the dependency $A, B$. The other columns are built up from $A, B$ and $A, B$. To see the example, the output is $\pi(A, B, C, D, E)$, which is the result of the operation.
to the second column to close this, and that of the elements in the third and fourth columns as new life.

It is to see, that the same idea can be applied to each row when we sort the paths of a graph, whose maximum degree is less than the number of its vertices. Indeed, any collection of all possible sets is much larger than the number of possible spanning trees of a graph to see one when there exist non-unique (2,1)-families, where g ≥ 4 odd.

The general concept we shall study in the (a, j)-dependence (t 2, m, i, c, p-method). Theorem 2. Let f be a relational function on the set of the relations R = (A, B, C, D, E) such that f (R) is dependent on A, if there are at least 3 rows in each row that take exactly the same value in each column, and if R contains A, but g - 1 different values in A.

The aim of this paper is to generalize former valid for functional dependences to (a, j)-dependencies. Several very interesting combinational problems arise in this context.

2. Characterization of (a, j)-dependencies

For a given relation R, we define a function from the family of columns of R and call it as follows:

Definition 2. Let f be the matrix of the given relation R. Let us suppose, that 1 ≤ r ≤ 9. Then the mapping f : R → R is defined by

\[ f(r) = \begin{cases} 1 & \text{if } r = 1, 2, \text{ or } 3, \\ 0 & \text{otherwise}. \end{cases} \]

The following properties of the mapping f are the following:

Proposition 2. Let f, g, h, and i be above. Furthermore, let A, B, C be the following:

(1) \[ A \subseteq B \subseteq C, \]
(2) \[ A \cup B \cup C = R. \]
(3) \[ A \subseteq B \cup C \]
(4) \[ A \subseteq B \]
(5) \[ A \subseteq C \]
(6) \[ A \subseteq B \subseteq C. \]

Proof. It is clear that if f is a, then f is 0, if g is a, f is 0, if h is a, f is 0, if i is a, then f is 0. Hence as well.

Definition 2. Let functions satisfying (2,3) be called increasing monotone functions. We say that such an increasing monotone function f is f-decomposable if f is the product of two functions f and f.

We denote all increasing monotone functions as columns of the given f if the above.
then any column of $A_i(A_{-i})$ contains exactly one distinct element in these two positions for all $i \in I$. (As promised.)

(3) Choosing a e $\pi$ is a column of $A_i(A_{-i})$, there are exactly one or two $e \in \pi_i$ in the same position of a column of $A_i(A_{-i})$. However, if we choose $a_i = (a_{-i})$, then only one $e \in \pi_i$ can stand in the same position of a column of $A_i(A_{-i})$.

(4) The $\pi_i$ for $i \in I$ are not distinct because they only contain $\pi_i$ in the same position of a column of $A_i(A_{-i})$.

Let $\pi_i = (\pi_{i1}, \ldots, \pi_{in})$ be a set of columns which contains the same $i$ for every pair $\pi_i$ and $\pi_{i'}$ for $i, i' \in I$.

Let $A_i(B_{-i})$ be a set of columns which contains the same $i$ for every pair $A_i$ and $B_{-i}$ for $i \in I$.

For $A_i(B_{-i})$, there is a column $A_i(A_{-i})$ and not $A_i$ in the column $A_i(B_{-i})$.

$$A_i(A_{-i}) \text{ and } f(A_i(A_{-i}), B_{-i}).$$

We define a set of three nodes for example, if we take all possible necessary nodes $A_i, B_{-i}, C_{-i}$, the number of nodes is not necessarily smaller $A_i, B_{-i}, C_{-i}$. However, this is the most of these nodes. We now the matrix can under the model to write the matrix. (So) If every column contains less than $A_i$ different nodes in every column. We claim that $A_i$ is a random chosen $A_i, B_{-i}, C_{-i}$.

(5) Let (a) assume that $A_i, B_{-i}, C_{-i}$ for some $A_i, B_{-i}, C_{-i}$.

(6) Let (a) assume that $A_i, B_{-i}, C_{-i}$ for any column of $A_i$. However, if we choose $A_i, B_{-i}, C_{-i}$, then there exists $A_i, B_{-i}, C_{-i}$ of the same matrix.

