On Inferences of Full Hierarchical Dependencies

Sven Hartmann **Sebastian Link**[*]

Department of Information Systems, Information Science Research Centre
Massey University, Palmerston North, New Zealand
E-mail: [S.Hartmann,S.Link]@massey.ac.nz

Abstract

Full hierarchical dependencies (FHDs) constitute a large class of relational dependencies. A relation exhibits an FHD precisely when it can be decomposed into *at least* two of its projections without loss of information. Therefore, FHDs generalise multivalued dependencies (MVDs) in which case the number of these projections is precisely two. The implication of FHDs has been defined in the context of some fixed finite universe.

This paper identifies a sound and complete set of inference rules for the implication of FHDs. This axiomatisation is very reminiscent of that for MVDs. Then, an alternative notion of FHD implication is introduced in which the underlying set of attributes is left undetermined. The main result proposes a finite axiomatisation for FHD implication in undetermined universes. Moreover, the result clarifies the role of the complementation rule as a mere means of database normalisation. In fact, an axiomatisation for FHD implication in fixed universes is proposed which allows to infer any FHDs either without using the complementation rule at all or only in the very last step of the inference. This also characterises the expressiveness of an incomplete set of inference rules in fixed universes. The results extend previous work on MVDs by Biskup.

1 Introduction

Relational databases still form the core of most database management systems, even after more than three decades following their introduction in (Codd 1970). The relational model organises data into a collection of relations. These structures permit the storage of inconsistent data, inconsistent in a semantic sense. Since this is not acceptable additional assertions, called dependencies, are formulated that every database is compelled to obey. There are many different classes of dependencies which can be utilised for improving the representation of the target database (Fagin & Vardi 1986, Thalheim 1991, Vardi 1987).

Multivalued dependencies (MVDs) (Fagin 1977, Zaniolo 1976) are an important class of dependencies. A relation exhibits an MVD precisely when it is decomposable into exactly two of its projections without loss of information (Fagin 1977). This property is fundamental to relational database design, in particular 4NF (Fagin 1977), and a lot of research has been devoted to studying the behaviour of these dependencies. Recently, extensions of multivalued dependencies have been found very useful for various design problems in advanced data models such as the nested relational data model (Fischer, Saxton, Thomas & Van Gucht 1985), the Entity-Relationship model (Thalheim 2000), data models that support nested lists (Hartmann & Link 2004, Hartmann, Link & Schewe 2006) and XML (Vincent & Liu 2003, Vincent, Liu & Liu 2003).

Full hierarchical dependencies (FHDs) (Delobel 1978) constitute a large class of relational dependencies that subsume MVDs. A relation exhibits an FHD precisely when it is the natural join over at least two of its projections. The classical notion of an FHD (Delobel 1978) is dependent on the underlying universe R. For MVDs (Fagin 1977) their dependence on the relation schema R is syntactically reflected by the R-complementation rule which is part of the axiomatisation of MVDs (Beeri, Fagin & Howard 1977). The R-complementation rule is special in the sense that it is the only inference rule which is dependent on R. Further research on this fact has led to an alternative notion of semantic implication in which the underlying universe is left undetermined (Biskup 1980). In the same paper Biskup shows that this notion can be captured syntactically by a sound and complete set of inference rules, denoted by $\mathfrak{S}_0$. If $R\mathfrak{S}_0$ results from adding the R-complementation rule to $\mathfrak{S}_0$, then $R\mathfrak{S}_0$ is sound and complete for the R-implication of MVDs. In fact, every inference of an MVD by $R\mathfrak{S}_0$ can be turned into an inference of the same MVD in which the R-complementation rule is applied at most once, and if it is applied, then in the last step of the inference ($R\mathfrak{S}_0$ is said to be R-complementary). This indicates that the R-complementation rule simply reflects a part of the decomposition process, and does not necessarily infer semantically meaningful consequences.

Interestingly, research has not been continued in this direction but focused on the original notion of R-implication. Since research on dependencies seems to experience a recent revival in the context of other data models (Fischer et al. 1985, Hartmann & Link 2004, Link 2006*a*, Thalheim 2003, Vincent & Liu 2003, Vincent et al. 2003) it seems desirable to further extend the knowledge on relational dependencies. An advancement of such knowledge may simplify the quest of finding suitable and comprehensible extensions of relational dependencies to currently popular data models.

Contributions. In this paper we will extend the work by Biskup (Biskup 1980) from MVDs to FHDs. First, we propose a minimal complete set of sound inference rules for the implication of FHDs in fixed universes. Almost all inference rules are extensions of familiar rules from the axiomatisation of MVDs (Beeri et al. 1977). In particular, the dependence of

---

[*] This research was supported by Marsden Funding, Royal Society of New Zealand

FHDs on the underlying set R of attributes is syntactically reflected by the R-complementation rule.

Example 1.1. *Suppose we design a database for a* DVD *collection. So far, we have decided to use attributes* Title*,* Actor*,* Feature *and* Language*. In order to model that the title of a DVD determines the set of actors independently from the rest of the information in any schema, and the title of a DVD also determines the set of DVD features independently from the rest of the information in any schema we specify the FHD*

$$\text{Title} : \{\{\text{Actor}\}, \{\text{Feature}\}\}.$$

Note that this FHD is equivalent to the two MVDs

$$\text{Title} \twoheadrightarrow \text{Actor} \qquad \textit{and} \qquad \text{Title} \twoheadrightarrow \text{Feature}.$$

If the underlying relation schema R consists of the four attributes above, then we may infer the FHDs

$$\text{Title}:\{\{\text{Actor}\},\{\text{Language}\}\} \textit{ and}$$
$$\text{Title}:\{\{\text{Feature}\},\{\text{Language}\}\}$$

since {Language} *is the DVD-complement of* {Title, Actor, Feature}*. However, if we add a further attribute such as* Production_Year *to the schema DVD, then neither*

$$\text{Title}:\{\{\text{Actor}\},\{\text{Language}\}\} \textit{ nor}$$
$$\text{Title}:\{\{\text{Feature}\},\{\text{Language}\}\}$$

are DVD-implied by Title : {{Actor}, {Feature}}*. The fundamental difference is that the latter FHD has been specified to establish a set-valued correspondence between DVD titles and the actors staring in the movie with that title (and the features available on the DVD with that title, respectively). The other two FHDs do not necessarily correspond to any semantic information, but simply result from the context in which* Title*,* Actor*, and* Feature *are considered. If this context is altered, then the respective FHDs disappear.* □

Example 1.1 suggests that the complementation rule is a mere means of database normalisation and does not always result in the inference of necessarily meaningful consequences. An axiomatisation for the R-implication of FHDs should therefore be complementary. The second contribution is the proposal of such a minimal axiomatisation for FHDs in fixed universes, i.e., no proper subset of inference rules is still both complete and complementary. Moreover, we investigate the impact of replacing the R-complementation rule by the so-called R-axiom. This clarifies the role of the R-complementation rule for FHDs further and extends previous work on MVDs (Biskup 1978).

One may argue that consequences that depend on the underlying universe are no consequences at all. Consequently, the notion of R-implication is not acceptable. Thus, we extend the alternative notion of implication from MVDs (Biskup 1980) to FHDs. This notion leaves the underlying set of attributes undetermined. The third contribution is the identification of a minimal complete set of sound inference rules for the implication of FHDs in undetermined universes. For instance, in Example 1.1 both FHDs

$$\textit{Title}:\{\{\textit{Actor}\},\{\textit{Language}\}\} \text{ and}$$
$$\textit{Title}:\{\{\textit{Feature}\},\{\textit{Language}\}\}$$

are DVD-implied by *Title*:{{*Actor*},{*Feature*}} where DVD={Title,Actor,Feature,Language}, but none of the two FHDs is implied.

Previous Work. The first axiomatisation for the R-implication of MVDs was given in (Beeri et al. 1977). The notion of implication in undetermined universes is from (Biskup 1980) in which an axiomatisation for this notion of MVD implication is given. Minimality of MVD axiomatisations are discussed in (Biskup 1978, Mendelzon 1979) in which (Biskup 1978) also introduces the R-axiom as a very weak form of the R-complementation rule. Recently, more minimal axiomatisations for both complete and complementary axiomatisations for the R-implication of MVDs, and complete axiomatisations for MVD implication in undetermined universes were studied (Hartmann & Link 2006, Link 2006*a*). MVDs have also been studied in the presence of the null value *no information*. In that case, (Lien 1982) was the first to propose an axiomatisation in fixed universes, and (Link 2006*b*) proposes a complementary axiomatisation in fixed universes, and an axiomatisation in undetermined universes. In other data models MVDs have also been investigated in the context of fixed universes. Full hierarchical dependencies were introduced in (Delobel 1978). An axiomatisation for the implication in fixed universes can be found in (Thalheim 1991, Thalheim 2000).

Organisation. The article is structured as follows. In Section 2 we will start with a brief summary of notions from the relational model of data. In particular, we repeat the notion of implication in fixed universes and summarise the axiomatisation for the implication of MVDs. Subsequently, we introduce full hierarchical dependencies as a generalisation of MVDs, and prove the completeness and minimality of a set of sound inference rules for the implication of FHDs in fixed universes. Section 3 examines the property of complementarity for FHDs. It turns out that an extension of the *subset rule* from MVDs to FHDs plays a key role in achieving complementarity. Section 4 discusses the implication of FHDs in undetermined universes. Minimality of the axiomatisations for fixed and undetermined universes are studied in Section 5. Finally, we briefly comment on possible future work in Section 6.

