Acta Cybernetica **20** (2011) 285–329.

Weak Functional Dependencies on Trees with Restructuring

Attila Sali* and Klaus-Dieter Schewe†

Abstract

We present an axiomatisation for weak functional dependencies, i.e. disjunctions of functional dependencies, in the presence of several constructors for complex values. The investigated constructors capture records, sets, multisets, lists, disjoint union and optionality, i.e. the complex values are indeed trees. The constructors cover the gist of all complex value data models including object oriented databases and XML. Functional and weak functional dependencies are expressed on a lattice of subattributes, which even carries the structure of a Brouwer algebra as long as the union-constructor is absent. Its presence, however, complicates all results and proofs significantly. The reason for this is that the union-constructor causes non-trivial restructuring rules to hold. In particular, if either the set- or the the union-constructor is absent, a subset of the rules is complete for the implication of ordinary functional dependencies, while in the general case no finite axiomatisation for functional dependencies exists.

Keywords: functional dependency, weak functional dependency, axiomatisation, complex values, restructuring, embedded dependency, rational tree

1 Introduction

In the relational data model (RDM) a lot of research has been spent on the theory of dependencies, i.e. first-order sentences that are supposed to hold for all database instances (see [3, 25]). Various classes of dependencies for the RDM have been introduced (see [32] for a survey), and large parts of database theory deals with the finite axiomatisation of these dependencies and the finite implication problem for them. That is to decide that a dependency φ is implied by a set of dependencies Σ, where implication refers to the fact that all finite models of Σ are also models of φ. The easiest, yet most important class of dependencies is the class of *functional*

*Alfréd Rényi Institute of Mathematics, Budapest, P.O.B.127, H-1364 Hungary, E-mail: sali@renyi.hu

†Software Competence Center Hagenberg, Hagenberg, Austria and Johannes-Kepler-University Linz, Research Institute for Applied Knowledge Processing, Linz, Austria, E-mail: kd.schewe@scch.at, kd.schewe@faw.at

DOI: 10.14232/actacyb.20.2.2011.5

dependencies (FDs). Armstrong (see [6]) was the first to give a finite axiomatisation for FDs.

Dependency theory is a cornerstone of database design, as the semantics of the application domain cannot be expressed only by structures. Database theory has to investigate the implications arising from the presence of dependencies. This means to describe semantically desirable properties of "well-designed" databases, e.g. the absence of redundancy, to characterise them (if possible) syntactically by in-depth investigation of the dependencies, and to develop algorithms to transform schemata into normal forms, which guarantee the desirable properties to be satisfied.

However, the field of databases is no longer the unique realm of the RDM. First, so called semantic data models have been developed (see e.g. [9, 22]), which were originally just meant to be used as design aids, as application semantics was assumed to be easier captured by these models (see the argumentation in [7, 10, 35]). Later on some of these models, especially the nested relational model (see e.g. [25]), object oriented models (see e.g. [30]) and object-relational models, the gist of which are captured by the higher-order Entity-Relationship model (HERM, see [33, 34]) have become interesting as data models in their own right and some dependency and normalisation theory has been carried over to these advanced data models (see [14, 23, 24, 25, 31] as samples of the many work done on this so far). Most recently, the major research interest is on the model of semi-structured data and XML (see e.g. [1]), which may also be regarded as some kind of object oriented model.

We refer to all these models as *higher-order data models*. This is, because the most important extension that came with these models was the introduction of constructors for complex values. These constructors usually comprise bulk constructors for sets, lists and multisets, a disjoint union constructor, and an optionality or null-constructor. In fact, all the structure of higher-order data models (including XML as far as XML can be considered a data model) is captured by the introduction of (some or all of) these constructors.

The key problem is to develop dependency theories (or preferably a unified theory) for the higher-order data models. The development of such a dependency theory will have a significant impact on understanding application semantics and laying the grounds for a logically founded theory of well-designed non-relational databases.

So far, mainly keys and FDs for advanced data models have been investigated (see [5, 8, 12, 13, 15, 19, 20, 26, 27, 37, 38]), and this has led to several normal form proposals (see [4, 5, 16, 37]). The work in [16] contains explicit definitions of redundancy and update anomalies and proves (in the spirit of the work in [36]) that the suggested higher-level normal form (HLNF) in the presence of FDs is indeed equivalent to the absence of redundancy and sufficient for the absence of update anomalies. The work in [18] deals with disjunctions of FDs leading to so-called weak functional dependencies (wFDs), while in [17], [21], [39] and [40] first attempts are made to generalise multi-valued dependencies.

The work in this article still deals with functional dependencies and weak functional dependencies, in particular with the axiomatisation problem. The motivation for this work is that all the approaches made so far only deal with part of the problem. In other words, we still do not have one coherent theory, but merely a patchwork of partial (though nevertheless non-trivial) results:

- The different approaches use different definitions of functional dependencies none of which subsumes the other ones. Arenas and Libkin (see [5]) and similarly Vincent and Liu (see [37]) formalise FDs using paths in XML trees, while Hartmann et al. (see [19]) exploit constructors for lists, disjoint unions and optionality. Despite some initial attempts (see e.g. [41]) so far no common framework subsuming all these different classes of FDs exists. In particular, the class of FDs in [19] has a finite axiomatisation, while the one investigated in [5] has not.

- No approach so far deals with all mentioned constructors at the same time. Hartmann et al. (see [20]) prove a finite axiomatisation taking all constructors into account except the disjoint union constructor. The proof exploits the underlying algebraic structure of Brouwer algebras. Hartmann et al. (see [19]) prove a finite axiomatisation taking all but the set and multiset constructors into account, but at the same time deal with embedded functional dependencies and recursion. Finally, Sali and Schewe (see [27]) take all constructors into account and prove a finite axiomatisation for a restricted class of FDs, which still subsumes the one in [20].

The first objective of the research reported in this article was to remove the remaining restrictions in previous work (see [27]) and to achieve a finite axiomatisation for FDs on models, in which all constructors are present. We will show that such an axiomatisation does not exist. More precisely, we show that we have non-axiomatisability, if the set and the union constructor are combined, whereas if one of them is absent, we obtain a finite axiomatisation. However, switching to the slightly extended class of weak functional dependencies we obtain a finite, though not k-ary axiomatisation. This axiomatisation contains a large number of structural axioms reflecting the non-trivial equivalences between subattributes, which caused significant challenges for the completeness proof. These equivalences result from restructuring rules, which were mostly known already long ago (see e.g. [2]).

Our second objective was to provide a framework that subsumes the existing approaches to dependency theory at outlined below. For this we extend the framework of nested attributes resulting from the various constructors, which in fact captures finite trees, to rational trees, i.e. we capture recursion. Furthermore, we deal with wFDs and FDs that are defined on embedded attributes. With these extensions the classes of FDs developed by Arenas, Libkin and Vincent, Liu, respectively, can be represented as special cases of the general class of FDs. The axiomatisation of the enlarged class of wFDs is straightforward, once the axiomatisation of wFDs in the presence of all constructors is known.

Overview

In Section 2 we define the preliminaries for our theory of wFDs. We start with the definition of nested attributes that are composed of simple attributes using the constructors that have been mentioned above. Each nested attribute defines a set of complex values called its domain, and each complex value can be represented as a finite tree. We then define subattributes, which give rise to canonical projection maps on the domains. The presence of the union constructor leads to restructuring rules, which define non-trivial equivalences the set of subattributes of a given nested attribute. Finally, we investigate the algebraic structure of the set of subattributes of a given nested attribute. We obtain a lattice, which is even a Brouwer algebra, if the union constructor is absent. Nevertheless, also in the general case it is advantageous to define the notion of relative pseudo-complement.

In Section 3 we study certain ideals in such lattices of subattributes, focusing on the set of subattributes, on which two complex values coincide. These ideals are therefore called *coincidence ideals*. The objective is to obtain a precise characterisation in the sense that whenever an ideal satisfies the given set of properties, we can guarantee the existence of two complex values that coincide exactly on the given ideal. This leads to the *Central Theorem* on coincidence ideals, which will be a cornerstone of the completeness proof. The proof of this result, however, appears in [28].

In Section 4 we introduce FDs and wFDs formally and first derive sound derivation rules, most of which are structural axioms reflecting the properties of coincidence ideals. The main result in this section will be the *Completeness Theorem* for the implication of wFDs. We then approach the simpler class of FDs and first show the completeness of a subset of the rules in case not both the set and the union constructors are used. If both appear together, we show non-axiomatisability. Thus, the results in Section 4 fulfil our first objective.

In Section 5 we approach our second objective. We first introduce embedded dependencies and show that they do not affect our axiomatisation of wFDs. In a second step we extend the definition of nested attributes capturing also rational tree values, as they are used in the object models (see e.g. [3] and [30]). We will show that the axiomatisation of wFDs will also be preserved by this extension.

In Section 6 we discuss the relationship with related work. We show that the classes of FDs defined by Arenas, Libkin and Vincent, Liu, respectively, are captured in our framework with all extensions discussed. We discuss the impact of this result.

Finally, we summarise our work and discuss conclusions in Section 7. This includes a brief discussion of additional restructuring rules, problems of keys and Armstrong instances, and an outlook on other classes of dependencies.

2 Algebras of Nested Attributes

In this section we define our model of nested attributes, which covers the gist of higher-order data models including XML. In particular, we investigate the structure of the set $\mathcal{S}(X)$ of subattributes of a given nested attribute X. We show that we obtain a lattice, which in general is non-distributive. This lattice becomes a Brouwer algebra, if the union constructor is not used.

2.1 Nested Attributes

We start with a definition of simple attributes and values for them.

Definition 1. A *universe* is a finite set $\mathcal{U}$ together with domains (i.e. sets of values) $dom(A)$ for all $A \in \mathcal{U}$. The elements of $\mathcal{U}$ are called *simple attributes*.

For the relational model a universe was sufficient, as a relation schema could be defined by a subset $R \subseteq \mathcal{U}$. For higher-order data models, however, we need nested attributes. In the following definition we use a set $\mathcal{L}$ of labels, and tacitly assume that the symbol λ is neither a simple attribute nor a label, i.e. $\lambda \notin \mathcal{U} \cup \mathcal{L}$, and that simple attributes and labels are pairwise different, i.e. $\mathcal{U} \cap \mathcal{L} = \emptyset$.

Definition 2. Let $\mathcal{U}$ be a universe and $\mathcal{L}$ a set of labels. The set $\mathcal{N}$ of *nested attributes* (over $\mathcal{U}$ and $\mathcal{L}$) is the smallest set with $\lambda \in \mathcal{N}$, $\mathcal{U} \subseteq \mathcal{N}$, and satisfying the following properties:

- for $X \in \mathcal{L}$ and $X'_1, \dots, X'_n \in \mathcal{N}$ we have $X(X'_1, \dots, X'_n) \in \mathcal{N}$;
- for $X \in \mathcal{L}$ and $X' \in \mathcal{N}$ we have $X\{X'\} \in \mathcal{N}$, $X[X'] \in \mathcal{N}$, and $X\langle X' \rangle \in \mathcal{N}$;
- for $X_1, \dots, X_n \in \mathcal{L}$ and $X'_1, \dots, X'_n \in \mathcal{N}$ we have $X_1(X'_1) \oplus \cdots \oplus X_n(X'_n) \in \mathcal{N}$.

We call λ a *null attribute*, $X(X'_1, \dots, X'_n)$ a *record attribute*, $X\{X'\}$ a *set attribute*, $X[X']$ a *list attribute*, $X\langle X' \rangle$ a *multiset attribute* and $X_1(X'_1) \oplus \cdots \oplus X_n(X'_n)$ a *union attribute*.

In the following we will overload the use of symbols such as X, Y, etc. for nested attributes and labels. As record, set, list and multiset attributes have a unique leading label, this will not cause problems anyway. In all other cases it is clear from the context, whether a symbol denotes a nested attribute in $\mathcal{N}$ or a label. Usually, labels never appear as stand-alone symbols.

We also take the freedom to change the leading label X in a set, list or multiset attribute to $X_{\{1,\dots,n\}}$, if the component attribute is a union attribute, say $X_1(X'_1) \oplus \cdots \oplus X_n(X'_n)$. This emphasises the factors in the union attribute. We will see in the next two subsections that this notation will become important, when restructuring is considered.

We can now extend the association *dom* from simple to nested attributes, i.e. for each $X \in \mathcal{N}$ we will define a set of values $dom(X)$.

Definition 3. For each nested attribute $X \in \mathcal{N}$ we get a *domain* $dom(X)$ as follows:

- $dom(\lambda) = \{\top\}$;

- $dom(X(X'_1,\dots,X'_n)) = \{(v_1,\dots,v_n) \mid v_i \in dom(X'_i) \text{ for } i=1,\dots,n\}$;
- $dom(X\{X'\}) = \{\{v_1,\dots,v_k\} \mid k \in \mathbb{N} \text{ and } v_i \in dom(X') \text{ for } i=1,\dots,k \text{ and } v_i \neq v_j \text{ for } i \neq j\}$, i.e. each element in $dom(X\{X'\})$ is a finite set with (pairwise different) elements in $dom(X')$;
- $dom(X[X']) = \{[v_1,\dots,v_k] \mid k \in \mathbb{N} \text{ and } v_i \in dom(X') \text{ for } i=1,\dots,k\}$, i.e. each element in $dom(X[X'])$ is a finite (ordered) list with (not necessarily different) elements in $dom(X')$;
- $dom(X\langle X'\rangle) = \{\langle v_1,\dots,v_k\rangle \mid k \in \mathbb{N} \text{ and } v_i \in dom(X') \text{ for } i=1,\dots,k\}$, i.e. each element in $dom(X\langle X'\rangle)$ is a finite multiset with elements in $dom(X')$, or in other words each $v \in dom(X')$ has a *multiplicity* $m(v) \in \mathbb{N}$ in a value in $dom(X\langle X'\rangle)$;
- $dom(X_1(X'_1) \oplus \cdots \oplus X_n(X'_n)) = \{(X_i : v_i) \mid v_i \in dom(X'_i) \text{ for } i=1,\dots,n\}$.

Note that the relational model is covered, if only the record constructor is used. Thus, instead of a relation schema R we will now consider a nested attribute X, assuming that the universe $\mathcal{U}$ and the set of labels $\mathcal{L}$ are fixed. Instead of an R-relation r we will consider a finite set $r \subseteq dom(X)$.

Further note that each complex value $v \in dom(X)$ for some nested attribute $X \in \mathcal{N}$ can be represented as a finite tree. This will be extended in Section 5 to rational trees.

2.2 Subattributes

In the relational model a functional dependency $X \to Y$ for $X, Y \subseteq R \subseteq \mathcal{U}$ is satisfied by an R-relation r iff any two tuples $t_1, t_2 \in r$ that coincide on all the attributes in X also coincide on the attributes in Y. Crucial to this definition is that we can project R-tuples to subsets of attributes.

Therefore, in order to define FDs on a nested attribute $X \in \mathcal{N}$ we need a notion of subattribute. For this we define a partial order $\geq$ on nested attributes in such a way that whenever $X \geq Y$ holds, we obtain a canonical projection $\pi^X_Y : dom(X) \to dom(Y)$. However, this partial order has to be defined on equivalence classes of attributes, as some domains may be identified.

Definition 4. $\equiv$ is the smallest *equivalence relation* on $\mathcal{N}$ satisfying the following properties:

- $\lambda \equiv X()$;
- $X(X'_1,\dots,X'_n) \equiv X(X'_1,\dots,X'_n,\lambda)$;
- $X(X'_1,\dots,X'_n) \equiv X(X'_{\sigma(1)},\dots,X'_{\sigma(n)})$ for any permutation $\sigma \in \mathbf{S}_n$;
- $X_1(X'_1) \oplus \cdots \oplus X_n(X'_n) \equiv X_{\sigma(1)}(X'_{\sigma(1)}) \oplus \cdots \oplus X_{\sigma(n)}(X'_{\sigma(n)})$ for any permutation $\sigma \in \mathbf{S}_n$;

- $X(X'_1, \dots, X'_n) \equiv X(Y_1, \dots, Y_n)$ if $X'_i \equiv Y_i$ for all $i = 1, \dots, n$;
- $X_1(X'_1) \oplus \cdots \oplus X_n(X'_n) \equiv X_1(Y_1) \oplus \cdots \oplus X_n(Y_n)$ if $X'_i \equiv Y_i$ for all $i = 1, \dots, n$;
- $X\{X'\} \equiv X\{Y\}$ if $X' \equiv Y$;
- $X[X'] \equiv X[Y]$ if $X' \equiv Y$;
- $X\langle X'\rangle \equiv X\langle Y\rangle$ if $X' \equiv Y$;
- $X(X'_1, \dots, Y_1(Y'_1) \oplus \cdots \oplus Y_m(Y'_m), \dots, X'_n) \equiv Y_1(X'_1, \dots, Y'_1, \dots, X'_n) \oplus \dots$ $\cdots \oplus Y_m(X'_1, \dots, Y'_m, \dots, X'_n)$;
- $X_{\{1,\dots,n\}}\{X_1(X'_1) \oplus \cdots \oplus X_n(X'_n)\} \equiv X_{\{1,\dots,n\}}(X_1\{X'_1\}, \dots, X_n\{X'_n\})$;
- $X_{\{1,\dots,n\}}\langle X_1(X'_1) \oplus \cdots \oplus X_n(X'_n)\rangle \equiv X_{\{1,\dots,n\}}(X_1\langle X'_1\rangle, \dots, X_n\langle X'_n\rangle)$.

Basically, the first four cases in this equivalence definition state that λ in record attributes can be added or removed, and that order in record and union attributes does not matter. The last three cases in Definition 4 cover restructuring rules, two of which were already introduced by Abiteboul and Hull (see [2]). Obviously, if we have a set of labelled elements with up to n different labels, we can split this set into n subsets, each of which contains just the elements with a particular label, and the union of these sets is the original set. The same holds for multisets. Of course, we can also split a list of labelled elements into lists containing only elements with the same label, thereby preserving the order, but in this case we cannot invert the splitting and thus cannot claim an equivalence.

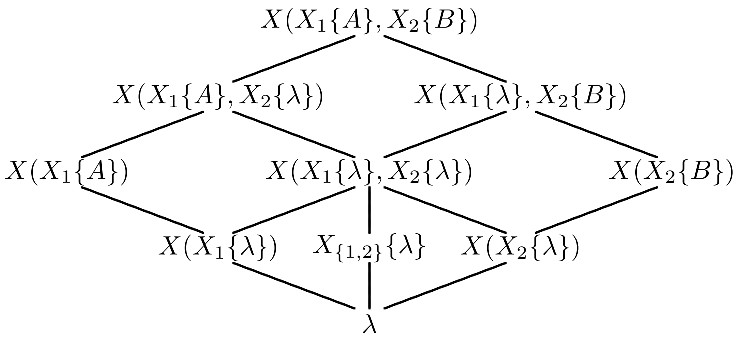

Figure 1: The lattice $\mathcal{S}(X\{X_1(A) \oplus X_2(B)\}) = \mathcal{S}(X(X_1\{A\}, X_2\{B\}))$

In the following we identify $\mathcal{N}$ with the set $\mathcal{N}/_\equiv$ of equivalence classes. In particular, we will write $=$ instead of $\equiv$, and in the following definition we should say that Y is a subattribute of X iff $\tilde{X} \geq \tilde{Y}$ holds for some $\tilde{X} \equiv X$ and $\tilde{Y} \equiv Y$. In particular, for $X \equiv Y$ we obtain $X \geq Y$ and $Y \geq X$.

Definition 5. For $X, Y \in \mathcal{N}$ we say that Y is a *subattribute* of X, iff $X \geq Y$ holds, where $\geq$ is the smallest partial order on $\mathcal{N}/_\equiv$ satisfying the following properties:

- $X \geq \lambda$ for all $X \in \mathcal{N}$;
- $X(Y_1, \dots, Y_n) \geq X(X'_{\sigma(1)}, \dots, X'_{\sigma(m)})$ for some injective $\sigma : \{1, \dots, m\} \to \{1, \dots, n\}$ and $Y_{\sigma(i)} \geq X'_{\sigma(i)}$ for all $i = 1, \dots, m$;
- $X_1(Y_1) \oplus \dots \oplus X_n(Y_n) \geq X_{\sigma(1)}(X'_{\sigma(1)}) \oplus \dots \oplus X_{\sigma(n)}(X'_{\sigma(n)})$ for some permutation $\sigma \in \mathbf{S}_n$ and $Y_i \geq X'_i$ for all $i = 1, \dots, n$;
- $X\{Y\} \geq X\{X'\}$ iff $Y \geq X'$;
- $X[Y] \geq X[X']$ iff $Y \geq X'$;
- $X\langle Y\rangle \geq X\langle X'\rangle$ iff $Y \geq X'$;
- $X_{\{1,\dots,n\}}[X_1(X'_1) \oplus \dots \oplus X_n(X'_n)] \geq X(X_1[X'_1], \dots, X_n[X'_n])$;
- $X_{\{1,\dots,k\}}[X_1(X'_1) \oplus \dots \oplus X_k(X'_k)] \geq X_{\{1,\dots,\ell\}}[X_1(X'_1) \oplus \dots \oplus X_\ell(X'_\ell)]$ for $k \geq \ell$;
- $X(X_{i_1}\{\lambda\}, \dots, X_{i_k}\{\lambda\}) \geq X_{\{i_1,\dots,i_k\}}\{\lambda\}$;
- $X(X_{i_1}\langle\lambda\rangle, \dots, X_{i_k}\langle\lambda\rangle) \geq X_{\{i_1,\dots,i_k\}}\langle\lambda\rangle$;
- $X(X_{i_1}[\lambda], \dots, X_{i_k}[\lambda]) \geq X_{\{i_1,\dots,i_k\}}[\lambda]$.

Note that the last four cases in Definition 5 cover further restructuring rules due to the union constructor. Obviously, if we are given a list of elements labelled with $X_1, \dots, X_n$, we can take the individual sublists – preserving the order – that contain only those elements labelled by X_i and build the tuple of these lists. In this case we can turn the label into a label for the whole sublist. This explains the first of the last four subattribute relationships.