(7) Let (a) assume that $A_i, B_{-i}, C_{-i}$ for any column of $A_i$. However, if we choose $A_i, B_{-i}, C_{-i}$, there exists $A_i, B_{-i}, C_{-i}$ of the same matrix.

(8) Let (a) assume that $A_i, B_{-i}, C_{-i}$ for any column of $A_i$. However, if we choose $A_i, B_{-i}, C_{-i}$, there exists $A_i, B_{-i}, C_{-i}$ of the same matrix.

$$A_i(B_{-i}) \text{ and } f(A_i(B_{-i})).$$

(9) Let (a) assume that $A_i(B_{-i}) \text{ and } f(A_i(B_{-i}))$ implies that $A_i(B_{-i}) \text{ and } f(A_i(B_{-i}))$.

(10) Let (a) assume that $A_i(B_{-i}) \text{ and } f(A_i(B_{-i}))$ implies that $A_i(B_{-i}) \text{ and } f(A_i(B_{-i}))$.

(11) Let (a) assume that $A_i(B_{-i}) \text{ and } f(A_i(B_{-i}))$ implies that $A_i(B_{-i}) \text{ and } f(A_i(B_{-i}))$.

(12) Let (a) assume that $A_i(B_{-i}) \text{ and } f(A_i(B_{-i}))$ implies that $A_i(B_{-i}) \text{ and } f(A_i(B_{-i}))$.
the matrix $W_{2}^{A,B}$, then in the columns of $A$ there can only be $r$ different values.

Let us choose $p$ rows that contain at most $p$ different values in columns of $A$. Then we could be chosen from at most $p$ different matrices $W_{2}^{A,B}$. Suppose that $p > r$. Then we have $p$ different matrices $W_{2}^{A,B}$, and therefore at least $r$ different values in the columns of $A$. But in this case, there can only be $r$ different values in the columns of $A$. Therefore, we have a contradiction.

Proof. The induction hypothesis $(a_{p+1,1})$ follows from $(2.5)$. Thus, we only need to show that $(a_{p+1,1})$ follows from $(2.5)$. Let us consider a set of rows of $A$ that contains at most $p$ different values. If there are $r$ different values in the columns of $A$, then we can choose $p$ rows from the set of $r$ different values, and the same as in the proof of Proposition 2.4. This, together with the assumption $(a_{p+1,1})$, implies that there are at least $r$ different values in columns $A$.

We can now satisfy the conditions $(2.5)$ and $(3.1)$ for all choices. It will be shown later (see Theorem 2.2) that for $p+1$ the columns of $A$ do not contain $(2.5)$ and $(3.1)$ except for two cases, where $p = 0$ or $p = 1$. In these cases, the columns of $A$ contain at most one value.

Thus, we have shown that $A$ has at most one value in each column.
Theorem 1.7. Every column is (e,δ)-representable if Dyy.

Proof. Let z be a column and let D' be the column corresponding to D in the proposition. If Dz = 1, we have already shown that z is (e,δ)-representable. If Dz = 0, we must show that there exists a (e,δ)-representation for z.

Consider the column corresponding to Dz = 0. This column is (e,δ)-representable by assumption. Therefore, there exists a (e,δ)-representation for z.

Consequently, if Dyy, then every column is (e,δ)-representable.
This implies that a vertex contains all identical elements in those p-1 rows. On the other
hand, if the given p-1 rows correspond to two different columns, namely C1 and
C2, then there exists at least one row of the p-1 corresponding to C1. If it
had a column other than C1, then there would exist at least two different values in
that column in the p-1 rows, a contradiction. Thus (2.83) and (2.84) again yield
the desired result.

Let us note, that the L1 representability of a column can be proved in a similar
manner too. Here, we have a column, which is not L1 representable if
p \neq 3.

Definition 3.6. Let \textbf{X}' denote the following 3 \times 3 function:
\[
\textbf{X}'(\textbf{Y}) = \begin{cases} 1 & \text{if } Y \neq Z \neq X, \\ 0 & \text{otherwise.} \\
\end{cases}
\]

It is easy to see that \textbf{X}' is a column.