2 Dependencies in Fixed Universes

Let $\mathfrak{A} = \{A_1, A_2, \ldots\}$ be a (countably) infinite set of symbols, called *attributes*. A *relation schema* is a finite set $R = \{A_1, \ldots, A_n\}$ of distinct *attributes* from $\mathfrak{A}$, which represent column names of a relation. Each attribute A_i of a relation schema is associated an infinite domain $dom(A_i)$ which represents the set of possible values that can occur in the column named A_i. If X and Y are sets of attributes, then we may write XY for $X \cup Y$. If $X = \{A_1, \ldots, A_m\}$, then we may write $A_1 \cdots A_m$ for X. In particular, we may write simply A to represent the singleton $\{A\}$. A *tuple* over $R = \{A_1, \ldots, A_n\}$ (R-tuple or simply tuple, if R is understood) is a function $t : R \to \bigcup_{i=1}^{n} dom(A_i)$ with $t(A_i) \in dom(A_i)$ for $i = 1, \ldots, n$. For $X \subseteq R$ let $t[X]$ denote the restriction of the tuple t over R on X, and $dom(X) = \prod_{A \in X} dom(A)$ the Cartesian product of the domains of attributes in X. A *relation* r over R is a finite set of tuples over R. The relation schema R is also called the domain $Dom(r)$ of the relation r over R. Let $r[X] = \{t[X] \mid t \in r\}$ denote the *projection* of the relation r over R on $X \subseteq R$. For $X, Y \subseteq R$, $r_1 \subseteq dom(X)$ and $r_2 \subseteq dom(Y)$ let $r_1 \bowtie r_2 = \{t \in dom(XY) \mid \exists t_1 \in r_1, t_2 \in r_2$ with $t[X] = t_1[X]$ and $t[Y] = t_2[Y]\}$ denote the *natural join* of r_1 and r_2. Note that the 0-ary relation $\{()\}$ is the projection $r[\emptyset]$ of r on $\emptyset$ as well as left and right identity of the natural join operator.

Functional dependencies (FDs) between sets of

Title	Actor	Feature	Language
King Kong	Naomi Watts	Deleted Scenes	English
King Kong	Jack Black	Photo Gallery	French
King Kong	Naomi Watts	Deleted Scenes	French
King Kong	Naomi Watts	Photo Gallery	French
King Kong	Naomi Watts	Photo Gallery	English
King Kong	Jack Black	Photo Gallery	English
King Kong	Jack Black	Deleted Scene	English
King Kong	Jack Black	Deleted Scene	French

Title	Actor
King Kong	Naomi Watts
King Kong	Jack Black

Title	Feature
King Kong	Deleted Scenes
King Kong	Photo Gallery

Title	Language
King Kong	English
King Kong	French

Table 1: A DVD-relation and its projections

attributes have played a central role in the study of relational databases (Beeri & Bernstein 1979, Bernstein 1976, Bernstein & Goodman 1980, Codd 1970, Codd 1972), and seem to be central for the study of database design in other data models as well (Arenas & Libkin 2004, Hara & Davidson 1999, Levene & Loizou 1998, Link 2006*a*, Tari, Stokes & Spaccapietra 1997, Weddell 1992, Wijsen 1999). The notion of a functional dependency is well-understood and the semantic interaction between these dependencies has been syntactically captured by Armstrong's well-known axioms (Armstrong 1974, Armstrong, Nakamura & Rudnicki 2002). A *functional dependency* (FD) (Codd 1972) on the relation schema R is an expression $X \to Y$ where $X, Y \subseteq R$. A relation r over R *satisfies* the FD $X \to Y$, denoted by $\models_r X \to Y$, if and only if every pair of tuples in r that agrees on each of the attributes in X also agrees on the attributes in Y. That is, $\models_r X \to Y$ if and only if $t_1[Y] = t_2[Y]$ whenever $t_1[X] = t_2[X]$ holds for any $t_1, t_2 \in r$.

FDs are incapable of modelling many important properties that database users have in mind. Multivalued dependencies (MVDs) provide a more general notion and offer a response to the shortcomings of FDs. A *multivalued dependency* (MVD) (Fagin 1977, Zaniolo 1976) on R is an expression $X \twoheadrightarrow Y$ where $X, Y \subseteq R$. A relation r over R *satisfies* the MVD $X \twoheadrightarrow Y$, denoted by $\models_r X \twoheadrightarrow Y$, if and only if for all $t_1, t_2 \in r$ with $t_1[X] = t_2[X]$ there is some $t \in r$ with $t[XY] = t_1[XY]$ and $t[X(R - Y)] = t_2[X(R - Y)]$. Informally, the relation r satisfies $X \twoheadrightarrow Y$ when the value on X determines the set of values on Y independently from the set of values on $R - Y$. This actually suggests that the relation schema R is overloaded in the sense that it carries two independent facts XY and $X(R - Y)$. More precisely, it is shown in (Fagin 1977) that MVDs "provide a necessary and sufficient condition for a relation to be decomposable into two of its projections without loss of information (in the sense that the original relation is guaranteed to be the join of the two projections)". This means that $\models_r X \twoheadrightarrow Y$ if and only if $r = r[XY] \bowtie r[X(R - Y)]$. This characteristic of MVDs is fundamental to relational database design and 4NF (Fagin 1977). A lot of research has therefore been devoted to studying the behaviour of these dependencies.

Full hierarchical dependencies generalise multivalued dependencies (Delobel 1978).

Definition 2.1. *A* full hierarchical dependency *on the relation schema R is an expression $X : S$ where $X \subseteq R$ and S is a non-empty set of pairwise disjoint subsets of R that are also disjoint from X, i.e., $S \neq \emptyset$, for all $Y \in S$ we have $Y \subseteq R$ and for all $Y, Z \in S \cup \{X\}$ we have $Y \cap Z = \emptyset$. An R-relation $r \subseteq dom(R)$ is said to* satisfy *(or said to be a* model *of) the full hierarchical dependency $X : \{Y_1, \ldots, Y_k\}$ on R, denoted by $\models_r X : \{Y_1, \ldots, Y_k\}$, if and only if for all $t_1, \ldots, t_{k+1} \in r$ the following condition is satisfied: if $t_i[X] = t_j[X]$ for all $1 \leq i, j \leq k+1$, then there is some $t \in r$ such that $t[XY_i] = t_i[XY_i]$ for $i = 1, \ldots, k$ and $t[X(R - XY_1 \cdots Y_k)] = t_{k+1}[X(R - XY_1 \cdots Y_k)]$.* ☐

Notice that Definition 2.1 reduces to the definition of MVDs in case that $k = 1$. Note that our definition of FHDs is slightly different from the original definition (Delobel 1978). There, an FHD is defined as an expression

$$X : Y_1 \mid \cdots \mid Y_k$$

such that $X, Y_1, \ldots, Y_k$ form a *partition* of the underlying relation schema R. Our definition is different in two aspects. Firstly, $X, Y_1, \ldots, Y_k, R - XY_1 \cdots Y_k$ form a partition of R in our definition. Secondly, the right-handed side of an FHD is a set system $\{Y_1, \ldots, Y_k\}$ over R in our definition, i.e., there is no sequence of attribute sets $Y_1, \ldots, Y_k$ as in the original definition (Delobel 1978). These two differences result in a simpler axiomatisation and the correspondence to the original definition of MVDs is much stronger (Fagin 1977). The next theorem illustrates the importance of FHDs for the removal of redundant data.

Theorem 2.1. *Let $X, Y_1, \ldots, Y_k \subseteq R$ be pairwise disjoint and $k \geq 1$. An R-relation r satisfies the FHD $X : \{Y_1, \ldots, Y_k\}$ on R if and only if $r = r[XY_1] \bowtie \cdots \bowtie r[XY_k] \bowtie r[X(R - XY_1 \cdots Y_k)]$.* ☐

Example 2.1. *Consider again the relation schema* DVD *comprising the four attributes* Title*,* Actor*,* Feature *and* Language*. Table 1 shows a relation r that satisfies the FHD*

$$\text{Title} : \{\{\text{Actor}\}, \{\text{Feature}\}\}$$

and the projections of r on {Title,Actor}*,* {Title,Feature} *and* {Title,Language}*, respectively. Indeed, r is the natural join of its projections on these three attribute sets.* ☐

Theorem 2.2. *Let $X, Y_1, \ldots, Y_k \subseteq R$ be pairwise disjoint and $k \geq 1$. An R-relation r satisfies the FHD $X : \{Y_1, \ldots, Y_k\}$ on R if and only if for all $i = 1, \ldots, k$, r satisfies the MVD $X \twoheadrightarrow Y_i$ on R.* ☐

Example 2.2. *Consider again the relation r in Table 1. In particular, one can verify that r satisfies the MVDs* Title $\twoheadrightarrow$ Actor *and* Title $\twoheadrightarrow$ Feature*.* ☐

For the design of a relational database schema dependencies are normally specified as semantic constraints on the relations which are intended to be instances of the schema. During the design process one usually needs to determine further dependencies which are logically implied by the given ones. In order to emphasise the dependence of implication from the underlying relation schema R we refer to R-implication.

Definition 2.2. *Let* $\Sigma = \{X_1 : \{Y_1^1, \dots, Y_{l_1}^1\}, \dots, X_n : \{Y_1^n, \dots, Y_{l_n}^n\}\}$ *and* $X : \{Y_1, \dots, Y_k\}$ *be FHDs on the relation schema* R*, i.e.,*

$$X \cup \bigcup_{i=1}^{k} Y_i \cup \bigcup_{j=1}^{n} \left(X_j \cup \bigcup_{s=1}^{l_j} Y_s^j \right) \subseteq R.$$

Then Σ *R-implies* $X : \{Y_1, \dots, Y_k\}$ *if and only if each relation* r *over* R *that satisfies all FHDs in* Σ *also satisfies* $X : \{Y_1, \dots, Y_k\}$*.* ☐

Notice that Definition 2.2 covers MVDs in case that $l_j = 1$ for $j = 1, \dots, n$ and $k = 1$ (Biskup 1980). In order to determine all logical consequences of a set of MVDs one can use the following set of inference rules for the R-implication of multivalued dependencies (Beeri et al. 1977). These *inference rules* have the form

$$\frac{\text{premise}}{\text{conclusion}}$$

and inference rules without a premise are called *axioms*. Note that we use the natural complementation rule (Biskup 1978) instead of the complementation rule that was originally proposed (Beeri et al. 1977).

$$\frac{}{X \twoheadrightarrow Y}\,{}^{Y \subseteq X} \qquad\qquad \frac{X \twoheadrightarrow Y}{XU \twoheadrightarrow YV}\,{}^{V \subseteq U}$$