For the other restructuring rules we have to add a little remark on notation here. As we identify $X\{X_1(X'_1) \oplus \dots \oplus X_n(X'_n)\}$ with $X(X_1\{X'_1\}, \dots, X_n\{X'_n\})$, we obtain subattributes $X(X_{i_1}\{X'_{i_1}\}, \dots \dots, X_{i_k}\{X'_{i_k}\})$ for each subset $I = \{i_1, \dots, i_k\} \subseteq \{1, \dots, n\}$. However, restructuring requires some care with labels. If we simply reused the label X in the last property in Definition 5, we would obtain

$$X\{X_1(X'_1) \oplus X_2(X'_2)\} \equiv X(X_1\{X'_1\}, X_2\{X'_2\}) \geq X(X_1\{X'_1\}) \geq X(X_1\{\lambda\}) \geq X\{\lambda\}.$$

However, the last step here is wrong, as the left hand side is an indicator for the subset containing the elements with label X_1 being empty or not, whereas the right hand side is the corresponding indicator for the whole set, i.e. elements with labels X_1 or X_2. No such mapping can be claimed. In fact, what we really have to do is to mark the set label in an attribute of the form $X\{X_1(X'_1) \oplus \dots \oplus X_n(X'_n)\}$ to indicate the inner union attribute, i.e. we should use $X_{\{1,\dots,n\}}$ (or even $X_{\{X_1,\dots,X_n\}}$) instead of X. As long as we are not dealing with subattributes of the form $X_{\{1,\dots,k\}}\{\lambda\}$, the additional index does not add any information and thus can be omitted to increase readability. The same applies to the multiset- and the list-constructor.

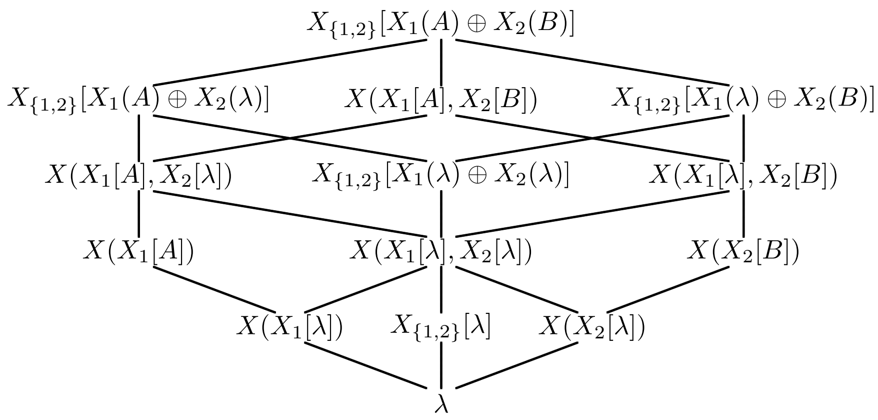

Figure 2: The lattice $\mathcal{S}(X[X_1(A) \oplus X_2(B)])$

Subattributes of the form $X_I\{\lambda\}$, $X_I[\lambda]$ and $X_I\{\lambda\}$ were called *counter attributes* in [27], because they can be considered as counters for the number of elements in a list or multiset or as flags that tell, whether sets are empty or not. Note that $X_\emptyset\{\lambda\} = \lambda$, $X_{\{1,\dots,n\}}\{\lambda\} = X\{\lambda\}$ and $X_{\{i\}}\{\lambda\} = X(X_i\{\lambda\})$. Analogous conventions apply to list and multiset attributes.

Further note that due to the restructuring rules in Definitions 4 and 5 we may have the case that a record attribute is a subattribute of a set attribute and vice versa. This cannot be the case, if the union-constructor is absent. However, the presence of the restructuring rules allows us to assume that the union-constructor only appears inside a set-constructor or as the outermost constructor. This will be frequently exploited in our proofs.

Obviously, $X \geq Y$ induces a projection map $\pi_Y^X : dom(X) \to dom(Y)$. For $X \equiv Y$ we have $X \geq Y$ and $Y \geq X$ and the projection maps π_Y^X and π_X^Y are inverse to each other.

We use the notation $\mathcal{S}(X) = \{Z \in \mathcal{N} \mid X \geq Z\}$ to denote the *set of subattributes* of a nested attribute X. Figure 1 shows the subattributes of $X\{X_1(A) \oplus X_2(B)\} = X(X_1\{A\}, X_2\{B\})$ together with the relation $\geq$ on them. Note that the subattribute $X_{\{1,2\}}\{\lambda\}$ would not occur, if we only considered the record-structure, whereas other subattributes such as $X(X_i\{\lambda\})$ would not occur, if we only considered the set-structure. This is a direct consequence of the restructuring rules.

Figure 2 shows the subattributes of $X[X_1(A) \oplus X_2(B)]$ together with the relation $\geq$ on them. The subattributes $X_{\{1,2\}}[\lambda]$ would not occur, if we only considered the list-structure, whereas other subattributes such as $X(X_i[\lambda])$ would not occur, if we ignored the restructuring rules. Figure 3 shows the subattributes of $X\{X_1(A) \oplus X_2(B) \oplus X_3(C)\}$ together with the relation $\geq$ on them. The subattribute $X_I\{\lambda\}$ for $|I| \geq 2$ would not occur, if we only considered the record-structure.

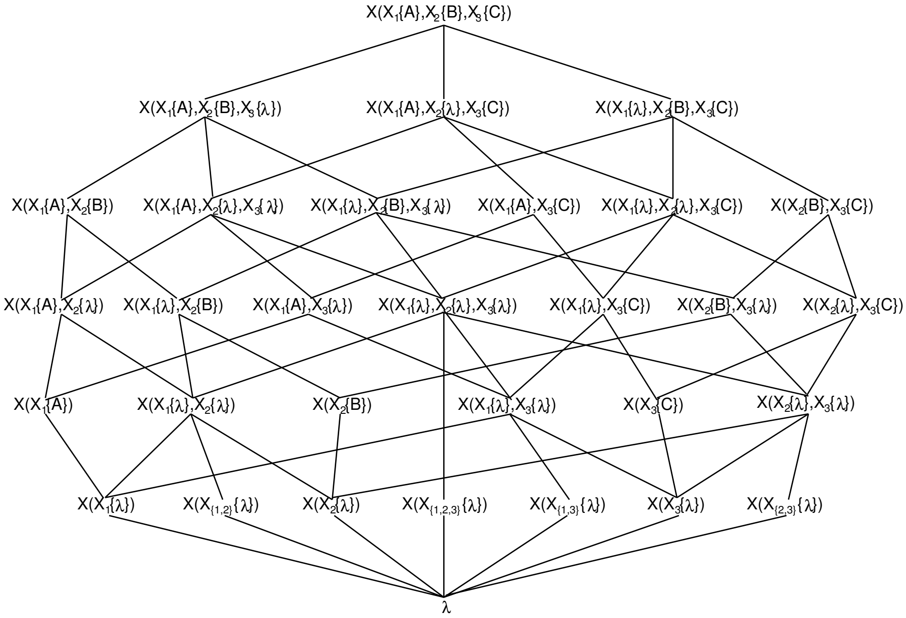

Figure 3: The subattribute lattice $\mathcal{S}(X\{X_1(A) \oplus X_2(B) \oplus X_3(C)\})$

2.3 The Lattice Structure

The set of subattributes $\mathcal{S}(X)$ of a nested attribute X plays the same role in the dependency theory for higher-order data models as the powerset $\mathcal{P}(R)$ for a relation schema R plays in the dependency theory for the relational model. $\mathcal{P}(R)$ is a Boolean algebra with order $\subseteq$, intersection $\cap$, union $\cup$ and the difference $-$. So, the question arises which algebraic structure $\mathcal{S}(X)$ carries.

Definition 6. Let $\mathcal{L}$ be a lattice with zero and one, partial order $\leq$, join $\sqcup$ and meet $\sqcap$. $\mathcal{L}$ has *relative pseudo-complements* iff for all $Y, Z \in \mathcal{L}$ the infimum $Y \leftarrow Z = \bigsqcap\{U \mid U \sqcup Y \geq Z\}$ exists. Then $Y \leftarrow 1$ (1 being the one in $\mathcal{L}$) is called the *relative complement* of Y.

If we have distributivity in addition, we call $\mathcal{L}$ a *Brouwer algebra*. In this case the relative pseudo-complements satisfy $U \geq (Y \leftarrow Z)$ iff $(U \sqcup Y \geq Z)$, but if we do not have distributivity this property may be violated though relative pseudo-complements exist.

Theorem 1. *The set $\mathcal{S}(X)$ of subattributes carries the structure of a lattice with zero and one and relative pseudo-complements, where the order $\geq$ is as defined in Definition 5, and λ and X are the zero and one, respectively. If X does not contain the union constructor, $\mathcal{S}(X)$ defines a Brouwer algebra.*

Proof. For $X = \lambda$ and simple attributes $X = A$ we obtain trivial lattices with only one or two elements. Applying the record constructor leads to a cartesian product of lattices, while the set, list and multiset constructors add a new zero element to a lattice. These extensions preserve the properties of a Brouwer algebra.

In the case of set, list and multiset constructors applied to a union attribute we add counter attributes. This preserves the properties of a lattice and the existence of relative pseudo-complement, while distributivity may be lost.

□

Example 1. Let $X = X\{X_1(A) \oplus X_2(B)\}$ with $\mathcal{S}(X)$ as illustrated in Figure 1, $Y_1 = X\{\lambda\}$, $Y_2 = X(X_2\{B\})$, and $Z = X(X_1\{A\})$. Then we have

$$Z \sqcap (Y_1 \sqcup Y_2) = X(X_1\{A\}) \sqcap (X\{\lambda\} \sqcup X(X_2\{B\})) =$$
$$X(X_1\{A\}) \sqcap X(X_1\{\lambda\}, X_2\{B\}) = X(X_1\{\lambda\}) \neq \lambda = \lambda \sqcup \lambda =$$
$$(X(X_1\{A\}) \sqcap X\{\lambda\}) \sqcup (X(X_1\{A\}) \sqcap X(X_2\{B\})) = (Z \sqcap Y_1) \sqcup (Z \sqcap Y_2) \ .$$

This shows that $\mathcal{S}(X)$ in general is not a distributive lattice. Furthermore, $Y' \sqcup Z \geq Y_1$ holds for all Y' except λ, $X(X_1\{\lambda\})$ and $X(X_1\{A\})$. So $Z \leftarrow Y_1 = \lambda$, but not all $Y' \geq \lambda$ satisfy $Y' \sqcup Z \geq Y_1$.

It is easy to determine explicit inductive definitions of the operations $\sqcap$ (meet), $\sqcup$ (join) and $\leftarrow$ (relative pseudo-complement). This can be done by boring technical verification of the properties of meets, joins and relative pseudo-complements and is therefore omitted here.

3 Coincidence Ideals

In this section we investigate sets of subattributes, on which two complex values coincide. It is rather easy to see that these turn out to be ideals in the lattice $\mathcal{S}(X)$, i.e. they are non-empty and downward-closed. Therefore, we will call them *coincidence ideals*. However, there are many other properties that hold for coincidence ideals.

There are two major reasons for looking at coincidence ideals. The first one is that properties of subattributes, on which two complex values coincide, may give rise to axioms for functional dependencies. We will indeed see that the properties of coincidence ideals in Definition 7 are very closely related to the sound axioms and rules that we will derive in Theorems 3, 5 and 6.

The second reason is that in the completeness proof we will have to construct two complex values that coincide exactly on a given set of attributes, so that a set of dependencies is satisfied by these values, while a non-derivable dependency is not. This step appears also in the corresponding completeness proof for the RDM, but in that case it is trivial, because it simply amounts to getting two tuples that coincide on a given set of attributes, but differ on all others.

Thus, what we want to achieve is a characterisation of a coincidence ideal that allows us to construct two complex values that coincide exactly on it. This will be the main result of this section, called the Central Theorem 2 on coincidence ideals.

The proof of this result in [28] is very technical. In a nutshell, what we did was to discover properties of coincidence ideals, "translate" them into axioms for (weak) functional dependencies, ensure that we can rediscover these properties from the particular set of subattributes that arises naturally in the completeness proof (see Lemma 2), which required to weaken the axioms as much as possible, and finally show that the properties are sufficient for the desired Central Theorem.

Definition 7. A subset $\mathcal{F} \subseteq \mathcal{S}(X)$ is called a *coincidence ideal* on $\mathcal{S}(X)$ iff there exist complex values $t_1, t_2 \in dom(X)$ such that $\mathcal{F} = \{Y \in \mathcal{S}(X) \mid \pi_Y^X(t_1) = \pi_Y^X(t_2)\} \subseteq \mathcal{S}(X)$ is the set of subattributes, on which they coincide.

In [18] and in [26] the term "SHL-ideal" was used instead; in [19] in a restricted setting the term "HL-ideal" was used. Note that in all these cases not all the conditions in Theorem 2 were yet present.

In order to characterise sufficient and necessary properties of coincidence ideals we will need the notion of reconsilable subattributes, which was already used in the axiomatisations of restricted cases (see [19, 20]). The following Definition 8 extends this notion to capture all constructors, in particular the union constructor.

Definition 8. Two subattributes $Y, Z \in \mathcal{S}(X)$ are called *reconsilable* iff one of the following holds:

1. $Y \geq Z$ or $Z \geq Y$;
2. $X = X[X']$, $Y = X[Y']$, $Z = X[Z']$ and $Y', Z' \in \mathcal{S}(X')$ are reconsilable;
3. $X = X(X_1, \dots, X_n)$, $Y = X(Y_1, \dots, Y_n)$, $Z = X(Z_1, \dots, Z_n)$ and $Y_i, Z_i \in \mathcal{S}(X_i)$ are reconsilable for all $i = 1, \dots, n$;
4. $X = X_1(X_1') \oplus \dots \oplus X_n(X_n')$, $Y = X_1(Y_1') \oplus \dots \oplus X_n(Y_n')$, $Z = X_1(Z_1') \oplus \dots \oplus X_n(Z_n')$ and $Y_i', Z_i' \in \mathcal{S}(X_i')$ are reconsilable for all $i = 1, \dots, n$;
5. $X = X[X_1(X_1') \oplus \dots \oplus X_n(X_n')]$, $Y = X(Y_1, \dots, Y_n)$ with $Y_i = X_i[Y_i']$ or $Y_i = \lambda = Y_i'$, $Z = X[X_1(Z_1') \oplus \dots \oplus X_n(Z_n')]$, and Y_i', Z_i' are reconsilable for all $i = 1, \dots, n$.

Note that for the set- and multiset-constructor we can only obtain reconsilability for subattributes in a $\geq$-relation.

Theorem 2 (Central Theorem)**.** *Let $X \in \mathcal{N}$ be a nested attribute. Then $\mathcal{F} \subseteq \mathcal{S}(X)$ is a coincidence ideal iff the following conditions are satisfied:*

1. $\lambda \in \mathcal{F}$;
2. *if $Y \in \mathcal{F}$ and $Z \in \mathcal{S}(X)$ with $Y \geq Z$, then $Z \in \mathcal{F}$;*
3. *if $Y, Z \in \mathcal{F}$ are reconsilable, then $Y \sqcup Z \in \mathcal{F}$;*
4. a) *if $X_I\{\lambda\} \in \mathcal{F}$ and $X_J\{\lambda\} \notin \mathcal{F}$ for $I \subsetneq J$, then $X(X_{i_1}\{X_{i_1}'\}, \dots, X_{i_k}\{X_{i_k}'\}) \in \mathcal{F}$ for $I = \{i_1, \dots, i_k\}$;*

b) *if* $X_I\{\lambda\} \in \mathcal{F}$ *and* $X(X_i\{\lambda\}) \notin \mathcal{F}$ *for all* $i \in I$*, then there is a partition* $I = I_1 \dot{\cup} I_2$ *with* $X_{I_1}\{\lambda\} \notin \mathcal{F}$, $X_{I_2}\{\lambda\} \notin \mathcal{F}$ *and* $X_{I'}\{\lambda\} \in \mathcal{F}$ *for all* $I' \subseteq I$ *with* $I' \cap I_1 \neq \emptyset \neq I' \cap I_2$*;*

c) *if* $X_{\{1,\dots,n\}}\{\lambda\} \in \mathcal{F}$ *and* $X_{I^-}\{\lambda\} \notin \mathcal{F}$ *(for* $I^- = \{i \in \{1,\dots,n\} \mid X(X_i\{\lambda\}) \notin \mathcal{F}\}$*), then there exists some* $i \in I^+ = \{i \in \{1,\dots,n\} \mid X(X_i\{\lambda\}) \in \mathcal{F}\}$ *such that for all* $J \subseteq I^-$ $X_{J\cup\{i\}}\{\lambda\} \in \mathcal{F}$ *holds;*

d) *if* $X_J\{\lambda\} \notin \mathcal{F}$ *and* $X_{\{j\}}\{\lambda\} \notin \mathcal{F}$ *for all* $j \in J$ *and for all* $i \in I$ *there is some* $J_i \subseteq J$ *with* $X_{J_i\cup\{i\}}\{\lambda\} \notin \mathcal{F}$*, then* $X_{I\cup J}\{\lambda\} \notin \mathcal{F}$*, provided* $I \cap J = \emptyset$*;*

e) *if* $X_{I^-}\{\lambda\} \in \mathcal{F}$ *and* $I' \subseteq I^+$ *such that for all* $i \in I'$ *there is some* $J \subseteq I^-$ *with* $X_{J\cup\{i\}}\{\lambda\} \notin \mathcal{F}$*, then* $X_{I'\cup J'}\{\lambda\} \notin \mathcal{F}$ *for all* $J' \subseteq I^-$ *with* $X_{J'}\{\lambda\} \notin \mathcal{F}$*;*

5. a) *if* $X_I\{\lambda\} \in \mathcal{F}$ *and* $X_J\{\lambda\} \in \mathcal{F}$ *with* $I \cap J = \emptyset$*, then* $X_{I\cup J}\{\lambda\} \in \mathcal{F}$*;*

   b) *if* $X_I[\lambda] \in \mathcal{F}$ *and* $X_J[\lambda] \in \mathcal{F}$ *with* $I \cap J = \emptyset$*, then* $X_{I\cup J}[\lambda] \in \mathcal{F}$*;*

   c) *if* $X_I\langle\lambda\rangle \in \mathcal{F}$ *and* $X_J\langle\lambda\rangle \in \mathcal{F}$ *with* $I \cap J = \emptyset$*, then* $X_{I\cup J}\langle\lambda\rangle \in \mathcal{F}$*;*

   d) *if* $X_I[\lambda] \in \mathcal{F}$ *and* $X_J[\lambda] \in \mathcal{F}$ *with* $J \subseteq I$*, then* $X_{I-J}[\lambda] \in \mathcal{F}$*;*

   e) *if* $X_I\langle\lambda\rangle \in \mathcal{F}$ *and* $X_J\langle\lambda\rangle \in \mathcal{F}$ *with* $J \subseteq I$*, then* $X_{I-J}\langle\lambda\rangle \in \mathcal{F}$*;*

   f) *if* $X_I[\lambda] \in \mathcal{F}$ *and* $X_J[\lambda] \in \mathcal{F}$*, then* $X_{I\cap J}[\lambda] \in \mathcal{F}$ *iff* $X_{(I-J)\cup(J-I)}[\lambda] \in \mathcal{F}$*;*

   g) *if* $X_I\langle\lambda\rangle \in \mathcal{F}$ *and* $X_J\langle\lambda\rangle \in \mathcal{F}$*, then* $X_{I\cap J}\langle\lambda\rangle \in \mathcal{F}$ *iff* $X_{(I-J)\cup(J-I)}\langle\lambda\rangle \in \mathcal{F}$*;*

6. a) *for* $X = X\{\bar{X}\{X_1(X'_1)\oplus\cdots\oplus X_n(X'_n)\}\}$*, whenever* $I \subseteq \{1,\dots,n\}$*, there is a partition* $I = I^- \cup I_{+-} \cup I_+ \cup I_-$ *such that*

      i. $X\{\bar{X}_{\{i\}}\{\lambda\}\} \in \mathcal{F}$ *iff* $i \notin I^-$*,*
      ii. $X\{\bar{X}_{I'}\{\lambda\}\} \in \mathcal{F}$*, whenever* $I' \cap I_+ \neq \emptyset$*,*
      iii. $X\{\bar{X}_{I'}\{\lambda\}\} \in \mathcal{F}$ *iff* $X\{\bar{X}_{I'\cap(I_{+-}\cup I^-)}\{\lambda\}\} \in \mathcal{F}$*, whenever* $I' \subseteq I_{+-} \cup I^- \cup I_-$*;*

   b) *for* $X = X\langle\bar{X}\{X_1(X'_1)\oplus\cdots\oplus X_n(X'_n)\}\rangle$*, whenever* $I \subseteq \{1,\dots,n\}$*, there is a partition* $I = I^- \cup I_{+-} \cup I_+ \cup I_-$ *such that*

      i. $X\langle\bar{X}_{\{i\}}\{\lambda\}\rangle \in \mathcal{F}$ *iff* $i \notin I^-$*,*
      ii. $X\langle\bar{X}_{I'}\{\lambda\}\rangle \in \mathcal{F}$*, whenever* $I' \cap I_+ \neq \emptyset$*,*
      iii. $X\langle\bar{X}_{I'}\{\lambda\}\rangle \in \mathcal{F}$ *iff* $X\langle\bar{X}_{I'\cap(I_{+-}\cup I^-)}\{\lambda\}\rangle \in \mathcal{F}$*, whenever* $I' \subseteq I_{+-} \cup I^- \cup I_-$*;*

7. a) *if* $X = X(X'_1,\dots,X'_n)$*, then* $\mathcal{F}_i = \{Y_i \in \mathcal{S}(X'_i) \mid X(\lambda,\dots,Y_i,\dots,\lambda) \in \mathcal{F}\}$ *is a coincidence ideal;*

   b) *if* $X = X[X']$*, such that* X' *is not a union attribute, and* $\mathcal{F} \neq \{\lambda\}$*, then* $\mathcal{G} = \{Y \in \mathcal{S}(X') \mid X[Y] \in \mathcal{F}\}$ *is a coincidence ideal;*

   c) *If* $X = X_1(X'_1) \oplus \cdots \oplus X_n(X'_n)$ *and* $\mathcal{F} \neq \{\lambda\}$*, then the set* $\mathcal{F}_i = \{Y_i \in \mathcal{S}(X'_i) \mid X_1(\lambda) \oplus \cdots \oplus X_i(Y_i) \oplus \cdots \oplus X_n(\lambda) \in \mathcal{F}\}$ *is a coincidence ideal;*

d) if $X = X\{X'\}$, such that X' is not a union attribute, and $\mathcal{F} \neq \{\lambda\}$, then $\mathcal{G} = \{Y \in \mathcal{S}(X') \mid X\{Y\} \in \mathcal{F}\}$ is a defect coincidence ideal;

e) if $X = X\langle X'\rangle$, such that X' is not a union attribute, and $\mathcal{F} \neq \langle\lambda\rangle$, then $\mathcal{G} = \{Y \in \mathcal{S}(X') \mid X\langle Y\rangle \in \mathcal{F}\}$ is a defect coincidence ideal.