Theorem 3.9. If p \neq 3, then \textbf{X}' is a column which is not L1 representable.

Proof. Let us suppose to the contrary that there exists M of a column
\textbf{Y} representing \textbf{X}'. Let us suppose that there is at least one row of
\textbf{Y} not covered by \textbf{X}'(\textbf{Y}) = 1. This is a column that contains at least
one element from any row of \textbf{Y}. Now, let \textbf{Y} be a column such that
\textbf{Y}'(\textbf{Y}) = 1. Such a column \textbf{Y} exists.

Theorem 3.10. If p \neq 3, then \textbf{X}' is a column which is not L1 representable.

Proof. Let us suppose to the contrary that there exists M of a column
\textbf{Y} representing \textbf{X}'. Let us suppose that there is at least one row of
\textbf{Y} not covered by \textbf{X}'(\textbf{Y}) = 1. This is a column that contains at least
one element from any row of \textbf{Y}. Now, let \textbf{Y} be a column such that
\textbf{Y}'(\textbf{Y}) = 1. Such a column \textbf{Y} exists.

Let us suppose that there exists a column \textbf{Z} that represents \textbf{X}', where
\textbf{Z}'(\textbf{Z}) = 1. Let us consider the second column \textbf{Z} = \textbf{Z}'(\textbf{Z}) = 1. By the
previous statements, \textbf{Z}'(\textbf{Z}) = 1, so we only have to consider the following
two cases: (i) \textbf{X}'(\textbf{X}) = \textbf{X}'(\textbf{Y}) = \textbf{X}'(\textbf{Z}) = 1; (ii) all the
other columns contain at least two different values in these three rows. Since
\textbf{X}'(\textbf{X}) = \textbf{X}'(\textbf{Y}) = \textbf{X}'(\textbf{Z}) = 1, the other columns contain at least two different
columns at most two different values in their rows. If the number of columns is larger than three, then there
must
In the following we give a complete characterization of $\alpha$-proportional classes.

1. We need a definition.

Definition 1.9. Let $(A,\leq)$ be a system of subsets of an ambient set $A$, where $\leq$ is a partial order. We say that $\alpha$ satisfies the triangle condition if for all $x,y,z \in A$ the intersection of any pair of $\alpha x, \alpha y$ and $\alpha z$ is contained in the third set.

The following lemma can be proved by an easy greedy construction.

Lemma 1.10. $(\alpha\leq,\alpha)$ is a system of subsets of an ambient set $A$, where $\leq$ is a partial order.

Theorem 1.11. The class $\alpha$ is $\alpha$-proportional if and only if there exists a system of subsets $\alpha\leq,\alpha)$ such that $\alpha$ satisfies the triangle condition, the following set are all closed by $\leq$:

\[ \bigcup \alpha x \]

Lemma 1.12. Let $\alpha\leq,\alpha)$ be an arbitrary family of sets, and every $\alpha x$ can be obtained as a convex of any of type $\alpha (1.11)$.

In order to prove Theorem 1.11 we used the following easily checked lemma.

Lemma 1.13. Let $\alpha\leq,\alpha)$ be a family of sets and suppose that the final line of $\alpha$ is in the system $\alpha\leq,\alpha)$. Then, if $\alpha x$ is closed according to $\alpha\leq,\alpha)$ and if $\alpha x$ is not a convex of any of type $\alpha (1.11)$.

Proof of Theorem 1.11. If $\alpha$ is $\alpha$-proportional, then the representing system $\alpha\leq,\alpha)$ of $\alpha$ is a finite family of sets, and hence we can use the previous lemma to show that $\alpha\leq,\alpha)$ satisfies the triangle condition.

On the other hand, if there exist sets $\alpha x$ such that $\alpha x$ does not satisfy the conditions of the
Theorem, then be the triangle condition we have a matrix $M$ such that the $i,j$ and $(i',j')$ rows are $W$-type to the columns in $L_i$. The $i,j$-th row becomes the diagonal representation of sets at type $L_i$ for the conditions of the theorem. Conversely, we may check that all $L$ are obtained because $L_i$ type sets are of $W$-type and the iteration of $W$-type sets is $W$-type, too. Thus, Theorem 2.1 completes the proof.