(reflexivity, $\mathcal{R}_{\text{MVD}}$) (augmentation, $\mathcal{A}_{\text{MVD}}$)

$$\frac{X \twoheadrightarrow Y, Y \twoheadrightarrow Z}{X \twoheadrightarrow Z - Y} \qquad\qquad \frac{X \twoheadrightarrow Y, W \twoheadrightarrow Z}{X \twoheadrightarrow Z - Y}\,{}^{Y \cap W = \emptyset}$$

(pseudo-transitivity, $\mathcal{T}_{\text{MVD}}$) (subset, $\mathcal{S}_{\text{MVD}}$)

$$\frac{X \twoheadrightarrow Y}{X \twoheadrightarrow R - Y} \qquad\qquad \frac{X \twoheadrightarrow Y, X \twoheadrightarrow Z}{X \twoheadrightarrow YZ}$$

(R-complementation, $\mathcal{C}_R^{\text{MVD}}$) (union, $\mathcal{U}_{\text{MVD}}$)

$$\frac{X \twoheadrightarrow Y, X \twoheadrightarrow Z}{X \twoheadrightarrow Z - Y} \qquad\qquad \frac{X \twoheadrightarrow Y, X \twoheadrightarrow Z}{X \twoheadrightarrow Y \cap Z}$$

(difference, $\mathcal{D}_{\text{MVD}}$) (intersection, $\mathcal{I}_{\text{MVD}}$)

The set $\{\mathcal{R}_{\text{MVD}}, \mathcal{A}_{\text{MVD}}, \mathcal{T}_{\text{MVD}}, \mathcal{C}_R^{\text{MVD}}, \mathcal{U}_{\text{MVD}}, \mathcal{D}_{\text{MVD}}, \mathcal{I}_{\text{MVD}}\}$ is both sound and complete for the R-implication of MVDs (Beeri et al. 1977). Let $\Sigma \cup \{\sigma\}$ be a set of dependencies from the class $\mathcal{C}$ on the relation schema R. Let $\Sigma \vdash_{\mathfrak{S}} \sigma$ denote the inference of σ from a set Σ of dependencies from $\mathcal{C}$ with respect to the set $\mathfrak{S}$ of inference rules. Let $\Sigma_{\mathfrak{S}}^+ = \{\sigma \mid \Sigma \vdash_{\mathfrak{S}} \sigma\}$ denote the *syntactic hull* of Σ under inference using only rules from $\mathfrak{S}$. In what follows we use the letter R to emphasise the fact that a set $R\mathfrak{S}$ refers to R-implication, i.e., to implication in fixed universes. An inference rule is called *R-sound* if the set of dependencies in the premise of the rule R-implies the dependency in the conclusion. It is well-known that all the rules above are R-sound for all R (i.e. if one restricts U, V, W, X, Y, Z to be subsets of R) (Beeri et al. 1977). The set $R\mathfrak{S}$ is called *R-sound* for the R-implication of dependencies from $\mathcal{C}$ if and only if for every set Σ of dependencies from $\mathcal{C}$ on the relation schema R we have $\Sigma_{R\mathfrak{S}}^+ \subseteq \Sigma_R^* = \{\sigma \in \mathcal{C} \mid \Sigma \ R\text{-implies } \sigma\}$. The set $R\mathfrak{S}$ is called *R-complete* for the R-implication of dependencies from $\mathcal{C}$ if and only if for every set Σ of dependencies from $\mathcal{C}$ on R we have $\Sigma_R^* \subseteq \Sigma_{R\mathfrak{S}}^+$. Furthermore, the set $R\mathfrak{S}$ is called complete (sound) for the R-implication of dependencies from $\mathcal{C}$ if and only if it is R-complete (R-sound) for the R-implication of dependencies from $\mathcal{C}$ for all relation schemata R.

An interesting question is now whether all the rules of a certain set of inference rules are really necessary to capture the R-implication of dependencies from $\mathcal{C}$ for every relation schema R. More precisely, an inference rule $\mathfrak{R}$ is said to be *R-independent* from the set $R\mathfrak{S}$ if and only if there is some set $\Sigma \cup \{\sigma\}$ of dependencies from $\mathcal{C}$ on the relation schema R such that $\sigma \notin \Sigma_{R\mathfrak{S}}^+$, but $\sigma \in \Sigma_{R\mathfrak{S} \cup \{\mathfrak{R}\}}^+$. Moreover, an inference rule $\mathfrak{R}$ is said to be *independent* from $R\mathfrak{S}$ if and only if there is some relation schema R such that $\mathfrak{R}$ is R-independent from $R\mathfrak{S}$. A complete set $R\mathfrak{S}$ is called *minimal* for the R-implication of dependencies from $\mathcal{C}$ if and only if every inference rule $\mathfrak{R} \in R\mathfrak{S}$ is independent from $R\mathfrak{S} - \{\mathfrak{R}\}$. This means that no proper subset of $R\mathfrak{S}$ is still complete.

2.1 Sound Inference Rules for FHDs in fixed Universes

We will denote the set of inference rules from Table 2 by $R\mathfrak{S}$.

Theorem 2.3. *The set* $R\mathfrak{S}$ *is sound for the R-implication of FHDs.*

Sketch. We need to show that for an arbitrary relation schema R, and an arbitrary set Σ of FHDs on R we have $\Sigma_{R\mathfrak{S}}^+ \subseteq \Sigma_R^*$. The proof will make extensive use of Theorem 2.2 and the R-soundness of the inference rules for MVDs (Beeri et al. 1977). Notice that it is sufficient to show the R-soundness of each inference rule. Subsequently, one can show by induction that each FHD occurring in an inference by $R\mathfrak{S}$ is also R-implied by Σ.

We only show how to prove the *augmentation rule* $\mathcal{A}$. Let r be some arbitrary R-relation that satisfies $X : \{Y_1, \dots, Y_k\}$. Theorem 2.2 tells us that r also satisfies $X \twoheadrightarrow Y_i$ for all $i = 1, \dots, k$. The soundness of the augmentation rule $\mathcal{A}_{\text{MVD}}$ shows that r also satisfies $XZ \twoheadrightarrow Y_i$ for all $i = 1, \dots, k$, and the soundness of the reflexivity axiom $\mathcal{R}_{\text{MVD}}$ identifies r as a model of $XZ \twoheadrightarrow Z$. We conclude by the soundness of the difference rule $\mathcal{D}_{\text{MVD}}$ that r satisfies $XZ \twoheadrightarrow Y_i - Z$ for all $i = 1, \dots, k$. Consequently, $\models_r XZ : \{Y_1 - Z, \dots, Y_k - Z\}$ by Theorem 2.2.

The proof for the remaining rules is similar. ☐

Lemma 2.1. *The following inference rules are derivable from* $R\mathfrak{S}$

$$\frac{X : \{Y_1, \dots, Y_k\}}{X : \{Y_1, \dots, Y_k, \emptyset\}} \qquad\qquad \frac{X : \{Y_1, \dots, Y_k, Y_{k+1}\}}{X : \{Y_1, \dots, Y_k Y_{k+1}\}}$$

(empty-set-introduction, $\mathcal{I}_\emptyset$*)* *(merging,* $\mathcal{M}$*)*

and thus are sound for the R-implication of FHDs.

Proof. The *empty-set-introduction rule* $\mathcal{I}_\emptyset$ is derivable from $R_\emptyset, \mathcal{A}$ and $\mathcal{T}$.

$$\dfrac{\dfrac{\dfrac{}{\emptyset : \{\emptyset\}}\,\mathcal{R}_\emptyset}{X : \{\emptyset\}}\,\mathcal{A} \qquad \dfrac{X : \{Y_1, \dots, Y_k\}}{X \cup \emptyset : \{Y_1, \dots, Y_k\}}\,\mathcal{A}}{X : \{Y_1, \dots, Y_k, \emptyset\}}\,\mathcal{T}$$

The *merging rule* $\mathcal{M}$ is derivable from $\mathcal{O}$ and $\mathcal{U}$.

$$\dfrac{\dfrac{X : \{Y_1, \dots, Y_k, Y_{k+1}\}}{X : \{Y_1, \dots, Y_k\}}\,\mathcal{O} \qquad \dfrac{X : \{Y_1, \dots, Y_k, Y_{k+1}\}}{X : \{Y_{k+1}\}}\,\mathcal{O}}{X : \{Y_1, \dots, Y_{k-1}, Y_k Y_{k+1}\}}\,\mathcal{U}$$

This concludes the proof. ☐

$$\frac{}{\emptyset : \{\emptyset\}}$$
(empty-set-axiom, $\mathcal{R}_\emptyset$)

$$\frac{X : \{Y_1, \dots, Y_k\}}{XZ : \{Y_1 - Z, \dots, Y_k - Z\}}$$
(augmentation, $\mathcal{A}$)

$$\frac{XY : \{Y_1, \dots, Y_k\}, X : \{Y\}}{X : \{Y_1, \dots, Y_k, Y\}}$$
(transitivity, $\mathcal{T}$)

$$\frac{X : \{Y_1, \dots, Y_k\}}{X : \{Y_1, \dots, Y_{k-1}, R - XY_1 \cdots Y_k\}}$$
(R-complementation, $\mathcal{C}_R$)

$$\frac{X : \{Y_1, \dots, Y_k, Y\}}{X : \{Y_1, \dots, Y_k\}}$$
(omission, $\mathcal{O}$)

$$\frac{X : \{Y_1, \dots, Y_k\}, X : \{Z\}}{X : \{Y_1 - Z, \dots, Y_{k-1} - Z, Y_k Z\}}$$
(union, $\mathcal{U}$)

$$\frac{X : \{Y_1, \dots, Y_k\}, X : \{Z\}}{X : \{Y_1, \dots, Y_{k-1}, Y_k - Z\}}$$
(difference, $\mathcal{D}$)

$$\frac{X : \{Y_1, \dots, Y_k\}, X : \{Z\}}{X : \{Y_1, \dots, Y_{k-1}, Y_k \cap Z\}}$$
(intersection, $\mathcal{I}$)

Table 2: Inference rules for the R-implication of FHDs

2.2 Completeness

Let R be some arbitrary relation schema, and let Σ be a set of FHDs on R. Let $Dep_R(X)$ be the set of all $W \subseteq R$ for which some FHD $X : S$ with $W \in S$ is derivable from Σ using the inference rules $R\mathfrak{S}$, i.e., $Dep_R(X) = \{W \subseteq R \mid X : S \in \Sigma^+_{R\mathfrak{S}}$ and $W \in S\}$. Note that $Dep_R(X)$ is finite, and $(Dep_R(X), \subseteq, \cup, \cap, (\cdot)^{\mathcal{C}}, \emptyset, R)$ constitutes a Boolean algebra due to the soundness of union, difference and intersection rule. Recall that an element $a \in P$ of a poset $(P, \sqsubseteq, 0)$ with least element 0 is called an *atom* of $(P, \sqsubseteq, 0)$ (Graetzer 1998) if and only if $a \neq 0$ and every element $b \in P$ with $b \sqsubseteq a$ satisfies $b = 0$ or $b = a$. $(P, \sqsubseteq, 0)$ is called *atomic* if and only if for every element $b \in P$ with $b \neq 0$ there is an atom $a \in P$ with $a \sqsubseteq b$. In particular, every finite Boolean algebra is atomic. The set $DepB_R(X)$ of all atoms of $(Dep_R(X), \subseteq, \emptyset)$ is called the *dependency basis* of X with respect to Σ (Beeri 1980).

Theorem 2.4. *Let $\Sigma \cup \{X : S\}$ be a set of FHDs on the relation schema R. Then $X : S \in \Sigma^+_{R\mathfrak{S}}$ if and only if for every $Y \in S$ there is some $\mathcal{Y} \subseteq DepB_R(X)$ such that $Y = \bigcup \mathcal{Y}$.*

Proof. Let $Y \in S$ for $X : S \in \Sigma^+_{R\mathfrak{S}}$. That is, $Y \in Dep_R(X)$, and since every element b of a Boolean algebra is the join over those atoms a with $a \leq b$ we know that $Y = \bigcup \mathcal{Y}$ for $\mathcal{Y} = \{W \in DepB_R(X) \mid W \subseteq Y\}$.

Vice versa, let $Y = \bigcup \mathcal{Y}$ for some $\mathcal{Y} \subseteq DepB_R(X)$. Since $DepB_R(X) \subseteq Dep_R(X)$ and $Dep_R(X)$ is closed under unions it follows that $Y \in Dep_R(X)$. As this is true for all $Y \in S$ and the sets in S are pairwise disjoint we can apply the *empty-set-introduction rule* $\mathcal{I}_\emptyset$ and the *union rule* $\mathcal{U}$ to infer $X : S \in \Sigma^+_{R\mathfrak{S}}$. ☐

Theorem 2.5. *The set $R\mathfrak{S}$ of inference rules is complete for the R-implication of FHDs.* ☐

2.3 A minimal Axiomatisation

The objective is to identify a minimal subset of $R\mathfrak{S}$, i.e., a set in which each single rule is essential and not derivable from the rest of the rules.

Theorem 2.6. *The set $\{\mathcal{R}_\emptyset, \mathcal{A}, \mathcal{T}, \mathcal{O}, \mathcal{C}_R\}$ consisting of the* empty-set axiom, *the* augmentation rule, *the* transitivity rule, *the* omission rule *and the* R-complementation rule *is minimal for the R-implication of FHDs.*

We prove Theorem 2.6 by a series of lemmata. First, it is shown that union rule $\mathcal{U}$, intersection rule $\mathcal{I}$ and difference rule $\mathcal{D}$ are derivable from $\{\mathcal{R}_\emptyset, \mathcal{A}, \mathcal{T}, \mathcal{O}, \mathcal{C}_R\}$. Subsequently, it is demonstrated that none of the rules in $\{\mathcal{R}_\emptyset, \mathcal{A}, \mathcal{T}, \mathcal{O}, \mathcal{C}_R\}$ can be omitted without losing completeness for the R-implication of FHDs.

Lemma 2.2. *The union rule $\mathcal{U}$ is derivable from $\{\mathcal{A}, \mathcal{T}, \mathcal{C}_R, \mathcal{O}\}$.* ☐

Lemma 2.3. *The intersection rule $\mathcal{I}$ is derivable from $\{\mathcal{A}, \mathcal{T}, \mathcal{C}_R, \mathcal{O}, \mathcal{U}\}$.* ☐

Lemma 2.4. *The difference rule $\mathcal{D}$ is derivable from $\{\mathcal{C}_R, \mathcal{I}\}$.*

Proof. Note that $Y_k \cap (R - XZ) = Y_k - XZ = Y_k - Z$ since $Y_k \cap X = \emptyset$.

$$\cfrac{X : \{Y_1, \dots, Y_k\} \qquad \cfrac{X : \{Z\}}{X : \{R - XZ\}}\,\mathcal{C}_R}{X : \{Y_1, \dots, Y_{k-1}, Y_k - Z\}}\,\mathcal{I}$$

This concludes the proof. ☐

Theorem 2.5 and Lemmata 2.2, 2.3 and 2.4 show that $\{\mathcal{R}_\emptyset, \mathcal{A}, \mathcal{T}, \mathcal{O}, \mathcal{C}_R\}$ is indeed complete for the R-implication of FHDs. In order to complete the proof of Theorem 2.6 it remains to verify the independence of every inference rule in $\{\mathcal{R}_\emptyset, \mathcal{A}, \mathcal{T}, \mathcal{O}, \mathcal{C}_R\}$ from the rest of the rules.

Lemma 2.5. *The empty-set-axiom $\mathcal{R}_\emptyset$ is independent from $\{\mathcal{A}, \mathcal{T}, \mathcal{O}, \mathcal{C}_R\}$.*

Proof. Let $R = \emptyset$, $\Sigma = \emptyset$ and $\sigma = \emptyset : \{\emptyset\}$. It follows that $\sigma \notin \Sigma^+_{\{\mathcal{A}, \mathcal{T}, \mathcal{O}, \mathcal{C}_R\}}$, but $\sigma \in \Sigma^+_{\{\mathcal{R}_\emptyset, \mathcal{A}, \mathcal{T}, \mathcal{O}, \mathcal{C}_R\}}$. ☐

Lemma 2.6. *The augmentation rule $\mathcal{A}$ is independent from $\{\mathcal{R}_\emptyset, \mathcal{T}, \mathcal{O}, \mathcal{C}_R\}$.*

Proof. Let $R = A$, $\Sigma = \emptyset$ and $\sigma = A : \{\emptyset\}$. The closure $\Sigma^+_{\mathfrak{S}}$ of a set $\mathfrak{S}$ of sound inference rules is represented as a table. The FHD $X : \{Y_1, \dots, Y_k\}$ belongs to $\Sigma^+_{\mathfrak{S}}$ if and only if the entry in row labelled X and column labelled $\{Y_1, \dots, Y_k\}$ is a cross ×. Notice that there are some cells which do not represent any FHD. These have the entry $\star$. The closure $\Sigma^+_{\{\mathcal{R}_\emptyset, \mathcal{T}, \mathcal{O}, \mathcal{C}_R\}}$ can be obtained as follows. The *empty-set-axiom* $\mathcal{R}_\emptyset$ yields the FHD $\emptyset : \{\emptyset\}$. A subsequent application of the *R-complementation rule* $\mathcal{C}_R$ allows us to obtain the FHD $\emptyset : \{A\}$. Finally, the *transitivity rule* $\mathcal{T}$ can be applied to $\emptyset : \{\emptyset\}$ and $\emptyset : \{A\}$ to infer $\emptyset : \{A, \emptyset\}$. This set is closed under derivation using $\{\mathcal{R}_\emptyset, \mathcal{T}, \mathcal{O}, \mathcal{C}_R\}$.

	$\{\emptyset\}$	$\{A\}$	$\{A, \emptyset\}$
$\emptyset$	×	×	×
A		$\star$	$\star$

It follows that $\sigma \notin \Sigma^+_{\{\mathcal{R}_\emptyset, \mathcal{T}, \mathcal{O}, \mathcal{C}_R\}}$ but $\sigma \in \Sigma^+_{\{\mathcal{R}_\emptyset, \mathcal{A}, \mathcal{T}, \mathcal{O}, \mathcal{C}_R\}}$. ☐

Lemma 2.7. *The transitivity rule $\mathcal{T}$ is independent from $\{\mathcal{R}_\emptyset, \mathcal{A}, \mathcal{O}, \mathcal{C}_R\}$.*

Proof. Let $R = A$, $\Sigma = \emptyset$ and $\sigma = \emptyset : \{A, \emptyset\}$. The closure $\Sigma^+_{\{\mathcal{R}_\emptyset, \mathcal{A}, \mathcal{O}, \mathcal{C}_R\}}$ can be obtained as follows. The *empty-set-axiom* $\mathcal{R}_\emptyset$ yields the FHD $\emptyset : \{\emptyset\}$. A subsequent application of the *R-complementation rule* $\mathcal{C}_R$ allows us to obtain the FHD $\emptyset : \{A\}$. Finally, the *augmentation rule* $\mathcal{A}$ can be applied to $\emptyset : \{\emptyset\}$ to infer $A : \{\emptyset\}$. This set is closed under derivation using $\{\mathcal{R}_\emptyset, \mathcal{A}, \mathcal{O}, \mathcal{C}_R\}$.

	$\{\emptyset\}$	$\{A\}$	$\{A, \emptyset\}$
$\emptyset$	×	×	
A	×	⋆	⋆

It follows that $\sigma \notin \Sigma^+_{\{\mathcal{R}_\emptyset, \mathcal{A}, \mathcal{O}, \mathcal{C}_R\}}$ but $\sigma \in \Sigma^+_{\{\mathcal{R}_\emptyset, \mathcal{A}, \mathcal{T}, \mathcal{O}, \mathcal{C}_R\}}$. ☐

Lemma 2.8. *The omission rule $\mathcal{O}$ is independent from $\{\mathcal{R}_\emptyset, \mathcal{A}, \mathcal{T}, \mathcal{C}_R\}$.*

Proof. In this case we can choose $R = AB$, $\Sigma = \{\emptyset : \{\emptyset, A\}\}$ and $\sigma = \emptyset : \{A\}$. ☐

Lemma 2.9. *The R-complementation rule $\mathcal{C}_R$ is independent from $\{\mathcal{R}_\emptyset, \mathcal{A}, \mathcal{T}, \mathcal{O}\}$.*

Proof. Let $R = A$, $\Sigma = \emptyset$ and $\sigma = \emptyset : \{A\}$. The closure $\Sigma^+_{\{\mathcal{R}_\emptyset, \mathcal{A}, \mathcal{T}, \mathcal{O}\}}$ can be obtained as follows. The *empty-set-axiom* $\mathcal{R}_\emptyset$ yields the FHD $\emptyset : \{\emptyset\}$. A subsequent application of the *augmentation rule* $\mathcal{A}$ allows us to derive the FHD $A : \{\emptyset\}$. This set is closed under derivation using $\{\mathcal{R}_\emptyset, \mathcal{A}, \mathcal{T}, \mathcal{O}\}$.