In property 7 of the theorem a *defect coincidence ideal* on $\mathcal{S}(X)$ is a subset $\mathcal{F} \subseteq \mathcal{S}(X)$ satisfying properties 1, 2, 4(a)-(d), 6(a),(b), 7(d)-(e) and

8. a) if $X = X(X'_1, \dots, X'_n)$, then $\mathcal{F}_i = \{Y_i \in \mathcal{S}(X'_i) \mid X(\lambda, \dots, Y_i, \dots, \lambda) \in \mathcal{F}\}$ is a defect coincidence ideal;

   b) if $X = X[X']$, such that X' is not a union attribute, and $\mathcal{F} \neq \{\lambda\}$, then $\mathcal{G} = \{Y \in \mathcal{S}(X') \mid X[Y] \in \mathcal{F}\}$ is a defect coincidence ideal;

   c) If $X = X_1(X'_1) \oplus \cdots \oplus X_n(X'_n)$ and $\mathcal{F} \neq \{\lambda\}$, then the set $\mathcal{F}_i = \{Y_i \in \mathcal{S}(X'_i) \mid X_1(\lambda) \oplus \cdots \oplus X_i(Y_i) \oplus \cdots \oplus X_n(\lambda) \in \mathcal{F}\}$ is a defect coincidence ideal.

The proof of Theorem 2, in particular, showing that the conditions are sufficient, is very technical and lengthy (see [28]). The general idea is to use structural induction extending the corresponding proofs in [19] and in [20]. However, a difficulty arises with the set and multiset constructors, as for them defect coincidence ideals have to be dealt with. The work in [20, Lemmata 21 and 24] contains a proof for the case that the union constructor does not appear at all. This has been generalised in [27, Lemma 4.3] to the general case but excluding counter attributes, i.e. attributes of the form $X_I\{\lambda\}$, $X_I\langle\lambda\rangle$ or $X_I[\lambda]$ with $|I| \geq 2$.

4 Functional Dependencies and Weak Functional Dependencies

In this section we will define functional and weak functional dependencies on $\mathcal{S}(X)$ and derive a sound and complete system of derivation rules for wFDs.

Definition 9. Let $X \in \mathcal{N}$. A *functional dependency* (FD) on $\mathcal{S}(X)$ is an expression $\mathcal{Y} \to \mathcal{Z}$ with $\mathcal{Y}, \mathcal{Z} \subseteq \mathcal{S}(X)$. A *weak functional dependency* (wFD) on $\mathcal{S}(X)$ is an expression $\{\!|\mathcal{Y}_i \to \mathcal{Z}_i \mid i \in I|\!\}$ with an index set I and $\mathcal{Y}_i, \mathcal{Z}_i \subseteq \mathcal{S}(X)$.

In the following we consider finite sets $r \subseteq dom(X)$, which we will call simply *instances* of X.

Definition 10. Let r be an instance of X. We say that r *satisfies the FD* $\mathcal{Y} \to \mathcal{Z}$ on $\mathcal{S}(X)$ (notation: $r \models \mathcal{Y} \to \mathcal{Z}$) iff for all $t_1, t_2 \in r$ with $\pi_Y^X(t_1) = \pi_Y^X(t_2)$ for all $Y \in \mathcal{Y}$ we also have $\pi_Z^X(t_1) = \pi_Z^X(t_2)$ for all $Z \in \mathcal{Z}$.

An instance $r \subseteq dom(X)$ *satisfies the wFD* $\{\!|\mathcal{Y}_i \to \mathcal{Z}_i \mid i \in I|\!\}$ on $\mathcal{S}(X)$ (notation: $r \models \{\!|\mathcal{Y}_i \to \mathcal{Z}_i \mid i \in I|\!\}$) iff for all $t_1, t_2 \in r$ there is some $i \in I$ with $\{t_1, t_2\} \models \mathcal{Y}_i \to \mathcal{Z}_i$.

According to this definition we identify a wFD $\{\!\!\{\mathcal{Y} \to \mathcal{Z}\}\!\!\}$, i.e. the index set contains exactly one element, with the "ordinary" FD $\mathcal{Y} \to \mathcal{Z}$.

Note that our notion of weak functional dependencies is indeed more general than the one used in [32, p.75] based on the work by Demetrovics and Gyepesi (see [11]). The straighforward generalisation of the dependencies introduced by Demetrovics and Gyepesi would only lead to wFDs of the form $\{\!\!\{\mathcal{Y} \to \{Z_i\} \mid i \in I\}\!\!\}$, i.e. the left hand side of all involved FDs is always the same, while the right hand side only contains a single subattribute. Our notion of wFDs covers also so called *dual functional dependencies* (dFDs) (see [11]), which would take the form $\{\!\!\{\{Y_i\} \to \{Z_j\} \mid i \in I, j \in J\}\!\!\}$.

Let Σ be a set of FDs and wFDs. A FD or wFD ψ is implied by Σ (notation: $\Sigma \models \psi$) iff all instances r with $r \models \varphi$ for all $\varphi \in \Sigma$ also satisfy ψ. As usual we write $\Sigma^* = \{\psi \mid \Sigma \models \psi\}$.

As usual we write Σ^+ for the set of all FDs and wFDs that can be derived from Σ by applying a system $\mathfrak{R}$ of axioms and rules, i.e. $\Sigma^+ = \{\psi \mid \Sigma \vdash_{\mathfrak{R}} \psi\}$. We omit the standard definitions of derivations with a given rule system, and also write simply $\vdash$ instead of $\vdash_{\mathfrak{R}}$, if the rule system is clear from the context.

Our goal is to find a finite axiomatisation, i.e. a finite rule system $\mathfrak{R}$ such that $\Sigma^* = \Sigma^+$ holds. The rules in $\mathfrak{R}$ are *sound* iff $\Sigma^+ \subseteq \Sigma^*$ holds, and *complete* iff $\Sigma^* \subseteq \Sigma^+$ holds.

4.1 Sound Derivation Rules

Let us first look only at FDs. In general, two complex values in $dom(X)$ that coincide on subattributes Y and Z of X need not coincide on $Y \sqcup Z$. So we cannot expect a simple generalisation of Armstrong's extension rule for FDs in the relational model. However, the notion of "reconsilability" introduced in Definition 8 will permit such a generalisation.

Theorem 3. *The following axioms and rules are sound for the implication of FDs on $\mathcal{S}(X)$:*

reflexivity axiom:

$$\frac{}{\mathcal{Y} \to \mathcal{Z}}\ \mathcal{Z} \subseteq \mathcal{Y} \tag{1}$$

subattribute axiom:

$$\frac{}{\{Y\} \to \{Z\}}\ Y \geq Z \tag{2}$$

join axiom:

$$\frac{}{\{Y, Z\} \to \{Y \sqcup Z\}}\ Y, Z \textit{ reconsilable} \tag{3}$$

λ axiom:

$$\frac{}{\emptyset \to \{\lambda\}} \tag{4}$$

extension rule:

$$\frac{\mathcal{Y} \rightarrow \mathcal{Z}}{\mathcal{Y} \rightarrow \mathcal{Y} \cup \mathcal{Z}} \tag{5}$$

transitivity rule:

$$\frac{\mathcal{Y} \rightarrow \mathcal{Z} \quad \mathcal{Z} \rightarrow \mathcal{U}}{\mathcal{Y} \rightarrow \mathcal{U}} \tag{6}$$

Proof. The soundness of the axioms (1), (2) and (4) is trivial.

For (3) let $t_1, t_2 \in r$ for some instance $r \subseteq dom(X)$ with $\pi^X_Y(t_1) = \pi^X_Y(t_2)$ and $\pi^X_Z(t_1) = \pi^X_Z(t_2)$ for reconsilable subattributes $Y, Z \in \mathcal{S}(X)$.

1. In case $Y \geq Z$ we have $Y \sqcup Z = Y$ and thus $\pi^X_{Y \sqcup Z}(t_1) = \pi^X_{Y \sqcup Z}(t_2)$.

2. In case $X = X[X']$ we must have $Y = X[Y']$ and $Z = X[Z']$ with reconsilable subattributes $Y', Z' \in \mathcal{S}(X')$. Furthermore, $t_1 = [t_{1,1}, \dots, t_{1,n}]$ and $t_2 = [t_{2,1}, \dots, t_{2,m}]$. This gives $n = m$, $\pi^{X'}_{Y'}(t_{1,j}) = \pi^{X'}_{Y'}(t_{2,j})$ and $\pi^{X'}_{Z'}(t_{1,j}) = \pi^{X'}_{Z'}(t_{2,j})$ for all $j = 1, \dots, n$. By induction we obtain $\pi^{X'}_{Y' \sqcup Z'}(t_{1,j}) = \pi^{X'}_{Y' \sqcup Z'}(t_{2,j})$ for all $j = 1, \dots, n$. From this and $Y \sqcup Z = X[Y' \sqcup Z']$ follows $\pi^X_{Y \sqcup Z}(t_1) = \pi^X_{Y \sqcup Z}(t_2)$.

3. In case $X = X(X_1, \dots, X_n)$ we must have $Y = X(Y_1, \dots, Y_n)$ and $Z = X(Z_1, \dots, Z_n)$ with reconsilable subattributes $Y_i, Z_i \in \mathcal{S}(X_i)$ for $i = 1, \dots, n$. Furthermore, $t_1 = (t_{1,1}, \dots, t_{1,n})$ and $t_2 = (t_{2,1}, \dots, t_{2,n})$, which implies $\pi^{X_i}_{Y_i}(t_{1,i}) = \pi^{X_i}_{Y_i}(t_{2,i})$ and $\pi^{X_i}_{Z_i}(t_{1,i}) = \pi^{X_i}_{Z_i}(t_{2,i})$ for all $i = 1, \dots, n$. By induction we obtain $\pi^{X_i}_{Y_i \sqcup Z_i}(t_{1,i}) = \pi^{X_i}_{Y_i \sqcup Z_i}(t_{2,i})$ for all $i = 1, \dots, n$. From this and $Y \sqcup Z = X(Y_1 \sqcup Z_1, \dots, Y_n \sqcup Z_n)$ follows $\pi^X_{Y \sqcup Z}(t_1) = \pi^X_{Y \sqcup Z}(t_2)$.

4. In case $X = X_1(X'_1) \oplus \cdots \oplus X_n(X'_n)$ we must have $Y = X_1(Y_1) \oplus \cdots \oplus X_n(Y_n)$ and $Z = X_1(Z_1) \oplus \cdots \oplus X_n(Z_n)$ with reconsilable subattributes $Y_i, Z_i \in \mathcal{S}(X'_i)$ for $i = 1, \dots, n$. Furthermore $t_1 = (X_i : t'_1)$ and $t_2 = (X_i : t'_2)$ for some $i \in \{1, \dots, n\}$, which implies $\pi^{X'_i}_{Y_i}(t'_1) = \pi^{X'_i}_{Y_i}(t'_2)$ and $\pi^{X'_i}_{Z_i}(t'_1) = \pi^{X'_i}_{Z_i}(t'_2)$. By induction we obtain $\pi^{X'_i}_{Y_i \sqcup Z_i}(t'_1) = \pi^{X'_i}_{Y_i \sqcup Z_i}(t'_2)$. Finally, $Y \sqcup Z = X_1(Y_1 \sqcup Z_1) \oplus \cdots \oplus X_n(Y_n \sqcup Z_n)$ implies $\pi^X_{Y \sqcup Z}(t_1) = \pi^X_{Y \sqcup Z}(t_2)$ as desired.

5. In case $X = X[X_1(X'_1) \oplus \cdots \oplus X_n(X'_n)]$ we must have $Y = X(Y_1, \dots, Y_n)$ with $Y_i = X_i[Y'_i]$ or $Y_i = \lambda = Y'_i$, and $Z = X[X_1(Z'_1) \oplus \cdots \oplus X_n(Z'_n)]$, such that Y'_i, Z'_i are reconsilable for all $i = 1, \dots, n$. We get $Y \sqcup Z = X[X_1(Y'_1 \sqcup Z'_1) \oplus \cdots \oplus X_n(Y'_n \sqcup Z'_n)]$. As $Z \geq X[\lambda]$, we also have $\pi^X_{X[\lambda]}(t_1) = \pi^X_{X[\lambda]}(t_2)$, so t_1 and t_2 are lists of equal length. Therefore, assume $t_j = [t_{j1}, \dots, t_{jm}]$ for $j = 1, 2$ and $t_{jk} = (X_\ell : t''_{jk})$. This gives $\pi^X_{Y \sqcup Z}(t_j) = [t'_{j1}, \dots, t'_{jm}]$ with $t'_{jk} = (X_\ell : \pi^{X'_\ell}_{Y'_\ell \sqcup Z'_\ell}(t''_{jk}))$. We know $\pi^{X'_\ell}_{Z'_\ell}(t''_{1k}) = \pi^{X'_\ell}_{Z'_\ell}(t''_{2k})$, so we are done for $Y_\ell = \lambda$. For $Y_\ell \neq \lambda$ the sublists containing all $(X_\ell : t''_{jk})$ coincide on Y'_ℓ. As Y'_ℓ and Z'_ℓ are semi-disjoint, we have $\pi^{X'_\ell}_{Y'_\ell \sqcup Z'_\ell}(t''_{1k}) = \pi^{X'_\ell}_{Y'_\ell \sqcup Z'_\ell}(t''_{2k})$ by induction, which implies $\pi^X_{Y \sqcup Z}(t_1) = \pi^X_{Y \sqcup Z}(t_2)$.

For the extension rule (5) let $t_1, t_2 \in r$ for some instance $r \subseteq dom(X)$ with $r \models \mathcal{Y} \to \mathcal{Z}$, and assume $\pi_Y^X(t_1) = \pi_Y^X(t_2)$ holds for all $Y \in \mathcal{Y}$. Then we must have as well $\pi_Z^X(t_1) = \pi_Z^X(t_2)$ for all $Z \in \mathcal{Z}$, which implies $\pi_Y^X(t_1) = \pi_Y^X(t_2)$ for all $Y \in \mathcal{Y} \cup \mathcal{Z}$, i.e. $r \models \mathcal{Y} \to \mathcal{Y} \cup \mathcal{Z}$.

For the transitivity rule (6) let $t_1, t_2 \in r$ for some instance $r \subseteq dom(X)$ with $r \models \mathcal{Y} \to \mathcal{Z}$ and $r \models \mathcal{Z} \to \mathcal{U}$, and assume $\pi_Y^X(t_1) = \pi_Y^X(t_2)$ holds for all $Y \in \mathcal{Y}$. Then we must have as well $\pi_Z^X(t_1) = \pi_Z^X(t_2)$ for all $Z \in \mathcal{Z}$ by the first premise, and hence $\pi_U^X(t_1) = \pi_U^X(t_2)$ for all $U \in \mathcal{U}$ by the second premise, which shows $r \models \mathcal{Y} \to \mathcal{U}$ as desired.

□

In [20] it was shown that the six of axioms and rules in Theorem 3, i.e. (1) – (6) are complete for the implication of FDs, if the union constructor is not present. In this case (2), (3) and (4) are axioms that deal with the Brouwer algebra structure on $\mathcal{S}(X)$, while (1), (5) and (6) are the well known Armstrong axioms and rules.

Theorem 4. *The following rules for the implication of FDs on* $\mathcal{S}(X)$ *can be derived from the rules in Theorem 3:*

union rule:

$$\frac{\mathcal{Y} \to \mathcal{Z} \quad \mathcal{Y} \to \mathcal{U}}{\mathcal{Y} \to \mathcal{Z} \cup \mathcal{U}} \tag{7}$$

fragmentation rule:

$$\frac{\mathcal{Y} \to \mathcal{Z}}{\mathcal{Y} \to \{Z\}}\ Z \in \mathcal{Z} \tag{8}$$

join rule:

$$\frac{\{Y\} \to \{Z\}}{\{Y\} \to \{Y \sqcup Z\}}\ Y, Z \textit{ reconsilable} \tag{9}$$

Proof. For the union rule (7) we use the following derivation:

$$\frac{\dfrac{\mathcal{Y} \to \mathcal{Z}}{\mathcal{Y} \to \mathcal{Y} \cup \mathcal{Z}} \quad \dfrac{\dfrac{\dfrac{\overline{\mathcal{Y} \cup \mathcal{Z} \to \mathcal{Y}} \quad \mathcal{Y} \to \mathcal{U}}{\mathcal{Y} \cup \mathcal{Z} \to \mathcal{U}}}{\mathcal{Y} \cup \mathcal{Z} \to \mathcal{Y} \cup \mathcal{Z} \cup \mathcal{U}} \quad \overline{\mathcal{Y} \cup \mathcal{Z} \cup \mathcal{U} \to \mathcal{Z} \cup \mathcal{U}}}{\mathcal{Y} \cup \mathcal{Z} \to \mathcal{Z} \cup \mathcal{U}}}{\mathcal{Y} \to \mathcal{Z} \cup \mathcal{U}}$$

For the fragmentation rule (8) we use the following derivation:

$$\frac{\mathcal{Y} \to \mathcal{Z} \quad \overline{\mathcal{Z} \to \{Z\}}}{\mathcal{Y} \to \{Z\}}$$

Finally, for the join-rule (9) we use the following derivation:

$$\frac{\dfrac{\{Y\} \to \{Z\}}{\{Y\} \to \{Y, Z\}} \quad \overline{\{Y, Z\} \to \{Y \sqcup Z\}}}{\{Y\} \to \{Y \sqcup Z\}}$$

□

If the union constructor is present, we obtain further subattributes, for which we obtain additional axioms. These will be set, multiset and list axioms (10) – (18) in the following Theorem 5. Furthermore, we obtain rules that derive FDs on $\mathcal{S}(X)$ from FDs on $\mathcal{S}(X')$ for *embedded attributes* X', i.e. X' results from X by stripping away the outermost constructor. The following definition clarifies in an exact way, how embedded attributes and induced instances for embedded attributes have to be understood. This will become important also for the extensions in Section 5.

Definition 11. Let $X \in \mathcal{N}$ be a nested attribute. The *set of embedded attributes* of X is the smallest set $\text{emb}(X)$ with $X \in \text{emb}(X)$ satisfying the following properties:

1. If $X = X(X_1, \dots, X_n)$ is a record attribute, then $\text{emb}(X_i) \subseteq \text{emb}(X)$ holds for all $i = 1, \dots, n$.
2. If $X = X_1(X_1') \oplus \cdots \oplus X_n(X_n')$ is a union attribute, then $\text{emb}(X_i') \subseteq \text{emb}(X)$ holds for all $i = 1, \dots, n$.
3. If $X = X\{X'\}$ is a set attribute, then $\text{emb}(X') \subseteq \text{emb}(X)$ holds.
4. If $X = X[X']$ is a list attribute, then $\text{emb}(X') \subseteq \text{emb}(X)$ holds.
5. If $X = X\langle X' \rangle$ is a multiset attribute, then $\text{emb}(X') \subseteq \text{emb}(X)$ holds.

If $r \subseteq dom(X)$ is an instance of X, then for each $Y \in \text{emb}(X)$ we obtain the *induced instance* $r \downarrow Y$ in the following way:

1. $r \downarrow X = r$;
2. $r \downarrow Z = (r \downarrow Y) \downarrow Z$ for $Z \in \text{emb}(Y)$ and $Y \in \text{emb}(X)$;
3. $r \downarrow X_i = \{t_i \in dom(X_i) \mid \exists t \in r . t = (t_1, \dots, t_i, \dots, t_n)\}$ for a record attribute $X = X(X_1, \dots, X_n)$;
4. $r \downarrow X_i = \{t_i \in dom(X_i) \mid \exists t \in r . t = (X_i : t_i)\}$ for a union attribute $X = X_1(X_1') \oplus \cdots \oplus X_n(X_n')$;
5. $r \downarrow X' = \{t' \in dom(X') \mid \exists t \in r . t' \in t\}$ for a set attribute $X = X\{X'\}$;
6. $r \downarrow X' = \{t' \in dom(X') \mid \exists t \in r . t' \in t\}$ for a multiset attribute $X = X\langle X' \rangle$;
7. $r \downarrow X' = \{t' \in dom(X') \mid \exists t \in r . t = [\dots, t', \dots]\}$ for a list attribute $X = X[X']$.

In dealing now with FDs $\mathcal{Y} \to \mathcal{Z}$ defined embedded attributes $U \in \text{emb}(X)$ we let $r \models \mathcal{Y} \to \mathcal{Z}$ mean $r \downarrow U \models \mathcal{Y} \to \mathcal{Z}$. This generalises canonically to wFDs.

Theorem 5. *In addition to the axioms and rules in Theorem 3 the following axioms and rules are sound for the implication of FDs on $\mathcal{S}(X)$:*

set axiom:

$$\frac{}{\{X_I\{\lambda\}, X_J\{\lambda\}\} \to \{X_{I\cup J}\{\lambda\}\}}\ I \cap J = \emptyset \tag{10}$$

multiset axioms:

$$\frac{}{\{X_I\langle\lambda\rangle, X_J\langle\lambda\rangle\} \to \{X_{I\cup J}\langle\lambda\rangle\}}\ I \cap J = \emptyset \tag{11}$$

$$\frac{}{\{X_I\langle\lambda\rangle, X_{I\cup J}\langle\lambda\rangle\} \to \{X_J\langle\lambda\rangle\}}\ I \cap J = \emptyset \tag{12}$$

$$\frac{}{\{X_I\langle\lambda\rangle, X_J\langle\lambda\rangle, X_{I\cap J}\langle\lambda\rangle\} \to \{X_{(I-J)\cup(J-I)}\langle\lambda\rangle\}} \tag{13}$$

$$\frac{}{\{X_I\langle\lambda\rangle, X_J\langle\lambda\rangle, X_{(I-J)\cup(J-I)}\langle\lambda\rangle\} \to \{X_{I\cap J}\langle\lambda\rangle\}} \tag{14}$$

list axioms:

$$\frac{}{\{X_I[\lambda], X_J[\lambda]\} \to \{X_{I\cup J}[\lambda]\}}\ I \cap J = \emptyset \tag{15}$$

$$\frac{}{\{X_I[\lambda], X_{I\cup J}[\lambda]\} \to \{X_J[\lambda]\}}\ I \cap J = \emptyset \tag{16}$$

$$\frac{}{\{X_I[\lambda], X_J[\lambda], X_{I\cap J}[\lambda]\} \to \{X_{(I-J)\cup(J-I)}[\lambda]\}} \tag{17}$$

$$\frac{}{\{X_I[\lambda], X_J[\lambda], X_{(I-J)\cup(J-I)}[\lambda]\} \to \{X_{I\cap J}[\lambda]\}} \tag{18}$$

set lifting rule:

$$\frac{\{Y\} \to \mathcal{Z}}{\{X\{Y\}\} \to \{X\{Z\} \mid Z \in \mathcal{Z}\}}\ X = X\{X'\},\ Y \in \mathcal{S}(X'), \mathcal{Z} \subseteq \mathcal{S}(X') \tag{19}$$

record lifting rule:

$$\frac{\mathcal{Y}_i \to \mathcal{Z}_i}{\{X(\lambda, \dots, Y_i, \dots, \lambda) \mid Y_i \in \mathcal{Y}_i\} \to \{X(\lambda, \dots, Z_i, \dots, \lambda) \mid Z_i \in \mathcal{Z}_i\}}\ \mathcal{C} \tag{20}$$

with conditions $\mathcal{C} : X = X(X_1, \dots, X_n)$ *and* $\mathcal{Y}_i, \mathcal{Z}_i \subseteq \mathcal{S}(X_i)$

union lifting rule:

$$\frac{\mathcal{Y}_i \to \mathcal{Z}_i}{\{\dots \oplus X_i(Y_i) \oplus \dots \mid Y_i \in \mathcal{Y}_i\} \to \{\dots \oplus X_i(Z_i) \oplus \dots \mid Z_i \in \mathcal{Z}_i\}}\ \mathcal{C} \tag{21}$$

with conditions $\mathcal{C} : X = X(X_1, \dots, X_n)$ *and* $\mathcal{Y}_i, \mathcal{Z}_i \subseteq \mathcal{S}(X_i'), \mathcal{Y}_i \neq \emptyset$

multiset lifting rule:

$$\frac{\{Y\} \to \mathcal{Z}}{\{X\langle Y\rangle\} \to \{X\langle Z\rangle \mid Z \in \mathcal{Z}\}}\; X = X\langle X'\rangle,\ Y \in \mathcal{S}(X'), \mathcal{Z} \subseteq \mathcal{S}(X') \tag{22}$$

list lifting rule:

$$\frac{\mathcal{Y} \to \mathcal{Z}}{\{X[Y] \mid Y \in \mathcal{Y}\} \to \{X[Z] \mid Z \in \mathcal{Z}\}}\; X = X[X'],\ \mathcal{Y}, \mathcal{Z} \subseteq \mathcal{S}(X'), \mathcal{Y} \neq \emptyset \tag{23}$$

Proof. For the set axiom (10) let $t_1, t_2 \in dom(X)$ with $\pi^X_{X_I\{\lambda\}}(t_1) = \pi^X_{X_I\{\lambda\}}(t_2)$ and $\pi^X_{X_J\{\lambda\}}(t_1) = \pi^X_{X_J\{\lambda\}}(t_2)$. In case $\pi^X_{X_I\{\lambda\}}(t_1) = \pi^X_{X_J\{\lambda\}}(t_1) = \emptyset$ there are no values of the form $(X_i : v_i)$ with $i \in I \cup J$ in t_1, hence also not in t_2. In case at least one of these projections leads to a non-empty set we must have $(X_i : v_i) \in t_1$ for at least one $i \in I \cup J$ and one value $v_i \in dom(X'_i)$. The same holds for t_2, hence in both cases $\pi^X_{X_{I\cup J}\{\lambda\}}(t_1) = \pi^X_{X_{I\cup J}\{\lambda\}}(t_2)$.

For the first list axiom (15) let $t_1, t_2 \in dom(X)$. Then $\pi^X_{X_I[\lambda]}(t_1) = \pi^X_{X_I[\lambda]}(t_2)$ means that t_1 and t_2 contain the same number of elements of the form $(X_i : v_i)$ with $i \in I$. If the same holds for $I \cup J$, then t_1 and t_2 must also contain the same number of elements of the form $(X_i : v_i)$ with $i \in J$, i.e. $\pi^X_{X_J[\lambda]}(t_1) = \pi^X_{X_J[\lambda]}(t_2)$. The soundness of the second list axiom (16) follows from the same argument.

Analogously, for the third list axiom (17) for $Y \in \{X_I[\lambda], X_J[\lambda], X_{I\cap J}[\lambda]\}$ $\pi^X_Y(t_1) = \pi^X_Y(t_2)$ means that t_1, t_2 contain the same number of elements with labels in I, J and $I \cap J$, respectively. So they also contain the same number of elements with labels in $(I - J) \cup (J - I)$. The soundness of the fourth list axiom (18) follows from the same argument.

The proof for the four multiset axioms (11) – (14) is completely analogous to the proof for the list axioms.

For the set lifting rule (19) let $t_1, t_2 \in dom(X)$ with $\pi^X_{X\{Y\}}(t_1) = \pi^X_{X\{Y\}}(t_2)$. Without loss of generality – repeat elements, if necessary – we may write $t_i = \{t_{i1}, \dots, t_{ik}\}$ $(i = 1, 2)$. Then for all $j = 1, \dots, k$ we have $\pi^{X'}_Y(t_{1j}) = \pi^{X'}_Y(t_{2j})$. From the premise of the rule we get $\pi^{X'}_Z(t_{1j}) = \pi^{X'}_Z(t_{2j})$ for all $j = 1, \dots, k$ and all $Z \in \mathcal{Z}$, which implies $\pi^X_{X\{Z\}}(t_1) = \pi^X_{X\{Z\}}(t_2)$ for all $X\{Z\}$ with $Z \in \mathcal{Z}$.

For the record lifting rule (20) let $t_1, t_2 \in dom(X)$ with $\pi^X_{X(\lambda,\dots,Y_i,\dots,\lambda)}(t_1) = \pi^X_{X(\lambda,\dots,Y_i,\dots,\lambda)}(t_2)$ for all $Y_i \in \mathcal{Y}_i$. If $t_j = (t_{1j}, \dots, t_{nj})$ for $j = 1, 2$, then it follows $\pi^{X_i}_{Y_i}(t_{i1}) = \pi^{X_i}_{Y_i}(t_{i2})$ for all $Y_i \in \mathcal{Y}_i$ and thus also $\pi^{X_i}_{Z_i}(t_{i1}) = \pi^{X_i}_{Z_i}(t_{i2})$ for all $Z_i \in \mathcal{Z}_i$ by the premise of the rule. This gives $\pi^X_{X(\lambda,\dots,Z_i,\dots,\lambda)}(t_1) = \pi^X_{X(\lambda,\dots,Z_i,\dots,\lambda)}(t_2)$ for all $Z_i \in \mathcal{Z}_i$ as desired.

For the union lifting rule (21) let $t_1, t_2 \in dom(X)$ with $\pi^X_{\dots\oplus X_i(Y_i)\oplus\dots}(t_1) = \pi^X_{\dots\oplus X_i(Y_i)\oplus\dots}(t_2)$ for all $Y_i \in \mathcal{Y}_i$. In particular, t_1 and t_2 must have the same label, and we can assume $t_j = (X_i : t'_j)$ for $j = 1, 2$. Then we get $\pi^{X_i}_{Y_i}(t'_1) = \pi^{X_i}_{Y_i}(t'_2)$ for all $Y_i \in \mathcal{Y}_i$ and thus also $\pi^{X_i}_{Z_i}(t'_1) = \pi^{X_i}_{Z_i}(t'_2)$ for all $Z_i \in \mathcal{Z}_i$ by the premise of the rule. This implies $\pi^X_{\dots\oplus X_i(Z_i)\oplus\dots}(t_1) = \pi^X_{\dots\oplus X_i(Z_i)\oplus\dots}(t_2)$ for all $Z_i \in \mathcal{Z}_i$ as desired.

For the multiset lifting rule (22) let $t_1, t_2 \in dom(X)$ with $\pi^X_{X\langle Y\rangle}(t_1) = \pi^X_{X\langle Y\rangle}(t_2)$. In particular, t_1 and t_2 must contain the same number of elements, so we may write $t_i = \langle t_{i1}, \dots, t_{ik}\rangle$ $(i = 1, 2)$. Then for all $j = 1, \dots, k$ we obtain $\pi^{X'}_Y(t_{1j}) = \pi^{X'}_Y(t_{2j})$. From the premise of the rule we get $\pi^{X'}_Z(t_{1j}) = \pi^{X'}_Z(t_{2j})$ for all $j = 1, \dots, k$ and all $Z \in \mathcal{Z}$, which implies $\pi^X_{X\langle Z\rangle}(t_1) = \pi^X_{X\langle Z\rangle}(t_2)$ for all $X\langle Z\rangle$ with $Z \in \mathcal{Z}$.

For the list lifting rule (23) let $t_1, t_2 \in dom(X)$ with $\pi^X_{X[Y]}(t_1) = \pi^X_{X[Y]}(t_2)$ for all $X[Y]$ with $Y \in \mathcal{Y}$. As $\mathcal{Y} \neq \emptyset$, it follows that t_1 and t_2 must have the same length, say $t_i = [t_{i1}, \dots, t_{ik}]$ $(i = 1, 2)$, and for all $j = 1, \dots, k$ and all $Y \in \mathcal{Y}$ we have $\pi^{X'}_Y(t_{1j}) = \pi^{X'}_Y(t_{2j})$. Hence $\pi^{X'}_Z(t_{1j}) = \pi^{X'}_Z(t_{2j})$ for all $j = 1, \dots, k$ and all $Z \in \mathcal{Z}$, which implies $\pi^X_{X[Z]}(t_1) = \pi^X_{X[Z]}(t_2)$ for all $X[Z]$ with $Z \in \mathcal{Z}$.

□

According to the observation made before we may still say that all axioms and rules in Theorem 5 arise from the lattice structure on $\mathcal{S}(X)$.

The axioms and rules in Theorem 3 only apply to "ordinary" FDs. For the implication of wFDs we need additional axioms and rules.

Theorem 6. *The following axioms and rules are sound for the implication of wFDs on $\mathcal{S}(X)$:*

weakening rule:

$$\frac{\{\!|\mathcal{Y}_i \to \mathcal{Z}_i \mid i \in I|\!\}}{\{\!|\mathcal{Y}_j \to \mathcal{Z}_j \mid j \in J|\!\}}\; I \subseteq J \tag{24}$$

left union rule:

$$\frac{\{\!|\mathcal{Y} \to \mathcal{Z}_i \mid i \in I|\!\}}{\{\!|\mathcal{Y}_i \to \mathcal{Z}_i \mid i \in I|\!\}}\; \mathcal{Y} = \bigcup_{i\in I} \mathcal{Y}_i \tag{25}$$

shift rule:

$$\frac{\{\!|\mathcal{Y} \cup \mathcal{U}_1 \to \{Z\} \mid Z \in \mathcal{Z} \cup (\mathcal{U} - \mathcal{U}_1)|\!\} \dots \{\!|\mathcal{Y} \cup \mathcal{U}_k \to \{Z\} \mid Z \in \mathcal{Z} \cup (\mathcal{U} - \mathcal{U}_k)|\!\}}{\{\!|\mathcal{Y} \to \{Z\} \mid Z \in \mathcal{Z}|\!\}}\; \mathcal{C} \tag{26}$$

with condition $\mathcal{C}: \ \mathcal{P}(\mathcal{U}) = \{\mathcal{U}_1, \dots, \mathcal{U}_k\}$

union axiom *for* $X = X\{X_1(X'_1) \oplus \cdots \oplus X_n(X'_n)\}$ *and* $I = \{i_1, \dots, i_k\}$*:*

$$\frac{}{\{\!|\{X_I\{\lambda\}\} \to \{X_J\{\lambda\}\}, \{X_I\{\lambda\}\} \to \{X(X_{i_1}\{X'_{i_1}\}, \dots, X_{i_k}\{X'_{i_k}\})\}|\!\}}\; I \subsetneq J \tag{27}$$

partition axiom *for* $X = X\{X_1(X'_1) \oplus \cdots \oplus X_n(X'_n)\}$ *and* $I \subseteq \{1, \dots, n\}$*:*

$$\frac{}{\begin{array}{l}\{\!|\{X_I\{\lambda\}\} \to \{X_{I'_1 \cup I'_2}\{\lambda\} \mid \emptyset \neq I'_1 \subseteq I_1, \emptyset \neq I'_2 \subseteq I_2\},\\ \quad \{X_I\{\lambda\}\} \to \{X(X_i\{\lambda\})\} \mid I = I_1 \cup I_2, I_1 \cap I_2 = \emptyset, I_1 \neq \emptyset \neq I_2, i \in I|\!\}\end{array}} \tag{28}$$

first plus/minus axiom *for* $X = X\{X_1(X_1') \oplus \cdots \oplus X_n(X_n')\}$:

$$\frac{}{\{\!\!\{\{\lambda\} \to \{X_{J\cup\{i\}}\{\lambda\} \mid J \subseteq I^-\}, \{X_{\{1,\dots,n\}}\{\lambda\}\} \to \{X_{I^-}\{\lambda\}\}, \\ \{X(X_j\{\lambda\})\} \to \{X\}, \{\lambda\} \to \{X(X_k\{\lambda\})\} \mid i,j \in I^+, k \in I^-\}\!\!\}}\ \mathcal{C} \tag{29}$$

with condition $\mathcal{C} : \{1,\dots,n\} = I^+ \,\dot{\cup}\, I^-$

second plus/minus axiom *for* $X = X\{X_1(X_1') \oplus \cdots \oplus X_n(X_n')\}$:

$$\frac{}{\{\!\!\{\{X_{I\cup J}\{\lambda\}\} \to \{X_J\{\lambda\}\},\ \{X_{I\cup J}\{\lambda\}\} \to \{X_{\{j\}}\{\lambda\}\}, \\ \{X_{I\cup J}\{\lambda\}\} \to \{X_{J'\cup\{i_0\}}\{\lambda\} \mid J' \subseteq J\} \mid i_0 \in I, j \in J\}\!\!\}} \tag{30}$$

with condition $I \cap J = \emptyset$

third plus/minus axiom *for* $X = X\{X_1(X_1') \oplus \cdots \oplus X_n(X_n')\}$:

$$\frac{}{\{\!\!\{\{X_{I^-}\{\lambda\}, X_{I'\cup J'}\{\lambda\}, X_{\{i\}}\{\lambda\} \mid i \in I^+\} \to \{X_{J'}\{\lambda\}\}, \\ \{X_{I^-}\{\lambda\}, X_{I'\cup J'}\{\lambda\}, X_{\{i\}}\{\lambda\} \mid i \in I^+\} \to \{X_{J\cup\{\ell\}}\{\lambda\} \mid J \subseteq I^-\}, \\ \{X_{I^-}\{\lambda\}, X_{I'\cup J'}\{\lambda\}, X_{\{i\}}\{\lambda\} \mid i \in I^+\} \to \{X_{\{k\}}\{\lambda\}\} \mid k \in I^-, \ell \in I'\}\!\!\}} \tag{31}$$

with conditions $I^+ \cup I^- = \{1,\dots,n\}$, $I^+ \cap I^- = \emptyset$, $I' \subseteq I^+$, $J' \subseteq I^-$

partition axiom for sets *for* $X = X\{\bar{X}\{X_1(X_1') \oplus \cdots \oplus X_n(X_n')\}\}$ *and* $P \subseteq \mathcal{P}(I)$:

$$\frac{}{\{\!\!\{\{\lambda\} \to \{X\{\bar{X}_{I'}\{\lambda\}\} \mid I' \cap I^+ \neq \emptyset\} \cup \\ \{X\{\bar{X}_{J_\cup J_-}\{\lambda\}\}, X\{\bar{X}_J\{\lambda\}\} \mid J_- \subseteq I_-, J \in Q, J \subseteq I_{+-} \cup I^-\}, \\ \{\lambda\} \to \{X\{\bar{X}_K\{\lambda\}\}\}, \{X\{\bar{X}_{K'}\{\lambda\}\}\} \to \{X\} \mid \\ I = I^- \,\dot{\cup}\, I_+ \,\dot{\cup}\, I_- \,\dot{\cup}\, I_{+-}, Q \subseteq \mathcal{P}(I), \\ K \in (\mathcal{P}(I_{+-} \cup I^-)) - Q, K' \in (\mathcal{P}(I_{+-} \cup I^-)) \cap Q\}\!\!\}} \tag{32}$$

partition axiom for multisets *for* $X = X\langle\bar{X}\{X_1(X_1') \oplus \cdots \oplus X_n(X_n')\}\rangle$ *and* $P \subseteq \mathcal{P}(I)$:

$$\frac{}{\{\!\!\{\{\lambda\} \to \{X\langle\bar{X}_{I'}\{\lambda\}\rangle \mid I' \cap I^+ \neq \emptyset\} \cup \\ \{X\langle\bar{X}_{J\cup J_-}\{\lambda\}\rangle, X\langle\bar{X}_J\{\lambda\}\rangle \mid J_- \subseteq I_-, J \in Q, J \subseteq I_{+-} \cup I^-\}, \\ \{\lambda\} \to \{X\langle\bar{X}_K\{\lambda\}\rangle\}, \{X\langle\bar{X}_{K'}\{\lambda\}\rangle\} \to \{X\} \mid \\ I = I^- \,\dot{\cup}\, I_+ \,\dot{\cup}\, I_- \,\dot{\cup}\, I_{+-}, Q \subseteq \mathcal{P}(I), \\ K \in (\mathcal{P}(I_{+-} \cup I^-)) - Q, K' \in (\mathcal{P}(I_{+-} \cup I^-)) \cap Q\}\!\!\}} \tag{33}$$

Proof. The soundness proof for the weakening rule (24) is trivial.

For the left union rule (25) assume $r \not\models \{\!\!\{\mathcal{Y}_i \to \mathcal{Z}_i \mid i \in I\}\!\!\}$, i.e. there exist $t_1, t_2 \in r$ such that for all $i \in I$ we get $\pi_Y^X(t_1) = \pi_Y^X(t_2)$ for all $Y \in \mathcal{Y}_i$ and $\pi_{Z_i}^X(t_1) \neq \pi_{Z_i}^X(t_2)$ for some $Z_i \in \mathcal{Z}_i$. In particular, $\pi_Y^X(t_1) = \pi_Y^X(t_2)$ for all $Y \in \mathcal{Y}$, hence $r \not\models \{\!\!\{\mathcal{Y} \to \mathcal{Z}_i \mid i \in I\}\!\!\}$.

For the shift rule (26) assume $r \not\models \{\!\!\{\mathcal{Y} \to \{Z\} \mid Z \in \mathcal{Z}\}\!\!\}$, i.e. there exist $t_1, t_2 \in r$ such that $\pi_Y^X(t_1) = \pi_Y^X(t_2)$ for all $Y \in \mathcal{Y}$ and $\pi_Z^X(t_1) \neq \pi_Z^X(t_2)$ for all $Z \in \mathcal{Z}$. Take a maximal $\mathcal{U}' \subseteq \mathcal{U}$ such that $\pi_U^X(t_1) = \pi_U^X(t_2)$ for all $U \in \mathcal{U}'$. If we had $r \models \{\!\!\{\mathcal{Y} \cup \mathcal{U}' \to \{Z\} \mid Z \in \mathcal{Z} \cup (\mathcal{U} - \mathcal{U}')\}\!\!\}$, we would have $\mathcal{U}' \subsetneq \mathcal{U}$, and there would exist some $V \in \mathcal{U} - \mathcal{U}'$ with $\pi_V^X(t_1) = \pi_V^X(t_2)$, which contradicts the maximality of $\mathcal{U}$.

Let $X = X\{X_1(X_1') \oplus \cdots \oplus X_n(X_n')\} = X(X_1\{X_1'\}, \ldots, X_n\{X_n'\})$, $Y = X_I\{\lambda\}$, $Z_1 = X_J\{\lambda\}$ and $Z_2 = X(X(X_{i_1}\{X_{i_1}'\}, \ldots, X_{i_k}\{X_{i_k}'\})$ for the union axiom (27). Let $t_1, t_2 \in r$ with $\pi_Y^X(t_1) = \pi_Y^X(t_2)$ and $\pi_{Z_1}^X(t_1) \neq \pi_{Z_1}^X(t_2)$. Thus, one of t_1 or t_2 — without loss of generality let this be t_2 — must not contain elements of the form $(X_j : v_j)$ with $j \in J$. On the other hand, either t_1 and t_2 both contain elements of the form $(X_i : v_i)$ with $i \in I$ or both do not. As $I \subsetneq J$, it follows $\pi_{X(X_i\{\lambda\})}^X(t_1) = \pi_{X(X_i\{\lambda\})}^X(t_2) = \emptyset$ for all $i \in I$, which implies $\pi_{Z_2}^X(t_1) = \pi_{Z_2}^X(t_2)$.

For the partition axiom (28) let $t_1, t_2 \in r$ with $\pi_{X_I\{\lambda\}}^X(t_1) = \pi_{X_I\{\lambda\}}^X(t_2)$ and $\pi_{X(X_i\{\lambda\})}^X(t_1) \neq \pi_{X(X_i\{\lambda\})}^X(t_2)$ for all $i \in I$. Let $I_j \subseteq I$ be such that t_j contains an element of the form $(X_i : v_i)$ for all $i \in I_j$ $(j = 1, 2)$. Obviously, $I = I_1 \dot\cup I_2$ and $\pi_{X_{I'}\{\lambda\}}^X(t_1) = \pi_{X_{I'}\{\lambda\}}^X(t_2)$ for all $I' \subseteq I$ with $I' \cap I_1 \neq \emptyset \neq I' \cap I_2$.

For the first plus/minus axiom in (29) let t_1, t_2 satisfy $\pi_{X_j\{\lambda\}}^X(t_1) = \pi_{X_j\{\lambda\}}^X(t_2)$ and $\pi_{X_k\{\lambda\}}^X(t_1) \neq \pi_{X_k\{\lambda\}}^X(t_2)$ for all $j \in I^+$ and $k \in I^-$. Assume that for all $i \in I^+$ there is some $J \subseteq I^-$ with $\pi_{X_{J \cup \{i\}}\{\lambda\}}^X(t_1) \neq \pi_{X_{J \cup \{i\}}\{\lambda\}}^X(t_2)$, i.e. one of these projections must be $\emptyset$. As we have $\pi_{X_i\{\lambda\}}^X(t_1) = \pi_{X_i\{\lambda\}}^X(t_2)$, these must both be $\emptyset$, which implies $\pi_{X_{I^+}\{\lambda\}}^X(t_j) = \emptyset$ for $j = 1, 2$. Now $\pi_{X_k\{\lambda\}}^X(t_1) \neq \pi_{X_k\{\lambda\}}^X(t_2)$ for all $k \in I^-$, so if $\pi_{X_{I^-}\{\lambda\}}^X(t_1) \neq \pi_{X_{I^-}\{\lambda\}}^X(t_2)$ holds, one of these projections must be $\emptyset$ again, which implies that one t_j is $\emptyset$, the other not empty. That is $\pi_{X_{\{1,\ldots,n\}}\{\lambda\}}^X(t_1) \neq \pi_{X_{\{1,\ldots,n\}}\{\lambda\}}^X(t_2)$.

For the second plus/minus axiom in (30) assume that it does not hold. Then we find two complex values t_1, t_2 that coincide on $X_{I \cup J}\{\lambda\}$, but differ on $X_J\{\lambda\}$ and all $X_{\{j\}}\{\lambda\}$ with $j \in J$. Furthermore, for each $i \in I$ there is at least one $J_i \subseteq J$ such that t_1, t_2 differ on $X_{J_i \cup \{i\}}\{\lambda\}$. It follows that one of the two complex values – without loss of generality let this be t_1 – contains values $(X_j : \tau_j)$ for all $j \in J$, while the other one does not contain such values. Then we obtain $\pi_{X_{J' \cup \{i\}}\{\lambda\}}^X(t_1) \neq \emptyset$ for all $J' \subseteq J$ and all $i \in I$. As t_1, t_2 coincide on $X_{I \cup J}\{\lambda\}$, this also gives $\pi_{X_{J' \cup \{i\}}\{\lambda\}}^X(t_2) \neq \emptyset$ for all $J' \subseteq J$ and at least one $i \in I$ contradicting the assumption that for at least one such $J' = J_i$ we have $\pi_{X_{J' \cup \{i\}}\{\lambda\}}^X(t_1) \neq \pi_{X_{J' \cup \{i\}}\{\lambda\}}^X(t_2)$.

For the third plus/minus axiom in (31) assume that it does not hold. Then we find two complex values t_1, t_2 that coincide on $X_{I^-}\{\lambda\}$, all $X_{\{i\}}\{\lambda\}$ for $i \in I^+$, and on $X_{I' \cup J'}\{\lambda\}$, but differ on $X_{J'}\{\lambda\}$ and all $X_{\{k\}}\{\lambda\}$ with $k \in I^-$. Furthermore, for each $\ell \in I'$ there is at least one $J_\ell \subseteq I^-$ such that t_1, t_2 differ on $X_{J_\ell \cup \{\ell\}}\{\lambda\}$. Define $I_j^- = \{i \in I^- \mid \pi^X_{X_{\{i\}}\{\lambda\}}(t_j) \neq \emptyset\}$ $(j = 1, 2)$ to define a partition $I^- = I_1^- \cup I_2^-$. As t_1, t_2 differ on $X_{J'}\{\lambda\}$, this implies $J' \subseteq I_1'$ or $J' \subseteq I_2'$. Without loss of generality we can assume the first of these possibilities. As t_1, t_2 coincide on $X_{I' \cup J'}\{\lambda\}$, we must have $\pi^X_{X_{I'}\{\lambda\}}(t_2) \neq \emptyset$, so also $\pi^X_{X_{\{i\}}\{\lambda\}}(t_2) \neq \emptyset$ for some $i \in I'$. Then also $\pi^X_{X_{\{i\}}\{\lambda\}}(t_1) \neq \emptyset$ due to $I' \subseteq I^+$. Hence we get $\pi^X_{X_{J \cup \{i\}}\{\lambda\}}(t_j) \neq \emptyset$ for $j = 1, 2$ and all $J \subseteq I^-$ contradicting the assumption that at least one such $J = J_i$ exists, such that t_1, t_2 differ on $X_{J_i \cup \{i\}}\{\lambda\}$.

For the set partition axiom in (32) take any $S_1, S_2 \in dom(X)$. In case $S_1 = S_2 = \emptyset$ we simple choose $I_+ = I$, so we must have $I^- = I_- = I_{+-} = \emptyset$. Further take $Q' = \{\emptyset\}$. In case exactly one of the S_i is empty, we choose $I^- = I$, $I_+ = I_- = I_{+-} = \emptyset$, and

$$Q' = \{J \subseteq I^- \mid \pi^X_{X\{\bar{X}_J\{\lambda\}\}}(S_1) = \pi^X_{X\{\bar{X}_J\{\lambda\}\}}(S_2)\}.$$

In both cases we immediately get the satisfaction of the first involved FD, if $Q \cap (\mathcal{P}(I_{+-} \cup I^-)) = Q'$. However, if there is some $K \in Q'$ with $K \notin Q$, the FD $\{\lambda\} \to \{X\{\bar{X}_K\{\lambda\}\}\}$ is satisfied. Similarly, if there is some $K' \in Q$ with $K' \notin Q'$, then the FD $\{X\{\bar{X}_{K'}\{\lambda\}\}\} \to \{X\}$ is satisfied.

In the remaining case with $S_1 \neq \emptyset \neq S_2$ we take

$$\begin{aligned}
I_+ &= \{i \in I \mid \pi^X_{X\{\bar{X}_{\{i\}}\{\lambda\}\}}(S_1) = \{\{\top\}\} = \pi^X_{\bar{X}\{X_{\{i\}}\{\lambda\}\}}(S_2)\},\\
I_- &= \{i \in I \mid \pi^X_{X\{\bar{X}_{\{i\}}\{\lambda\}\}}(S_1) = \{\emptyset\} = \pi^X_{\bar{X}\{X_{\{i\}}\{\lambda\}\}}(S_2)\},\\
I^- &= \{i \in I \mid \pi^X_{X\{\bar{X}_{\{i\}}\{\lambda\}\}}(S_1) \neq \pi^X_{\bar{X}\{X_{\{i\}}\{\lambda\}\}}(S_2)\},
\end{aligned}$$

and $I_{+-} = I - I^- - I_+ - I_-$. Then S_1, S_2 obviously coincide on all $X\{\bar{X}_{I'}\{\lambda\}\}$ with $I' \cap I^+ \neq \emptyset$. If we take again

$$Q' = \{J \subseteq I^- \cup I_{+-} \mid \pi^X_{X\{\bar{X}_J\{\lambda\}\}}(S_1) = \pi^X_{X\{\bar{X}_J\{\lambda\}\}}(S_2)\},$$

then S_1, S_2 coincide on all $X\{\bar{X}_{J \cup J_-}\{\lambda\}\}$ and all $X\{\bar{X}_J\{\lambda\}\}$ with $J_- \subseteq I_-$ and $J \in Q'$. As in the previous two cases we obtain the satisfaction of the first involved FD, if $Q \cap (\mathcal{P}(I_{+-} \cup I^-)) = Q'$ holds. If this is not the case, one of the other FDs will be satisfied by $\{S_1, S_2\}$.

Finally, for the multiset partition axiom in (33) we proceed analogously. Let $M_1, M_2 \in dom(X)$. In case $M_1 = M_2 = \langle\rangle$ we simple choose $I_+ = I$, so we must have $I^- = I_- = I_{+-} = \emptyset$. Further take $Q' = \{\emptyset\}$. In case exactly one of the M_i is the empty multiset, we choose $I^- = I$, $I_+ = I_- = I_{+-} = \emptyset$, and $Q' = \{J \subseteq I^- \mid \pi^X_{X\langle\bar{X}_J\{\lambda\}\rangle}(M_1) = \pi^X_{X\langle\bar{X}_J\{\lambda\}\rangle}(M_2)\}$.

In case $M_1 \neq \langle\rangle \neq M_2$ we take

$$I_+ = \{i \in I \mid \pi^X_{X\langle\bar{X}_{\{i\}}\{\lambda\}\rangle}(M_1) = \langle \underbrace{\{\top\}}_{x \text{ times}} \rangle = \pi^X_{X\langle\bar{X}_{\{i\}}\{\lambda\}\rangle}(M_2)\},$$

$$I_- = \{i \in I \mid \pi^X_{X\langle\bar{X}_{\{i\}}\{\lambda\}\rangle}(M_1) = \langle \underbrace{\emptyset}_{x \text{ times}} \rangle = \pi^X_{X\langle\bar{X}_{\{i\}}\{\lambda\}\rangle}(M_2)\},$$

$$I^- = \{i \in I \mid \pi^X_{X\langle\bar{X}_{\{i\}}\{\lambda\}\rangle}(M_1) \neq \pi^X_{X\langle\bar{X}_{\{i\}}\{\lambda\}\rangle}(M_2)\},$$

and $I_{+-} = I - I^- - I_+ - I_-$. As before we define $Q' = \{J \subseteq I^- \cup I_{+-} \mid \pi^X_{X\langle\bar{X}_J\{\lambda\}\rangle}(M_1) = \pi^X_{X\langle\bar{X}_J\{\lambda\}\rangle}(M_2)\}$.

In all three cases M_1, M_2 coincide on all $X\langle\bar{X}_{I'}\{\lambda\}\rangle$ with $I' \cap I^+ \neq \emptyset$, on all $X\langle\bar{X}_{J\cup J_-}\{\lambda\}\rangle$ and all $X\langle\bar{X}_J\{\lambda\}\rangle$ with $J_- \subseteq I_-$ and $J \in Q'$. Hence $\{M_1, M_2\}$ satisfies the first involved FD, if $Q \cap (\mathcal{P}(I_{+-} \cup I^-)) = Q'$ holds, while for other Q one of the other FDs will be satisfied.

□

Note that the first three rules (24), (25) and (26) in Theorem 6 are a slight generalisation of the rules used for wFDs in the RDM (see e.g. [32, p.100f.]). The other axioms (27) – (33) arise again from the structure of the subattribute lattice.

4.2 The Completeness Theorem for the Derivation of wFDs

We now want to show that the axioms and rules for the implication of wFDs in Theorems 3, 5 and 6 are also complete. This gives our main result. Before we come to the proof let us make a little observation on the union-constructor.

If $X = X_1(X'_1) \oplus \cdots \oplus X_n(X'_n)$, then each instance r of X can be partioned into r_i $(i = 1, \dots, n)$, where r_i contains exactly the X_i-labelled elements of r. Then r satisfies a FD $\varphi \equiv \mathcal{Y} \to \mathcal{Z}$ iff each r_i satifies the i'th projection φ_i of φ, which results by replacing all subattributes $Y = X_1(Y_1) \oplus \cdots \oplus X_n(Y_n)$ in $\mathcal{Y}$ or $\mathcal{Z}$ by $X_i(Y_i)$. Similarly, we see $\varphi \in \Sigma^+$ iff $\varphi_i \in \Sigma_i^+$ for all $i = 1, \dots, n$.

Lemma 1. *Let $r \subseteq dom(X)$ be an instance of $X = X_1(X'_1) \oplus \cdots \oplus X_n(X'_n)$ and let $\mathcal{Y} \to \mathcal{Z}$ be a FD on $\mathcal{S}(X)$. Define $r_i = \{(X_i : v_i) \mid (X_i : v_i) \in r\}$, $\mathcal{Y}_i = \{X_i(Y_i) \mid X_1(Y_1) \oplus \cdots \oplus X_n(Y_n) \in \mathcal{Y}$ for some $Y_j (j = 1, \dots, n, j \neq i)\}$, and $\mathcal{Z}_i = \{X_i(Z_i) \mid X_1(Z_1) \oplus \cdots \oplus X_n(Z_n) \in \mathcal{Z}$ for some $Z_j (j = 1, \dots, n, j \neq i)\}$ $(i = 1, \dots, n)$. Furthermore, for a set Σ of FDs on $\mathcal{S}(X)$ let $\Sigma_i = \{\mathcal{Y}_i \to \mathcal{Z}_i \mid \mathcal{Y} \to \mathcal{Z} \in \Sigma\}$. Then the following holds:*

1. *$r \models \mathcal{Y} \to \mathcal{Z}$ iff $r_i \models \mathcal{Y}_i \to \mathcal{Z}_i$ holds for all $i = 1, \dots, n$;*
2. *$\mathcal{Y} \to \mathcal{Z} \in \Sigma^+$ iff $\mathcal{Y}_i \to \mathcal{Z}_i \in \Sigma_i^+$ for all $i = 1, \dots, n$.*

Proof. For the first claim let us first assume $r \models \mathcal{Y} \to \mathcal{Z}$. Take $t_1 = (X_i : t'_1) \in r_i$ and $t_2 = (X_i : t'_2) \in r_i$ with $\pi^{X'_i}_{Y_i}(t'_1) = \pi^{X'_i}_{Y_i}(t'_2)$ for all Y_i with $X_i(Y_i) \in \mathcal{Y}_i$. Then $\pi^X_{X_1(Y_1)\oplus\cdots\oplus X_n(Y_n)}(t_1) = \pi^X_{X_1(Y_1)\oplus\cdots\oplus X_n(Y_n)}(t_2)$ holds for all $X_1(Y_1)\oplus\cdots\oplus X_n(Y_n) \in$

$\mathcal{Y}$, and thus $r \models \mathcal{Y} \to \mathcal{Z}$ implies $\pi^X_{X_1(Z_1)\oplus\cdots\oplus X_n(Z_n)}(t_1) = \pi^X_{X_1(Z_1)\oplus\cdots\oplus X_n(Z_n)}(t_2)$ for all $X_1(Z_1) \oplus \cdots \oplus X_n(Z_n) \in \mathcal{Z}$. This gives $\pi^{X'_i}_{Z_i}(t'_1) = \pi^{X'_i}_{Z_i}(t'_2)$ for all Z_i with $X_i(Z_i) \in \mathcal{Z}_i$, hence $r_i \models \mathcal{Y}_i \to \mathcal{Z}_i$ holds for all $i = 1, \dots, n$.

Conversely, assume $r_i \models \mathcal{Y}_i \to \mathcal{Z}_i$ holds for all $i = 1, \dots, n$, and take $t_1 = (X_i : t'_1) \in r$ and $t_2 = (X_j : t'_2) \in r$. If $i \neq j$, then t_1, t_2 differ on all subattributes except λ, i.e. $\{t_1, t_2\} \models \mathcal{Y} \to \mathcal{Z}$. So assume $j = i$, i.e. $t_1, t_2 \in r_i$, and $\pi^X_Y(t_1) = \pi^X_Y(t_2)$ for all $Y \in \mathcal{Y}$, i.e. for $Y = X_1(Y_1) \oplus \cdots \oplus X_n(Y_n)$ we obtain $\pi^{X'_i}_{Y_i}(t'_1) = \pi^{X'_i}_{Y_i}(t'_2)$. As $X_i(Y_i) \in \mathcal{Y}_i$, the premise implies $\pi^{X'_i}_{Z_i}(t'_1) = \pi^{X'_i}_{Z_i}(t'_2)$ for all Z_i with $X_i(Z_i) \in \mathcal{Z}_i$ and further $\pi^X_Z(t_1) = \pi^X_Z(t_2)$ for all $Z \in \mathcal{Z}$, hence $r \models \mathcal{Y} \to \mathcal{Z}$.

For the second claim first assume $\mathcal{Y}_i \to \mathcal{Z}_i \in \Sigma_i^+$ for all $i = 1, \dots, n$. Then $\mathcal{Y} \to \mathcal{Z} \in \Sigma^+$ results from successive applications of the union lifting rule (21) together with the subattribute axiom (2), the reflexivity axiom (1) and the transitivity rule (6).

Conversely, in a derivation of $\mathcal{Y} \to \mathcal{Z}$ from Σ all involved subattributes other than λ will have the form $X_1(U_1) \oplus \cdots \oplus X_n(U_n)$ with $X'_i \geq U_i$. Reducing this to $X_i(U_i)$ in each step gives a valid derivation of $\mathcal{Y}_i \to \mathcal{Z}_i$ from Σ_i.

□

Theorem 7 (Completeness Theorem)**.** *The set of axioms and rules in Theorems 3, 5 and 6 is complete for the implication of wFDs on $\mathcal{S}(X)$.*

Proof. Let Σ be a set of wFDs on $\mathcal{S}(X)$ and assume $\{\!\!\{\mathcal{Y}_i \to \mathcal{Z}_i \mid i \in I\}\!\!\} \notin \Sigma^+$. Due to the union rule (7) we must have $\{\!\!\{\mathcal{Y}_i \to \{Z_i\} \mid i \in I\}\!\!\} \notin \Sigma^+$ for some selected $Z_i \in \mathcal{Z}_i$. Furthermore, due to the left union rule (25) we get $\{\!\!\{\mathcal{Y} \to \{Z_i\} \mid i \in I\}\!\!\} \notin \Sigma^+$ with $\mathcal{Y} = \bigcup_{i \in I} \mathcal{Y}_i$.

Let $\mathcal{Z} = \{Z \mid Z \geq Z_i \text{ for some } i \in I\}$ and $\mathcal{U} = \mathcal{S}(X) - \mathcal{Y} - \mathcal{Z}$. Due to the reflexivity axiom (1) we obviously have $Z_i \notin \mathcal{Y}$, and then $\mathcal{Y} \sqcap \mathcal{Z} = \emptyset$ due to the subattribute axiom (2). Due to the shift rule (26) there must exist some $\mathcal{U}' \subseteq \mathcal{U}$ with $\{\!\!\{\mathcal{Y} \cup \mathcal{U}' \to \{Z\} \mid Z \in \mathcal{Z} \cup (\mathcal{U} - \mathcal{U}')\}\!\!\} \notin \Sigma^+$. Otherwise we could derive $\{\!\!\{\mathcal{Y} \to \{Z\} \mid Z \in \mathcal{Z}\}\!\!\}$, and thus $\{\!\!\{\mathcal{Y} \to \{Z_i\} \mid i \in I\}\!\!\} \in \Sigma^+$ contradicting our assumption.

Lemma 2. *Let $\mathcal{U}'$ be maximal with the given property. Then $\mathcal{F} = \mathcal{Y} \cup \mathcal{U}'$ is a coincidence ideal.*

We first prove Lemma 2, then continue the proof of Theorem 7.

of Lemma 2.
1. Assume $\mathcal{F} = \emptyset$. This implies $\mathcal{Z} \cup \mathcal{U} = \mathcal{S}(X)$ and thus $\{\!\!\{\emptyset \to \{Z\} \mid Z \in \mathcal{S}(X)\}\!\!\} \notin \Sigma^+$. This wFD, however, can be derived from $\emptyset \to \{\lambda\} \in \Sigma^+$ (due to the λ-axiom (4)) using the weakening rule (24). Thus, $\mathcal{F}$ is not empty.

2. Now let $Y \in \mathcal{F}$ and $Y \geq Y'$ Assume $Y' \notin \mathcal{F}$. So $Y' \in \mathcal{U}$, otherwise we get $Y' \in \mathcal{Y} \cup \mathcal{Z}$, which implies $Y' \geq Z_i$ for some $i \in I$ and furtheron $Y \geq Y' \geq Z_i$, which gives the contradiction $Y \in \mathcal{Z}$.

Now take $\mathcal{U}'' = \mathcal{U}' \cup \{Y'\}$. The subattribute axiom (2) together with the extension and transitivity rules (5) and (6) implies $\mathcal{Y} \cup \mathcal{U}' \to \mathcal{Y} \cup \mathcal{U}'' \in \Sigma^+$. As $\mathcal{U}'$ was chosen maximal, we also have $\{\!|\mathcal{Y} \cup \mathcal{U}'' \to \{Z\} \mid Z \in \mathcal{Z} \cup (\mathcal{U} - \mathcal{U}'')|\!\} \in \Sigma^+$. Using the transitivity rule (6) again, this gives $\{\!|\mathcal{Y} \cup \mathcal{U}' \to \{Z\} \mid Z \in \mathcal{Z} \cup (\mathcal{U} - \mathcal{U}'')|\!\} \in \Sigma^+$. Then the weakening rule (24) leads to the contradiction $\{\!|\mathcal{Y} \cup \mathcal{U}' \to \{Z\} \mid Z \in \mathcal{Z} \cup (\mathcal{U} - \mathcal{U}')|\!\} \in \Sigma^+$.

3. Let $Y_1, Y_2 \in \mathcal{F}$ be reconsilable. Assume $Y = Y_1 \sqcup Y_2 \notin \mathcal{F}$. If $Y \in \mathcal{U}$, we take $\mathcal{U}'' = \mathcal{U}' \cup \{Y\}$. Due to the maximality of $\mathcal{U}'$ we get $\{\!|\mathcal{Y} \cup \mathcal{U}'' \to \{Z\} \mid Z \in \mathcal{Z} \cup (\mathcal{U} - \mathcal{U}'')|\!\} \in \Sigma^+$, thus by the weakening rule (24) also $\{\!|\mathcal{Y} \cup \mathcal{U}'' \to \{Z\} \mid Z \in \mathcal{Z} \cup (\mathcal{U} - \mathcal{U}')|\!\} \in \Sigma^+$.

   On the other hand, the join axiom (3) implies $\{Y_1, Y_2\} \to \{Y\} \in \Sigma^+$. Using the reflexivity axiom (1), the extension rule (5) and the transitivity rule (6) we obtain $\mathcal{Y} \cup \mathcal{U}' \to \mathcal{Y} \cup \mathcal{U}'' \in \Sigma^+$, from which we get the contradiction $\{\!|\mathcal{Y} \cup \mathcal{U}' \to \{Z\} \mid Z \in \mathcal{Z} \cup (\mathcal{U} - \mathcal{U}')|\!\} \in \Sigma^+$ by another application of the transitivity rule.

   If $Y \notin \mathcal{U}$, we get $Y \in \mathcal{Z}$, thus $\{Y\}$ is among the right hand sides in $\{\!|\mathcal{Y} \cup \mathcal{U}' \to \{Z\} \mid Z \in \mathcal{Z} \cup (\mathcal{U} - \mathcal{U}')|\!\} \notin \Sigma^+$. However, the join rule (9) together with the reflexivity axiom and the transitivity rule imply $\mathcal{Y} \cup \mathcal{U}' \to \{Y\} \in \Sigma^+$, hence the weakening rule leads to the contradiction $\{\!|\mathcal{Y} \cup \mathcal{U}' \to \{Z\} \mid Z \in \mathcal{Z} \cup (\mathcal{U} - \mathcal{U}')|\!\} \in \Sigma^+$.

4. a) Assume $X_I\{\lambda\} \in \mathcal{F}$, but $(X_J\{\lambda\} \notin \mathcal{F}$ for $\{i_1, \dots, i_k\} = I \subsetneq J$. As $\mathcal{S}(X)$ is partitioned into $\mathcal{Z} \cup (\mathcal{U} - \mathcal{U}')$ and $\mathcal{F} = \mathcal{Y} \cup \mathcal{U}'$, we must have $X_J\{\lambda\} \in \mathcal{Z} \cup (\mathcal{U} - \mathcal{U}')$. From the union axiom (27), the transitivity rule and $X(X_I\{\lambda\}) \in \mathcal{F}$ we conclude

   $$\{\!|\mathcal{Y} \cup \mathcal{U}' \to \{Z\} \mid Z \in \{X_J\{\lambda\}, X(X_{i_1}\{X'_{i_1}\}, \dots, X_{i_k}\{X'_{i_k}\})\}|\!\} \in \Sigma^+.$$

   Due to the weakening rule (24) it follows $\{\!|\mathcal{Y} \cup \mathcal{U}' \to \{Z\} \mid Z \in \mathcal{W}|\!\} \in \Sigma^+$ for all $\mathcal{W} \subseteq \mathcal{S}(X)$ with $X_J\{\lambda\}, X(X_{i_1}\{X'_{i_1}\}, \dots, X_{i_k}\{X'_{i_k}\}) \in \mathcal{W}$. According to the definition of $\mathcal{U}'$ we must have either $X_J\{\lambda\} \notin \mathcal{Z} \cup (\mathcal{U} - \mathcal{U}')$ or $X(X_{i_1}\{X'_{i_1}\}, \dots, X_{i_k}\{X'_{i_k}\}) \notin \mathcal{Z} \cup (\mathcal{U} - \mathcal{U}')$, which implies $X(X_{i_1}\{X'_{i_1}\}, \dots, X_{i_k}\{X'_{i_k}\}) \in \mathcal{F}$.

   b) Assume $X_I\{\lambda\} \in \mathcal{F}$, but $X(X_i\{\lambda\}) \notin \mathcal{F}$ for all $i \in I$. In particular $X(X_i\{\lambda\}) \in \mathcal{Z} \cup (\mathcal{U} - \mathcal{U}')$. Using the partition axiom (28), the transitivity rule and $X_I\{\lambda\} \in \mathcal{F}$ we conclude $\{\!|\mathcal{Y} \cup \mathcal{U}' \to \{X_{I'_1 \cup I'_2}\{\lambda\} \mid \emptyset \neq I'_1 \subseteq I_1, \emptyset \neq I'_2 \subseteq I_2\}, \mathcal{Y} \cup \mathcal{U}' \to \{X(X_i\{\lambda\})\} \mid I = I_1 \dot\cup I_2, i \in I, I_1 \neq \emptyset \neq I_2|\!\} \in \Sigma^+$.

   If for all partitions $I = I_1 \dot\cup I_2$ we had at least one $X_{I'_1 \cup I'_2}\{\lambda\} \in \mathcal{Z} \cup (\mathcal{U} - \mathcal{U}')$, we can apply the reflexivity axiom, the transitivity rule and the weakening rule to derive $\{\!|\mathcal{Y} \cup \mathcal{U}' \to \{Z\} \mid Z \in \mathcal{Z} \cup (\mathcal{U} - \mathcal{U}')|\!\} \in \Sigma^+$ contradicting the assumption on $\mathcal{U}'$. Therefore, there is a partition $I = I_1 \dot\cup I_2$ with $\{X_{I'_1 \cup I'_2}\{\lambda\} \mid \emptyset \neq I'_1 \subseteq I_1, \emptyset \neq I'_2 \subseteq I_2\} \subseteq \mathcal{F}$.

Choose such a partition. If we had $X_{I_1}\{\lambda\} \in \mathcal{F}$, we could choose a maximal $J \subsetneq I_1$ with $X_J\{\lambda\} \notin \mathcal{F}$, and $X_{\{j\}}\{\lambda\} \notin \mathcal{F}$ for all $j \in J$. So, we can partition I_1 into J and $I' = I_1 - J$. Now use property 4(d) – which we prove soon not using 4(b). Due to this property we find some $i \in I'$ such that $X_{J' \cup \{i\}}\{\lambda\} \in \mathcal{F}$ holds for all $J' \subseteq J$. In particular, for $J' = \emptyset$ we obtain a contradiction. Hence we must have $X_{I_1}\{\lambda\} \notin \mathcal{F}$ and by symmetry also $X_{I_2}\{\lambda\} \notin \mathcal{F}$.

c) Assume $X_{\{1,\dots,n\}}\{\lambda\} \in \mathcal{F}$, $X_{I^-}\{\lambda\} \notin \mathcal{F}$ and for all $i \in I^+$ there is some $J \subseteq I^-$ with $X_{J\cup\{i\}}\{\lambda\} \notin \mathcal{F}$. Let this J be denoted as J_i. Taking the first plus/minus axiom (29), the left hand side of the FDs are always in $\mathcal{F}$. Therefore, using the reflexivity axiom and the transitivity rule we derive $\{\!|\mathcal{F} \to \{X_{J_i\cup\{i\}}\{\lambda\}\}, \mathcal{F} \to \{X\}, \mathcal{F} \to \{X_j\{\lambda\}\}, \mathcal{F} \to \{X_{I^-}\{\lambda\}\} \mid i \in I^+, j \in I^-|\!\} \in \Sigma^+$. Now the right hand sides of the FDs are all not in $\mathcal{F}$, so the weakening rule implies $\{\!|\mathcal{F} \to \{Z\} \mid Z \notin \mathcal{F}|\!\} \in \Sigma^+$ contradicting the construction of $\mathcal{F}$, according to which $\mathcal{S}(X) - \mathcal{F} = \mathcal{Z} \cup (\mathcal{U} - \mathcal{U}')$. and $\{\!|\mathcal{F} \to \{Z\} \mid Z \in \mathcal{Z} \cup (\mathcal{U} - \mathcal{U}')|\!\} \notin \Sigma^+$.

d) Let $I \cap J = \emptyset$ and $X_J\{\lambda\} \notin \mathcal{F}$, $X_{\{j\}}\{\lambda\} \notin \mathcal{F}$ for all $j \in J$, and for all $i \in I$ there is some $J_i \subseteq J$ with $X_{J_i\cup\{i\}}\{\lambda\} \notin \mathcal{F}$. Furthermore, assume $X_{I\cup J}\{\lambda\} \in \mathcal{F}$. Then from the second plus/minus axiom (30), the transitivity rule (6) and the reflexivity axiom (1) we derive $\{\!|\mathcal{F} \to \{X_J\{\lambda\}\}, \mathcal{F} \to \{X_{\{j\}}\{\lambda\}\}, \mathcal{F} \to \{X_{J_i\cup\{i\}}\{\lambda\}\} \mid i \in I, j \in J|\!\} \in \Sigma^+$. Here the right hand sides of all involved FDs have the form $\{Z\}$ with $Z \notin \mathcal{F}$, so the weakening rule (24) gives $\{\!|\mathcal{F} \to \{Z\} \mid Z \notin \mathcal{F}|\!\} \in \Sigma^+$ contradicting the construction of $\mathcal{F}$. Hence we must have $X_{I\cup J}\{\lambda\} \notin \mathcal{F}$.

e) Assume $X_{I^-}\{\lambda\} \in \mathcal{F}$, and let $I' \subseteq I^+$ such that for all $i \in I'$ there is some $J_i \subseteq I^-$ with $X_{J_i\cup\{i\}}\{\lambda\} \notin \mathcal{F}$. Let $J' \subseteq I^-$ with $X_{J'}\{\lambda\} \notin \mathcal{F}$ and assume $X_{I'\cup J'}\{\lambda\} \in \mathcal{F}$. Then for $i \in I'$, $k \in I^-$ using the the third plus/minus axiom (31), the reflexivity axiom (1) and the transitivity rule (6) we derive $\{\!|\mathcal{F} \to \{X_{J'}\{\lambda\}\}, \mathcal{F} \to \{X_{J_i\cup\{i\}}\{\lambda\}\}, \mathcal{F} \to \{X_{\{k\}}\{\lambda\}\}|\!\} \in \Sigma^+$. Again the right hand sides of all involved FDs have the form $\{Z\}$ with $Z \notin \mathcal{F}$ leading to the contradiction $\{\!|\mathcal{F} \to \{Z\} \mid Z \notin \mathcal{F}|\!\} \in \Sigma^+$ by applying the weakening rule (24). Hence we must have $X_{I'\cup J'}\{\lambda\} \notin \mathcal{F}$ for all $J' \subseteq I^-$ with $X_{J'}\{\lambda\} \notin \mathcal{F}$.

5. a) Let $X_I\{\lambda\}, X_J\{\lambda\} \in \mathcal{F}$ with $I \cap J = \emptyset$, but assume $X_{I\cup J}\{\lambda\} \notin \mathcal{F}$, i.e. $X_{I\cup J}\{\lambda\} \in \mathcal{Z} \cup (\mathcal{U} - \mathcal{U}')$. From the set axiom (10), the reflexivity axiom (1) and the transitivity rule (6) we derive $\mathcal{F} \to \{X_{I\cup J}\{\lambda\}\} \in \Sigma^+$ and further $\{\!|\mathcal{F} \to \{Z\} \mid Z \notin \mathcal{F}|\!\} \in \Sigma^+$ by the weakening rule (24). This contradicts the construction of $\mathcal{F}$, so we must have $X_{I\cup J}\{\lambda\} \in \mathcal{F}$.

   The proof of properties 5(b) and (c) is completely analogous using (15) and (11), respectively, instead of (10).

   d) Let $X_I[\lambda], X_J[\lambda] \in \mathcal{F}$ with $I \subseteq J$, but assume $X_{J-I}[\lambda] \notin \mathcal{F}$. From the second list axiom (16), the reflexivity axiom (1) and the transitivity rule (6) we derive $\mathcal{F} \to \{X_{J-I}[\lambda]\} \in \Sigma^+$. Applying the weakening rule (24) leads to the contradiction $\{\!\!|\mathcal{F} \to \{Z\} \mid Z \notin \mathcal{F}|\!\!\} \in \Sigma^+$. Hence we must have $X_{J-I}[\lambda] \in \mathcal{F}$.

The proof of property 5(e) is completely analogous using (12) instead of (16).

f) Let $X_I[\lambda], X_J[\lambda] \in \mathcal{F}$ and assume $X_{I\cap J}[\lambda] \in \mathcal{F}$, but $X_{(I\cup J)-(I\cap J)}[\lambda] \notin \mathcal{F}$. The the third list axiom (17), the reflexivity axiom (1) and the transitivity rule (6) allow us to derive $\mathcal{F} \to \{X_{(I\cup J)-(I\cap J)}[\lambda]\} \in \Sigma^+$. Further application of the weakening rule (24) leads to the contradiction $\{\!\!|\mathcal{F} \to \{Z\} \mid Z \notin \mathcal{F}|\!\!\} \in \Sigma^+$. Hence we must have $X_{(I\cup J)-(I\cap J)}[\lambda] \in \mathcal{F}$.

Analogously, assuming $X_{I\cap J}[\lambda] \notin \mathcal{F}$ and $X_{(I\cup J)-(I\cap J)}[\lambda] \in \mathcal{F}$ leads to the same contradiction using the fourth list axiom (18) instead of (17). The proof of property 5(g) is completely analogous using (13) and (14) instead of (17) and (18), respectively.

6. If property 6(a) were not satisfied, then for all partitions $I \cap I^+ = I_+ \cup I_- \cup I_{+-}$ one of the properties ii or iii in Definition 7 6(a) would be violated. In case property ii is violated there is some I' with $I' \cap I_+ \neq \emptyset$ and $X\{\bar{X}_{I'}\{\lambda\}\} \notin \mathcal{F}$. In case property iii is violated there exists some $I' \subseteq I_{+-} \cup I^- \cup I_-$ such that either $X\{\bar{X}_{I'}\{\lambda\}\} \in \mathcal{F}$ and $X\{\bar{X}_{I'\cap(I_{+-}\cup I^-)}\{\lambda\}\} \notin \mathcal{F}$ or $X\{\bar{X}_{I'}\{\lambda\}\} \notin \mathcal{F}$ and $X\{\bar{X}_{I'\cap(I_{+-}\cup I^-)}\{\lambda\}\} \in \mathcal{F}$. Then define $J_- = I' - I_{+-} - I^-$ and $J = I' - I_-$, which gives $X\{\bar{X}_J\{\lambda\}\} \notin \mathcal{F}$ in the first case and $X\{\bar{X}_{J_-\cup J}\{\lambda\}\} \notin \mathcal{F}$ in the second case.

   Let $Q = \{J \subseteq I \mid X\{\bar{X}_J\{\lambda\}\} \in \mathcal{F}\}$. Then the right hand side of the first FD in the set partition axiom (32) contains a subattribute $Z \notin \mathcal{F}$, and the same holds for the involved FDs of the form $\{\lambda\} \to \{X\{\bar{X}_K\{\lambda\}\}\}$. For the remaining involved FDs we can replace the right hand side by $\{Z\}$ with some $Z \notin \mathcal{F}$ using the subattribute axiom (2) and the transitivity rule. Thus, using (32), (2), the reflexivity axiom, the transitivity rule, and the weakening rule, we derive the contradiction $\{\!\!|\mathcal{F} \to \{Z\} \mid Z \notin \mathcal{F}|\!\!\} \in \Sigma^+$. Hence there is a partition $I = I^- \cup I_+ \cup I_- \cup I_{+-}$ satisfying the properties i, ii or iii in Definition 7 6(a).

   The proof of property 6(b) is completely analogous using the multiset partition axiom (33) instead of (32).

7. For the proof of property 7 observe that the proofs of properties 1 - 6 follow a simple pattern. Assuming that the property does not hold we obtain an instance of a particular axiom, which together with the reflexivity axiom (1), the transitivity rule (6) and the weakening rule (24) allows us to derive the contradiction $\{\!\!|\mathcal{F} \to \{Z\} \mid Z \notin \mathcal{F}|\!\!\} \in \Sigma^+$.

   To be precise, we used the λ axiom (4) for property 1, the subattribute axiom (2) for property 2, the join axiom (3) for property 3, the union axiom (27) for property 4(a), the partition axiom (28) for property 4(b), the first plus/minus axiom (29) for property 4(c), the second plus/minus axiom (30) for property 4(d), the third plus/minus axiom (31) for property 4(e), the set axiom (10) for property 5(a), the four list axioms (15) – (18) for properties 5(b),(d) and (f), the four multiset axioms (11) – (14) for properties 5(c),(e) and (g), the set partition axiom (32) for property 6(a), and the multiset partition axiom (33) for property 6(b).

We can apply the record lifting rule (20), the union lifting rule (21) and the list lifting rule (23) to all these axioms to derive additional axioms, and we can apply the set lifting rule (19) and the multiset lifting rule (22) to the axioms except (3), (31) and (10) – (18). The resulting axioms differ from the original ones only by "wrapping" constructors around the involved attributes. Then using exactly the same arguments as before, we obtain additional properties for $\mathcal{F}$ that correspond to the required properties for the embedded ideals $\mathcal{F}_i$ or $\mathcal{G}$ used in properties 7(a)-(e) and 8(a)-(c), which completes the proof.

□

Proof of Theorem 7 (continued): Due to the restructuring rules in Definition 4 we may assume that the union-constructor appears in X only inside a set-, list- or multiset-constructor or as the outermost constructor.

Let us first assume that the outermost constructor is not the union-constructor. Then we can apply the Central Theorem 2, which gives us $r = \{t_1, t_2\} \subseteq dom(X)$ with $\pi_Y^X(t_1) = \pi_Y^X(t_2)$ iff $Y \in \mathcal{F} = \mathcal{Y} \cup \mathcal{U}'$. In particular, $\pi_Y^X(t_1) = \pi_Y^X(t_2)$ for all $i \in I$ and $Y \in \mathcal{Y}_i$, and $\pi_{Z_i}^X(t_1) \neq \pi_{Z_i}^X(t_2)$ for all $i \in I$. That is, $r \not\models \{\!\!\{\mathcal{Y}_i \rightarrow \{Z_i\} \mid i \in I\}\!\!\}$. From the soundness of the fragmentation rule (8) we conclude $r \not\models \{\!\!\{\mathcal{Y}_i \rightarrow \mathcal{Z}_i \mid i \in I\}\!\!\}$.

Now assume that the outermost constructor of X is the union-constructor, say $X = X_1(X_1') \oplus \cdots \oplus X_n(X_n')$. We know from Lemma 2 that $\mathcal{F} = \mathcal{Y} \cup \mathcal{U}'$ is a coincidence ideal on $\mathcal{S}(X)$. If $\mathcal{F} = \{\lambda\}$, then take $t_1 = (X_1 : t_1')$ and $t_2 = (X_2 : t_2')$ with arbitrary $t_j' \in dom(X_j')$. Then $\pi_Y^X(t_1) = \pi_Y^X(t_2)$ iff $Y = \lambda$. As before this implies $r \not\models \{\!\!\{\mathcal{Y}_i \rightarrow \mathcal{Z}_i \mid i \in I\}\!\!\}$ with $r = \{t_1, t_2\}$.

For $\mathcal{F} \neq \{\lambda\}$ take the embedded coincidence ideal $\mathcal{F}_i$ on $\mathcal{S}(X_i')$ according to Definition 7. Using the Central Theorem 2 we find $t_{i1}, t_{i2} \in dom(X_i')$ with $\pi_{Y_i}^{X_i'}(t_{i1}) = \pi_{Y_i}^{X_i'}(t_{i2})$ iff $Y_i \in \mathcal{F}_i$.

As we have $\{\!\!\{\mathcal{F} \rightarrow \{Z\} \mid Z \in \mathcal{Z} \cup (\mathcal{U} - \mathcal{U}')\}\!\!\} \notin \Sigma^+$, we must also have $\{\!\!\{\mathcal{F}_j \rightarrow \{Z\} \mid Z \in (\mathcal{Z} \cup (\mathcal{U} - \mathcal{U}'))_j\}\!\!\} \notin \Sigma_j^+$ for at least one j according to Lemma 1. In particular, for $Z_i = X_1(Z_{i1}') \oplus \cdots \oplus X_n(Z_{in}')$ we find some j such that $Z_{ij}' \notin \mathcal{F}_j$ for all $i \in I$.

Now take $r = \{(X_j : t_{j1}), (X_j : t_{j2})\}$. Then for all $i \in I$ and all $Y = X_1(Y_1) \oplus \cdots \oplus X_n(Y_n) \in \mathcal{Y}_i \subseteq \mathcal{F}$ we have $Y_j \in \mathcal{F}_j$, and we obtain

$$\pi_Y^X((X_j : t_{j1})) = (X_j : \pi_{Y_j}^{X_j'}(t_{j1})) = (X_j : \pi_{Y_j}^{X_j'}(t_{j2})) = \pi_Y^X((X_j : t_{j2})) .$$

On the other hand, $Z_{ij}' \notin \mathcal{F}_j$ implies

$$\pi_{Z_i}^X((X_j : t_{j1})) = (X_j : \pi_{Z_{ij}'}^{X_j'}(t_{j1})) \neq (X_j : \pi_{Z_{ij}'}^{X_j'}(t_{j2})) = \pi_{Z_i}^X((X_j : t_{j2}))$$

for all $i \in I$. That is $r \not\models \{\!\!\{\mathcal{Y}_i \to \{Z_i\} \mid i \in I\}\!\!\}$, and hence $r \not\models \{\!\!\{\mathcal{Y}_i \to \mathcal{Z}_i \mid i \in I\}\!\!\}$ by the soundness of the fragmentation rule (8).

The next Lemma 3 shows $r \models \Sigma$ in both cases. This implies $r \models \Sigma^*$, and thus $\{\!\!\{\mathcal{Y}_i \to \mathcal{Z}_i \mid i \in I\}\!\!\} \notin \Sigma^*$, which completes the proof of Theorem 7.

□

Lemma 3. $r \models \Sigma$.

Proof. First assume again that the outermost constructor is not the union-constructor. Let $\{\!\!\{\mathcal{V}_j \to \mathcal{W}_j \mid j \in J\}\!\!\} \in \Sigma$.

1. If $\mathcal{V}_j \not\subseteq \mathcal{Y} \cup \mathcal{U}'$ for some $j \in J$, we get $\pi_V^X(t_1) \neq \pi_V^X(t_2)$ for some $V \in \mathcal{V}_j$. Thus $r \models \mathcal{V}_j \to \mathcal{W}_j$ and due to the soundness of the weakening rule also $r \models \{\!\!\{\mathcal{V}_j \to \mathcal{W}_j \mid j \in J\}\!\!\}$.
2. If $\mathcal{V}_j \subseteq \mathcal{Y} \cup \mathcal{U}'$ for all $j \in J$, we get $\mathcal{Y} \cup \mathcal{U}' \to \mathcal{V}_j \in \Sigma^+$ from the reflexivity axiom, $\{\!\!\{\mathcal{Y} \cup \mathcal{U}' \to \mathcal{W}_j \mid j \in J\}\!\!\} \in \Sigma^+$ from the transitivity rule, and $\{\!\!\{\mathcal{Y} \cup \mathcal{U}' \to \{W_j\} \mid j \in J\}\!\!\} \in \Sigma^+$ for any choices $W_j \in \mathcal{W}_j$ from the fragmentation rule.

   Assume we could select $W_j \in \mathcal{W}_j - \mathcal{Y} - \mathcal{U}'$ for all $j \in J$. Then the weakening rule implies $\{\!\!\{\mathcal{Y} \cup \mathcal{U}' \to \{W\} \mid W \in \mathcal{S}(X) - \mathcal{Y} - \mathcal{U}'\}\!\!\} \in \Sigma^+$. However, $\mathcal{S}(X) - \mathcal{Y} - \mathcal{U}' = \mathcal{Z} \cup (\mathcal{U} - \mathcal{U}')$, so we get a contradiction to the choice of $\mathcal{U}'$.

   Therefore, we must have $\mathcal{W}_j \subseteq \mathcal{Y} \cup \mathcal{U}'$ for some $j \in J$. By construction of r we get $\pi_W^X(t_1) = \pi_W^X(t_2)$ for all $W \in \mathcal{W}_j$, thus $r \models \mathcal{V}_j \to \mathcal{W}_j$. This implies $r \models \{\!\!\{\mathcal{V}_j \to \mathcal{W}_j \mid j \in J\}\!\!\}$ due to the soundness of the weakening rule.

If the outermost constructor is the union-constructor, then according to Lemma 1 we have to show $r_j \models \Sigma_j$. The proof is analogous to the case before.

□

4.3 The Case of Functional Dependencies

Theorem 7 shows the axiomatisation of wFDs. If Σ is a set of "ordinary" FDs, we can apply the axioms and rules to Σ and then the FDs in Σ^+ will be the implied FDs. Of course, we would like to have an axiomatisation for FDs that avoids such a detour via the wFDs.

We first observe that most of the axioms and rules for wFDs in Theorem 6 depend on the joint occurrence of the set and the union constructor. Only the weakening rule (24), the left union rule (25) and the shift rule (26) do not make such a special assumption. In particular, in a derivation of FDs for a nested attribute X that does not contain both the union and the set constructor, the special axioms in Theorem 6 will not be needed. We now show that indeed none of the rules for the derivation of wFDs are needed either, i.e. the set of axioms and rules in Theorems 3 and 5 excluding the set axiom (10) are sound and complete for the derivation of FDs in this case.

In order to prove this, observe that properties 4, 5(a) and 6 in Theorem 2 can be ignored, if the union and the set constructor do not appear jointly.

Theorem 8. *Let $X \in \mathcal{N}$ be a nested attribute not containing both the set and the union constructor. Then the set of axioms and rules in Theorems 3, 5 excluding the set axiom (10) is complete for the implication of FDs on $\mathcal{S}(X)$.*

Proof. Let Σ be a set of FDs on $\mathcal{S}(X)$ and assume $\mathcal{Y} \rightarrow \mathcal{Z} \notin \Sigma^+$. Then due to the union rule (7) there exists a subattribute $Z \in \mathcal{Z}$ with $\mathcal{Y} \rightarrow \{Z\} \notin \Sigma^+$. Thus, $Z \notin \bar{\mathcal{Y}} = \{Z' \mid \mathcal{Y} \rightarrow \{Z'\} \in \Sigma^+\}$. We show that $\mathcal{F} = \bar{\mathcal{Y}}$ is a coincidence ideal on $\mathcal{S}(X)$:

1. $\lambda \in \mathcal{F}$ follows immediately from the reflexivity axiom (1), the λ axiom (4), and the transitivity rule (6).
2. For $Z_1 \in \mathcal{F}$ and $Z_1 \geq Z_2$ the subattribute axiom (2) and the transitivity rule (6) imply $Z_2 \in \mathcal{F}$.
3. For reconsilable $Z_1, Z_2 \in \mathcal{F}$ the join axiom (3) and the transitivity rule (6) imply $Z_1 \sqcup Z_2 \in \mathcal{F}$.
5. Property (b)–(g) result immediately from applying the multiset axioms (11) – (14) and the list axioms (15) – (18) together with the transitivity rule (6).
7. The proof of property 7 in Definition 7 is analogous to the corresponding proof for Lemma 2. We apply lifting rules (19) – (23) to the axioms used in the proof of properties 1, 2, 3 and 5, then apply the same argument as before.

If the outermost constructor is not the union-constructor, we can apply the Central Theorem 2, which gives us $r = \{t_1, t_2\} \subseteq dom(X)$ with $\pi_Y^X(t_1) = \pi_Y^X(t_2)$ iff $Y \in \mathcal{F}$. In particular, $\pi_Y^X(t_1) = \pi_Y^X(t_2)$ for all $Y \in \mathcal{Y}$, and $\pi_Z^X(t_1) \neq \pi_Z^X(t_2)$. That is, $r \not\models \mathcal{Y} \rightarrow \{Z\}$. From the soundness of the fragmentation rule (8) we conclude $r \not\models \mathcal{Y} \rightarrow \mathcal{Z}$.

If the outermost constructor of X is the union-constructor, say $X = X_1(X_1') \oplus \cdots \oplus X_n(X_n')$ with $n \geq 2$, then either $\mathcal{F} = \{\lambda\}$ or we obtain embedded coincidence ideals $\mathcal{F}_i$ on $\mathcal{S}(X_i')$ $(i = 1, \ldots, n)$ according to Definition 7. In the first case take $t_1 = (X_1 : t_1')$ and $t_2 = (X_2 : t_2')$ with arbitrary $t_j' \in dom(X_j')$. Then $\pi_U^X(t_1) = \pi_U^X(t_2)$ iff $U = \lambda$. As before $Z \notin \mathcal{F}$ implies $r \not\models \mathcal{Y} \rightarrow \mathcal{Z}$ with $r = \{t_1, t_2\}$.

In the second case the Central Theorem 2 gives us $t_{i1}, t_{i2} \in dom(X_i')$ with $\pi_{Y_i}^{X_i'}(t_{i1}) = \pi_{Y_i}^{X_i'}(t_{i2})$ iff $Y_i \in \mathcal{F}_i$. As we have $\mathcal{F} \rightarrow \{Z\} \notin \Sigma^+$, we must also have $\mathcal{F}_j \rightarrow \{Z_j\} \notin \Sigma_j^+$ for at least one j according to Lemma 1, in particular $Z_j \notin \mathcal{F}_j$.

Now take $r = \{(X_j : t_{j1}), (X_j : t_{j2})\}$. Then for all $Y = X_1(Y_1) \oplus \cdots \oplus X_n(Y_n) \in \mathcal{Y} \subseteq \mathcal{F}$ we have $Y_j \in \mathcal{F}_j$, and we obtain

$$\pi_Y^X((X_j : t_{j1})) = (X_j : \pi_{Y_j}^{X_j'}(t_{j1})) = (X_j : \pi_{Y_j}^{X_j'}(t_{j2})) = \pi_Y^X((X_j : t_{j2})).$$

On the other hand, $Z_j \notin \mathcal{F}_j$ implies

$$\pi_Z^X((X_j : t_{j1})) = (X_j : \pi_{Z_j}^{X_j'}(t_{j1})) \neq (X_j : \pi_{Z_j}^{X_j'}(t_{j2})) = \pi_Z^X((X_j : t_{j2})).$$

That is $r \not\models \mathcal{Y} \to \{Z\}$, and hence $r \not\models \mathcal{Y} \to \mathcal{Z}$ by the soundness of the fragmentation rule (8).

We finally show $r \models \Sigma$ in both cases, which proves the theorem. We show this for the case that the outermost constructor is the not union-constructor. If the outermost constructor is the union-constructor, then according to Lemma 1 we have to show $r_j \models \Sigma_j$ for all $j = 1, \ldots, n$, the proof of which is analogous to the first case. So let $\mathcal{U} \to \mathcal{V} \in \Sigma$. We distinguish two cases:

- If $\mathcal{U} \subseteq \mathcal{F}$, then $\pi_U^X(t_1) = \pi_U^X(t_2)$ for all $U \in \mathcal{U}$. The reflexivity axiom and the transitivity rule allow us to derive $\mathcal{Y} \to \mathcal{V} \in \Sigma^+$, which means $\mathcal{V} \subseteq \mathcal{F}$ and thus $\pi_V^X(t_1) = \pi_V^X(t_2)$ for all $V \in \mathcal{V}$, i.e. $r \models \mathcal{U} \to \mathcal{V}$.
- If $\mathcal{U} \subseteq \mathcal{F}$, then there is some $U \in \mathcal{U}$ with $\pi_U^X(t_1) \neq \pi_U^X(t_2)$, which immediately implies $r \models \mathcal{U} \to \mathcal{V}$.

□

In fact, the proof shows a bit more than claimed. We only needed that properties 4, 5(a) and 6 of Theorem 2 are immediately satisfied, because the corresponding attributes can both appear as subattributes of an attribute $X' \in \text{emb}(X)$. This gives the following theorem.

Theorem 9 (Completeness Theorem for FDs)**.** *Let $X \in \mathcal{N}$ be a nested attribute such that no subattribute $Y \in \mathcal{S}(X')$ of an embedded attribute $X' \in emb(X)$ has the form $X'_I\{\lambda\}$ with $|I| \geq 2$. Then the set of axioms and rules in Theorems 3, 5 excluding the set axiom (10) is complete for the implication of FDs on $\mathcal{S}(X)$.*

Let us now investigate the question, whether the restriction on the attribute X in Theorem 9 can be dropped. Unfortunately, this is not the case, i.e. if both the union and the set constructor are present, more precisely, if the union constructor does appear immediately inside a set constructor, then there is no finite axiomatisation.

Theorem 10. *If $X \in \mathcal{N}$ is a nested attribute such that there exists a subattribute $X'_I\{\lambda\} \in \mathcal{S}(X')$ with $|I| \geq 2$ of an embedded attribute $X' \in emb(X)$, then there does not exist a finite, sound and complete system of axioms and rules for the implication of FDs on $\mathcal{S}(X)$.*

The proof will exploit a general result about closures under k-ary implication, which was proven in [3, Proposition 9.3.2]. We first define the necessary notions for this result.

Definition 12. Let $X \in \mathcal{N}$ be a nested attribute, Γ a class of dependencies on $\mathcal{S}(X)$, and $k \geq 0$.

A set $\Sigma \subseteq \Gamma$ of dependencies on $\mathcal{S}(X)$ is *closed under implication* with respect to Γ iff $\Sigma \models \varphi$ implies $\varphi \in \Sigma$ for all $\varphi \in \Gamma$.

Σ is *closed under k-ary implication* with respect to Γ iff for all $\varphi \in \Gamma$ whenever $\Sigma' \models \varphi$ holds for some $\Sigma' \subseteq \Sigma$ with $|\Sigma'| \leq k$, then this implies $\varphi \in \Sigma$.

Furthermore, we will exploit *ground derivation rules* that result from the derivation rules we used so far by instantiating the variables in the premise and the conclusion in such a way that the side conditions are satisfied.

Theorem 11. *Let $X \in \mathcal{N}$ be a nested attribute, Γ a class of dependencies on $\mathcal{S}(X)$, and let $k \geq 0$. Then there exists a k-ary ground axiomatisation for Γ iff each $\Sigma \subseteq \Gamma$ that is closed under k-ary implication is also closed under implication.*

The proof of Theorem 11 was given in [3, Proposition 9.3.2]. In fact, the proposition was formulated for the relational model, but the proof does not depend on that.

of Theorem 10. If for the class Γ of FDs on $\mathcal{S}(X)$ we had a finite axiomatisation, then there would exist some $k \geq 0$ such that Γ has a k-ary ground axiomatisation. According to Theorem 11 $\Sigma \subseteq \Gamma$ that is closed under k-ary implication would also be closed under implication. So take $X = X\{X_1(X_1') \oplus \cdots \oplus X_{k+2}(X_{k+2}')\}$ and the set

$$\Sigma_k = \{\{X_{\{1,\dots,k+1\}}\{\lambda\}, X_{\{i,k+2\}}\{\lambda\}\} \rightarrow \{X\} \mid i = 1, \dots, k+1\}.$$

By looking at instances that satisfy only k of these $k+1$ FDs we see that there is no k-ary implication of $\varphi_k = \{X_{\{1,\dots,k+1\}}\{\lambda\}\} \rightarrow \{X_{\{k+1\}}\{\lambda\}\}$ from Σ_k. So the k-ary closure of Σ_k will not contain φ_k.

On the other hand we obviously have $\Sigma_k \models \varphi_k$, so the closure of Σ_k will contain φ_k. That is, the k-ary closure of Σ_k is not closed under implication contradicting our assumption.

□

5 Extensions

In this section we extend the work on FDs and wFDs in several directions. First we will consider dependencies also on embedded attributes that were introduced in Section 3. We will see that this has very little impact on the theory, as we can show a completeness result also for these dependencies without extending the set of axioms and rules. Secondly, we will abandon the restriction on the trees to be finite and look at rational trees. Also this extension will not require additional rules.

5.1 Embedded Dependencies

The set $emb(X)$ of embedded attributes of a nested attribute X is simply characterised by: $X' \in emb(X)$ iff X' occurs somewhere within the nested structure of X. Embedded attributes are used in the proof of Theorem 2 in [28], but the theory of FDs and wFDs in the previous section did not make much further use of these attributes.

However, the lifting rules (19)-(23) implicitly contained FDs on embedded attributes. Nevertheless, we only looked at sets Σ of dependencies on $\mathcal{S}(X)$, so only trivial dependencies on embedded attributes played a role. We now make dependencies on embedded attributes explicit.

Definition 13. Let $X \in \mathcal{N}$. An *embedded functional dependency* (eFD) on $\mathcal{S}(X)$ is an expression $X' : \mathcal{Y} \to \mathcal{Z}$ with $X' \in emb(X)$ and $\mathcal{Y}, \mathcal{Z} \subseteq \mathcal{S}(X')$. An *embedded weak functional dependency* (ewFD) on $\mathcal{S}(X)$ is an expression $X' : \{\!|\mathcal{Y}_i \to \mathcal{Z}_i \mid i \in I|\!\}$ with $X' \in emb(X)$, an index set I and $\mathcal{Y}_i, \mathcal{Z}_i \subseteq \mathcal{S}(X')$.

In the following we consider again instances of X, i.e. finite sets $r = r(X) \subseteq dom(X)$. For each embedded attribute $X' \in emb(X)$, r induces an instance $r(X') \subseteq dom(X')$ in the obvious way: $v' \in r(X')$ iff there exists some $v \in r(X)$ such that v' occurs in v at the position indicated by X' in the nesting of X. Using this extension of instances, the satisfaction definition for eFDs and ewFDs is straightforward.

Definition 14. Let r be an instance of X. We say that r *satisfies the eFD* $X' : \mathcal{Y} \to \mathcal{Z}$ on $\mathcal{S}(X)$ (notation: $r \models X' : \mathcal{Y} \to \mathcal{Z}$) iff for all $t_1, t_2 \in r(X')$ with $\pi_Y^X(t_1) = \pi_Y^X(t_2)$ for all $Y \in \mathcal{Y}$ we also have $\pi_Z^X(t_1) = \pi_Z^X(t_2)$ for all $Z \in \mathcal{Z}$.

An instance $r \subseteq dom(X)$ *satisfies the ewFD* $X' : \{\!|\mathcal{Y}_i \to \mathcal{Z}_i \mid i \in I|\!\}$ on $\mathcal{S}(X)$ (notation: $r \models X' : \{\!|\mathcal{Y}_i \to \mathcal{Z}_i \mid i \in I|\!\}$) iff for all $t_1, t_2 \in r(X')$ there is some $i \in I$ with $\{t_1, t_2\} \models \mathcal{Y}_i \to \mathcal{Z}_i$.

According to this definition we may again identify an ewFD $X' : \{\!|\mathcal{Y} \to \mathcal{Z}|\!\}$, i.e. the index set contains exactly one element, with an "ordinary" eFD $X' : \mathcal{Y} \to \mathcal{Z}$.

If Σ is a set of eFDs or ewFDs on $\mathcal{S}(X)$, we write again $\Sigma \models \psi$, if the eFD or ewFD ψ is implied by Σ, and $\Sigma \vdash \psi$, if the eFD or ewFD ψ can be derived from Σ by means of some set $\mathfrak{R}$ of axioms and rules. In this way we retain the definition of Σ^* and Σ^+ for a set of eFDS or ewFDs on $\mathcal{S}(X)$.

We may further introduce another extension to FDs and wFDs by means of *contexts*. A *context* is a set of embedded attributes, i.e. $C \subseteq emb(X)$. A context C is *non-trivial* for $X' \in emb(X)$ iff no $X'' \in C$ is a subattribute of X' nor can it be rewritten as a record attribute with X' as one of its components.

Definition 15. Let $X \in \mathcal{N}$. A *contextual functional dependency* (cFD) on $\mathcal{S}(X)$ is an expression $C \mid X' : \mathcal{Y} \to \mathcal{Z}$ with $X' \in emb(X)$, a non-trivial context C, and $\mathcal{Y}, \mathcal{Z} \subseteq \mathcal{S}(X')$. A *contextual weak functional dependency* (cwFD) on $\mathcal{S}(X)$ is an expression $C \mid X' : \{\!|\mathcal{Y}_i \to \mathcal{Z}_i \mid i \in I|\!\}$ with $X' \in emb(X)$, a non-trivial context C, an index set I and $\mathcal{Y}_i, \mathcal{Z}_i \subseteq \mathcal{S}(X')$.

A context C partitions an instance $r(X)$ into disjoint instances using an equivalence relation $\sim_C$ defined as follows: $v_1 \sim_C v_2$ iff for each $Y \in C$ there exists some $v \in dom(Y)$ appearing in both v_1 and v_2 as the only value with this property. An equivalence class of $r(X)$ with respect to $\sim_C$ is called a *C-restricted fragment* of $r(X)$.

Definition 16. Let r be an instance of X. We say that r *satisfies the cFD* $C \mid X' : \mathcal{Y} \to \mathcal{Z}$ on $\mathcal{S}(X)$ (notation: $r \models C \mid X' : \mathcal{Y} \to \mathcal{Z}$) iff each C-restricted fragment of $r(X)$ satisfies the eFD $X' : \mathcal{Y} \to \mathcal{Z}$.

r satisfies the cwFD $C \mid X' : \{\!|\mathcal{Y}_i \to \mathcal{Z}_i \mid i \in I|\!\}$ on $\mathcal{S}(X)$ (notation: $r \models C \mid X' : \{\!|\mathcal{Y}_i \to \mathcal{Z}_i \mid i \in I|\!\}$) iff each C-restricted fragment of $r(X)$ satisfies the ewFD $X' : \{\!|\mathcal{Y}_i \to \mathcal{Z}_i \mid i \in I|\!\}$.

5.2 Extended Completeness Result

Let us first look at the derivation rules in Theorems 3, 5 and 6. In all these rules except the lifting rules (19)-(23) all dependencies are defined on $\mathcal{S}(X)$ with X left implicit, and the soundness proofs use arbitrary instances of X. In making X explicit, we turn the rules into derivation rules for eFDs and ewFDs. We can even turn them into derivation rules for cFDs and cwFDs by adding the prefix $C \mid X$ to all occurring dependencies. The soundness proof remains in all cases the same, because the notion of satisfaction defined in Definitions 14 and 16 only requires to consider only specific instances of X.

For the lifting rules the situation is similar; the only difference is that the dependencies in the rule conclusions are defined on some attribute X, while those in the premises are defined on some $X' \in emb(X)$. More precisely, we have $X = \{X'\}$, $X = X(X_1, \dots, X_n)$ with $X_i = X'$, $X = X_1(X_1') \oplus \cdots \oplus X_n(X_n')$ with $X_i' = X'$, $X = \langle X' \rangle$, and $X = [X']$, respectively. Nevertheless, making these attributes explicit and adding a context prefix defines derivation rules for eFDs, ewFDs, cFDs and cwFDs. These rules are obviously sound, as the soundness proof (Theorem 5) does not require any change except the mentioned syntactic modifications.

For cwFDs (and hence also for cFDs) we can add another derivation rule linking different contexts together. For this we define that context C' is *more restrictive* than context C (notation: $C' \trianglelefteq C$) iff $\sim_{C'} \subseteq \sim_C$ holds. Obviously, the empty context is the least restrictive one. Then we get the following derivation rule (*context rule*):

$$\frac{C \mid X' : \{\!\!\{ \mathcal{Y}_i \to \mathcal{Z}_i \mid i \in I \}\!\!\}}{C' \mid X' : \{\!\!\{ \mathcal{Y}_i \to \mathcal{Z}_i \mid i \in I \}\!\!\}} \; C' \trianglelefteq C \tag{34}$$

Theorem 12. *The rules in Theorems 3, 5 and 6 (with the syntactic modifications above) and rule (34) are sound for the implication of eFDs, ewFDs, cFDs and cwFDs.*

Proof. We only have to show the soundness of the context rule (34). So assume that $r = r(X)$ satisfies $C \mid X' : \{\!\!\{ \mathcal{Y}_i \to \mathcal{Z}_i \mid i \in I \}\!\!\}$. Take a C'-restricted fragment $r_{C'}(X)$ of $r(X)$ and $t_1, t_2 \in r_{C'}(X')$. As $\sim_{C'} \subseteq \sim_C$ holds, t_1 and t_2 must be in the same C-restricted fragment $r_C(X')$ of $r(X)$. Furthermore, there is some $i \in I$ such that $\{t_1, t_2\}$ satisfies the FD $\mathcal{Y}_i \to \mathcal{Z}_i$, hence $r \models C' \mid X' : \{\!\!\{ \mathcal{Y}_i \to \mathcal{Z}_i \mid i \in I \}\!\!\}$.

□

Now we can even extend the completeness result in Theorem 7 to cwFDs.

Theorem 13 (Completeness Theorem)**.** *The set of axioms and rules in Theorems 3, 5 and 6 (with the syntactic modifications above) and rule (34) is complete for the implication of cwFDs on $\mathcal{S}(X)$.*

Proof. Let Σ be a set of cwFDs on $\mathcal{S}(X)$ and assume $C \mid X' : \{\!\!\{ \mathcal{Y}_i \to \mathcal{Z}_i \mid i \in I \}\!\!\} \notin \Sigma^+$. Due to the context rule also $\emptyset \mid X' : \{\!\!\{ \mathcal{Y}_i \to \mathcal{Z}_i \mid i \in I \}\!\!\} \notin \Sigma^+$ holds, so we actually have to deal with an ewFD. Due to the union rule (7) we must have $X' : \{\!\!\{ \mathcal{Y}_i \to \{Z_i\} \mid i \in I \}\!\!\} \notin \Sigma^+$ for some selected $Z_i \in \mathcal{Z}_i$. Furthermore, due to the left union rule (25) we get $X' : \{\!\!\{\mathcal{Y} \to \{Z_i\} \mid i \in I\}\!\!\} \notin \Sigma^+$ with $\mathcal{Y} = \bigcup_{i \in I} \mathcal{Y}_i$. Due to the lifting rules we may assume that there is no cwFD $C' \mid X'' : \{\!\!\{\mathcal{Y}'_j \to \mathcal{Z}'_j \mid j \in J\}\!\!\} \in \Sigma$ with $X'' \in emb(X')$; otherwise we could use a restricted set of dependencies.

Let $\mathcal{Z} = \{Z \mid Z \geq Z_i \text{ for some } i \in I\}$ and $\mathcal{U} = \mathcal{S}(X') - \mathcal{Y} - \mathcal{Z}$. Due to the reflexivity axiom (1) we obviously have $Z_i \notin \mathcal{Y}$, and then $\mathcal{Y} \sqcap \mathcal{Z} = \emptyset$ due to the subattribute axiom (2). Due to the shift rule (26) there must exist some $\mathcal{U}' \subseteq \mathcal{U}$ with $\{\!\!\{\mathcal{Y} \cup \mathcal{U}' \to \{Z\} \mid Z \in \mathcal{Z} \cup (\mathcal{U} - \mathcal{U}')\}\!\!\} \notin \Sigma^+$. Otherwise we could derive $X' : \{\!\!\{\mathcal{Y} \to \{Z\} \mid Z \in \mathcal{Z}\}\!\!\}$, and thus $X' : \{\!\!\{\mathcal{Y} \to \{Z_i\} \mid i \in I\}\!\!\} \in \Sigma^+$ contradicting our assumption.

Let $\mathcal{U}'$ be maximal with the given property. Then using Lemma 2 we obtain that $\mathcal{F} = \mathcal{Y} \cup \mathcal{U}'$ is a coincidence ideal.

Without loss of generality we may assume $X' \neq X$; otherwise we are back to the case that was already handled in the proof of Theorem 7. Therefore, due to the restructuring rules we can assume that the outermost constructor in X' is not the union constructor. Then we can apply the Central Theorem 2, which gives us $r' = \{t'_1, t'_2\} \subseteq dom(X')$ with $\pi^X_Y(t'_1) = \pi^X_Y(t'_2)$ iff $Y \in \mathcal{F} = \mathcal{Y} \cup \mathcal{U}'$. In particular, $\pi^X_Y(t'_1) = \pi^X_Y(t'_2)$ for all $i \in I$ and $Y \in \mathcal{Y}_i$, and $\pi^X_{Z_i}(t'_1) \neq \pi^X_{Z_i}(t'_2)$ for all $i \in I$. That is, $r' \not\models \{\!\!\{\mathcal{Y}_i \to \{Z_i\} \mid i \in I\}\!\!\}$. From the soundness of the fragmentation rule (8) we conclude $r' \not\models \{\!\!\{\mathcal{Y}_i \to \mathcal{Z}_i \mid i \in I\}\!\!\}$.

We now "lift" r' to an instance r of X such that $r(X') = r'$ holds. For this take a chain $X_0, \dots, X_k$ of maximal length with $X_0 = X$, $X_k = X'$, and $X_i \in emb(X_{i-1}) - \{X_i\}$ for $i = 1, \dots, k$. Then for $i = k, \dots, 0$ define inductively $t_{i1}, t_{i2} \in dom(X_i)$ starting with $t_{kj} = t'_j$ for $j = 1, 2$.

- For $X_i = X_i(X'_1, \dots, X'_\ell)$ define $t_{ij} = (t^1_{ij}, \dots, t^\ell_{ij})$ with

$$t^h_{ij} = \begin{cases} t_{(i+1)j} & \text{if } X'_h = X_{i+1} \\ \tau^{X'_h}_\lambda & \text{else} \end{cases} .$$

- For $X_i = X_i\{X_{i+1}\}$ define $t_{ij} = \{t_{(i+1)j}\}$, if $X_i \notin C$ holds; otherwise take $t_{i1} = t_{i2} = \{t_{(i+1)1}, t_{(i+1)2}\}$.
- For $X_i = X_i\langle X_{i+1}\rangle$ define $t_{ij} = \langle t_{(i+1)j}\rangle$, if $X_i \notin C$ holds; otherwise take $t_{i1} = t_{i2} = \langle t_{(i+1)1}, t_{(i+1)2}\rangle$.
- For $X_i = X_i\{X_{i+1}\}$ define $t_{ij} = [t_{(i+1)j}]$, if $X_i \notin C$ holds; otherwise take $t_{i1} = t_{i2} = [t_{(i+1)1}, t_{(i+1)2}]$.

Finally, take $t_j = t_{0j}$ for $j = 1, 2$ and $r = \{t_1, t_2\}$. Then obviously $r(X') = r'$ holds. Consequently, $r \not\models X' : \{\!\!\{\mathcal{Y}_i \to \mathcal{Z}_i \mid i \in I\}\!\!\}$.

If C contains an attribute $\bar{X}$ with $X' \in emb(\bar{X})$, then r only contains one element, so it is its only C-restricted fragment. In this case we obviously get $r \not\models C \mid X' : \{\!\!\{\mathcal{Y}_i \to \mathcal{Z}_i \mid i \in I\}\!\!\}$. However, if C does not contain such an attribute, then due to our construction we also get that r equals its only C-restricted fragment, hence again $r \not\models C \mid X' : \{\!\!\{\mathcal{Y}_i \to \mathcal{Z}_i \mid i \in I\}\!\!\}$.

Now take $C' \mid X'' : \{\!\!\{\mathcal{V}_j \to \mathcal{W}_j \mid j \in J\}\!\!\} \in \Sigma$. If $X' \notin emb(X'')$ holds, then due to our construction $r(X'')$ will only contain one element, hence r trivially satisfies this cwFD. Thus we can assume $X' \in emb(X'')$. In this case we unnest the attributes in $\mathcal{V}_j$ and $\mathcal{W}_j$, until we obtain $\mathcal{V}'_j.\mathcal{W}'_j \subseteq \mathcal{S}(X')$. As in Lemma 3 we consider two cases:

1. If $\mathcal{V}'_j \not\subseteq \mathcal{Y} \cup \mathcal{U}'$ for some $j \in J$, we get $\pi^{X'}_{V'}(t'_1) \neq \pi^{X'}_{V'}(t'_2)$ for some $V' \in \mathcal{V}'_j$. We can assume that $r(X'')$ contains two elements, say $r(X'') = \{t''_1, t''_2\}$. Then we get $\pi^{X''}_{V}(t''_1) \neq \pi^{X''}_{V}(t''_2)$ for some $V \in \mathcal{V}_j$. Thus $r \models X'' : \{\!\!\{\mathcal{V}_j \to \mathcal{W}_j \mid j \in J\}\!\!\}$, hence also $r \models C' \mid X'' : \{\!\!\{\mathcal{V}_j \to \mathcal{W}_j \mid j \in J\}\!\!\}$.
2. If $\mathcal{V}'_j \subseteq \mathcal{Y} \cup \mathcal{U}'$ for all $j \in J$, then using the same arguments as in the proof of Lemma 3 we must have $\mathcal{W}'_j \subseteq \mathcal{Y} \cup \mathcal{U}'$ for some $j \in J$. By construction of r' we get $\pi^{X'}_{W'}(t'_1) = \pi^{X'}_{W'}(t'_2)$ for all $W' \in \mathcal{W}'_j$. Due to our construction oif r this implies $\pi^{X''}_{W}(t''_1) = \pi^{X''}_{W}(t''_2)$ for all $W \in \mathcal{W}_j$. Thus $r \models X'' : \mathcal{V}_j \to \mathcal{W}_j$. This implies $r \models C' \mid X'' : \{\!\!\{\mathcal{V}_j \to \mathcal{W}_j \mid j \in J\}\!\!\}$ due to the soundness of the weakening rule and the context rule.

So $r \models \Sigma$, hence also $r \models \Sigma^*$, from which we get $C \mid X' : \{\!\!\{\mathcal{Y}_i \to \mathcal{Z}_i \mid i \in I\}\!\!\} \notin \Sigma^*$. This completes the proof of the theorem.

□

Note that this completeness proof is only a slight modification of the proof of Theorem 7 exploiting the additional context rule, while the major arguments remain the same. Therefore, it is straightforward to apply these modifications also to the proof of Theorem 9, which leads to the following theorem on the completeness of cFDs (with the necessary syntactic modifications of the rules).

Theorem 14. *Let $X \in \mathcal{N}$ be a nested attribute such that no subattribute $Y \in \mathcal{S}(X')$ of an embedded attribute $X' \in emb(X)$ has the form $X'_I\{\lambda\}$ with $|I| \geq 2$. Then the set of axioms and rules in Theorems 3, 5 excluding the set axiom (10) together with the context rule (34) is complete for the implication of cFDs on $\mathcal{S}(X)$.*

5.3 Rational Trees

So far, all nested attributes had a fixed depth, and all complex values were representable as finite trees. In order to capture object oriented structures as in [30] and XML as in [1], we have to allow recursively defined attributes that take *rational trees* as their values, i.e. trees with only finitely many distinct subtrees. The notion of nested attributes has already been extended in this direction in [19]; we simply have to add $\mathcal{L} \subseteq \mathcal{N}$ to Definition 2 of nested attributes.

Definition 17. Let $\mathcal{U}$ be a universe and $\mathcal{L}$ a set of labels. The set $\mathcal{N}$ of *nested attributes* (over $\mathcal{U}$ and $\mathcal{L}$) is the smallest set with $\lambda \in \mathcal{N}$, $\mathcal{U} \subseteq \mathcal{N}$, $\mathcal{L} \subseteq \mathcal{N}$, and satisfying the following properties:

- for $X \in \mathcal{L}$ and $X'_1, \ldots, X'_n \in \mathcal{N}$ we have $X(X'_1, \ldots, X'_n) \in \mathcal{N}$;

- for $X \in \mathcal{L}$ and $X' \in \mathcal{N}$ we have $X\{X'\} \in \mathcal{N}$, $X[X'] \in \mathcal{N}$, and $X\langle X' \rangle \in \mathcal{N}$;
- for $X_1, \dots, X_n \in \mathcal{L}$ and $X'_1, \dots, X'_n \in \mathcal{N}$ we have $X_1(X'_1) \oplus \cdots \oplus X_n(X'_n) \in \mathcal{N}$.

We say that a label $Y \in \mathcal{L}$ occurring inside a nested attribute X, is a *defining label* iff it is introduced by one of the three cases in Definition 2. Otherwise it is a *referencing label*. We require that each label Y appears at most once as a defining label in a nested attribute X, and that each referencing label also occurs as a defining label. In other words, if we represent a nested attribute by a labelled tree, a defining label is the label of a non-leaf node, and a referencing label is the label of a leaf node.

We still have to extend Definition 3. For this assume $X \in \mathcal{N}$ and let Y be a referencing label in X. If we replace Y by the nested attribute that is defined by Y within X, we call the result an *expansion* of X. Note that in such an expansion a label may now appear more than once as a defining label, but all the nested attributes defined by a label can be identified, as the corresponding sets of expansions are identical.

In order to define domains assume set of *label variables* $\psi(Y)$ for each $Y \in \mathcal{L}$. Then for each expansion X' of a nested attribute X we define $dom(X')$ as in Definition 3 with the following modifications:

- for a referencing label Y we take $dom(Y) = \psi(Y)$;
- for a label Y defining the nested attribute Y' take $dom(Y) = \{y : v \mid y \in \psi(Y), v \in dom(Y')\}$;
- allow only such values v in $dom(X')$, for which the values of referencing labels also occur inside v exactly once at the position of a defining label.

Finally, define $dom(X) = \bigcup_{X'} dom(X')$, where the union spans over all expansions X' of X.

There is no need to change the definition of subattributes. We only have to be aware of the fact that now a nested attribute has several expansions, and they all can be used to define subattributes. Also the definitions of FDs and wFDs do not require more than the tiny addition that the sets of subattributes used in them must be finite (which they were automatically so far).

With these modifications we can easily repeat the whole theory of coincidence ideals and dependencies. The decisive property we exploit is the finiteness of a set Σ of wFDs. Then we can always find an expansion of X that is large enough such that the remaining referencing labels can actually be treated in the same way as simple attributes. In particular, the domain associated with these labels is infinite. This leads immediately to the following result.

Theorem 15. *The soundness and completeness theorems 3, 5, 6, 7 and 9 also hold for nested attributes X with the extensions from Definition 17.*

The same arguments also apply to embedded and contextual FDs and wFDs. We only have to be careful with the notation of embedded attributes in their definition, as these are no longer unique. Thus, instead of $X' \in emb(X)$ we consider *embedding paths* $X_0, \dots, X_k$ of maximal length with $X_0 = X$, $X_k = X'$ and $X_i \in emb(X_{i-1}) - \{X_i\}$ for $i = 1, \dots, k$. We also define $\mathcal{S}(X_0, \dots, X_k) = \mathcal{S}(X_k)$ as the associated set of subattributes.

Definition 18. Let $X \in \mathcal{N}$. An *embedded functional dependency* (eFD) on $\mathcal{S}(X)$ is an expression $P : \mathcal{Y} \to \mathcal{Z}$ with an embedding path P and $\mathcal{Y}, \mathcal{Z} \subseteq \mathcal{S}(P)$. An *embedded weak functional dependency* (ewFD) on $\mathcal{S}(X)$ is an expression $P : \{\!\{\mathcal{Y}_i \to \mathcal{Z}_i \mid i \in I\}\!\}$ with an embedding path P, an index set I and $\mathcal{Y}_i, \mathcal{Z}_i \subseteq \mathcal{S}(P)$.

This definition carries over naturally to contextual dependencies. Using the same argument as for wFDs we can also generalise the soundness and completeness results for contextual dependencies.

Theorem 16. *The soundness and completeness theorems 12, 13 and 14 also hold for nested attributes X with the extensions from Definition 17.*

6 Related Work

Apart from previous work by us and our colleagues Link and Hartmann that has been intensively used in this article there are two major related research groups working on dependencies on trees. Both Arenas and Libkin (see [5]) and Vincent, Liu and Liu (see [37]) place their work directly in the context of XML, while we take a more general approach using various constructors and rational trees. This implies that depending on the choice of incorporating order or not, these related approaches only handle one of the three bulk constructors, either lists or sets, while we take all three into account simultaneously. In fact, both Arenas and Libkin and Vincent et al. do not consider order, so the related case in our work refers to the use of the set constructor, apparently exactly the case, for which FDs cannot be finitely axiomatised. Furthermore, none of the other groups handles weak functional dependencies.

As emphasised in [37], but not proven, the different notions of XML FDs in the work by Arenas and Libkin and Vincent et al., respectively, coincide in case of complete information. Vincent, Liu and Liu claim that their notion of FDs actually captures incomplete information, while Arenas's and Libkin's work does. In our work, incomplete information is captured by the null attribute λ, so it boils down to the question, whether our definition of FDs can capture those defined by the other groups.

As emphasised in Section 5 the notion of FD from Definition 9 is bound to finite trees of fixed depth, while the work by the others deal with the variable depth of XML trees. So, without the extension to rational trees our notion of FDs cannot capture the other ones nor vice versa, because our definition of FDs involves complex subattributes, so equality is "generated" even on sets. However, taking cFDs on rational trees, it is not too difficult to see that the XFDs defined in [5] are actually representable in our framework. We may always restrict ourselves to XFDs $p_1 \dots p_k \rightarrow p$, i.e. the right hand side is a singleton. Then the right hand side defines an embedded attribute X', while the paths on the left hand side then give rise to either a subattribute of X' or the context subattributes. We illustrate this relation by a final example referring to the DTD in [5, Example 1.1] and the XFDs in [5, Example 4.1].

Example 2. The DTD in [5, Example 1.1] can be represented by the nested attribute

$$courses\{course(CNO, title(S), taken\_by\{student(SNO, name(S), grade(S))\})\}.$$

Then the following eFDs and cFDs represent the XFDs in [5, Example 4.1]:

$$\begin{aligned} &course : \{course(CNO)\} \rightarrow \\ &\qquad \{course(CNO, title(S), taken\_by\{student(SNO, name(S), grade(S))\})\} \\ &course \mid student : \{student(SNO)\} \rightarrow \{student(SNO, name(S), grade(S))\} \\ &student : \{student(SNO)\} \rightarrow \{student(name(S))\} \end{aligned}$$

7 Conclusions

In this article we completed our work on the axiomatisation of functional dependencies and weak functional dependencies on trees with restructuring. These trees arise from constructors for complex values comprising arbitrarily nesting of finite sets, multisets, lists, disjoint unions and records and a "null" attribute. Restructuring, i.e. non-trivial equivalence between these attributes are mainly due to the presence of the union constructor. While our previous work in [27] captured the case, where so called counter-attributes were excluded, we now were able to provide a sound and complete set of derivation rules for weak functional dependencies without this restriction. The price for this result was a very deep and very technical investigation of certain ideals in the algebra of subattributes leading to the central theorem on coincidence ideals, which gives an exact characterisation of sets of subattributes, on which two complex values coincide. We were further able to generalise the axiomatisation to capture dependencies on embedded attributes thereby including classes of FDs defined by others (see e.g. [5]).

Though our results require quite a heavy mathematical machinery, the technical characterisation of coincidence ideals in [28] to remove a seemingly not severe restriction in our previous results, we should emphasise that the unrestricted classes of FDs and wFDS treated in this article capture counting by means of subattributes. That is, whenever we have a multiset or list attribute, the projection of a complex value to a counter-attribute tells us how many values of a certain kind appear in this multiset or list. This is a concept that has not been handled in the context of functional dependencies before.

Unfortunately, for set attributes this is slightly different, as the counter-attributes in this case merely function as flags indicating, whether the subset of values of a certain kind is empty or not. This shows us that there is still more work needed to capture counting completely. In [29] we started work in this direction by deliberately adding more restructuring rules – so far, only intrinsic, unavoidable equivalences have been used. However, we may even take a list and forget the order of its elements, thus mapping it to a multiset, or map a multiset to its set of elements, i.e. we obtain an extension of the subattribute order by adding $X[Y] \geq X\langle Y\rangle \geq X\{Y\}$. Similarly, we could treat a set attribute as a multiset attribute, and then define FDs on it by using the subattributes of this corresponding multiset attribute.

The work in [29] only contains the first step in this direction, as only functional dependencies not involving the union constructor are handled. That is, the more interesting counter-attributes and the intrinsic restructuring rules are absent. The natural question is, how our results in this article can be generalised to deal also with these extensions to restructuring in general. Other open problem to be addressed in future are linked to other classes of dependencies, e.g. multi-valued and join dependencies as in [21] and [40] and to the existence of Armstrong instances (see e.g. [27]).

References

[1] S. Abiteboul, P. Buneman, and D. Suciu. *Data on the Web: From Relations to Semistructured Data and XML.* Morgan Kaufmann Publishers, 2000.

[2] S. Abiteboul and R. Hull. Restructuring hierarchical database objects. *Theoretical Computer Science*, 62(1-2):3–38, 1988.

[3] S. Abiteboul, R. Hull, and V. Vianu. *Foundations of Databases.* Addison-Wesley, 1995.

[4] M. Arenas and L. Libkin. A normal form for XML documents. In *PODS 2002.* ACM, 2002.

[5] M. Arenas and L. Libkin. A normal form for XML documents. *ACM Transactions on Database Systems*, 29(1):195–232, 2004.

[6] W. W. Armstrong. Dependency structures of database relationships. *Information Processing*, pages 580–583, 1974.

[7] C. Batini, S. Ceri, and S. B. Navathe. *Conceptual Database Design: An Entity-Relationship Approach.* Benjamin Cummings, 1992.

[8] P. Buneman, S. Davidson, W. Fan, C. Hara, and W. Tan. Keys for XML. In *Tenth WWW Conference.* IEEE, 2001.

[9] P. P. Chen. The Entity-Relationship model: Towards a unified view of data. *ACM Transactions Database Systems*, 1:9–36, 1976.

[10] P. P. Chen. English sentence structure and Entity-Relationship diagrams. *Information Science*, 29:127–149, 1983.

[11] J. Demetrovics and G. Gyepesi. On the functional dependency and some generalizations of it. *Acta Cybernetica*, 5:295–305, 1981.

[12] W. Fan and L. Libkin. On XML integrity constraints in the presence of DTDs. In *PODS 2001*. ACM, 2001.

[13] W. Fan and J. Siméon. Integrity constraints for XML. In *PODS 2000*. ACM, 2000.

[14] S. Hartmann. Decomposing relationship types by pivoting and schema equivalence. *Data & Knowledge Engineering*, 39:75–99, 2001.

[15] S. Hartmann and S. Link. On functional dependencies in advanced data models. *Electronic Notes in Theoretical Computer Science*, 84, 2003.

[16] S. Hartmann, S. Link, and K.-D. Schewe. Generalizing Boyce-Codd normal form to conceptual databases. In *Information Modelling and Knowledge Bases XV*, volume 105 of *Frontiers in Artificial Intelligence and Applications*, pages 88–105. IOS Press, 2004.

[17] S. Hartmann, S. Link, and K.-D. Schewe. Reasoning about functional and multi-valued dependencies in the presence of lists. In D. Seipel and J. M. Turull Torres, editors, *Foundations of Information and Knowledge Systems*, volume 2942 of *LNCS*, pages 134–154. Springer Verlag, 2004.

[18] S. Hartmann, S. Link, and K.-D. Schewe. Weak functional dependencies in higher-order datamodels. In D. Seipel and J. M. Turull Torres, editors, *Foundations of Information and Knowledge Systems*, volume 2942 of *LNCS*, pages 116–133. Springer Verlag, 2004.

[19] S. Hartmann, S. Link, and K.-D. Schewe. Functional dependencies over XML documents with DTDs. *Acta Cybernetica*, 17(1):153–171, 2005.

[20] S. Hartmann, S. Link, and K.-D. Schewe. Axiomatisation of functional dependencies in the presence of records, lists, sets and multisets. *Theoretical Computer Science*, 355:167–196, 2006.

[21] S. Hartmann, S. Link, and K.-D. Schewe. Functional and multi-valued dependencies in nested databases generated by record and list constructor. *Annals of Mathematics and Artificial Intelligence*, 46:111–164, 2006.

[22] R. Hull and R. King. Semantic database modeling: Survey, applications and research issues. *ACM Computing Surveys*, 19(3), 1987.

[23] W. Y. Mok, Y. K. Ng, and D. W. Embley. A normal form for precisely characterizing redundancy in nested relations. *ACM Transactions on Database Systems*, 21:77–106, 1996.

[24] Z. M. Özsoyoglu and L. Y. Yuan. A new normal form for nested relations. *ACM Transactions on Database Systems*, 12:111–136, 1987.

[25] J. Paredaens, P. De Bra, M. Gyssens, and D. Van Gucht. *The Structure of the Relational Database Model*. Springer-Verlag, 1989.

[26] A. Sali. Minimal keys in higher-order datamodels. In D. Seipel and J. M. Turull Torres, editors, *Foundations of Information and Knowledge Systems*, volume 2942 of *LNCS*, pages 242–251. Springer Verlag, 2004.

[27] A. Sali and K.-D. Schewe. Counter-free keys and functional dependencies in higher-order datamodels. *Fundamenta Informaticae*, 70(3):277–301, 2006.

[28] A. Sali and K.-D. Schewe. A characterisation of coincidence ideals for complex values. *Journal of Universal Computer Science*, 15(1):304–354, 2009.

[29] K.-D. Schewe. Functional dependencies with counting on trees. *Journal of Universal Computer Science*, 11(12):2063–2075, 2005.

[30] K.-D. Schewe and B. Thalheim. Fundamental concepts of object oriented databases. *Acta Cybernetica*, 11(4):49–85, 1993.

[31] Z. Tari, J. Stokes, and S. Spaccapietra. Object normal forms and dependency constraints for object-oriented schemata. *ACM Transactions on Database Systems*, 22:513–569, 1997.

[32] B. Thalheim. *Dependencies in Relational Databases*. Teubner-Verlag, 1991.

[33] B. Thalheim. Foundations of entity-relationship modeling. *Annals of Mathematics and Artificial Intelligence*, 6:197–256, 1992.

[34] B. Thalheim. *Entity-Relationship Modeling: Foundations of Database Technology*. Springer-Verlag, 2000.

[35] A. M. Tjoa and L. Berger. Transformation of requirement specifications expressed in natural language into an EER model. In *Entity-Relationship Approach*, volume 823 of *LNCS*. Springer-Verlag, 1993.

[36] M. Vincent. *The semantic justification for normal forms in relational database design*. PhD thesis, Monash University, Melbourne, Australia, 1994.

[37] M. Vincent, J. Liu, and C. Liu. Strong functional dependencies and their application to normal forms in XML. *ACM Transactions on Database Systems*, 29(3):445–462, 2004.

[38] M. W. Vincent and J. Liu. Functional dependencies for XML. In *Web Technologies and Applications: 5th Asia-Pacific Web Conference*, volume 2642 of *LNCS*, pages 22–34. Springer-Verlag, 2003.

[39] M. W. Vincent and J. Liu. Multivalued dependencies and a 4NF for XML. In *Advanced Information Systems Engineering: 15th International Conference CAiSE 2003*, volume 2681 of *LNCS*, pages 14–29. Springer-Verlag, 2003.

[40] M. W. Vincent and J. Liu. Multivalued dependencies in XML. In *British National Conference on Database Systems: BNCOD 2003*, volume 2712 of *LNCS*, pages 4–18. Springer-Verlag, 2003.

[41] J. Wang and R. Topor. Removing XML data redundancies using functional and equality-generating dependencies. In G. Dobbie and H. Williams, editors, *Database Technologies 2005 – Sixteenth Australasian Database Conference*, volume 39 of *CRPIT*, pages 65–74. Australian Computer Society, 2005.

Received 11th September 2009