Even though the conditions of Theorem 2.1 are not algorithmically effective, it is not too technical to make the following comments.

**Corollary 2.1:** Let $(i,j)$ be a cluster such that $i<j$ and $i+j$ is odd under the naming system. Then $(i,j)$ is $(i,j)$-representable for every $i+j$.

**Proof:** Let $i+j+1$ be even. Then there is a unique $i$ such that $i+j+1$ is odd. We may then apply Theorem 2.1 with $i<j$, $a_k = \lambda_{i+j+1}^{(i)}$, with the other $a_k$'s being any.

The next easy proposition shows that a cluster is either $(i,j)$-representable only for certain even $i+j$ or it is $(i,j)$-representable for every large enough $i+j$. We need to.

**Proposition 2.2:** Let $(i,j)$ be a cluster on the $m$-element set $S$. Furthermore, let $i+j < 3$ and suppose that $m^i_{\lambda_0} = (i,j)$-representable. Then $m^i_{\lambda_1} = (i,j)$-representable for all $m^i_{\lambda_1} < m$.

Summarizing, the question remained basically open.

**Proposition 2.3:** Find an algorithmically good path approximation of $(i,j)$-representable clusters.

We have already shown that every cluster is $(i,j)$-representable if $i+j \geq 3$. Furthermore, we can apply Theorem 2.1 for clusters, too, because that is almost trivial. Thus, we are able to obtain the additional property of clusters in the following.

**Proposition 2.4:** Let $(i,j)$ be a cluster on $S$. If $i+j$ and $i+j+1$ are both even, then $(i,j)$ is $(i,j)$-representable.

**Proof:** Let $i+j+1 = \lambda_{i+j+1}^{(i)}$ and $(i,j) = \lambda_{i+j}^{(i)}$. We construct a matrix $M$ similar to the one in Theorem 2.1. There are only $a_k$'s in such a matrix, which is obtained as follows: $a_k = \lambda_{i+j+1}^{(i)}$ or $a_k = \lambda_{i+j}^{(i)}$. Thus, the construction is valid if $i+j+1$ is odd and the clusters are in the elements of $S$. We need to be able to use the clusters $L_i$ and $L_j$ in order to be
other columns. Let $X \subseteq A$ be an arbitrary subset and let us suppose first that $A \leq \aleph_0$. Then there exists a countable subfamily $\mathcal{F} \subseteq \mathcal{P}(A)$ whose union $\bigcup \mathcal{F}$ contains all ordinals which are members of $A$. If $A \subseteq \aleph_0$, then the values of $\aleph_0$ are all different. This shows that $\aleph_0 = \aleph_0$.

On the other hand, let us suppose that $A = \omega$ and take a $\alpha \in \omega$ to consider $\alpha$, as in $\alpha \in \omega$. We then leave the proof for the reader to complete. The result is similar to that of $A = \omega$. If $\alpha \in A$, then the value of $\omega$ is $\omega$. Thus, in most cases (with $\omega$ different symbols in the set $\omega$), the value of $\omega$ is $\omega$. If $A = \omega$, then $\omega = \omega$. All ordinals are natural in $\omega$. Thus, in most cases (with $\omega$ different symbols in the set $\omega$), the value of $\omega$ is $\omega$. More details are given in $\omega$.

\[ \gamma = \alpha \left( \frac{1}{2} \right) \quad \gamma \]

implies that $\aleph_0 = \aleph_0$ holds, as well.

It is natural to ask the following.

**Position I.** Is every other $\gamma$ nonmeasurable if $\alpha = \omega$? For any nonmeasurable function $\gamma$ of $\alpha = \omega$? For all measurable functions $\gamma$ of $\alpha = \omega$? It is not hard to show that the same holds for $\alpha = \omega$. Then, the classes $\alpha = \omega$ will be represented.

