Ann Math Artif Intell (2007) 50:195–226
DOI 10.1007/s10472-007-9067-0

Full hierarchical dependencies in fixed and undetermined universes

Sven Hartmann · Henning Köhler · Sebastian Link

Published online: 19 July 2007

Abstract Full hierarchical dependencies (FHDs) constitute a large class of relational dependencies. A relation exhibits an FHD precisely when it is the natural join over *at least* two of its projections that all share the same join attributes. Therefore, FHDs generalise multivalued dependencies (MVDs) in which case the number of these projections is precisely two. The implication of FHDs has originally 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 first main result establishes a finite axiomatisation for FHD implication in undetermined universes. It is then formally clarified that the complementation rule is only a mere means for database normalisation. In fact, the second main result establishes a finite axiomatisation for FHD implication in fixed universes which allows to infer 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.

A preliminary version of this paper was presented at FoIKS 2006, Budapest, Hungary.

This research was supported by Marsden Funding, Royal Society of New Zealand.

S. Hartmann (✉) · H. Köhler · S. Link
Information Science Research Centre, Dept. of Information Systems, Massey University,
Palmerston North, New Zealand
e-mail: s.hartmann@massey.ac.nz

H. Köhler
e-mail: h.koehler@massey.ac.nz

S. Link
e-mail: s.link@massey.ac.nz

Keywords Relational database · Full hierarchical dependency · Implication · Inference

Mathematics Subject Classification (2000) 68P15

1 Introduction

Relational databases still form the core of most database management systems, even after more than three decades following their introduction in Codd [13]. 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 [17, 32, 35].

Multivalued dependencies (MVDs) [16, 40] are an important class of dependencies. A relation exhibits an MVD precisely when it is the natural join over exactly two of its projections that both share the same join attributes [16]. This property is fundamental to relational database design, in particular 4NF [16], and a lot of research has been devoted to studying the behaviour of these dependencies. Recently, extensions of MVDs have been found very useful for various design problems in advanced data models such as the nested relational data model [18], the Entity-Relationship model [33], data models that support nested lists [21, 23] and XML [36, 37].

Full hierarchical dependencies (FHDs), called full first-order hierarchical decompositions in [15], 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 that all share the same join attributes. The classical notion of an FHD [15] is dependent on the underlying universe R. For MVDs [16] their dependence on the relation schema R is syntactically reflected by the R-complementation rule which is part of the axiomatisation of MVDs [6]. The R-complementation rule is special in the sense that it is the only inference rule in this axiomatisation 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 [10]. 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 $\mathfrak{S}_C$ results from adding the R-complementation rule to $\mathfrak{S}_0$, then $\mathfrak{S}_C$ is R-sound and R-complete for the R-implication of MVDs for all relation schemata R. In fact, every inference of an MVD by $\mathfrak{S}_C$ 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 ($\mathfrak{S}_C$ 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 [18, 21, 26, 34, 36, 37] 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.

1.1 Contributions

In this paper we will extend the work by Biskup [10] 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 [6]. In particular, the dependence of 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*. We would like to model that the title of a DVD determines the set of actors independently from the rest of the information in any schema, and that the title of a DVD also determines the set of DVD features independently from the rest of the information in any schema. In order to enforce such 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) we decide to specify the FHD

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

Table 1 shows an example of an R-relation that satisfies this FHD and where the underlying relation schema R consists of the four attributes above. Note that the FHD is equivalent to the two MVDs

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

If R is fixed as above, then we may infer the FHDs

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

Table 1 A relation satisfying the FHD Title:{{Actor},{Feature}}

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

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

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

are R-implied by $Title : \{\{Actor\}, \{Feature\}\}$. In sharp contrast to the latter FHD 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 [9].

One may argue that consequences that depend on the underlying universe are no consequences at all. This, however, implies that the notion of R-implication is not acceptable. Thus, we extend the alternative notion of implication from MVDs [10] 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

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

are R-implied by

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

where R={Title,Actor,Feature,Language}, but none of the two FHDs is implied.

1.2 Previous work

The first axiomatisation for the R-implication of MVDs was given in Beeri et al. [6]. The notion of implication in undetermined universes is from Biskup [10] in which an axiomatisation for this notion of MVD implication is given. Minimality of MVD axiomatisations are discussed in Biskup [9] and Mendelzon [28] in which Biskup [9] 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 [22, 26]. MVDs have also been studied in the presence of the null value *no information*. In that case, Lien [25] was the first to propose an axiomatisation in fixed universes, and Link [27] 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. These data formats include Entity-Relationship models [34], nested data models [23], fuzzy data models [29, 30] and XML [36, 37]. FHDs were introduced in [15]. An axiomatisation for the implication in fixed universes can be found in Thalheim [32, 33].

1.3 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 FHDs 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 $X, Y \subseteq R$, finite $r_1 \subseteq dom(X)$ 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 attributes have played a central role in the study of relational databases [5, 7, 8, 12–14], and seem to be central for the study of database design in other data models as well [1, 20, 24, 26, 31, 38, 39]. The notion of a FD is well-understood and the semantic interaction between these dependencies has been syntactically captured by Armstrong's well-known axioms [2, 3]. A FD [14] 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. MVDs provide a more general notion and offer a response to the shortcomings of FDs. A MVD [16, 40] 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 $t[X(R - XY)] = t_2[X(R - XY)]$. 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 $R - XY$. This actually suggests that the relation schema R is overloaded in the sense that it carries two independent facts XY and $X(R - XY)$. More precisely, it is shown in [16] 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 $r = r[XY] \bowtie r[X(R - XY)]$. This characteristic of MVDs is fundamental to relational database design and 4NF [16]. A lot of research has therefore been devoted to studying the behaviour of these dependencies.

FHDs generalise MVDs [15].

Definition 2.1 A FHD on a 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 FHD $X : \{Y_1, \ldots, Y_k\}$ 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 what Delobel originally introduced as full first-order hierarchical decompositions [15]. In fact, a full first-order hierarchical decomposition over the relation schema R 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 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. In this sense, our FHDs are implicitly full. In particular, notice the equivalence between the full first-order hierarchical decomposition $X : Y \mid R - XY$, the FHD $X : \{Y\}$ and the MVD $X \twoheadrightarrow Y$. Secondly, the right-hand 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 $Y_1, \ldots, Y_k$ that indicates the order in which R is successively decomposed, as there is in [15]. These two differences result in a simpler axiomatisation and the correspondence to the original definition of MVDs is much stronger [16]. The following result is a straightforward generalisation from the MVD case [16]. We omit the proof.

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 R from Example 1.1 comprising the four attributes *Title*, *Actor*, *Feature* and *Language*. Recall the R-relation r from Table 1 satisfying the FHD

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

Table 2 shows 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.

The next result shows essentially that FHDs summarise MVDs with the same left-hand side [15]. Since we represent FHDs differently from their original definition [15] and since we will make extensive use of this result we include a proof.

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 $l = 1, \ldots, k$, r satisfies the MVD $X \twoheadrightarrow Y_l$ on R.*

Proof Let r satisfy the FHD $X : \{Y_1, \ldots, Y_k\}$ on R. We show that r also satisfies $X \twoheadrightarrow Y_l$ for all $l = 1, \ldots, k$. Let l_0 be arbitrary with $1 \leq l_0 \leq k$, and let $t_1, t_2 \in r$ such that $t_1[X] = t_2[X]$ holds. For $j = 1, \ldots, k+1$ let

$$t'_j := \begin{cases} t_1 & , \text{if } j = l_0 \\ t_2 & \text{otherwise} \end{cases} .$$

Table 2 Some projections of the relation from Table 1

Title	Feature	Language
King Kong	Deleted Scenes	English
King Kong	Photo Gallery	French
King Kong	Deleted Scenes	French
King Kong	Photo Gallery	French

Title	Actor	Language
King Kong	Naomi Watts	English
King Kong	Jack Black	French
King Kong	Naomi Watts	French
King Kong	Jack Black	English

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

Then $t'_i[X] = t'_j[X]$ for all $1 \leq i \leq j \leq k+1$. Since r satisfies $X : \{Y_1, \ldots, Y_k\}$ there is some $t \in r$ such that $t[XY_j] = t'_j[XY_j]$ holds for $j = 1, \ldots, k$, and $t[X(R - XY_1 \cdots Y_k)] = t_{k+1}[X(R - XY_1 \cdots Y_k)]$. According to the definition of t'_j we have $t[XY_{l_0}] = t'_{l_0}[XY_{l_0}] = t_1[XY_{l_0}]$ and $t[X(R - XY_{l_0})] = t_2[X(R - XY_{l_0})]$. Therefore, r also satisfies $X \twoheadrightarrow Y_{l_0}$.

Let r satisfy the MVD $X \twoheadrightarrow Y_l$ on R, for $l = 1, \ldots, k$. We show that r also satisfies the FHD $X : \{Y_1, \ldots, Y_k\}$. Therefore, let $t_1, \ldots, t_{k+1} \in r$ such that $t_i[X] = t_j[X]$ holds for $1 \leq i, j \leq k$. Since $t_1[X] = t_2[X]$ holds and r satisfies $X \twoheadrightarrow Y_1$ there is some $t'_1 \in r$ such that $t'_1[XY_1] = t_1[XY_1]$ and $t'_1[X(R - XY_1)] = t_2[X(R - XY_1)]$. Since $t'_1[X] = t_3[X]$ holds and r satisfies $X \twoheadrightarrow Y_1Y_2$ (by the soundness of the union rule for MVDs) there is some $t'_2 \in r$ such that $t'_2[XY_1Y_2] = t'_1[XY_1Y_2]$ and $t'_2[X(R - XY_1Y_2)] = t_3[X(R - XY_1Y_2)]$. Continuing this we know that $t'_{k-1}[X] = t_{k+1}[X]$ holds. Since r satisfies $X \twoheadrightarrow Y_1 \cdots Y_k$ there is some $t'_k \in r$ such that $t'_k[XY_1 \cdots Y_k] = t'_{k-1}[XY_1 \cdots Y_k]$ and $t'_k[X(R - XY_1 \cdots Y_k)] = t_{k+1}[X(R - XY_1 \cdots Y_k)]$. Let $t := t'_k$. Then, $t[XY_l] = t_l[XY_l]$ for all $l = 1, \ldots, k$ and $t[X(R - XY_1 \cdots Y_k)] = t_{k+1}[X(R - XY_1 \cdots Y_k)]$. Consequently, r also satisfies the FHD $X : \{Y_1, \ldots, Y_k\}$. □

Notice that hierarchical dependencies permit a *stepwise* decomposition of the underlying relation schema by splitting one current component into two components. This feature distinguishes hierarchical dependencies from general join dependencies which do not have this property [32].

Example 2.2 Suppose we have the DVD schema

$$R = \{Title, Actor, Feature, Language\}$$

together with the FHD $Title : \{\{Actor\}, \{Feature\}\}$. In a first step we may decompose R into

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

based on the set-valued correspondence between actors and DVD titles. The corresponding projections of the relation r in Table 1 can be seen in Table 2. In a second step we may decompose $\{Title, Feature, Language\}$ into $\{Title, Feature\}$ and $\{Title, Language\}$.

Alternatively, one may first choose to decompose R into

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

based on the set-valued correspondence between features and DVD titles.

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 $\Sigma = \{X_1 : \{Y_1^1, \ldots, Y_{l_1}^1\}, \ldots, X_n : \{Y_1^n, \ldots, Y_{l_n}^n\}\}$ and $X : \{Y_1, \ldots, Y_k\}$ be FHDs on the relation schema R, in particular

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

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

Notice that Definition 2.2 covers MVDs in case that $l_j = 1$ for $j = 1, \ldots, n$ and $k = 1$ [10]. In order to determine the logical consequences of a set of MVDs with respect to R-implication one can use the following set of inference rules [6]. 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 [9] instead of the complementation rule that was originally proposed [6].

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

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

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

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

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

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

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

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

Let R be some arbitrary relation schema. The set

$$\left\{\mathcal{R}_{\text{MVD}}, \mathcal{A}_{\text{MVD}}, \mathcal{T}_{\text{MVD}}, \mathcal{C}_{\text{MVD}}^R, \mathcal{U}_{\text{MVD}}, \mathcal{D}_{\text{MVD}}, \mathcal{I}_{\text{MVD}}\right\}$$

is both R-sound and R-complete for the R-implication of MVDs [6]. In what follows we use $\mathcal{C}$ to denote either the class of MVDs or the class of FHDs. Furthermore, we use $\mathfrak{S}$ to denote a set of inference rules. Let $\Sigma \cup \{\sigma\}$ be a set of dependencies from $\mathcal{C}$ on the relation schema R. Let $\Sigma \vdash_{\mathfrak{S}} \sigma$ denote the inference of σ from Σ with respect to $\mathfrak{S}$. Let $\Sigma_{\mathfrak{S}}^+ = \{\sigma \mid \Sigma \vdash_{\mathfrak{S}} \sigma\}$ denote the *syntactic hull* of Σ under inference using only rules from $\mathfrak{S}$. 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 [6]. The set $\mathfrak{S}$ 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 $\Sigma^+_{\mathfrak{S}} \subseteq \Sigma^*_R = \{\sigma \in \mathcal{C} \mid \Sigma\ R\text{-implies}\ \sigma\}$. The set $\mathfrak{S}$ 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 $\Sigma^*_R \subseteq \Sigma^+_{\mathfrak{S}}$.

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 all relation schemata R. More precisely, an inference rule $\mathfrak{R}$ is said to be *independent* from the set $\mathfrak{S}$ if and only if there is some relation schema R and some finite set $\Sigma \cup \{\sigma\}$ of dependencies from $\mathcal{C}$ on R such that $\sigma \notin \Sigma^+_{\mathfrak{S}}$, but $\sigma \in \Sigma^+_{\mathfrak{S}\cup\{\mathfrak{R}\}}$. Let $\mathfrak{S}$ be a set of inference rules that is R-complete for the R-implication of dependencies from $\mathcal{C}$ for all relation schemata R. Then $\mathfrak{S}$ is said to be *minimal* for the R-implication of dependencies from $\mathcal{C}$ if and only if every inference rule $\mathfrak{R} \in \mathfrak{S}$ is independent of $\mathfrak{S} - \{\mathfrak{R}\}$. This means that no proper subset of $\mathfrak{S}$ is still R-complete for the R-implication of dependencies in $\mathcal{C}$ for all relation schemata R. It was shown by Mendelzon [28] that $\mathfrak{M} = \{\mathcal{R}_{\text{MVD}}, \mathcal{C}^R_{\text{MVD}}, \mathcal{T}_{\text{MVD}}\}$ forms such a minimal set for the R-implication of MVDs. In the same paper, Mendelzon further motivates the study of the independence of inference rules and comments in more detail on the special role of the R-complementation rule.

2.1 Sound inference rules for FHDs in fixed universes

We will denote the set of the following inference rules by $\mathfrak{H}$.

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

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

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

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

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

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

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

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

Theorem 2.3 *For all relation schemata R the set $\mathfrak{H}$ is R-sound for the R-implication of FHDs.*

Proof We need to show that for an arbitrary relation schema R, and an arbitrary set Σ of FHDs on R we have $\Sigma^+_{\mathfrak{H}} \subseteq \Sigma^*_R$. The proof will make extensive use of Theorem 2.2 and the R-soundness of the inference rules for MVDs [6]. 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 $\mathfrak{H}$ is also R-implied by Σ.

The soundness of the *empty-set-axiom* $\mathcal{R}_\emptyset$ follows from Theorem 2.2 and the soundness of the reflexivity axiom $\mathcal{R}_{\text{MVD}}$ for MVDs.

Consider now the *augmentation rule* $\mathcal{A}$, and let r be some arbitrary R-relation that satisfies $X : \{Y_1, \ldots, Y_k\}$. Theorem 2.2 tells us that r also satisfies $X \twoheadrightarrow Y_i$ for all $i = 1, \ldots, k$. The soundness of the augmentation rule $\mathcal{A}_{\text{MVD}}$ shows that r also satisfies $XZ \twoheadrightarrow Y_i$ for all $i = 1, \ldots, k$, 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 $i = 1, \ldots, k$. Consequently, $\models_r XZ : \{Y_1 - Z, \ldots, Y_k - Z\}$ by Theorem 2.2.

Next we show the soundness of the *transitivity rule* $\mathcal{T}$. Let r be some arbitrary R-relation that satisfies $XY : \{Y_1, \ldots, Y_k\}$ and $X : Y$. Note that every Y_i is disjoint from XY. Theorem 2.2 shows that r satisfies $XY \twoheadrightarrow Y_i$ for all $i = 1, \ldots, k$ and r satisfies $X \twoheadrightarrow Y$. We conclude by soundness of the augmentation rule $\mathcal{A}_{\text{MVD}}$ that $\models_r X \twoheadrightarrow XY$. We apply the soundness of the pseudo-transitivity rule $\mathcal{T}_{\text{MVD}}$ and conclude that r satisfies $X \twoheadrightarrow Y_i - XY$ for every $i = 1, \ldots, k$. However, $Y_i - XY = Y_i$ since Y_i and XY are disjoint. It is now a consequence of Theorem 2.2 that r satisfies $X : \{Y_1, \ldots, Y_k, Y\}$.

Next we show the soundness of the *omission rule* $\mathcal{O}$. Let r be some arbitrary R-relation that satisfies $X : \{Y_1, \ldots, Y_k, Y\}$. Theorem 2.2 tells us that r must satisfy $X \twoheadrightarrow Y_i$ for all $i = 1, \ldots, k$, and r must satisfy $X \twoheadrightarrow Y$. Consequently, $\models_r X : \{Y_1, \ldots, Y_k\}$ by Theorem 2.2.

The soundness of the *R-complementation rule* $\mathcal{C}_R$ is an immediate consequence of Theorem 2.2. In fact, if r satisfies the FHD $X : \{Y_1, \ldots, Y_k\}$, then this is equivalent to r satisfying

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

However, this means r is a model of $X : \{Y_1, \ldots, Y_{k-1}, R - XY_1 \cdots Y_k\}$ since $R - (XY_1 \cdots Y_{k-1}(R - XY_1 \cdots Y_k)) = Y_k$.

Consider now the *union rule* $\mathcal{U}$, and let r be some arbitrary R-relation that satisfies $X : \{Y_1, \ldots, Y_k\}$ and $X : \{Z\}$. We know by Theorem 2.2 that r satisfies $X \twoheadrightarrow Y_i$ for all $i = 1, \ldots, k$, and r satisfies $X \twoheadrightarrow Z$. We conclude by the soundness of the union rule $\mathcal{U}_{\text{MVD}}$ that $\models_r X \twoheadrightarrow Y_k Z$. Moreover, the soundness of the difference rule $\mathcal{D}_{\text{MVD}}$ implies that r satisfies $X \twoheadrightarrow Y_i - Z$ for all $i = 1, \ldots, k-1$. In view of Theorem 2.2 this shows that r satisfies $X : \{Y_1 - Z, \ldots, Y_{k-1} - Z, Y_k Z\}$.

For the soundness of the *difference rule* $\mathcal{D}$ assume that the arbitrary R-relation r satisfies both $X : \{Y_1, \ldots, Y_k\}$ and $X : \{Z\}$. We know by Theorem 2.2 that r satisfies $X \twoheadrightarrow Y_i$ for all $i = 1, \ldots, k$, and r satisfies $X \twoheadrightarrow Z$. According to the soundness of the difference rule $\mathcal{D}_{\text{MVD}}$ we also have $\models_r X \twoheadrightarrow Y_k - Z$. Theorem 2.2 shows now that r is a model of $X : \{Y_1, \ldots, Y_{k-1}, Y_k - Z\}$.

Finally, we prove the soundness of the *intersection rule* $\mathcal{I}$. Let r be some arbitrary R-relation that satisfies $X : \{Y_1, \ldots, Y_k\}$ and $X : \{Z\}$. Theorem 2.2 implies that r satisfies $X \twoheadrightarrow Y_i$ for all $i = 1, \ldots, k$, and r satisfies $X \twoheadrightarrow Z$. According to the soundness of the intersection rule $\mathcal{I}_{\text{MVD}}$ we also have $\models_r X \twoheadrightarrow Y_k \cap Z$. Theorem 2.2 shows then that r is indeed a model of $X : \{Y_1, \ldots, Y_{k-1}, Y_k \cap Z\}$. □

Lemma 2.1 *The following inference rules are derivable from* $\mathfrak{H}$

$$\frac{X:\{Y_1,\ldots,Y_k\}}{X:\{Y_1,\ldots,Y_k,\emptyset\}}\ \textit{(empty-set-introduction, } \mathcal{I}_\emptyset\textit{)} \qquad \frac{X:\{Y_1,\ldots,Y_k,Y_{k+1}\}}{X:\{Y_1,\ldots,Y_kY_{k+1}\}}\ \textit{(merging, } \mathcal{M}\textit{)}$$

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

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

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

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

$$\dfrac{\dfrac{X:\{Y_1,\ldots,Y_k,Y_{k+1}\}}{X:\{Y_1,\ldots,Y_k\}}\,\mathcal{O} \quad \dfrac{X:\{Y_1,\ldots,Y_k,Y_{k+1}\}}{X:\{Y_{k+1}\}}\,\mathcal{O}}{X:\{Y_1,\ldots,Y_{k-1},Y_kY_{k+1}\}}\,\mathcal{U}$$

This concludes the proof. □

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 $W\subseteq R-X$ for which some FHD $X:S$ with $W\in S$ is derivable from Σ using the inference rules $\mathfrak{H}$, i.e., $Dep_R(X)=\{W\subseteq R-X\mid \exists S \text{ such that } X:S\in\Sigma^+_{\mathfrak{H}} \text{ and } W\in S\}$. Note that $Dep_R(X)$ is finite, and $(Dep_R(X),\subseteq,\cup,\cap,(\cdot)^{\mathcal{C}},\emptyset,R-X)$ 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 $(P,\sqsubseteq,0)$ [19] 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 [4].

Theorem 2.4 *Let* $\Sigma\cup\{X:S\}$ *be a set of FHDs on the relation schema R. Then* $X:S\in\Sigma^+_{\mathfrak{H}}$ *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 $X:S\in\Sigma^+_{\mathfrak{H}}$. 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 $Y\in S$ be arbitrary and suppose that $Y=\bigcup\mathcal{Y}$ 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 $Y\in Dep_R(X)$. Let $S=\{Y_1,\ldots,Y_k\}$. Then we know by definition of

$Dep_R(X)$ that $X : \{Y_i\} \in \Sigma^+_{\mathfrak{H}}$ holds for all $i = 1, \ldots, k$. For the following inference notice that the Y_i are mutually disjoint.

$$\cfrac{\cfrac{\cfrac{X : \{Y_1\}}{X : \{Y_1, \emptyset\}}\mathcal{I}_\emptyset \quad X : \{Y_2\}}{X : \{Y_1, Y_2\}}\mathcal{U}}{\vdots}$$

$$\cfrac{\cfrac{}{X : \{Y_1, \ldots, Y_{k-1}, \emptyset\}}\mathcal{I}_\emptyset \quad X : \{Y_k\}}{X : \{Y_1, \ldots, Y_{k-1}, Y_k\}}\mathcal{U}$$

We can therefore conclude that $X : S \in \Sigma^+_{\mathfrak{H}}$. □

Theorem 2.5 *For all relation schemata R the set* $\mathfrak{H}$ *of inference rules is R-complete for the R-implication of FHDs.*

Proof Let R be an arbitrary relation schema, and let $\Sigma \cup \{X : S\}$ be an arbitrary set of FHDs on R. We need to show that $\Sigma^*_R \subseteq \Sigma^+_{\mathfrak{H}}$ holds. Let $DepB_R(X) = \{Y_1, \ldots, Y_k\}$, i.e., $\{X, Y_1, \ldots, Y_k\}$ forms a partition of R. We choose k tuples $t_1, \ldots, t_k$ that all have the same value on all attributes in X, and any pair of distinct tuples has different values on all the attributes in every Y_i. More precisely, $t_i[X] = t_j[X]$ for all $1 \leq i, j \leq k$, and for all $1 \leq i < j \leq k$ and for all $l = 1, \ldots, k$ we have $t_i[B] \neq t_j[B]$ for all $B \in Y_l$. Such a relation may look as follows:

X	Y_1	Y_2	$\cdots$	Y_k
$0\ldots0$	$1\ldots1$	$1\ldots1$	$\cdots$	$1\ldots1$
$0\ldots0$	$2\ldots2$	$2\ldots2$	$\cdots$	$2\ldots2$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$0\ldots0$	$k\ldots k$	$k\ldots k$	$\cdots$	$k\ldots k$

We will now use these k tuples to construct the following R-relation r consisting of k^k different tuples. In fact, for each list $\mathcal{L} = [j_1, \ldots, j_k]$ of k numbers from $\{1, \ldots, k\}$ we define the tuple $t_{\mathcal{L}}$ by

$$t_{[j_1,\ldots,j_k]}[XY_i] := t_{j_i}[XY_i] \qquad \text{for } i = 1, \ldots, k.$$

In particular, the original tuples $t_1, \ldots, t_k$ also occur in the relation r, namely as tuples $t_{[1,1,\ldots,1]}, \ldots, t_{[k,k,\ldots,k]}$, respectively. The relation r may look as follows:

	X	Y_1	Y_2	$\cdots$	Y_k
$t_{[1,1,\ldots,1]}$:	$0\ldots0$	$1\ldots1$	$1\ldots1$	$\cdots$	$1\ldots1$
$t_{[2,2,\ldots,2]}$:	$0\ldots0$	$2\ldots2$	$2\ldots2$	$\cdots$	$2\ldots2$
	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$t_{[k,k,\ldots,k]}$:	$0\ldots0$	$k\ldots k$	$k\ldots k$	$\cdots$	$k\ldots k$
$t_{[1,1,\ldots,2]}$:	$0\ldots0$	$1\ldots1$	$1\ldots1$	$\cdots$	$2\ldots2$
	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$t_{[1,1,\ldots,k]}$:	$0\ldots0$	$1\ldots1$	$1\ldots1$	$\cdots$	$k\ldots k$
	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$t_{[k,k,\ldots,k-1]}$:	$0\ldots0$	$k\ldots k$	$k\ldots k$	$\cdots$	$k-1\ldots k-1$

We will now show that r satisfies Σ. Let $U : \{V_1, \ldots, V_l\} \in \Sigma$. We need to show that if there are $t_1, \ldots, t_{l+1} \in r$ such that $t_i[U] = t_j[U]$ holds for all $1 \leq i, j \leq l+1$, then there is some $t \in r$ such that $t[UV_i] = t_i[UV_i]$ for $i = 1, \ldots, l$ and $t[U(R - UV_1 \cdots V_l)] = t_{l+1}[U(R - UV_1 \cdots V_l)]$. Let

$$W := \bigcup \{Y_i \in DepB_R(X) \mid U \cap Y_i \neq \emptyset\}.$$

One can see from the construction above that for all $i = 1, \ldots, k$ the following holds: two tuples in r either coincide on all the attributes of Y_i or differ on every single attribute of Y_i. Therefore, any two tuples in r that coincide on U must also coincide on every Y_i that U has non-empty intersection with. Consequently, $t_i[UW] = t_j[UW]$ for all $1 \leq i, j \leq l+1$. Notice that every attribute of U that is an attribute of some Y_i must also be an attribute of W by the definition of W. Furthermore, every attribute of U that is not an attribute of any of the Y_i must be an attribute of X since $\bigcup\{X, Y_1, \ldots, Y_k\} = R$. Therefore, $U \subseteq XW$. We apply the *augmentation rule* $\mathcal{A}$ to $U : \{V_1, \ldots, V_l\} \in \Sigma$ to derive $U(XW - U) : \{V_1 - (XW - U), \ldots, V_l - (XW - U)\} \in \Sigma^+_{\mathfrak{H}}$. However, $U(XW - U) = XW$ and since $U, V_1, \ldots, V_l$ are mutually disjoint we also have $V_i - (XW - U) = V_i - XW$ for $i = 1, \ldots, l$. That is, $XW : \{V_1 - XW, \ldots, V_l - XW\} \in \Sigma^+_{\mathfrak{H}}$. Moreover, since W is the union of elements from $DepB_R(X)$ we have $X : \{W\} \in \Sigma^+_{\mathfrak{H}}$ by Theorem 2.4. Subsequently, we can apply the *transitivity rule* $\mathcal{T}$ to

$$X : \{W\},\ XW : \{V_1 - XW, \ldots, V_l - XW\} \in \Sigma^+_{\mathfrak{H}}$$

and derive that $X : \{V_1 - XW, \ldots, V_l - XW, W\} \in \Sigma^+_{\mathfrak{H}}$, too. According to Theorem 2.4 it follows that $V_i - XW$ is the union of some Y_j for each $i = 1, \ldots, l$. As $XW, V_1 - XW, \ldots, V_l - XW$ are also mutually disjoint the construction of r ensures the existence of some tuple $t \in r$ that satisfies $t[XW(V_i - XW)] = t_i[XW(V_i - XW)]$ for $i = 1, \ldots, l$ and $t[XW(R - XWV_1 \cdots V_l)] = t_{l+1}[XW(R - XWV_1 \cdots V_l)]$. Since $UV_i \subseteq XW(V_i - XW)$ holds for $i = 1, \ldots, l$ this shows that $t[UV_i] = t_i[UV_i]$ for $i = 1, \ldots, l$. Since $U(R - UV_1 \cdots V_l) \subseteq XW(R - XWV_1 \cdots V_l)$ we also have $t[U(R - V_1 \cdots V_l)] = t_{l+1}[U(R - V_1 \cdots V_l)]$. This, however, shows that r satisfies Σ.

Now suppose that $X : S \in \Sigma^*_R$. In order to verify the completeness of the rule set $\mathfrak{H}$ we will show that $X : S \in \Sigma^+_{\mathfrak{H}}$ holds, too. Since r satisfies Σ it follows that r also satisfies Σ^*_R. In particular, $\models_r X : S$. Assume that there is some member Y of S which is not the union of any members of $DepB_R(X) = \{Y_1, \ldots, Y_k\}$. We conclude that Y intersects some Y_i without subsuming it, i.e., $Y \cap Y_i \neq \emptyset$ and $Y_i \cap (R - XY) \neq \emptyset$. Take, for instance, the two tuples $t_1 = t_{[1,1,\ldots,1]}$ and $t_2 = t_{[2,2,\ldots,2]}$ of r which coincide on X and differ on every attribute in $R - X$. From the construction of r we know that any two tuples of r either coincide on Y_i or differ on every single attribute of Y_i. Assume there is some tuple $t \in r$ with $t[XY] = t_1[XY]$ and $t[X(R - XY)] = t_2[X(R - XY)]$. Since $Y \cap Y_i \neq \emptyset$ holds $t[XY] = t_1[XY]$ implies $t[Y_i] = t_1[Y_i]$. Moreover, since $Y_i \cap (R - XY) \neq \emptyset$ holds $t[X(R - XY)] = t_2[X(R - XY)]$ implies $t[Y_i] = t_2[Y_i]$, and therefore $t_1[Y_i] = t_2[Y_i]$. This, however, is a contradiction since t_1 and t_2 even differ on every attribute of Y_i by definition. Hence, we conclude that there is no tuple $t \in r$ with $t[XY] = t_1[XY]$ and $t[X(R - XY)] = t_2[X(R - XY)]$. This means r does not satisfy $X : \{Y\}$. However, this contradicts $\models_r X : S$ according to Theorem 2.4. Hence, every member of S is the union of some members of $DepB_R(X)$. Thus, $X : S \in \Sigma^+_{\mathfrak{H}}$ by Theorem 2.4. □

2.3 A minimal axiomatisation

The objective is to identify a minimal subset $\mathfrak{H}_{\min}$ of $\mathfrak{H}$, 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* $\mathfrak{H}_{min} = \{\mathcal{R}_\emptyset, \mathcal{A}, \mathcal{T}, \mathcal{O}, \mathcal{C}_R\}$ *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 $\mathfrak{H}_{\min}$. Subsequently, it is demonstrated that none of the rules in $\mathfrak{H}_{\min}$ can be omitted without losing R-completeness for the R-implication of FHDs for some relation schemata R.

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

Proof

$$
\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{X:\{Y_1,\ldots,Y_k\}}{X:\{Y_1,\ldots,Y_{k-1},R-XY_1\cdots Y_k\}}\,\mathcal{C}_R}{XZ:\{Y_1-Z,\ldots,Y_{k-1}-Z,R-XZY_1\cdots Y_k\}}\,\mathcal{A} \qquad X:\{Z\}}{X:\{Y_1-Z,\ldots,Y_{k-1}-Z,R-XZY_1\cdots Y_k,Z\}}\,\mathcal{T}}{X:\{Y_1-Z,\ldots,Y_{k-1}-Z,R-XZY_1\cdots Y_k\}}\,\mathcal{O}}{X:\{Y_1-Z,\ldots,Y_{k-1}-Z,\underbrace{R-(X(Y_1-Z)\cdots(Y_{k-1}-Z)(R-XZY_1\cdots Y_k))}_{=Y_kZ}\}}\,\mathcal{C}_R
$$

Note that $R-(X(Y_1-Z)\cdots(Y_{k-1}-Z)(R-XZY_1\cdots Y_k)) = Y_kZ$ since $X \cap Y_k = \emptyset$, $X \cap Z = \emptyset$ and $Y_i \cap Y_k = \emptyset$ for $i = 1, \ldots, k-1$. □

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

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

$$
\dfrac{\dfrac{\dfrac{X:\{Y_1,\ldots,Y_k\}}{X:\{Y_1,\ldots,Y_{k-1}\}}\,\mathcal{O}}{X:\{Y_1,\ldots,Y_{k-1},\emptyset\}}\,\mathcal{I}_\emptyset \qquad \dfrac{\dfrac{\dfrac{\dfrac{X:\{Y_1,\ldots,Y_k\}}{X:\{Y_k\}}\,\mathcal{O}}{X(R-XZ):\{Y_k-(R-XZ)\}}\,\mathcal{A} \qquad \dfrac{X:\{Z\}}{X:\{R-XZ\}}\,\mathcal{C}_R}{X:\{Y_k\cap Z, R-XZ\}}\,\mathcal{T}}{X:\{Y_k\cap Z\}}\,\mathcal{O}}{X:\{Y_1,\ldots,Y_{k-1},Y_k\cap Z\}}\,\mathcal{U}
$$

Note that the left-most branch becomes necessary if $k \geq 2$. □

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 $Y_k \cap X = \emptyset$.

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

This concludes the proof. □

Theorem 2.5 and Lemmata 2.2, 2.3 and 2.4 show that $\mathfrak{H}_{\min}$ is indeed R-complete for the R-implication of FHDs for all relation schemata R. In order to complete the proof of Theorem 2.6 it remains to verify the independence of every inference rule in $\mathfrak{H}_{\min}$ from the rest of the rules.

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

Proof Let $\Sigma = \emptyset$ and $\sigma = \emptyset : \{\emptyset\}$. Since $\{\mathcal{A}, \mathcal{T}, \mathcal{O}, \mathcal{C}_R\}$ does not contain any axiom we conclude that $\Sigma^+_{\{\mathcal{A},\mathcal{T},\mathcal{O},\mathcal{C}_R\}} = \emptyset$ and hence $\sigma \notin \Sigma^+_{\{\mathcal{A},\mathcal{T},\mathcal{O},\mathcal{C}_R\}}$. However, $\sigma \in \Sigma^+_{\{\mathcal{R}_\emptyset,\mathcal{A},\mathcal{T},\mathcal{O},\mathcal{C}_R\}}$. □

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

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 $X : \{Y_1, \ldots, Y_k\}$ belongs to $\Sigma^+_{\mathfrak{S}}$ if and only if the entry in row labelled X and column labelled $\{Y_1, \ldots, Y_k\}$ is a cross $\times$. 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 $\{\mathcal{R}_\emptyset, \mathcal{T}, \mathcal{O}, \mathcal{C}_R\}$. Notice that the only inference rule capable of producing an FHD with left-hand side A is the *augmentation rule* $\mathcal{A}$ which has been dropped.

	$\{\emptyset\}$	$\{A\}$	$\{A, \emptyset\}$
$\emptyset$	$\times$	$\times$	$\times$
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 of* $\{\mathcal{R}_\emptyset, \mathcal{A}, \mathcal{O}, \mathcal{C}_R\}$.

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 $\{\mathcal{R}_\emptyset, \mathcal{A}, \mathcal{O}, \mathcal{C}_R\}$. Notice that the only inference rule capable of producing an FHD with right-hand side $\{A, \emptyset\}$ is the *transitivity rule* $\mathcal{T}$ which has been dropped.

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

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 of* $\{\mathcal{R}_\emptyset, \mathcal{A}, \mathcal{T}, \mathcal{C}_R\}$.

Proof Let $R = AB$, $\Sigma = \{\emptyset : \{\emptyset, A\}\}$ and $\sigma = \emptyset : \{A\}$. The closure $\Sigma^+_{\{\mathcal{R}_\emptyset, \mathcal{A}, \mathcal{T}, \mathcal{C}_R\}}$ can be obtained as follows. The *empty-set-axiom* $\mathcal{R}_\emptyset$ yields the FHD $\emptyset : \{\emptyset\}$. Subsequent applications of the *augmentation rule* $\mathcal{A}$ yield the $\{\emptyset\}$-column. One may apply the *R-complementation rule* $\mathcal{C}_R$ to each of the infered FHDs to derive $\emptyset : \{AB\}$, $A : \{B\}$, $B : \{A\}$, respectively. Three applications of the *transitivity rule* $\mathcal{T}$ yield $\emptyset : \{AB, \emptyset\}$, $A : \{B, \emptyset\}$, $B : \{A, \emptyset\}$, respectively. We may then enter the FHD $\emptyset : \{\emptyset, A\}$, and apply the *R-complementation rule* $\mathcal{C}_R$ to infer $\emptyset : \{\emptyset, B\}$ and $\emptyset : \{\emptyset, AB\}$. Finally, the *transitivity rule* $\mathcal{T}$ yields $\emptyset : \{\emptyset, A, B\}$. This set is closed under derivation using $\{\mathcal{R}_\emptyset, \mathcal{A}, \mathcal{T}, \mathcal{C}_R\}$. Notice that an inference of σ must somehow involve the FHD in Σ, but the only inference rule capable of removing the empty set $\emptyset$ from $\{\emptyset, A\}$ is the *omission rule* which has been dropped.

	$\{\emptyset\}$	$\{A\}$	$\{B\}$	$\{AB\}$	$\{\emptyset, A\}$	$\{\emptyset, B\}$	$\{A, B\}$	$\{\emptyset, AB\}$	$\{\emptyset, A, B\}$
$\emptyset$	$\times$			$\times$	$\times$	$\times$	$\times$	$\times$	$\times$
A	$\times$	$\star$	$\times$	$\star$	$\star$	$\times$	$\star$	$\star$	$\star$
B	$\times$	$\times$	$\star$	$\star$	$\times$	$\star$	$\star$	$\star$	$\star$
AB	$\times$	$\star$	$\star$	$\star$	$\star$	$\star$	$\star$	$\star$	$\star$

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

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

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 $\{\mathcal{R}_\emptyset, \mathcal{A}, \mathcal{T}, \mathcal{O}\}$. Notice that the only inference rule capable of inferring any FHD with a right-hand side different that is different from $\emptyset$ is the *R-complementation rule* $\mathcal{C}_R$ which has been dropped.

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

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 [9] has replaced the R-complementation rule $\mathcal{C}^R_{\text{MVD}}$ by the so-called *R-axiom* $\dfrac{}{\emptyset \twoheadrightarrow R}$ and still obtained an R-complete axiomatisation for the R-implication of MVDs for all relation schemata R. That is, *R-axiom*, *augmentation rule* $\mathcal{A}_{\text{MVD}}$ and *pseudo-transitivity rule* $\mathcal{T}_{\text{MVD}}$ form an R-sound and R-complete set of inference rules for the R-implication of MVDs for all relation schema R [9].

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

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

Proof We 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.

$$\dfrac{\dfrac{\dfrac{\dfrac{X : \{Y_1,\ldots,Y_k\}}{X : \{Y_1,\ldots,Y_{k-1}\}}\,\mathcal{O}}{XY_1\cdots Y_{k-2} : \{Y_{k-1},\emptyset\}}\,\mathcal{A}}{XY_1\cdots Y_{k-2} : \{Y_{k-1}\}}\,\mathcal{O} \qquad \dfrac{\dfrac{\dfrac{\dfrac{X : \{Y_1,\ldots,Y_k\}}{XY_1\cdots Y_{k-1} : \{Y_k,\emptyset\}}\,\mathcal{A}}{XY_1\cdots Y_{k-1} : \{Y_k\}}\,\mathcal{O} \qquad \dfrac{\dfrac{}{\emptyset : \{R\}}\,R\text{-axiom}}{XY_1\cdots Y_k : \{R - XY_1\cdots Y_k\}}\,\mathcal{A}}{XY_1\cdots Y_{k-1} : \{R - XY_1\cdots Y_k, Y_k\}}\,\mathcal{T}}{XY_1\cdots Y_{k-1} : \{R - XY_1\cdots Y_k\}}\,\mathcal{O}}{XY_1\cdots Y_{k-2} : \{R - XY_1\cdots Y_k, Y_{k-1}\}}\,\mathcal{T}$$

$$\vdots$$

$$\dfrac{\dfrac{X : \{Y_1,\ldots,Y_k\}}{X : \{Y_1\}}\,\mathcal{O} \qquad XY_1 : \{R - XY_1\cdots Y_k, Y_{k-1},\ldots,Y_2\}}{X : \{R - XY_1\cdots Y_k, Y_{k-1},\ldots,Y_1\}}\,\mathcal{T}$$

This concludes the proof. □

3 A complementary axiomatisation

We have seen before that the set $\mathfrak{H}_{\min} = \{\mathcal{R}_\emptyset, \mathcal{A}, \mathcal{T}, \mathcal{O}, \mathcal{C}_R\}$ 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. An R-complete set $\mathfrak{S}$ is said to be *R-complementary* if and only if for every set $\Sigma \cup \{\sigma\}$ of FHDs on R the inference of σ from Σ using $\mathfrak{S}$ can be turned into an inference of σ from Σ using $\mathfrak{S}$ 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 set $\mathfrak{H}_{\min}$ is not R-complementary for all relation schemata R.

Example 3.1 Let Σ consist of the two FHDs

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

An inspection of the syntactical inference rules $\mathcal{R}_\emptyset, \mathcal{A}, \mathcal{T}, \mathcal{O}$ shows that $Title : \{\{Actor\}, \{Feature, Language\}\} \notin \Sigma^+_{\{\mathcal{R}_\emptyset, \mathcal{A}, \mathcal{T}, \mathcal{O}\}}$. Moreover, Lemma 4.1 shows that $Title : \{\{Actor\}, Y\} \notin \Sigma^+_{\{\mathcal{R}_\emptyset, \mathcal{A}, \mathcal{T}, \mathcal{O}\}}$ for any Y such that

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

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

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

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

$$\text{DVD} - \{Title, Actor, Feature, Language\} = \{Crew\}$$

$\mathcal{C}_{\text{DVD}}$ 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 an axiomatisation for FHDs that is R-complementary for all relation schemata R. This objective has been successfully achieved in the case of MVDs [10, 27]. A key role played the so-called *subset rule* $\mathcal{S}_{\text{MVD}}$ which we now generalise into the framework of FHDs.

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

The reason for including the set $Y - Y_1 \cdots Y_k$ in the conclusion of the *subset rule* is purely technical. In fact, this set is needed to shift applications of the R-complementation rule to the end of an inference (see the proof of Theorem 3.1 for details).

Lemma 3.1 *For all relation schemata R the subset rule $\mathcal{S}$ is R-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 $i = 1, \ldots, k$. Notice that $Y - (Y_i \cap Y) = Y - Y_i$ holds for all $i = 1, \ldots, k$. We may apply the difference rule $\mathcal{D}_{\text{MVD}}$ to $X \twoheadrightarrow Y$ and $X \twoheadrightarrow (Y_1 \cap Y)$ and conclude that r satisfies $X \twoheadrightarrow Y - Y_1$. In a second step we apply the difference rule $\mathcal{D}_{\text{MVD}}$ to $X \twoheadrightarrow Y - Y_1$ and $X \twoheadrightarrow (Y_2 \cap Y)$ and conclude that r satisfies $X \twoheadrightarrow (Y - Y_1Y_2)$ since $(Y - Y_1) - (Y \cap Y_2) = Y - Y_1Y_2$ holds. Successively applying the difference rule $\mathcal{D}_{\text{MVD}}$ in this way leads us to conclude that r satisfies $X \twoheadrightarrow (Y - Y_1 \cdots Y_k)$. A further application of Theorem 2.2 shows that r satisfies $X : \{Y_1 \cap Y, \ldots, Y_k \cap Y, Y - Y_1 \cdots Y_k\}$. □

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 $\mathfrak{S}$. The following theorem shows constructively how to yield an R-complementary set $\mathfrak{H}_C$ by adding the subset rule $\mathcal{S}$ and merging rule $\mathcal{M}$ to $\mathfrak{H}_{\min}$.

Theorem 3.1 *Let* Σ *be a set of FHDs. For each inference* γ *from* Σ *by the system* $\mathfrak{H}_{min} = \{\mathcal{R}_\emptyset, \mathcal{A}, \mathcal{T}, \mathcal{O}, \mathcal{C}_R\}$ *there is an inference* ξ *from* Σ *by the system* $\mathfrak{H}_C = \{\mathcal{R}_\emptyset, \mathcal{A}, \mathcal{T}, \mathcal{O}, \mathcal{S}, \mathcal{M}, \mathcal{C}_R\}$ *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*
- *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 $\mathfrak{H}_{\min}$ which has length l. We consider five cases according to which inference rule in $\mathfrak{H}_{\min}$ was applied to infer σ_l from $[\sigma_1, \ldots, \sigma_{l-1}]$. The subset rule $\mathcal{S}$ is used in Cases 3.3 and 3.4 while the merging rule $\mathcal{M}$ is used in Case 4 to shift applications of the R-complementation rule $\mathcal{C}_R$ towards the end of the inference.

Case 1 The FHD σ_l is either $\emptyset : \{\emptyset\}$ (*empty-set-axiom* $\mathcal{R}_\emptyset$) or $\sigma_l \in \Sigma$. In this case $\xi = [\sigma_l]$ has the desired properties.

Case 2 We infer σ_l by applying the *augmentation rule* $\mathcal{A}$ 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 are of the following form:

$$\cfrac{\cfrac{X : \{Y_1, \ldots, Y_k\}}{X : \{Y_1, \ldots, Y_{k-1}, R - XY_1 \cdots Y_k\}}\,\mathcal{C}_R}{XZ : \{Y_1 - Z, \ldots, Y_{k-1} - Z, R - XZY_1 \cdots Y_k\}}\,\mathcal{A}.$$

However, these steps can be replaced as follows:

$$\cfrac{\cfrac{X : \{Y_1, \ldots, Y_k\}}{XZ : \{Y_1 - Z, \ldots, Y_k - Z\}}\,\mathcal{A}}{XZ : \{Y_1 - Z, \ldots, Y_{k-1} - Z, \underbrace{R - XZ(Y_1 - Z) \cdots (Y_k - Z)}_{= R - XZY_1 \cdots Y_k}\}}\,\mathcal{C}_R.$$

The result of this replacement is an inference with the desired properties.

Case 3 We infer σ_l by applying the *transitivity rule* $\mathcal{T}$ to the premises σ_i and σ_j with $i, j < l$. Let ξ_i (ξ_j) be obtained by using the induction hypothesis for $\gamma_i := [\sigma_1, \ldots, \sigma_i]$ ($\gamma_j := [\sigma_1, \ldots, \sigma_j]$).

Consider the inference $\xi := [\xi_i, \xi_j, \sigma_l]$. Then we distinguish between four cases according to the occurrence of the R-complementation rule $\mathcal{C}_R$ in ξ_i and ξ_j.

Case 3.1 If $\mathcal{C}_R$ is applied neither in ξ_i nor in ξ_j, then ξ has the desired properties.

Case 3.2 If $\mathcal{C}_R$ is applied in ξ_i (as the last rule) but not in ξ_j, then the last step of ξ_i and the last step of ξ are of the following form:

$$\frac{X:\{Y\} \quad \dfrac{XY:\{Y_1,\ldots,Y_k\}}{XY:\{Y_1,\ldots,Y_{k-1},R-XYY_1\cdots Y_k\}}\,\mathcal{C}_R}{X:\{Y_1,\ldots,Y_{k-1},R-XYY_1\cdots Y_k,Y\}}\,\mathcal{T}.$$

However, these steps can be replaced as follows:

$$\dfrac{\dfrac{X:\{Y\} \qquad XY:\{Y_1,\ldots,Y_k\}}{X:\{Y_1,\ldots,Y_k,Y\}}\,\mathcal{T}}{X:\{Y_1,\ldots,Y_{k-1},Y,R-XYY_1\cdots Y_k\}}\,\mathcal{C}_R$$

The result of this replacement is an inference with the desired properties.

Case 3.3 If $\mathcal{C}_R$ is applied in ξ_j (as the last rule) but not in ξ_i, then the last step of ξ_j and the last step of ξ are of the following form:

$$\frac{\dfrac{X:\{Y\}}{X:\{R-XY\}}\,\mathcal{C}_R \qquad X(R-XY):\{Y_1,\ldots,Y_k\}}{X:\{Y_1,\ldots,Y_k,R-XY\}}\,\mathcal{T}.$$

However, these steps can be replaced as follows:

$$\dfrac{\dfrac{X:\{Y\} \qquad X(R-XY):\{Y_1,\ldots,Y_k\}}{X:\{Y_1\cap Y,\ldots,Y_k\cap Y,Y-Y_1\cdots Y_k\}}\,\mathcal{S}}{X:\{Y_1,\ldots,Y_k,\underbrace{R-XY_1\cdots Y_k(Y-Y_1\cdots Y_k)}_{=R-XYY_1\cdots Y_k=R-XY}\}}\,\mathcal{C}_R.$$

Notice that Y and $X(R-XY)$ are disjoint, and $Y_1\cdots Y_k$ and $X(R-XY)$ are disjoint as well. It follows that $Y_1\cdots Y_k\subseteq Y$, and therefore $Y_i\cap Y=Y_i$ for all $i=1,\ldots,k$.

The result of this replacement is an inference with the desired properties.

Case 3.4 If $\mathcal{C}_R$ is applied in both ξ_i and ξ_j (as the last rule), then the last steps of ξ_i, ξ_j and ξ are of the following form:

$$\frac{\dfrac{X:\{Y\}}{X:\{R-XY\}}\,\mathcal{C}_R \quad \dfrac{X(R-XY):\{Y_1,\ldots,Y_k\}}{X(R-XY):\{Y_1,\ldots,Y_{k-1},R-X(R-XY)Y_1\cdots Y_k\}}\,\mathcal{C}_R}{X:\{Y_1,\ldots,Y_{k-1},\underbrace{R-X(R-XY)Y_1\cdots Y_k}_{=Y-Y_1\cdots Y_k},R-XY\}}\,\mathcal{T}.$$

Notice that $R - X(R - XY)Y_1 \cdots Y_k = Y - Y_1 \cdots Y_k$ since $Y \cap X = \emptyset$. However, these steps can be replaced as follows:

$$\dfrac{\dfrac{X : \{Y\} \qquad X(R - XY) : \{Y_1, \ldots, Y_k\}}{X : \{Y_1, \ldots, Y_k, Y - Y_1 \ldots Y_k\}}\, \mathcal{S}}{X : \{Y_1, \ldots, Y_{k-1}, Y - Y_1 \ldots Y_k, \underbrace{R - XY_1 \cdots Y_k(Y - Y_1 \cdots Y_k)}_{=R-XYY_1\cdots Y_k = R-XY}\}}\, \mathcal{C}_R$$

Notice that Y and $X(R - XY)$ are disjoint, and $Y_1 \cdots Y_k$ and $X(R - XY)$ are disjoint as well. That is, $Y_1 \cdots Y_k \subseteq Y$, and therefore $Y_i \cap Y = Y_i$ for all $i = 1, \ldots, k$.

The result of this replacement is an inference with the desired properties.

Case 4 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.

Case 5 We infer σ_l by applying the *R-complementation rule* $\mathcal{C}_R$ 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 are of the following form:

$$\cfrac{\cfrac{X : \{Y_1, \ldots, Y_k\}}{X : \{Y_1, \ldots, Y_{k-1}, R - XY_1 \cdots Y_k\}}\,\mathcal{C}_R}{X : \{Y_1, \ldots, Y_{k-1}, \underbrace{R - XY_1 \cdots Y_{k-1}(R - XY_1 \cdots Y_k)}_{=Y_k}\}}\,\mathcal{C}_R.$$

Hence the inference obtained by deleting these two steps from ξ has the desired properties. □

Corollary 3.1 *For all relation schemata R the set $\mathfrak{H}_C$ of inference rules is R-complementary for the R-implication of FHDs.* □

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 [10, 27].

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 [10, 27], and Biskup introduced an alternative notion of MVD implication [10] in which the underlying set of attributes remains undetermined. We will now extend this notion to FHDs.

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

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

According to Theorem 2.1 the fixed relation schema R has now simply been replaced by $Dom(r)$ which varies with the relation r.

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.

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 a relation schema such that $X \cup \bigcup_{i=1}^{k} Y_i \cup \bigcup_{j=1}^{n} \left(X_j \cup \bigcup_{s=1}^{l_j} Y_s^j \right) \subseteq R$. Then $\Sigma = \left\{X_1 : \left\{Y_1^1, \ldots, Y_{l_1}^1\right\}, \ldots, X_n : \left\{Y_1^n, \ldots, Y_{l_n}^n\right\}\right\}$ 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 $\mathcal{C}_R$. However, *Title* : {{*Actor*}, {*Language*}} is not implied by Σ as the following counterexample relation r shows.

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

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

The notions of *soundness*, *completeness*, *independence* and *minimality* are simply adapted to the context of undetermined universes by dropping the reference to the underlying relation schema R from the corresponding notions in the context of fixed universes.

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 now explore the power of the common part of the systems $\mathfrak{H}_C$, namely $\mathfrak{H}_U = \{\mathcal{R}_\emptyset, \mathcal{A}, \mathcal{T}, \mathcal{O}, \mathcal{S}, \mathcal{M}\}$ which can be obtained from the systems $\mathfrak{H}_C$ by removing the R-complementation rule $\mathcal{C}_R$. Hence $\mathfrak{H}_U$ does not permit the possibly semantically meaningless inference of complementation. Theorem 3.1 states that $\mathfrak{H}_U$ is almost R-complete for the R-implication of FHDs for all relation schemata R.

Corollary 4.1 *Let $R \subseteq \mathfrak{A}$ be a finite set of attributes. Then for all finite sets $\Sigma = \left\{X_1 : \left\{Y_1^1, \ldots, Y_{l_1}^1\right\}, \ldots, X_n : \left\{Y_1^n, \ldots, Y_{l_n}^n\right\}\right.$ 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:*

$$X : \{Y_1, \ldots, Y_k\} \in \Sigma^+_{\mathfrak{H}_C} \qquad \textit{if and only if}$$
$$X : \{Y_1, \ldots, Y_k\} \in \Sigma^+_{\mathfrak{H}_U} \textit{ or there is some } i \textit{ such that } 1 \leq i \leq k \textit{ and}$$
$$X : \{Y_1, \ldots, Y_{i-1}, Y_{i+1}, \ldots, , Y_k, R - XY_1 \cdots Y_k\} \in \Sigma^+_{\mathfrak{H}_U}.$$

Corollary 4.1 indicates that by the subsystem $\mathfrak{H}_U$ of $\mathfrak{H}_C$ we can infer those consequences of a given set of FHDs which are independent of the underlying relation schema R.

Below we shall prove that the system $\mathfrak{H}_U$ is actually sound and complete for the implication of FHDs in the sense of Definition 4.1. That is, $\mathfrak{H}_U$ can generate exactly the implications in an undetermined universe. Before that we shall prove two lemmata.

The correctness of the first lemma can easily be observed by inspecting the syntactic definitions of the inference rules in $\mathfrak{H}_U$. For each of the rules, the right-hand side of the conclusion does not contain any attribute that did not already occur in the right-hand side of at least one of the premises.

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

For the next lemma one may notice that attributes outside of W can always be introduced only in the last step of the inference by utilising the *augmentation rule* $\mathcal{A}$.

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

Proof Let $\bar{\xi} = [R_1 : \{S_1^1, \ldots, S_{t_1}^1\}, \ldots, R_{m-1} : \{S_1^{m-1}, \ldots, S_{t_{m-1}}^{m-1}\}]$ be any inference of $X : \{Y_1, \ldots, Y_k\}$ from Σ by $\mathfrak{H}_U$. Consider the sequence ξ

$$[R_1 \cap W : \{S_1^1 \cap W, \ldots, S_{t_1}^1 \cap W\}, \ldots, R_{m-1} \cap W : \{S_1^{m-1} \cap W, \ldots, S_{t_{m-1}}^{m-1} \cap W\}].$$

We claim that ξ is an inference of $X \cap W : \{Y_1 \cap W, \ldots, Y_k \cap W\}$ from Σ by $\mathfrak{H}_U$. For if $R_i : \{S_1^i, \ldots, S_{t_1}^i\}$ is an element of Σ or $\emptyset : \{\emptyset\}$, then $R_i \cap W : \{S_1^i \cap W, \ldots, S_{t_1}^i \cap W\} = R_1 : \{S_1^1, \ldots, S_{t_1}^1\}$. On the other hand one can easily verify that if $R_i : \{S_1^i, \ldots, S_{t_1}^i\}$ is the result of applying one of the rules $\mathcal{A}, \mathcal{T}, \mathcal{O}, \mathcal{M}, \mathcal{S}$ in $\bar{\xi}$, then $R_1 \cap W : \{S_1^1 \cap W, \ldots, S_{t_1}^1 \cap W\}$ is the result of the same rule applied to the corresponding premises in ξ.

Now by Lemma 4.1 we know that for $i = 1, \ldots, k$, $Y_i \subseteq \bigcup_{j=1}^{n} \bigcup_{s=1}^{l_j} Y_s^j \subseteq W$, hence $Y_i \cap W = Y_i$. However, this implies that we can infer $X : \{Y_1, \ldots, Y_k\}$ from $X \cap W : \{Y_1 \cap W, \ldots, Y_k \cap W\}$ by the *augmentation rule* $\mathcal{A}$:

$$\frac{X \cap W : \{Y_1 \cap W, \ldots, Y_k \cap W\}}{\underbrace{(X \cap W) \cup X}_{=X} : \{\underbrace{Y_1 - X}_{=Y_1}, \ldots, \underbrace{Y_k - X}_{=Y_k}\}}\ \mathcal{A}.$$

Hence the inference $[\xi, X : \{Y_1, \ldots, Y_k\}]$ has the desired properties. □

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

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

$$\Sigma \text{ implies } X : \{Y_1, \ldots, Y_k\} \quad \text{if and only if} \quad X : \{Y_1, \ldots, Y_k\} \in \Sigma^+_{\mathfrak{H}_U}. \tag{4.1}$$

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 $\mathfrak{H}_U$ we assume $X : \{Y_1, \ldots, Y_k\} \in \Sigma^+_{\mathfrak{H}_U}$. Let r be any relation such that $T \subseteq Dom(r)$ and such that all FHDs $X_i : \left\{Y_1^i, \ldots, Y_{l_i}^i\right\} \in \Sigma$ are satisfied by r. Then we need to show that r also satisfies $X : \{Y_1, \ldots, Y_k\}$.

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

In order to prove the completeness of $\mathfrak{H}_U$ we assume $X : \{Y_1, \ldots, Y_k\} \notin \Sigma^+_{\mathfrak{H}_U}$. 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

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

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

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

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

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

by Corollary 4.1.

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

Since $\mathfrak{H}_U$ is sound and complete for the implication of FHDs and the Boolean rules are sound for the implication of FHDs one can conclude that the Boolean rules are derivable from $\mathfrak{H}_U$. Notice that this is in sharp contrast to the set $\mathfrak{H}_{\min} - \{\mathcal{C}_R\} = \{\mathcal{R}_\emptyset, \mathcal{A}, \mathcal{T}, \mathcal{O}\}$ as the following lemma shows.

Lemma 4.3 *The* union rule $\mathcal{U}$ *is independent of* $\{\mathcal{R}_\emptyset, \mathcal{A}, \mathcal{T}, \mathcal{O}\}$.

Proof Let $\Sigma = \{\emptyset : \{A\}, \emptyset : \{B\}\}$ and $\sigma = \emptyset : \{AB\}$. We obtain the closure $\Sigma^+_{\{\mathcal{R}_\emptyset, \mathcal{A}, \mathcal{T}, \mathcal{O}\}}$ 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 $\{\mathcal{R}_\emptyset, \mathcal{A}, \mathcal{T}, \mathcal{O}\}$. Notice that the *union rule* $\mathcal{U}$ is the only inference rule capable of inferring FHDs with a right-hand side in which the attributes A and B both belong to the same set, but the union rule has been dropped.

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

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

The following lemmata show how the Boolean rules can be derived from $\mathfrak{H}_U$.

Lemma 4.4 *The* union rule $\mathcal{U}$ *is derivable from* $\{\mathcal{A}, \mathcal{T}, \mathcal{M}\}$.

Proof

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

□

Lemma 4.5 *The* intersection rule $\mathcal{I}$ *is derivable from* $\mathfrak{H}_U = \{\mathcal{R}_\emptyset, \mathcal{A}, \mathcal{T}, \mathcal{M}, \mathcal{O}, \mathcal{S}\}$.

Proof

$$\dfrac{\dfrac{\dfrac{X:\{Y_1,\ldots,Y_k\}}{X:\{Y_1,\ldots,Y_k,\emptyset\}}\mathcal{I}_\emptyset}{X:\{Y_1,\ldots,Y_{k-1},\emptyset\}}\mathcal{O} \qquad \dfrac{\dfrac{X:\{Y_1,\ldots,Y_k\} \qquad X:\{Z\}}{X:\{Y_1\cap Z,\ldots,Y_k\cap Z, Z-Y_1\cdots Y_k\}}\mathcal{S}}{X:\{Y_k\cap Z\}}\mathcal{O}}{X:\{Y_1,\ldots,Y_{k-1},Y_k\cap Z\}}\mathcal{U}$$

The lemma is now a consequence of Lemma 4.4 and Lemma 2.1. □

Lemma 4.6 *The* difference rule $\mathcal{D}$ *is derivable from* $\{\mathcal{R}_\emptyset, \mathcal{A}, \mathcal{T}, \mathcal{M}, \mathcal{O}\}$.

Proof

$$\dfrac{\dfrac{\dfrac{X:\{Y_1,\ldots,Y_k\}}{X:\{Y_1,\ldots,Y_k,\emptyset\}}\mathcal{I}_\emptyset}{X:\{Y_1,\ldots,Y_{k-1},\emptyset\}}\mathcal{O} \qquad \dfrac{\dfrac{X:\{Z\} \qquad \dfrac{X:\{Y_1,\ldots,Y_k\}}{XZ:\{Y_1-Z,\ldots,Y_k-Z\}}\mathcal{A}}{X:\{Y_1-Z,\ldots,Y_k-Z\}}\mathcal{T}}{X:\{Y_k-Z\}}\mathcal{O}}{X:\{Y_1,\ldots,Y_{k-1},Y_k-Z\}}\mathcal{U}$$

The lemma is now a consequence of Lemma 4.4 and Lemma 2.1. □

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

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

For the following independence proofs we will utilise Lemma 4.1 which restricts the number of possible FHDs that can be inferred from Σ by $\mathfrak{H}_U$.

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

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 of* $\{R_\emptyset, \mathcal{T}, \mathcal{O}, \mathcal{S}, \mathcal{M}\}$.

Proof Let $\Sigma = \emptyset$ and $\sigma = A : \{\emptyset\}$. It follows 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 of* $\{R_\emptyset, \mathcal{A}, \mathcal{O}, \mathcal{S}, \mathcal{M}\}$.

Proof Let $\Sigma = \{A : \{B\}, AB : \{C\}\}$, and let $\sigma = A : \{B, C\}$. According to Lemma 4.1 only FHDs $X : S$ can be inferred from Σ for which all $Y \in S$ are subsets of $\{B, C\}$. The closure $\Sigma^+_{\{\mathcal{R}_\emptyset,\mathcal{A},\mathcal{O},\mathcal{S},\mathcal{M}\}}$ can be obtained as follows (omitting all those FHDs in which an attribute occurs that does not occur in Σ). The $\{\emptyset\}$-column can be obtained by an application of the *empty-set-axiom* $R_\emptyset$ followed by successive applications of the *augmentation rule* $\mathcal{A}$. We may then enter the FHDs from Σ and infer $AC : \{B\}$ from $A : \{B\}$ by means of the *augmentation rule* $\mathcal{A}$. Finally, the FHDs $A : \{B, \emptyset\}$, $AB : \{C, \emptyset\}$ and $AC : \{B, \emptyset\}$ can be inferred by applying the *subset rule* $\mathcal{S}$ to $A : \{B\}$, $AB : \{C\}$ and $AC : \{B\}$, respectively, and $\emptyset : \{\emptyset\}$. Further applications of any inference rules do not generate new FHDs.

	$\{\emptyset\}$	$\{B\}$	$\{C\}$	$\{BC\}$	$\{B, \emptyset\}$	$\{C, \emptyset\}$	$\{BC, \emptyset\}$	$\{B, C\}$	$\{B, C, \emptyset\}$
$\emptyset$	×								
A	×	×			×				
B	×	⋆		⋆	⋆		⋆	⋆	⋆
C	×		⋆	⋆		⋆	⋆	⋆	⋆
AB	×	⋆	×	⋆	⋆	×	⋆	⋆	⋆
AC	×	×	⋆	⋆	×	⋆	⋆	⋆	⋆
BC	×	⋆	⋆	⋆	⋆	⋆	⋆	⋆	⋆
ABC	×	⋆	⋆	⋆	⋆	⋆	⋆	⋆	⋆

Consequently, $\sigma \notin \Sigma^+_{\{R_\emptyset,\mathcal{A},\mathcal{O},\mathcal{S},\mathcal{M}\}}$ but $\sigma \in \Sigma^+_{\mathfrak{H}_c}$. □

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

Proof Let $\Sigma = \{\emptyset : \{A, B\}\}$, and let $\sigma = \emptyset : \{A\}$. According to Lemma 4.1 only FHDs $X : S$ can be inferred from Σ for which all $Y \in S$ are subsets of $\{A, B\}$. The closure $\Sigma^+_{\{\mathcal{R}_\emptyset,\mathcal{A},\mathcal{T},\mathcal{S},\mathcal{M}\}}$ can be obtained as follows (omitting all those FHDs in which an attribute occurs that does not occur in Σ). The $\{\emptyset\}$-column can be obtained by an application of the *empty-set-axiom* $R_\emptyset$ followed by successive applications of the *augmentation rule* $\mathcal{A}$. After entering the FHD from Σ one can apply the *merging rule* $\mathcal{M}$ to infer $\emptyset : \{AB\}$. One can then apply the *augmentation rule* $\mathcal{A}$ to $\emptyset : \{AB\}$ in order to infer both $A : \{B\}$ and $B : \{A\}$. Applying the *augmentation rule* $\mathcal{A}$ to $\emptyset : \{A, B\}$ results in $A : \{\emptyset, B\}$ and $B : \{A, \emptyset\}$. Finally, the *subset rule* $\mathcal{S}$ can be applied to $\emptyset : \{AB\}$ and $\emptyset : \{\emptyset\}$ to infer $\emptyset : \{AB, \emptyset\}$, and to $\emptyset : \{AB\}$ and $\emptyset : \{A, B\}$ to infer $\emptyset : \{A, B, \emptyset\}$. Further applications of any inference rules do not generate new FHDs.

	$\{\emptyset\}$	$\{A\}$	$\{B\}$	$\{AB\}$	$\{A, B\}$	$\{A, \emptyset\}$	$\{B, \emptyset\}$	$\{AB, \emptyset\}$	$\{A, B, \emptyset\}$
$\emptyset$	$\times$			$\times$	$\times$			$\times$	$\times$
A	$\times$	$\star$	$\times$	$\star$	$\star$	$\star$	$\times$	$\star$	$\star$
B	$\times$	$\times$	$\star$	$\star$	$\star$	$\times$	$\star$	$\star$	$\star$
AB	$\times$	$\star$	$\star$	$\star$	$\star$	$\star$	$\star$	$\star$	$\star$

Consequently, $\sigma \notin \Sigma^+_{\{R_\emptyset, \mathcal{A}, \mathcal{T}, \mathcal{S}, \mathcal{M}\}}$ but $\sigma \in \Sigma^+_{\mathfrak{H}_c}$. □

For the following three independence proofs we will utilise the fact that if $\mathcal{A}, \mathcal{T}, \mathcal{O} \in \mathfrak{S}$, then an FHD $X : \{Y_1, \ldots, Y_k\} \in \Sigma^+_{\mathfrak{S}}$ if and only if $X : \{Y_1, \ldots, Y_k, \emptyset\} \in \Sigma^+_{\mathfrak{S}}$, see Lemma 2.1. Therefore, we will only consider FHDs $X : S$ in which $\emptyset \notin S$.

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

Proof Let $\Sigma = \{\emptyset : \{AB\}, C : \{B\}\}$, and let $\sigma = \emptyset : \{B\}$. According to Lemma 4.1 only FHDs $X : S$ can be inferred from Σ for which all $Y \in S$ are subsets of $\{A, B\}$. The closure $\Sigma^+_{\{\mathcal{R}_\emptyset, \mathcal{A}, \mathcal{T}, \mathcal{O}, \mathcal{M}\}}$ can be obtained as follows (omitting all those FHDs in which an attribute occurs that does not occur in Σ). The $\{\emptyset\}$-column can be obtained by an application of the *empty-set-axiom* $R_\emptyset$ followed by successive applications of the *augmentation rule* $\mathcal{A}$. After entering the FHDs from Σ one can apply the *augmentation rule* $\mathcal{A}$ to $\emptyset : \{AB\}$ to infer $A : \{B\}$, $B : \{A\}$, $C : \{AB\}$, $BC : \{A\}$. applying the *augmentation rule* $\mathcal{A}$ to $C : \{B\}$ results in $AC : \{B\}$. Subsequently, the *transitivity rule* $\mathcal{T}$ is applied to $C : \{B\}$ and $BC : \{A\}$ to yield $C : \{A, B\}$. An application of the *omission rule* $\mathcal{O}$ to $C : \{A, B\}$ results in $C : \{A\}$. Further applications of any inference rules do not generate new FHDs.

	$\{\emptyset\}$	$\{A\}$	$\{B\}$	$\{AB\}$	$\{A, B\}$
$\emptyset$	$\times$			$\times$	
A	$\times$	$\star$	$\times$	$\star$	$\star$
B	$\times$	$\times$	$\star$	$\star$	$\star$
C	$\times$	$\times$	$\times$	$\times$	$\times$
AB	$\times$	$\star$	$\star$	$\star$	$\star$
AC	$\times$	$\star$	$\times$	$\star$	$\star$
BC	$\times$	$\times$	$\star$	$\star$	$\star$
ABC	$\times$	$\star$	$\star$	$\star$	$\star$

Consequently, $\sigma \notin \Sigma^+_{\{R_\emptyset, \mathcal{A}, \mathcal{T}, \mathcal{O}, \mathcal{M}\}}$ but $\sigma \in \Sigma^+_{\mathfrak{H}_c}$. □

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

Proof Let $\Sigma = \{\emptyset : \{A, B\}\}$, and let $\sigma = \emptyset : \{AB\}$. According to Lemma 4.1 only FHDs $X : S$ can be inferred from Σ for which all $Y \in S$ are subsets of $\{B, C\}$. The closure $\Sigma^+_{\{\mathcal{R}_\emptyset,\mathcal{A},\mathcal{T},\mathcal{O},\mathcal{S}\}}$ can be obtained as follows (omitting all those FHDs in which an attribute occurs that does not occur in Σ). The $\{\emptyset\}$-column can be obtained by an application of the *empty-set-axiom* $R_\emptyset$ followed by successive applications of the *augmentation rule* $\mathcal{A}$. After entering the FHD from Σ one may apply the *omission rule* $\mathcal{O}$ to $\emptyset : \{A, B\}$ to infer both $\emptyset : \{A\}$ and $\emptyset : \{B\}$. Subsequently, one may apply the *augmentation rule* $\mathcal{A}$ to infer $B : \{A\}$ from $\emptyset : \{A\}$, and to infer $A : \{B\}$ from $\emptyset : \{B\}$. Further applications of any inference rules do not generate new FHDs.

	$\{\emptyset\}$	$\{A\}$	$\{B\}$	$\{AB\}$	$\{A, B\}$
$\emptyset$	$\times$	$\times$	$\times$		$\times$
A	$\times$	$\star$	$\times$	$\star$	$\star$
B	$\times$	$\times$	$\star$	$\star$	$\star$
AB	$\times$	$\star$	$\star$	$\star$	$\star$

Consequently, $\sigma \notin \Sigma^+_{\{R_\emptyset,\mathcal{A},\mathcal{T},\mathcal{O},\mathcal{S}\}}$ but $\sigma \in \Sigma^+_{\mathfrak{H}_c}$. □

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

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

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 [6] the combined class of FDs and FHDs should also be studied in both fixed and undetermined universes.

According to Link [26] 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* [25, 27]. 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 FDs [24]. 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 and FHDs, respectively, that would extend the well-known synthesis algorithm for the class of FDs [7, 11, 12]. The notion of implication in undetermined universes provides an alternative basis for the formulation of such an algorithm.

Acknowledgements We thank the anonymous referees whose thorough comments and suggestions improved the readability of this paper.

References

1. Arenas, M., Libkin, L.: A normal form for XML documents. ACM Trans. Database Syst. **29**(1), 195–232 (2004)
2. Armstrong, W.W.: Dependency structures of database relationships. Inf. Process. **74**, 580–583 (1974)
3. Armstrong, W.W., Nakamura, Y., Rudnicki, P.: Armstrong's axioms. J. Formaliz. Math. **14** (2002)
4. Beeri, C.: On the membership problem for functional and multivalued dependencies in relational databases. ACM Trans. Database Syst. **5**(3), 241–259 (1980)
5. Beeri, C., Bernstein, P.A.: Computational problems related to the design of normal form relational schemata. ACM Trans. Database Syst. **4**(1), 30–59 (1979)
6. Beeri, C., Fagin, R., Howard, J.H.: A complete axiomatization for functional and multivalued dependencies in database relations. In: Proceedings of the SIGMOD International Conference on Management of Data, pp. 47–61. ACM, New York (1977)
7. Bernstein, P.: Synthesizing third normal form relations from functional dependencies. ACM Trans. Database Syst. **1**(4), 277–298 (1976)
8. Bernstein, P.A., Goodman, N.: What does Boyce-Codd normal form do? In: Proceedings of the 6th International Conference on Very Large Data Bases, pp. 245–259. IEEE Computer Society, Los Alamitos (1980)
9. Biskup, J.: On the complementation rule for multivalued dependencies in database relations. Acta Inform. **10**(3), 297–305 (1978)
10. Biskup, J.: Inferences of multivalued dependencies in fixed and undetermined universes. Theor. Comput. Sci. **10**(1), 93–106 (1980)
11. Biskup, J., Meyer, R.: Design of relational database schemes by deleting attributes in the canonical decomposition. J. Comput. Syst. Sci. **35**(1), 1–22 (1987)
12. Biskup, J., Dayal, U., Bernstein, P.: Synthesizing independent database schemas. In: SIGMOD Conference, pp. 143–151. ACM, New York (1979)
13. Codd, E.F.: A relational model of data for large shared data banks. Commun. ACM **13**(6), 377–387 (1970)
14. Codd, E.F.: Further normalization of the database relational model. In: Courant Computer Science Symposia 6: Data Base Systems, pp. 33–64. Prentice-Hall, Englewood Cliffs (1972)
15. Delobel, C.: Normalisation and hierarchical dependencies in the relational data model. ACM Trans. Database Syst. **3**(3), 201–222 (1978)
16. Fagin, R.: Multivalued dependencies and a new normal form for relational databases. ACM Trans. Database Syst. **2**(3), 262–278 (1977)
17. Fagin, R., Vardi, M.Y.: The theory of data dependencies: a survey. In: Mathematics of Information Processing: Proceedings of Symposia in Applied Mathematics, pp. 19–71. American Mathematical Society, Providence (1986)
18. Fischer, P.C., Saxton, L.V., Thomas, S.J., Van Gucht, D.: Interactions between dependencies and nested relational structures. J. Comput. Syst. Sci. **31**(3), 343–354 (1985)
19. Graetzer, G.: General Lattice Theory. Birkhauser, Boston (1998)
20. Hara, C., Davidson, S.: Reasoning about nested functional dependencies. In: Proceedings of the 18th SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems, pp. 91–100. ACM, New York (1999)

21. Hartmann, S., Link, S.: Multi-valued dependencies in the presence of lists. In: Proceedings of the 23rd SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems, pp. 330–341. ACM, New York (2004)
22. Hartmann, S., Link, S.: On a problem of Fagin concerning multivalued dependencies in relational databases. Theor. Comput. Sci. **353**(1–3), 53–62 (2006)
23. Hartmann, S., Link, S., Schewe, K.-D.: Functional and multivalued dependencies in nested databases generated by record and list constructor. Ann. Math. Artif. Intell. **46**(1–2), 114–164 (2006)
24. Levene, M., Loizou, G.: Axiomatisation of functional dependencies in incomplete relations. Theor. Comput. Sci. **206**(1–2), 283–300 (1998)
25. Lien, Y.E.: On the equivalence of data models. J. ACM **29**(2), 333–363 (1982)
26. Link, S.: On multivalued dependencies in fixed and undetermined universes. In: Fourth International Symposium on Foundations of Information and Knowledge Systems, number 3861 in Lecture Notes in Computer Science, pp. 257–276. Springer, Berlin Heidelberg New York (2006)
27. Link, S.: On the logical implication of multivalued dependencies with null values. In: Twelfth Computing: The Australasian Theory Symposium, number 51 in Conferences in Research and Practice in Information Technology, Hobart, 16–19 January 2006
28. Mendelzon, A.: On axiomatising multivalued dependencies in relational databases. J. ACM **26**(1), 37–44 (1979)
29. Nakata, M.: Dependencies in fuzzy databases: multivalued dependency. In: Proceedings of the Fifth IEEE International Conference on Fuzzy Systems, pp. 1591–1597. IEEE Computer Society, Los Alamitos (1996)
30. Sözat, M., Yazici, A.: A complete axiomatization for fuzzy functional and multivalued dependencies in fuzzy database relations. Fuzzy Sets Syst. **117**(2), 161–181 (2001)
31. Tari, Z., Stokes, J., Spaccapietra, S.: Object normal forms and dependency constraints for object-oriented schemata. ACM Trans. Database Syst. **22**, 513–569 (1997)
32. Thalheim, B.: Dependencies in Relational Databases. Teubner, Leipzig (1991)
33. Thalheim, B.: Entity-Relationship Modeling: Foundations of Database Technology. Springer, Berlin Heidelberg New York (2000)
34. Thalheim, B.: Conceptual treatment of multivalued dependencies. In: Conceptual Modeling—ER 2003, Proceedings, number 2813 in Lecture Notes in Computer Science, pp. 363–375. Springer, Berlin Heidelberg New York (2003)
35. Vardi, M.Y.: Fundamentals of dependency theory. In: Börger, E. (ed.) Trends in Theoretical Computer Science, pp. 171–224. Computer Science Press, Rockville (1987)
36. Vincent, M., Liu, J.: Multivalued dependencies in XML. In: Proceedings of the 20th British National Conference on Databases, number 2712 in Lecture Notes in Computer Science, pp. 4–18. Springer, Berlin Heidelberg New York (2003)
37. Vincent, M., Liu, J., Liu, C.: A redundancy free 4NF for XML. In: Proceedings of the 1st International XML Database Symposium, number 2824 in Lecture Notes in Computer Science, pp. 254–266. Springer, Berlin Heidelberg New York (2003)
38. Weddell, G.: Reasoning about functional dependencies generalized for semantic data models. ACM Trans. Database Syst. **17**(1), 32–64 (1992)
39. Wijsen, J.: Temporal FDs on complex objects. ACM Trans. Database Syst. **24**(1), 127–176 (1999)
40. Zaniolo, C.: Analysis and Design of Relational Schemata for Database Systems. Ph.D. thesis, UCLA, Tech. Rep. UCLA-ENG-7769 (1976)