	$\{\emptyset\}$	$\{A\}$	$\{A, \emptyset\}$
$\emptyset$	×		
A	×	⋆	⋆

It follows that $\sigma \notin \Sigma^+_{\{\mathcal{R}_\emptyset, \mathcal{A}, \mathcal{T}, \mathcal{O}\}}$ but $\sigma \in \Sigma^+_{\{\mathcal{R}_\emptyset, \mathcal{A}, \mathcal{T}, \mathcal{O}, \mathcal{C}_R\}}$. ☐

2.4 A weaker version of R-complementation

Biskup (Biskup 1978) has replaced the R-complementation rule $\mathcal{C}_R$ by the so-called *R-axiom* $\frac{}{\emptyset \twoheadrightarrow R}$ and still obtained a complete axiomatisation for the R-implication of MVDs. That is, *R-axiom*, *augmentation rule* and *pseudo-transitivity rule* form a sound and complete set of inference rules for the R-implication of MVDs (Biskup 1978).

Let $\frac{}{\emptyset : \{R\}}$ be the R-axiom for FHDs. This inference rule is sound for the R-implication of FHDs. As it turns out we can simply replace the *R-complementation rule* $\mathcal{C}_R$ by the *R-axiom* and still maintain completeness.

Theorem 2.7. *The set $\{\mathcal{R}_\emptyset, \mathcal{A}, \mathcal{T}, \mathcal{O}$,R-axiom$\}$ consisting of the* empty-set axiom*, the* augmentation rule*, the* transitivity rule*, the* omission rule *and the* R-axiom *is sound and complete for the R-implication of FHDs.*

Proof. One can show that the *R-complementation rule* $\mathcal{C}_R$ is derivable from the *R-axiom*, the *augmentation rule*, the *transitivity rule*, and the *omission rule*. The statement is then a consequence of Theorem 2.5. ☐

3 A complementary Axiomatisation

We have seen before that the set $\{\mathcal{R}, \mathcal{A}, \mathcal{T}, \mathcal{O}, \mathcal{C}_R\}$ is a minimal axiomatisation for the R-implication of FHDs. However, Example 1.1 has brought up a possible difficulty with this system. If the R-complementation rule simply represents the database normalisation process, then this should be reflected by any axiomatisation. More precisely, an application of the R-complementation rule $\mathcal{C}_R$ during any inferences should be restricted to the very last step of this inference (if needed at all). This would ensure that no possibly semantically meaningless information could be derived. A complete set $R\mathfrak{S}$ of inference rules is said to be *complementary* for the R-implication of FHDs if and only if it is R-complementary for every relation schema R. That is, for every set $\Sigma \cup \{\sigma\}$ of FHDs on R the inference of σ from Σ using $R\mathfrak{S}$ can be turned into an inference of σ from Σ using $R\mathfrak{S}$ in which the R-complementation rule $\mathcal{C}_R$ is applied at most once, and if it is applied, then it is applied in the last step of the inference. As it turns out the system $\{\mathcal{R}, \mathcal{A}, \mathcal{T}, \mathcal{O}, \mathcal{C}_R\}$ is not complementary.

Example 3.1. *Let Σ consist of the two FHDs*

Title : {Actor, Feature} *and* Title : {Actor, Language}.

It can be shown that

$$\text{Title} : \{\text{Actor}, \text{Feature Language}\} \notin \Sigma^+_{\{\mathcal{R}, \mathcal{A}, \mathcal{T}, \mathcal{O}\}},$$

see Lemma 5.7, and

$$\text{Title} : \{\text{Actor}, Y\} \notin \Sigma^+_{\{\mathcal{R}, \mathcal{A}, \mathcal{T}, \mathcal{O}\}}$$

for any Y such that

$$Y - \{\text{Title, Actor, Feature, Language}\} \neq \emptyset,$$

see Lemma 4.1.

For DVD={Title, Actor, Feature, Language, Crew} *we have*

$$\text{Title} : \{\text{Actor}, \text{Feature Language}\} \in \Sigma^+_{\{\mathcal{R}, \mathcal{A}, \mathcal{T}, \mathcal{O}, \mathcal{C}_R\}}.$$

Hence, in any such inference the DVD-*complementation rule* $\mathcal{C}_{\text{DVD}}$ *must be applied at least once. However, since*

DVD-{Title,Actor,Feature,Language}={Crew}

$\mathcal{C}_{\text{DVD}}$ *is not just used as the last rule.* ☐

The goal is to find an axiomatisation for the R-implication of FHDs that soundly reflects the role of the R-complementation rule $\mathcal{C}_R$ as a mere means of database normalisation. That is, we would like to identify a complementary axiomatisation for FHDs. This objective has been successfully achieved in the case of MVDs (Biskup 1980, Link 2006*b*). A key role played the so-called *subset rule* $\mathcal{S}_{\text{MVD}}$ which we now generalise into the framework of FHDs.

$$\frac{X : Y, W : \{Y_1, \ldots, Y_k\}}{X : \{Y_1 \cap Y, \ldots, Y_k \cap Y, Y - Y_1 \cdots Y_k\}} \; Y \cap W = \emptyset$$
$$\text{(subset, } \mathcal{S}\text{)}$$

It is not obvious why the set $Y - Y_1 \cdots Y_k$ is included in the conclusion of the *subset rule*. In fact, this set is needed in the proof of Theorem 3.1 to shift applications of the R-complementation rule to the end of an inference.

Lemma 3.1. *The subset rule $\mathcal{S}$ is sound for the R-implication of FHDs.*

Proof. Let r be some arbitrary relation such that $X \cup Y \cup W \cup \bigcup_{i=1}^{k} Y_i \subseteq R$. Let r satisfy $X : Y$ and $W : \{Y_1, \ldots, Y_k\}$ where $Y \cap W = \emptyset$ holds. We know by Theorem 2.2 that r satisfies $X \twoheadrightarrow Y$ and $W \twoheadrightarrow Y_i$ for all $i = 1, \ldots, k$. We conclude by the soundness of the subset rule $\mathcal{S}_{\text{MVD}}$ that $\models_r X \twoheadrightarrow Y_i \cap Y$ for all $i = 1, \ldots, k$. Moreover, the soundness of the difference rule $\mathcal{D}_{\text{MVD}}$ shows that r satisfies $X \twoheadrightarrow Y - Y_i$ for all $i = 1, \ldots, k$ since $Y - (Y_i \cap Y) = Y - Y_i$ holds. A further application of Theorem 2.2 shows that r satisfies $X : \{Y_1 \cap Y, \ldots, Y_k \cap Y, Y - Y_1 \cdots Y_k\}$. $\square$

Let Σ be a set of FHDs, and let $\mathfrak{S}$ be a set of inference rules. A finite sequence of FHDs $\gamma = [\sigma_1, \ldots, \sigma_n]$ is called an *inference from* Σ *by* $\mathfrak{S}$ if and only if each σ_i is either an element of Σ or is obtained by applying one of the rules of $\mathfrak{S}$ to appropriate elements of $\{\sigma_1, \ldots, \sigma_{i-1}\}$. We say that the inference γ infers σ_n, i.e., the last element of the sequence γ. In fact, $\Sigma^+_{\mathfrak{S}}$ denotes the set of all FHDs which are inferred by some inference from Σ by $\mathfrak{S}$.

Theorem 3.1. *Let* Σ *be a set of FHDs. For each inference* γ *from* Σ *by the system* $R\mathfrak{H} = \{\mathcal{R}_\emptyset, \mathcal{A}, \mathcal{T}, \mathcal{O}, \mathcal{C}_R\}$ *there is an inference* ξ *from* Σ *by the system* $R\mathfrak{H}_C = \{\mathcal{R}_\emptyset, \mathcal{A}, \mathcal{T}, \mathcal{O}, \mathcal{S}, \mathcal{M}, \mathcal{C}_R\}$ *with the following properties:*

- γ *and* ξ *infer the same FHD,*
- *in* ξ *the R-complementation rule* $\mathcal{C}_R$ *is applied at most once, and*
- *if* $\mathcal{C}_R$ *is applied in* ξ*, then* $\mathcal{C}_R$ *is applied as the last rule.*

Proof. We proceed by induction on the length l of γ. If $l = 1$, then $\xi := \gamma$ has the desired properties. Let $l > 1$, and $\gamma = [\sigma_1, \ldots, \sigma_l]$ be an inference from Σ by $R\mathfrak{H}$ which has length l. One would need to consider five cases according to which inference rule in $R\mathfrak{H}$ was applied to infer σ_l from $[\sigma_1, \ldots, \sigma_{l-1}]$. However, due to lack of space we only consider the case where we infer σ_l by applying the *omission rule* $\mathcal{O}$ to the premise σ_i with $i < l$. Let ξ_i be obtained by using the induction hypothesis for $\gamma_i := [\sigma_1, \ldots, \sigma_i]$.

Consider the inference $\xi := [\xi_i, \sigma_l]$. If $\mathcal{C}_R$ is not applied in ξ_i, then ξ has the desired properties. If $\mathcal{C}_R$ is applied in ξ_i (as the last rule), then the last two steps of ξ either have the form:

$$\dfrac{\dfrac{X : \{Y_1, \ldots, Y_k, Y\}}{X : \{Y_1, \ldots, Y_k, R - XYY_1 \cdots Y_k\}}\,\mathcal{C}_R}{X : \{Y_1, \ldots, Y_k\}}\,\mathcal{O}.$$

or

$$\dfrac{\dfrac{X : \{Y_1, \ldots, Y_k, Y\}}{X : \{Y_1, \ldots, Y_k, R - XYY_1 \cdots Y_k\}}\,\mathcal{C}_R}{XZ : \{Y_1, \ldots, Y_{i-1}, Y_{i+1}, \ldots, Y_k, R - XYY_1 \cdots Y_k\}}\,\mathcal{O}.$$

In the first case these steps may be simply replaced by

$$\dfrac{X : \{Y_1, \ldots, Y_k, Y\}}{X : \{Y_1, \ldots, Y_k\}}\,\mathcal{O}.$$

In the second case, these steps can be replaced as follows:

$$\dfrac{\dfrac{X : \{Y_1, \ldots, Y_k, Y\}}{X : \{Y_1, \ldots, Y_{i-1}, Y_{i+1}, \ldots, Y_k, Y_iY\}}\,\mathcal{M}}{X : \{Y_1, \ldots, Y_{i-1}, Y_{i+1}, \ldots, Y_k, R - XYY_1 \cdots Y_k\}}\,\mathcal{C}_R.$$

In both cases the result of these replacements is an inference with the desired properties. $\square$

Corollary 3.1. *The axiomatisation* $R\mathfrak{H}_C = \{\mathcal{R}_\emptyset, \mathcal{A}, \mathcal{T}, \mathcal{O}, \mathcal{S}, \mathcal{M}, \mathcal{C}_R\}$ *is complementary for the R-implication of FHDs.* $\square$

Theorem 3.1 states that $R\mathfrak{H}_C$ is almost complete for the R-implication of FHDs.

Corollary 3.2. *Let* $R \subseteq \mathfrak{A}$ *be a finite set of attributes. Then for all finite sets* $\Sigma = \{X_1 : \{Y_1^1, \ldots, Y_{l_1}^1\}, \ldots, X_n : \{Y_1^n, \ldots, Y_{l_n}^n\}$ *of FHDs, for all FHDs* $X : \{Y_1, \ldots, Y_k\}$ *such that* $X \cup \bigcup_{i=1}^{k} Y_i \cup \bigcup_{j=1}^{n} \left(X_j \cup \bigcup_{s=1}^{l_j} Y_s^j \right) \subseteq R$ *we have:* $X : \{Y_1, \ldots, Y_k\} \in \Sigma^+_{R\mathfrak{H}_C}$ *if and only if* $X : \{Y_1, \ldots, Y_k\} \in \Sigma^+_{\mathfrak{H}_C}$ *or there is some* i *such that* $1 \leq i \leq k$ *and* $X : \{Y_1, \ldots, Y_{i-1}, Y_{i+1}, \ldots, Y_k, R - XY_1 \cdots Y_k\} \in \Sigma^+_{\mathfrak{H}_C}$. $\square$

4 FHDs in undetermined universes

We will now investigate the alternative notion of implication in which the underlying set of attributes is undetermined. Notice that this form of implication has been studied for MVDs (Biskup 1980, Link 2006*a*).

According to Example 1.1 it may be argued that consequences which are dependent on the underlying relation schema are in fact no consequences at all. This implies, however, that the notion of R-implication is not suitable. This observation already applies to MVDs (Biskup 1980, Link 2006*a*), and Biskup introduced an alternative notion of MVD implication (Biskup 1980) in which the underlying set of attributes remains undetermined. We will now extend this notion to full hierarchical dependencies.

An FHD is a syntactic expression $X : \{Y_1, \ldots, Y_k\}$ with $X, Y_1, \ldots, Y_k \subseteq \mathfrak{A}$. The FHD $X : \{Y_1, \ldots, Y_k\}$ is *satisfied* by some relation r if and only if $X \cup \bigcup_{i=1}^{k} Y_i \subseteq Dom(r)$ and

$$r = r[XY_1] \bowtie \cdots \bowtie r[XY_k] \bowtie r[X \cup (Dom(r) - XYY_1 \cdots Y_k)].$$

Definition 4.1. *The set* $\Sigma = \{X_1 : \{Y_1^1, \ldots, Y_{l_1}^1\}, \ldots, X_n : \{Y_1^n, \ldots, Y_{l_n}^n\}$ *of FHDs implies the single FHD* $X : \{Y_1, \ldots, Y_k\}$ *if and only if for each relation* r *with* $X \cup \bigcup_{i=1}^{k} Y_i \cup \bigcup_{j=1}^{n} \left(X_j \cup \bigcup_{s=1}^{l_j} Y_s^j \right) \subseteq Dom(r)$ *the FHD* $X : \{Y_1, \ldots, Y_k\}$ *is satisfied by* r *whenever* r *already satisfies all FHDs in* Σ. $\square$

In this definition, the underlying relation schema is left undetermined. The only requirement is that the FHDs must apply to the relations.

Fact 1. *Let* R *be some relation schema and let* $X \cup \bigcup_{i=1}^{k} Y_i \cup \bigcup_{j=1}^{n} \left(X_j \cup \bigcup_{s=1}^{l_j} Y_s^j \right) \subseteq R$. *Then* $\Sigma = \{X_1 : \{Y_1^1, \ldots, Y_{l_1}^1\}, \ldots, X_n : \{Y_1^n, \ldots, Y_{l_n}^n\}$ *R-implies* $X : \{Y_1, \ldots, Y_k\}$ *whenever* Σ *implies* $X : \{Y_1, \ldots, Y_k\}$.

The converse of Fact 1, however, is false as the following example shows.

Example 4.1. *Let* Σ *consist of the single FHD* Title : {Actor, Feature}*, and let* R *be the relation schema* *with the four attributes* Title, Actor, Feature, Language. *Then* Title : {Actor, Language} *is R-implied by Σ by the soundness of the R-complementation rule. However,* Title : {Actor, Language} *is not implied by Σ as the following counterexample relation r shows.*

Title	*Actor*	*Feature*	*Language*	*Crew*
Musashi	*T. Mifune*	*Trailer*	*English*	*H. Hinagaki*
Musashi	*T. Mifune*	*Trailer*	*Japanese*	*H. Hojo*

While $r = r[\text{Title Actor}] \bowtie r[\text{Title Feature}] \bowtie r[\text{Title Language Crew}]$ *we have* $r \neq r[\text{Title Actor}] \bowtie r[\text{Title Language}] \bowtie r[\text{Title Feature Crew}]$. □

An inference rule is called *sound* if the set of dependencies in the premise of the rule implies the dependency in the conclusion. A set $\mathfrak{S}$ of inference rules is called *sound* for the implication of FHDs if and only if for every finite set Σ of FHDs we have $\Sigma^+_{\mathfrak{S}} \subseteq \Sigma^* = \{\sigma \mid \Sigma \text{ implies } \sigma\}$. The set $\mathfrak{S}$ is called *complete* for the implication of FHDs if and only if for every finite set Σ of FHDs we have $\Sigma^* \subseteq \Sigma^+_{\mathfrak{S}}$. An inference rule $\mathfrak{R}$ is said to be *independent* from the set $\mathfrak{S}$ if and only if there is some finite set $\Sigma \cup \{\sigma\}$ of FHDs such that $\sigma \notin \Sigma^+_{\mathfrak{S}}$, but $\sigma \in \Sigma^+_{\mathfrak{S}\cup\{\mathfrak{R}\}}$. A complete set $\mathfrak{S}$ of inference rules is called *minimal* if and only if every inference rule $\mathfrak{R}$ in $\mathfrak{S}$ is independent from $\mathfrak{S} - \{\mathfrak{R}\}$. This means that no proper subset of $\mathfrak{S}$ is still complete for the implication of FHDs.

While $\mathcal{R}_\emptyset, \mathcal{A}, \mathcal{T}, \mathcal{S}, \mathcal{O}, \mathcal{U}, \mathcal{D}, \mathcal{I}, \mathcal{I}_\emptyset, \mathcal{M}$ are all sound inference rules (since they are R-sound for all R), the *R-axiom* and the R-complementation rule $\mathcal{C}_R$ are R-sound but not sound, see Example 4.1.

We will verify in this section that the set $\mathfrak{H}_C = \{\mathcal{R}_\emptyset, \mathcal{A}, \mathcal{T}, \mathcal{O}, \mathcal{S}, \mathcal{M}\}$ is sound and complete for the implication of FHDs. That is, $\mathfrak{H}_C$ can generate exactly the implications in an undetermined universe.

Lemma 4.1. *Let* $\Sigma = \{X_1 : \{Y_1^1,\dots,Y_{l_1}^1\},\dots,X_n : \{Y_1^n,\dots,Y_{l_n}^n\}$ *be a finite set of FHDs. If* $X : \{Y_1,\dots,Y_k\} \in \Sigma^+_{\mathfrak{H}_C}$, *then for all* $i = 1,\dots,k$ *we have* $Y_i \subseteq \bigcup_{j=1}^{n}\bigcup_{s=1}^{l_j} Y_s^j$.

Proof. We will show by induction on the length l of the inference $\gamma = [\sigma_1,\dots,\sigma_m]$ from Σ by $\mathfrak{H}_C$ such that γ infers $\sigma_m = X : \{Y_1,\dots,Y_k\}$, then for all $i = 1,\dots,k$ we have $Y_i \subseteq \bigcup_{j=1}^{n}\bigcup_{s=1}^{l_j} Y_s^j$.

Let $m = 1$. Then $X : \{Y_1,\dots,Y_k\} = \sigma_1 = \emptyset : \{\emptyset\}$ or $\sigma_1 \in \Sigma$. In each of these cases the desired property is indeed satisfied.

Let $m > 1$. One would need to consider six cases by which σ_m is inferred from $[\sigma_1,\dots,\sigma_{m-1}]$. We only consider the case where we obtain σ_m by applying the *transitivity rule* $\mathcal{T}$ to the premises $\sigma_{i_1}, \sigma_{i_2}$ where $i_1, i_2 < m$. Then the last step of γ has the following form:

$$\frac{X' : \{Y'\}, X'Y' : \{Y_1',\dots,Y_{k-1}'\}}{X' : \{Y_1',\dots,Y_{k-1}',Y'\}}.$$

The induction hypothesis shows $Y_i' \subseteq \bigcup_{j=1}^{n}\bigcup_{s=1}^{l_j} Y_s^j$ for all $i = 1,\dots,k-1$ and $Y' \subseteq \bigcup_{j=1}^{n}\bigcup_{s=1}^{l_j} Y_s^j$. This already proves the condition in this case. The remaining cases are similar. □

Lemma 4.2. *Let* $\Sigma = \{X_1 : \{Y_1^1,\dots,Y_{l_1}^1\},\dots,X_n : \{Y_1^n,\dots,Y_{l_n}^n\}$ *be a finite set of FHDs. Let* $W := \bigcup_{j=1}^{n}(X_j \cup \bigcup_{s=1}^{l_j} Y_s^j)$. *If* $X : \{Y_1,\dots,Y_k\} \in \Sigma^+_{\mathfrak{H}_C}$, *then there is an inference* $\gamma = [\sigma_1,\dots,\sigma_m]$ *of* $X : \{Y_1,\dots,Y_k\}$ *from* Σ *by* $\Sigma^+_{\mathfrak{H}_C}$ *such that any attribute occurring in* $\sigma_1,\dots,\sigma_{m-1}$ *is an element of* W. □

Theorem 4.1. *The set* $\mathfrak{H}_C = \{\mathcal{R}_\emptyset, \mathcal{A}, \mathcal{T}, \mathcal{O}, \mathcal{S}, \mathcal{M}\}$ *is sound and complete for the implication of FHDs.*

Proof. Let $\Sigma = \{X_1 : \{Y_1^1,\dots,Y_{l_1}^1\},\dots,X_n : \{Y_1^n,\dots,Y_{l_n}^n\}$ be a finite set of FHDs, and let $X : \{Y_1,\dots,Y_k\}$ be an FHD. We need to prove that

$$\begin{aligned}&\Sigma \text{ implies } X : \{Y_1,\dots,Y_k\} \text{ if and only if}\\ &X : \{Y_1,\dots,Y_k\} \in \Sigma^+_{\mathfrak{H}_C}\end{aligned} \tag{4.1}$$

For convenience let us define $T := X \cup \bigcup_{i=1}^{k} Y_i \cup \bigcup_{j=1}^{n}\left(X_j \cup \bigcup_{s=1}^{l_j} Y_s^j\right)$. In order to prove the soundness of $\mathfrak{H}_C$ we assume $X : \{Y_1,\dots,Y_k\} \in \Sigma^+_{\mathfrak{H}_C}$. Let r be any relation such that $T \subseteq \mathit{Dom}(r)$ and such that all FHDs $X_i : \{Y_1^i,\dots,Y_{l_i}^i\} \in \Sigma$ are satisfied by r. Then we need to show that r also satisfies $X : \{Y_1,\dots,Y_k\}$.

According to Lemma 4.2 there is an inference γ of $X : \{Y_1,\dots,Y_k\}$ from Σ by $\mathfrak{H}_C$ such that $R \cup \bigcup_{i=1}^{t} S_i \subseteq T \subseteq \mathit{Dom}(r)$ for each FHD $R : \{S_1,\dots,S_t\}$ occurring in γ. Since each rule of $\mathfrak{H}_C$ is sound we can conclude by induction that each FHD occurring in γ is satisfied by r. In particular, r satisfies also $X : \{Y_1,\dots,Y_k\}$.

In order to prove the completeness of $\mathfrak{H}_C$ we assume $X : \{Y_1,\dots,Y_k\} \notin \Sigma^+_{\mathfrak{H}_C}$. Let $R \subseteq \mathfrak{A}$ be a finite set of attributes such that T is a proper subset of R, i.e. $T \subset R$.

Then $R - XY_1\cdots Y_k$ is not a subset of T. Hence, by Lemma 4.1

$$X : \{Y_1,\dots,Y_{i-1},Y_{i+1},\dots,Y_k,R - XY_1\cdots Y_k\} \notin \Sigma^+_{\mathfrak{H}_C}$$

for all $i = 1,\dots,k$. Now from $X : \{Y_1,\dots,Y_k\} \notin \Sigma^+_{\mathfrak{H}_C}$ and

$$X : \{Y_1,\dots,Y_{i-1},Y_{i+1},\dots,Y_k,R - XY_1\cdots Y_k\} \notin \Sigma^+_{\mathfrak{H}_C}$$

for all $i = 1,\dots,k$ we conclude that

$$X : \{Y_1,\dots,Y_k\} \notin \Sigma^+_{R\mathfrak{H}_C}$$

by Corollary 3.2.

Since $R\mathfrak{H}_C$ is complete for the R-implication of FHDs it follows that Σ does not R-imply $X : \{Y_1,\dots,Y_k\}$. Consequently, Σ does not imply $X : \{Y_1,\dots,Y_k\}$ by Fact 1. □

The following lemmata show that the Boolean rules are indeed derivable from $\mathfrak{H}_C$.

Lemma 4.3. *The* union rule $\mathcal{U}$ *is derivable from* $\{\mathcal{A}, \mathcal{T}, \mathcal{M}\}$.

Proof.

$$\cfrac{X : \{Z\} \qquad \cfrac{X : \{Y_1,\dots,Y_k\}}{XZ : \{Y_1 - Z,\dots,Y_k - Z\}}\,\mathcal{A}}{\cfrac{X : \{Y_1 - Z,\dots,Y_k - Z, Z\}}{X : \{Y_1 - Z,\dots,Y_{k-1} - Z, Y_kZ\}}\,\mathcal{M}}\,\mathcal{T}$$

□

Lemma 4.4. *The* intersection rule $\mathcal{I}$ *is derivable from* $\{\mathcal{A}, \mathcal{T}, \mathcal{M}, \mathcal{O}, \mathcal{S}\}$. ☐

Lemma 4.5. *The* difference rule $\mathcal{D}$ *is derivable from* $\{\mathcal{A}, \mathcal{T}, \mathcal{M}, \mathcal{O}\}$. ☐

5 The minimality of $\mathfrak{H}_C$ and $R\mathfrak{H}_C$

It is now the goal to demonstrate the minimality of $\mathfrak{H}_C = \{\mathcal{R}_\emptyset, \mathcal{A}, \mathcal{T}, \mathcal{O}, \mathcal{S}, \mathcal{M}\}$ for the implication of FHDs. As a consequence the system $R\mathfrak{H}_C$ is minimal for the R-implication of FHDs in the sense that none of its subsets is still both complete and complementary for the R-implication of FHDs.

For the following independence proof sketches one can utilise Lemma 4.1 which restricts the number of possible FHDs that can be inferred from Σ by $\mathfrak{H}_C$.

Lemma 5.1. *The* empty-set-axiom $R_\emptyset$ *is independent from* $\{\mathcal{A}, \mathcal{T}, \mathcal{O}, \mathcal{S}, \mathcal{M}\}$.

Proof. Let $\Sigma = \emptyset$ and $\sigma = \emptyset : \{\emptyset\}$. Then $\sigma \notin \Sigma^+_{\{\mathcal{A}, \mathcal{T}, \mathcal{O}, \mathcal{S}, \mathcal{M}\}}$, but $\sigma \in \Sigma^+_{\mathfrak{H}_C}$. ☐

Lemma 5.2. *The* augmentation rule $\mathcal{A}$ *is independent from* $\{R_\emptyset, \mathcal{T}, \mathcal{O}, \mathcal{S}, \mathcal{M}\}$.

Proof. Let $\Sigma = \emptyset$ and $\sigma = A : \{\emptyset\}$. One can show that $\sigma \notin \Sigma^+_{\{R_\emptyset, \mathcal{T}, \mathcal{O}, \mathcal{S}, \mathcal{M}\}} = \{\emptyset : \{\emptyset\}\}$ but $\sigma \in \Sigma^+_{\mathfrak{H}_C}$. ☐

Lemma 5.3. *The* transitivity rule $\mathcal{T}$ *is independent from* $\{R_\emptyset, \mathcal{A}, \mathcal{O}, \mathcal{S}, \mathcal{M}\}$.

Proof. In this case we choose $\Sigma = \{A : \{B\}, AB : \{C\}\}$, and $\sigma = A : \{C\}$. One can show that $\sigma \notin \Sigma^+_{\{R_\emptyset, \mathcal{A}, \mathcal{O}, \mathcal{S}, \mathcal{M}\}}$ but $\sigma \in \Sigma^+_{\mathfrak{H}_c}$. ☐

Lemma 5.4. *The* omission rule $\mathcal{O}$ *is independent from* $\{R_\emptyset, \mathcal{A}, \mathcal{T}, \mathcal{S}, \mathcal{M}\}$.

Proof. We choose $\Sigma = \{\emptyset : \{A, B\}\}$, and $\sigma = \emptyset : \{A\}$. One can show that $\sigma \notin \Sigma^+_{\{R_\emptyset, \mathcal{A}, \mathcal{T}, \mathcal{S}, \mathcal{M}\}}$ but $\sigma \in \Sigma^+_{\mathfrak{H}_c}$. ☐

Lemma 5.5. *The* subset rule $\mathcal{S}$ *is independent from* $\{R_\emptyset, \mathcal{A}, \mathcal{T}, \mathcal{O}, \mathcal{M}\}$.

Proof. We choose $\Sigma = \{\emptyset : \{AB\}, C : \{B\}\}$ and $\sigma = \emptyset : \{B\}$. One can show that $\sigma \notin \Sigma^+_{\{R_\emptyset, \mathcal{A}, \mathcal{T}, \mathcal{O}, \mathcal{M}\}}$ but $\sigma \in \Sigma^+_{\mathfrak{H}_c}$. ☐

Lemma 5.6. *The* merging rule $\mathcal{M}$ *is independent from* $\{R_\emptyset, \mathcal{A}, \mathcal{T}, \mathcal{O}, \mathcal{S}\}$.

Proof. Let $\Sigma = \{\emptyset : \{A, B\}\}$ and $\sigma = \emptyset : \{AB\}$. One can show that $\sigma \notin \Sigma^+_{\{R_\emptyset, \mathcal{A}, \mathcal{T}, \mathcal{O}, \mathcal{S}\}}$ but $\sigma \in \Sigma^+_{\mathfrak{H}_c}$. ☐

Theorem 5.1. *The set* $\mathfrak{H}_C = \{\mathcal{R}_\emptyset, \mathcal{A}, \mathcal{T}, \mathcal{O}, \mathcal{S}, \mathcal{M}\}$ *is minimal for the implication of FHDs.* ☐

Since the *R-complementation rule* $\mathcal{C}_R$ is independent from the rules in $\mathfrak{H}_C$ Theorem 5.1 implies that none of the proper subsets of $R\mathfrak{H}_C$ can still be complete and complementary for the R-implication of FHDs.

Corollary 5.1. *None of the proper subsets of* $R\mathfrak{H}_C = \{\mathcal{R}_\emptyset, \mathcal{A}, \mathcal{T}, \mathcal{O}, \mathcal{S}, \mathcal{M}, \mathcal{C}_R\}$ *is still both complete and complementary for R-implication of FHDs.* ☐

Finally, we show the independence of the *union rule* $\mathcal{U}$ from $\{\mathcal{R}_\emptyset, \mathcal{A}, \mathcal{T}, \mathcal{O}\}$. This completes Example 3.1.

Lemma 5.7. *The* union rule $\mathcal{U}$ *is independent from* $\{\mathcal{R}_\emptyset, \mathcal{A}, \mathcal{T}, \mathcal{O}\}$.

Proof. Let $\Sigma = \{\emptyset : \{A\}, \emptyset : \{B\}\}$ and $\sigma = \emptyset : \{AB\}$. The closure $\Sigma^+_{\{\mathcal{R}_\emptyset, \mathcal{A}, \mathcal{T}, \mathcal{O}\}}$ can be obtained as follows. The *empty-set-axiom* $\mathcal{R}_\emptyset$ yields the FHD $\emptyset : \{\emptyset\}$. A subsequent applications of the *augmentation rule* $\mathcal{A}$ allows us to derive the $\emptyset$-column. One may then enter the FHDs from Σ and apply the *augmentation rule* $\mathcal{A}$ subsequently to derive the $\{A\}$-column and $\{B\}$-column. One can then apply the *transitivity rule* $\mathcal{T}$ to $\emptyset : \{A\}$ and $A : \{B\}$ to infer $\emptyset : \{A, B\}$. This set is closed under derivation using $\{\mathcal{R}_\emptyset, \mathcal{A}, \mathcal{T}, \mathcal{O}\}$.

	$\{\emptyset\}$	$\{A\}$	$\{B\}$	$\{AB\}$	$\{A, B\}$
$\emptyset$	×	×	×		×
A	×	⋆	×	⋆	⋆
B	×	×	⋆	⋆	⋆
AB	×	⋆	⋆	⋆	⋆

It follows that $\sigma \notin \Sigma^+_{\{\mathcal{R}_\emptyset, \mathcal{A}, \mathcal{T}, \mathcal{O}\}}$ but $\sigma \in \Sigma^+_{\{\mathcal{R}_\emptyset, \mathcal{A}, \mathcal{T}, \mathcal{O}, \mathcal{U}\}}$. ☐

6 Future Work

We conclude this paper by listing some related problems that warrant further research.

The *subset rule* plays a key role in achieving complementarity for MVDs and FHDs. It would be interesting to see whether there are any complete sets of inference rules for the implication of MVDs (FHDs) in undetermined universes that do not feature the subset rule.

While FDs and MVDs have been investigated before and an axiomatisation is well-known for the class of both types of dependencies in fixed universes (Beeri et al. 1977) the combined class of FDs and FHDs should also be studied in both fixed and undetermined universes.

According to (Link 2006*a*) there seems to be a trade-off between minimality and complementarity for MVDs. That is, so far no minimal complete set of inference rules for the R-implication of MVDs has been identified that is also complementary. The question is whether there is any such system.

MVDs have been studied in the presence of null values, for instance with interpretation *no information* (Lien 1982, Link 2006*b*). Interestingly, the soundness of the *transitivity rule* fails when database relations are allowed to be incomplete. FHDs should therefore also be studied in the presence of null values. An interesting approach to incomplete data has been applied to the class of functional dependencies (Levene & Loizou 1998). There, a possible world semantics is used to explore all possible extensions of an incomplete database to a complete database. While weak FDs are satisfied by some possible world, strong FDs must be satisfied by all possible worlds. It would be interesting to generalise this work to MVDs (FHDs).

An open problem is the lack of a synthesis algorithm for MVDs that would extend the well-known synthesis algorithm for the class of FDs (Bernstein 1976, Biskup, Dayal & Bernstein 1979). The notion of MVD implication in undetermined universes provides an alternative basis for the formulation of such an algorithm.

References

Arenas, M. & Libkin, L. (2004), 'A normal form for XML documents', *ToDS* **29**(1), 195–232.

Armstrong, W. W. (1974), 'Dependency structures of database relationships', *Information Processing* **74**, 580–583.

Armstrong, W. W., Nakamura, Y. & Rudnicki, P. (2002), 'Armstrong's axioms', *Journal of formalized Mathematics* **14**.

Beeri, C. (1980), 'On the membership problem for functional and multivalued dependencies in relational databases', *ToDS* **5**(3), 241–259.

Beeri, C. & Bernstein, P. A. (1979), 'Computational problems related to the design of normal form relational schemata', *ToDS* **4**(1), 30–59.

Beeri, C., Fagin, R. & Howard, J. H. (1977), A complete axiomatization for functional and multivalued dependencies in database relations, *in* 'SIGMOD', ACM, pp. 47–61.

Bernstein, P. (1976), 'Synthesizing third normal form relations from functional dependencies', *ToDS* **1**(4), 277–298.

Bernstein, P. A. & Goodman, N. (1980), What does Boyce-Codd normal form do?, *in* 'VLDB', IEEE Computer Society, pp. 245–259.

Biskup, J. (1978), 'On the complementation rule for multivalued dependencies in database relations', *Acta Informatica* **10**(3), 297–305.

Biskup, J. (1980), 'Inferences of multivalued dependencies in fixed and undetermined universes', *TCS* **10**(1), 93–106.

Biskup, J., Dayal, U. & Bernstein, P. (1979), Synthesizing independent database schemas, *in* 'SIGMOD', pp. 143–151.

Codd, E. F. (1970), 'A relational model of data for large shared data banks', *Commun. ACM* **13**(6), 377–387.

Codd, E. F. (1972), Further normalization of the database relational model, *in* 'Courant Computer Science Symposia 6: Data Base Systems', Prentice-Hall, pp. 33–64.

Delobel, C. (1978), 'Normalisation and hierarchical dependencies in the relational data model', *ToDS* **3**(3), 201–222.

Fagin, R. (1977), 'Multivalued dependencies and a new normal form for relational databases', *ToDS* **2**(3), 262–278.

Fagin, R. & Vardi, M. Y. (1986), The theory of data dependencies: a survey, *in* 'Mathematics of Information Processing: Proceedings of Symposia in Applied Mathematics', American Mathematical Society, pp. 19–71.

Fischer, P. C., Saxton, L. V., Thomas, S. J. & Van Gucht, D. (1985), 'Interactions between dependencies and nested relational structures', *J. Comput. Syst. Sci.* **31**(3), 343–354.

Graetzer, G. (1998), *General Lattice Theory*, Birkhauser.

Hara, C. & Davidson, S. (1999), Reasoning about nested functional dependencies, *in* 'PoDS', ACM, pp. 91–100.

Hartmann, S. & Link, S. (2004), Multi-valued dependencies in the presence of lists, *in* 'PoDS', ACM, pp. 330–341.

Hartmann, S. & Link, S. (2006), 'On a problem of Fagin concerning multivalued dependencies in relational databases', *TCS* **353**(1-3), 53–62.

Hartmann, S., Link, S. & Schewe, K.-D. (2006), 'Functional and multivalued dependencies in nested databases generated by record and list constructor', *AMAI* **46**(1-2), 114–164.

Levene, M. & Loizou, G. (1998), 'Axiomatisation of functional dependencies in incomplete relations', *TCS* **206**(1-2), 283–300.

Lien, Y. E. (1982), 'On the equivalence of data models', *Journal of the ACM* **29**(2), 333–363.

Link, S. (2006*a*), On multivalued dependencies in fixed and undetermined universes, *in* 'FoIKS', number 3861 *in* 'Lecture Notes in Computer Science', pp. 257–276.

Link, S. (2006*b*), On the logical implication of multivalued dependencies with null values, *in* 'Twelfth Computing: The Australasian Theory Symposium', number 51 *in* 'CRPIT', pp. 113–122.

Mendelzon, A. (1979), 'On axiomatising multivalued dependencies in relational databases', *J. ACM* **26**(1), 37–44.

Tari, Z., Stokes, J. & Spaccapietra, S. (1997), 'Object normal forms and dependency constraints for object-oriented schemata', *ToDS* **22**, 513–569.

Thalheim, B. (1991), *Dependencies in Relational Databases*, Teubner-Verlag.

Thalheim, B. (2000), *Entity-Relationship Modeling: Foundations of Database Technology*, Springer.

Thalheim, B. (2003), Conceptual treatment of multivalued dependencies, *in* 'Conceptual Modeling - ER 2003, Proceedings', number 2813 *in* 'Lecture Notes in Computer Science', Springer, pp. 363–375.

Vardi, M. Y. (1987), Fundamentals of dependency theory, *in* E. Börger, ed., 'Trends in Theoretical Computer Science', Computer Science Press, pp. 171–224.

Vincent, M. & Liu, J. (2003), Multivalued dependencies in XML, *in* 'BNCOD', number 2712 *in* 'Lecture Notes in Computer Science', Springer, pp. 4–18.

Vincent, M., Liu, J. & Liu, C. (2003), A redundancy free 4NF for XML, *in* 'XMLSym', number 2824 *in* 'Lecture Notes in Computer Science', Springer, pp. 254–266.

Weddell, G. (1992), 'Reasoning about functional dependencies generalized for semantic data models', *ToDS* **17**(1), 32–64.

Wijsen, J. (1999), 'Temporal FDs on complex objects', *ToDS* **24**(1), 127–176.

Zaniolo, C. (1976), Analysis and Design of Relational Schemata for Database Systems, PhD thesis, UCLA, Tech. Rep. UCLA-ENG-7769.