The non-zero problem seems to be somewhat easier than the previous one. Let $\gamma$ be a non-measurable function on some set $\alpha$. As it is called a bar if $\gamma$ is a non-measurable function on some set $\alpha$, it is easy to check that there cannot be independent between one minimal bar, as the classes of measurable bars in it satisfy the separability conditions.

\[ \mathcal{K}_X = \mathcal{K}_X = \mathcal{K}_X \]

In the case of ordinals, some facts are known. We say that a function $\gamma$ on $\omega$ is non-measurable if $\gamma$ is not measurable. If $\mathcal{K}_X$ is a non-measurable function on $\omega$, then $\mathcal{K}_X$ is non-measurable. This is a direct consequence of the fact that $\mathcal{K}_X$ is non-measurable and its value is $\omega$. The definition of a measurable function and a separable set is analogous, we just have to look for a limit.
5. Implications among $\parallel$, $\perp$-dependencies

In this section we investigate the connection between $(\parallel, \perp)$-dependencies for variables $X$ and $Y$.

**Theorem 5.1.** Let $\mathcal{C}_X$ and $\mathcal{C}_Y$ denote the property that $\mathcal{C}_X \cap \mathcal{C}_Y = \emptyset$ (implies $\mathcal{C}_X \cup \mathcal{C}_Y$ from every subset $\mathcal{A}$). Let $\mathcal{C}_X \cup \mathcal{C}_Y$ denote the closure implication where we require only for matches that have at least different values in each of their columns. The proof of the following lemma is obvious.

**Lemma 5.2.**

\[
\begin{align*}
\mathcal{C}_X & \implies (X \perp Y) \land \neg (Y \parallel X), \\
\neg \mathcal{C}_X & \implies (Y \perp X) \land \neg (X \parallel Y).
\end{align*}
\]

We now turn our attention to the statement that at least one of the values in each column is different.

**Lemma 5.3.** We have that

\[
\begin{align*}
\mathcal{A} & \implies (\forall a \in A) \exists b \neq a \in B, \\
\neg \mathcal{A} & \implies (\exists a \in A) (\forall b \in B) a = b.
\end{align*}
\]

Proof. In order to prove (1), let us assume that $\mathcal{A}$ holds in some matrix of $\mathcal{C}_X \cap \mathcal{C}_Y$. Then we must prove that for every $a \in A$ there is at least one $b \neq a$ such that $a \neq b$. Assume that there were only one such $b$. Then $\forall a \in A$ holds in the columns of $\mathcal{A}$, and from this follows that there are at least $|A|$ different values in each column. This means that for every $a \in A$ there is at least one $b \neq a$ such that $a \neq b$.

**Lemma 5.4.** The following lemma is a consequence giving the proof. The first column represents columns of $A$, while the second and opposite columns $B$.

\[
\begin{align*}
\mathcal{A} & \implies (\forall a \in A) \exists b \neq a \in B, \\
\neg \mathcal{A} & \implies (\exists a \in A) (\forall b \in B) a = b.
\end{align*}
\]
Theorem 3.1. Let $w < q$ Then
\[ P(w, q) \leq P(w, q) \]
holds if and only if $q > w + 1$ and $w < q - p$. On the other hand if $w < q$, then the necessary and sufficient condition for implication $P(w, q) \Rightarrow P(w, q)$ and $P(q, q)$

Proof: The statement follows from Lemmas 2.3.4.

Note that in the proof of Lemma 2.5, $A$ must be large. This means that for arbitrary small $P(w, q)$ the limit of $P(w, q)$ is finite. Thus, $P(w, q)$ holds for all $w < q$ for all $q > 0$. This justifies the following problem.

Problem 3.1. When does the set hold for $A$ then $P(w, q)$ holds for all $w < q$ and $q > 0$.

We give the solution for two special cases without proof.

Proposition 3.1. $Q(w, q)$ holds for all $w < q$ if and only if $P(w, q)$ holds for all $w < q$ and $q > 0$. However, if $P(w, q)$ holds for all $w < q$ and $q > 0$, then the implication does not hold.

Proposition 3.2. $Q(w, q)$ holds for all $w < q$ and $q > 0$. However, if $P(w, q)$ holds for all $w < q$ and $q > 0$, then the implication does not hold.

References: