ARTICLE IN PRESS
YINCO:4175

Information and Computation ●●● (●●●●) ●●●–●●●

Contents lists available at ScienceDirect

Information and Computation

www.elsevier.com/locate/yinco

A finite axiomatization of conditional independence and inclusion dependencies ☆

Miika Hannula [a,b,*], Juha Kontinen [a]

[a] University of Helsinki, Department of Mathematics and Statistics, P.O. Box 68, 00014 Helsinki, Finland
[b] The University of Auckland, Department of Computer Science, Private Bag 92019, Auckland 1142, New Zealand

ARTICLE INFO

Article history:
Received 16 May 2014
Received in revised form 15 September 2015
Available online xxxx

ABSTRACT

We present a complete finite axiomatization of the unrestricted implication problem for inclusion and conditional independence atoms in the context of dependence logic. For databases, this result implies a finite axiomatization of the unrestricted implication problem for inclusion, functional, and embedded multivalued dependencies in the unirelational case. We also indicate the generality of our approach by showing the analogous result for inclusion and embedded join dependencies.

1. Introduction

We formulate a finite axiomatization of the implication problem for inclusion and conditional independence atoms (dependencies) in the dependence logic context. The input of this problem is given by a finite set $\Sigma \cup \{\phi\}$ consisting of conditional independence atoms and inclusion atoms, and the question to decide is whether the following logical consequence holds

$$\Sigma \models \phi. \tag{1}$$

Independence logic [14] and inclusion logic [8] are recent variants of dependence logic the semantics of which are defined over sets of assignments (teams) rather than a single assignment as in first-order logic. By viewing a team X with domain $\{x_1, \dots, x_k\}$ as a relation of the schema $\{x_1, \dots, x_k\}$, our results provide a finite axiomatization for the unrestricted implication problem of inclusion (INDs), functional (FDs), and embedded multivalued database dependencies (EMVDs). To demonstrate the generality of our approach, we also present an analogous finite axiomatization of the implication problem for inclusion and embedded join dependencies.

Dependence logic [29] extends first-order logic by dependence atomic formulas

$$=\!(x_1, \dots, x_n) \tag{2}$$

the meaning of which is that the value of x_n is functionally determined by the values of $x_1, \dots, x_{n-1}$. Independence logic replaces the dependence atoms by independence atoms

$$\boldsymbol{y} \perp_{\boldsymbol{x}} \boldsymbol{z},$$

☆ This article is an extended version of [17].

* Corresponding author.
E-mail addresses: m.hannula@auckland.ac.nz (M. Hannula), juha.kontinen@helsinki.fi (J. Kontinen).

http://dx.doi.org/10.1016/j.ic.2016.04.001

the intuitive meaning of which is that, with respect to any fixed value of $\boldsymbol{x}$, the variables $\boldsymbol{y}$ are totally independent of the variables $\boldsymbol{z}$. Furthermore, inclusion logic is based on inclusion atoms of the form

$$\boldsymbol{x} \subseteq \boldsymbol{y},$$

with the meaning that all the values of $\boldsymbol{x}$ appear also as values for $\boldsymbol{y}$. In order to generalize our results to embedded join dependencies [28], we consider its analogue a so-called join atom in the team semantics setting:

$$*(\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n),$$

where $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$ are tuples of variables listing $V_1, \ldots, V_n$, which is satisfied by a team X if

$$X \upharpoonright \bigcup_{i=1}^{n} V_i = \bowtie_{i=1}^{n} \{X \upharpoonright V_i\},$$

that is, the projection of X to the variables $\bigcup_{i=1}^{n} V_i$ is equal to the natural join of X's projections to V_i.

By viewing a team X of assignments with domain $\{x_1, \ldots, x_k\}$ as a relation over the schema $\{x_1, \ldots, x_k\}$, the atoms $=(\boldsymbol{x})$ and $\boldsymbol{x} \subseteq \boldsymbol{y}$ correspond to functional and inclusion dependencies. On the other hand, the atoms $\boldsymbol{y} \perp_{\boldsymbol{x}} \boldsymbol{z}$ and $*(\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n)$ correspond to (embedded) multivalued and join dependencies. It is worth noting that, for $\boldsymbol{b}$ and $\boldsymbol{c}$ sharing no variables, $\boldsymbol{b} \perp_{\boldsymbol{a}} \boldsymbol{c}$ can be expressed as $*(\boldsymbol{ab}, \boldsymbol{ac})$. Furthermore, the atom $=(x_1, \ldots, x_n)$ can be alternatively expressed as

$$x_n \perp_{x_1 \ldots x_{n-1}} x_n,$$

hence our results for independence atoms cover also the case where dependence atoms are present.

The team semantics of dependence logic is a very flexible logical framework in which various notions of dependence and independence can be formalized. Dependence logic and its variants have turned out to be applicable in various areas. For example, Arrow's theorem from social choice theory has recently been axiomatized and formally proved in this framework [25]. Also, the pure independence atom $\boldsymbol{y} \perp \boldsymbol{z}$ and its axioms has various concrete interpretations such as independence $X \perp\!\!\!\perp Y$ between two sets of random variables [13], and independence in vector spaces and algebraically closed fields [26].

Dependence logic is equi-expressive with existential second-order logic (ESO). Furthermore, the set of valid formulas of dependence logic has the same complexity as that of full second-order logic, hence it is not possible to give a complete axiomatization of dependence logic [29]. However, by restricting attention to syntactic fragments [31,15,20] or by modifying the semantics [9] complete axiomatizations have recently been obtained. The axiomatization presented in this article is based on the classical characterization of logical implication between dependencies in terms of the *Chase* procedure [22]. The novelty in our approach is the use of the so-called *Lax* team semantics of independence logic to simulate the chase on the logical level using only inclusion and independence atoms and existential quantification.

In database theory, the implication problems for various types of database dependencies have been extensively studied starting from Armstrong's axiomatization for functional dependencies [2]. Inclusion dependencies were axiomatized in [5], and an axiomatization for pure independence atoms is also known (see [27,13,21]). On the other hand, the implication problem for embedded multivalued dependencies (also embedded join dependencies), and for inclusion dependencies and functional dependencies together, are known to be undecidable [18,19,6]. Although undecidability renders finite (Armstrong-type) axiomatizations impossible, certain strategies remain:

Strategy 1. One possibility is to consider inference rules for more general classes of dependencies. For instance, join dependencies and functional dependencies both belong to the class of *typed dependencies* (TDs) that enjoys chase-based complete axiomatizations [3,4]. In particular, TDs have a representation in algebraic terms that bears a striking similarity to the approach of this article [32,1]. A *project-join* (PJ) expression is an algebraic expression using projection, natural join and inclusion as its building blocks. For example, a PJ of the form $\pi(AB) \bowtie \pi(BC) \subseteq \pi(DEF)$ indicates that projection to DEF subsumes natural join of projections to AB and BC. Extending the scope of PJs to so-called *extended relations*, defined over *extended relation schemata* of the form $R_e = R \cup \{A_i : A \in R, i \in \mathbb{N}\}$, gives rise to *extended project-join* (EPJ) expressions that capture the class of typed dependencies. For instance, FDs can be expressed in this setting: a relation r satisfies an FD of the form $A \to B$ if and only if its extended relation r^*, obtained by extending each tuple $t \in r$ with $A_i \mapsto t(A)$ for $A \in R$ and $i \in \mathbb{N}$, satisfies $\pi(AB) \bowtie \pi(AB_1) \subseteq \pi(ABB_1)$. Analogously, any TD over a schema R can be represented as an EPJ over R_e which in turn can be interpreted as an existentially quantified conjunction of PJs and equality atoms. The use of implicit existential quantification introduced in this paper is somewhat analogous to that in EPJs; here only a team semantics version is being applied. Furthermore, the approach of this paper generalizes to an axiomatization of TDs where only inclusion and embedded join dependencies appear in the intermediate steps of deductions [16].

Strategy 2. Another possibility is to remain within the original language but instead add rules that introduce new attributes, implicitly bound by some quantification. This approach was taken by Mitchell in [23] where a finite axiomatization was presented for FDs and INDs, the key rule being so-called *Attribute Introduction* of the form:

$$\frac{U \subseteq V \quad V \to B}{UA \subseteq VB} \tag{3}$$

Note that the prerequisite for applying (3) is that A is not allowed to appear in assumptions or in any earlier proof step, i.e., A is to be thought of as implicitly existentially quantified. The axiomatization in [23] is sound and complete if no "existentially quantified" attributes are allowed to appear in the final stage of a deduction. A version of (3) was also used in another finite axiomatization of FDs and INDs where dependencies were given in a graph representation [7]. In this paper we follow Mitchell's approach and involve implicit existential quantification in some of the axioms. Indeed, our *Inclusion Introduction* is a non-deterministic version of (3). *Start Axiom* (see Definition 3) in contrast requires Lax team semantics interpretation of existential quantification where new attributes may be assigned to multiple values instead of a single value. As in Mitchell's axiomatization, we obtain a sound and complete proof procedure with the prerequisite that new attributes must be eventually discharged in deductions.

Organization. This paper is organized as follows. In Section 2 we define team semantics and introduce some basic properties of team-based logics. In Section 3 we present a deduction system for conditional independence and inclusion dependencies, proving its completeness in Section 4 and 5, respectively. In Section 6 we show an extension of the axiomatization to the class of embedded join dependencies and inclusion dependencies. Section 7 finishes the paper with conclusive remarks.

2. Preliminaries

In this section we define team semantics and introduce dependence, independence and inclusion atoms. The version of team semantics presented here is the Lax one, originally introduced in [8], which will turn out to be valuable for our purposes due to its interpretation of existential quantification.

2.1. Team semantics

The semantics is formulated using sets of assignments called teams instead of single assignments. Let $\mathcal{M}$ be a model[1] with domain M. An *assignment* s of $\mathcal{M}$ is a finite mapping from a set of variables into M. A *team* X over $\mathcal{M}$ with domain $\mathrm{Dom}(X) = V$ is a set of assignments from V to M. For a subset W of V, we write $X \upharpoonright W$ for the team obtained by restricting all the assignments of X to the variables in W.

If s is an assignment, x a variable, and $a \in A$, then $s[a/x]$ denotes the assignment (with domain $\mathrm{Dom}(s) \cup \{x\}$) that agrees with s everywhere except that it maps x to a. For an assignment s, and a tuple of variables $\boldsymbol{x} = (x_1, \dots, x_n)$, we sometimes denote the tuple $(s(x_1), ..., s(x_n))$ by $s(\boldsymbol{x})$. For a formula ϕ, $\mathrm{Var}(\phi)$ and $\mathrm{Fr}(\phi)$ denote the sets of variables that appear in ϕ and appear free in ϕ, respectively. For a finite set of formulas $\Sigma = \{\phi_1, \dots, \phi_n\}$, we write $\mathrm{Var}(\Sigma)$ for $\mathrm{Var}(\phi_1) \cup \ldots \cup \mathrm{Var}(\phi_n)$, and define $\mathrm{Fr}(\Sigma)$ analogously. For tuples of variables $\boldsymbol{x}$ and $\boldsymbol{y}$, the concatenation of the tuples is denoted by $\boldsymbol{xy}$, and when using set operations $\boldsymbol{x} \cup \boldsymbol{y}$ and $\boldsymbol{x} \setminus \boldsymbol{y}$, then these tuples are interpreted as the sets of elements of these tuples.

Team semantics is defined for first-order logic formulas as follows:

Definition 1 *(Team Semantics).* Let $\mathcal{M}$ be a model and let X be any team over it. Then:

- If ϕ is a first-order atomic or negated atomic formula, then $\mathcal{M} \models_X \phi$ if and only if for all $s \in X$, $\mathcal{M} \models_s \phi$ (in Tarski semantics).
- $\mathcal{M} \models_X \psi \vee \theta$ if and only if there are Y and Z such that $X = Y \cup Z$ and $\mathcal{M} \models_Y \psi$ and $\mathcal{M} \models_Z \theta$.
- $\mathcal{M} \models_X \psi \wedge \theta$ if and only if $\mathcal{M} \models_X \psi$ and $\mathcal{M} \models_X \theta$.
- $\mathcal{M} \models_X \exists v \psi$ if and only if there is a function $F : X \to \mathcal{P}(M) \setminus \{\emptyset\}$ such that $\mathcal{M} \models_{X[F/v]} \psi$, where $X[F/v] = \{s[m/v] : s \in X, m \in F(s)\}$.
- $\mathcal{M} \models_X \forall v \psi$ if and only if $\mathcal{M} \models_{X[M/v]} \psi$, where $X[M/v] = \{s[m/v] : s \in X, m \in M\}$.

The following lemma is an immediate consequence of Definition 1.

Lemma 1. *Let $\mathcal{M}$ be a model, X a team and $\exists x_1 \ldots \exists x_n \phi$ a formula in team semantics setting where $x_1, \dots, x_n$ is a sequence of variables. Then*

$$\mathcal{M} \models_X \exists x_1 \ldots \exists x_n \phi \text{ iff for some function } F : X \to \mathcal{P}(M^n) \setminus \{\emptyset\},\ \mathcal{M} \models_{X[F/x_1 \ldots x_n]} \phi$$

where $X[F/x_1 \ldots x_n] := \{s[a_1/x_1] \ldots [a_n/x_n] \mid (a_1, \dots, a_n) \in F(s)\}$.

If $\mathcal{M} \models_X \phi$, then we say that X *satisfies* ϕ in $\mathcal{M}$. If ϕ is a sentence (i.e. a formula with no free variables), then we say that ϕ is *true* in $\mathcal{M}$, and write $\mathcal{M} \models \phi$, if $\mathcal{M} \models_{\{\emptyset\}} \phi$ where $\{\emptyset\}$ is the team consisting of the empty assignment. Note that $\{\emptyset\}$ is different from the *empty team* $\emptyset$ containing no assignments.

[1] In this article we consider only formulae over the empty vocabulary.

In the team semantics setting, formula ψ is a *logical consequence* of ϕ, written $\phi \Rightarrow \psi$, if for all models $\mathcal{M}$ and teams X, with $\mathrm{Fr}(\phi) \cup \mathrm{Fr}(\psi) \subseteq \mathrm{Dom}(X)$,

$$\mathcal{M} \models_X \phi \Rightarrow \mathcal{M} \models_X \psi.$$

Formulas ϕ and ψ are said to be *logically equivalent* if $\phi \Rightarrow \psi$ and $\psi \Rightarrow \phi$. Logics $\mathcal{L}$ and $\mathcal{L}'$ are said to be equivalent, $\mathcal{L} = \mathcal{L}'$, if every $\mathcal{L}$-sentence ϕ is equivalent to some $\mathcal{L}'$-sentence ψ, and vice versa.

2.2. Dependencies in team semantics

Dependence, independence and inclusion atoms are given the following semantics.

Definition 2. Let $\boldsymbol{x}$ be a tuple of variables and y a variable. Then $=\!(\boldsymbol{x}, y)$ is a *dependence atom* with the semantic rule

– $\mathcal{M} \models_X =\!(\boldsymbol{x}, y)$ if and only if for any $s, s' \in X$ with $s(\boldsymbol{x}) = s'(\boldsymbol{x})$, $s(y) = s'(y)$.

Let $\boldsymbol{x}$, $\boldsymbol{y}$ and $\boldsymbol{z}$ be tuples of variables. Then $\boldsymbol{y} \perp_{\boldsymbol{x}} \boldsymbol{z}$ is a *conditional independence atom* with the semantic rule

– $\mathcal{M} \models_X \boldsymbol{y} \perp_{\boldsymbol{x}} \boldsymbol{z}$ if and only if for any $s, s' \in X$ with $s(\boldsymbol{x}) = s'(\boldsymbol{x})$ there is a $s'' \in X$ such that $s''(\boldsymbol{x}) = s(\boldsymbol{x})$, $s''(\boldsymbol{y}) = s(\boldsymbol{y})$ and $s''(\boldsymbol{z}) = s'(\boldsymbol{z})$.

Furthermore, we will write $\boldsymbol{x} \perp \boldsymbol{y}$ as a shorthand for $\boldsymbol{x} \perp_{\emptyset} \boldsymbol{y}$, and call it a *pure independence atom*.

Let $\boldsymbol{x}$ and $\boldsymbol{y}$ be two tuples of variables of the same length. Then $\boldsymbol{x} \subseteq \boldsymbol{y}$ is an *inclusion atom* with the semantic rule

– $\mathcal{M} \models_X \boldsymbol{x} \subseteq \boldsymbol{y}$ if and only if for any $s \in X$ there is a $s' \in X$ such that $s(\boldsymbol{x}) = s'(\boldsymbol{y})$.

Note that in the definition of an inclusion atom $\boldsymbol{x} \subseteq \boldsymbol{y}$, the tuples $\boldsymbol{x}$ and $\boldsymbol{y}$ may both have repetitions. Also in the definition of a conditional independence atom $\boldsymbol{y} \perp_{\boldsymbol{x}} \boldsymbol{z}$, the tuples $\boldsymbol{x}$, $\boldsymbol{y}$ and $\boldsymbol{z}$ are not necessarily pairwise disjoint. Thus any dependence atom $=\!(\boldsymbol{x}, y)$ can be expressed as a conditional independence atom $y \perp_{\boldsymbol{x}} y$. Also any conditional independence atom $\boldsymbol{y} \perp_{\boldsymbol{x}} \boldsymbol{z}$ can be expressed as a conjunction of dependence atoms and a conditional independence atom $\boldsymbol{y}^* \perp_{\boldsymbol{x}} \boldsymbol{z}^*$ where $\boldsymbol{x}$, $\boldsymbol{y}^*$ and $\boldsymbol{z}^*$ are pairwise disjoint. For disjoint tuples $\boldsymbol{x}$, $\boldsymbol{y}$ and $\boldsymbol{z}$, independence atom $\boldsymbol{y} \perp_{\boldsymbol{x}} \boldsymbol{z}$ corresponds to the embedded multivalued dependency $\boldsymbol{x} \twoheadrightarrow \boldsymbol{y}|\boldsymbol{z}$. Hence the class of conditional independence atoms corresponds to the class of functional dependencies and embedded multivalued dependencies in database theory.

Proposition 1. *(See [12].) Let $\boldsymbol{y} \perp_{\boldsymbol{x}} \boldsymbol{z}$ be a conditional independence atom where $\boldsymbol{x}$, $\boldsymbol{y}$ and $\boldsymbol{z}$ are tuples of variables. If $\boldsymbol{y}^*$ lists the variables in $\boldsymbol{y} - \boldsymbol{x} \cup \boldsymbol{z}$, $\boldsymbol{z}^*$ lists the variables in $\boldsymbol{z} - \boldsymbol{x} \cup \boldsymbol{y}$, and $\boldsymbol{u}$ lists the variables in $\boldsymbol{y} \cap \boldsymbol{z} - \boldsymbol{x}$, then*

$$\mathcal{M} \models_X \boldsymbol{y} \perp_{\boldsymbol{x}} \boldsymbol{z} \Leftrightarrow \mathcal{M} \models_X \boldsymbol{y}^* \perp_{\boldsymbol{x}} \boldsymbol{z}^* \wedge \bigwedge_{u \in \boldsymbol{u}} =\!(\boldsymbol{x}, u).$$

The extension of first-order logic by dependence atoms, conditional independence atoms and inclusion atoms is called *dependence logic* ($\mathrm{FO}(=\!(\ldots))$), *independence logic* ($\mathrm{FO}(\perp_c)$) and *inclusion logic* ($\mathrm{FO}(\subseteq)$), respectively. The fragment of independence logic containing only pure independence atoms is called *pure independence logic*, written $\mathrm{FO}(\perp)$. For a collection of atoms $\mathcal{C} \subseteq \{=\!(\ldots), \perp_c, \subseteq\}$, we will write $\mathrm{FO}(\mathcal{C})$ (omitting the set parenthesis of $\mathcal{C}$) for first-order logic with these atoms.

We end this section with a list of properties of which only the first two are applied in the sequel.

Proposition 2. *For $\mathcal{C} = \{=\!(\ldots), \perp_c, \subseteq\}$, the following hold.*

1. *(Empty Team Property) For all models $\mathcal{M}$ and formulas $\phi \in \mathrm{FO}(\mathcal{C})$*

$$\mathcal{M} \models_{\emptyset} \phi.$$

2. *(Locality [8]) If $\phi \in \mathrm{FO}(\mathcal{C})$ is such that $\mathrm{Fr}(\phi) \subseteq V$, then for all models $\mathcal{M}$ and teams X,*

$$\mathcal{M} \models_X \phi \Leftrightarrow \mathcal{M} \models_{X \upharpoonright V} \phi.$$

3. *[8] An inclusion atom $\boldsymbol{x} \subseteq \boldsymbol{y}$ is logically equivalent to the pure independence logic formula*

$$\forall v_1 v_2 \boldsymbol{z}((\boldsymbol{z} \neq \boldsymbol{x} \wedge \boldsymbol{z} \neq \boldsymbol{y}) \vee (v_1 \neq v_2 \wedge \boldsymbol{z} \neq \boldsymbol{y}) \vee ((v_1 = v_2 \vee \boldsymbol{z} = \boldsymbol{y}) \wedge \boldsymbol{z} \perp v_1 v_2))$$

where v_1, v_2 and $\boldsymbol{z}$ are new variables.

4. *[11] Any independence logic formula is logically equivalent to some pure independence logic formula.*
5. *[29,14] Any dependence (or independence) logic sentence ϕ is logically equivalent to some existential second-order sentence ϕ^*, and vice versa.*
6. *[10] Any inclusion logic sentence ϕ is logically equivalent to some positive greatest fixpoint logic sentence ϕ^*, and vice versa.*

3. Deduction system

In this section we present a sound and complete axiomatization for the implication problem of inclusion and independence atoms. The implication problem is given by a finite set $\Sigma \cup \{\phi\}$ consisting of conditional independence and inclusion atoms, and the question is to decide whether $\Sigma \models \phi$.

Definition 3. In addition to the usual introduction and elimination rules for conjunction, we adopt the following rules for conditional independence and inclusion atoms. Note that in Identity Rule and Start Axiom, the new variables should be thought of as implicitly existentially quantified.

1. Reflexivity:

   $$\boldsymbol{x} \subseteq \boldsymbol{x}.$$

2. Projection and Permutation:

   $$\text{if } x_1 \ldots x_n \subseteq y_1 \ldots y_n, \text{ then } x_{i_1} \ldots x_{i_k} \subseteq y_{i_1} \ldots y_{i_k},$$

   for each sequence $i_1, \ldots, i_k$ of integers from $\{1, \ldots, n\}$.
3. Transitivity:

   $$\text{if } \boldsymbol{x} \subseteq \boldsymbol{y} \wedge \boldsymbol{y} \subseteq \boldsymbol{z}, \text{ then } \boldsymbol{x} \subseteq \boldsymbol{y}.$$

4. Identity Rule:

   $$\text{if } ab \subseteq cc \wedge \phi, \text{ then } \phi',$$

   where ϕ' is obtained from ϕ by replacing any number of occurrences of a by b.
5. Inclusion Introduction:

   $$\text{if } \boldsymbol{a} \subseteq \boldsymbol{b}, \text{ then } \boldsymbol{a}x \subseteq \boldsymbol{b}c,$$

   where x is a *new* variable.
6. Start Axiom:

   $$\boldsymbol{ac} \subseteq \boldsymbol{ax} \wedge \boldsymbol{b} \perp_{\boldsymbol{a}} \boldsymbol{x} \wedge \boldsymbol{ax} \subseteq \boldsymbol{ac}$$

   where $\boldsymbol{x}$ is a sequence of pairwise distinct *new* variables.
7. Chase Rule:

   $$\text{if } \boldsymbol{y} \perp_{\boldsymbol{x}} \boldsymbol{z} \wedge \boldsymbol{ab} \subseteq \boldsymbol{xy} \wedge \boldsymbol{ac} \subseteq \boldsymbol{xz}, \text{ then } \boldsymbol{abc} \subseteq \boldsymbol{xyz}.$$

8. Final Rule:

   $$\text{if } \boldsymbol{ac} \subseteq \boldsymbol{ax} \wedge \boldsymbol{b} \perp_{\boldsymbol{a}} \boldsymbol{x} \wedge \boldsymbol{abx} \subseteq \boldsymbol{abc}, \text{ then } \boldsymbol{b} \perp_{\boldsymbol{a}} \boldsymbol{c}.$$

In an application of Inclusion Introduction, the variable x is called the new variable of the deduction step. Similarly, in an application of Start Axiom, the variables of $\boldsymbol{x}$ are called the new variables of the deduction step. A deduction from Σ is a sequence of formulas $(\phi_1, \ldots, \phi_n)$ such that:

1. Each ϕ_i is either an element of Σ, an instance of Reflexivity or Start Axiom, or follows from one or more formulas of $\Sigma \cup \{\phi_1, \ldots, \phi_{i-1}\}$ by one of the rules presented above.
2. If ϕ_i is an instance of Start Axiom (or follows from $\Sigma \cup \{\phi_1, \ldots, \phi_{i-1}\}$ by Inclusion Introduction), then the new variables of $\boldsymbol{x}$ (or the new variable x) must not appear in $\Sigma \cup \{\phi_1, \ldots, \phi_{i-1}\}$.

We say that ϕ is provable from Σ, written $\Sigma \vdash \phi$, if there is a deduction $(\phi_1, \ldots, \phi_n)$ from Σ with $\phi = \phi_n$ and such that no variables in ϕ are new in $\phi_1, \ldots, \phi_n$.

4. Soundness

First we prove the soundness of these axioms.

Lemma 2. *Let $(\phi_1, \ldots, \phi_n)$ be a deduction from Σ, and let $\boldsymbol{y}$ list all the new variables of the deduction steps. Let $\mathcal{M}$ and X be such that $\mathcal{M} \models_X \Sigma$ and $Var(\Sigma_n) \setminus \boldsymbol{y} \subseteq Dom(X)$ where $\Sigma_n := \Sigma \cup \{\phi_1, \ldots, \phi_n\}$. Then*

$$\mathcal{M} \models_X \exists \boldsymbol{y} \bigwedge \Sigma_n.$$

Proof. We show the claim by induction on n. So assume that the claim holds for any deduction of length n. We prove that the claim holds for deductions of length $n+1$ also. Let $(\phi_1, \ldots, \phi_{n+1})$ be a deduction from Σ, and let $\boldsymbol{y}$ and $\boldsymbol{z}$ list all the new variables of the deduction steps $\phi_1, \ldots, \phi_n$ and ϕ_{n+1}, respectively. Note that ϕ_{n+1} might not contain any new variables in which case $\boldsymbol{z}$ is empty. Assume that $\mathcal{M} \models_X \Sigma$ for some $\mathcal{M}$ and X, where $\mathrm{Var}(\Sigma_{n+1}) \setminus \boldsymbol{yz} \subseteq \mathrm{Dom}(X)$. By Proposition 2.2 we may assume that $\mathrm{Var}(\Sigma_{n+1}) \setminus \boldsymbol{yz} = \mathrm{Dom}(X)$. We need to show that

$$\mathcal{M} \models_X \exists \boldsymbol{y} \exists \boldsymbol{z} \bigwedge \Sigma_{n+1}.$$

By the induction assumption,

$$\mathcal{M} \models_X \exists \boldsymbol{y} \bigwedge \Sigma_n,$$

and hence by Lemma 1 there i s a function $F : X \to \mathcal{P}(M^{|\boldsymbol{y}|}) \setminus \{\emptyset\}$ such that

$$\mathcal{M} \models_{X'} \bigwedge \Sigma_n \tag{4}$$

where $X' := X[F/\boldsymbol{y}]$. It suffices to show that

$$\mathcal{M} \models_{X'} \exists \boldsymbol{z} \bigwedge \Sigma_{n+1}.$$

If ϕ_{n+1} is an instance of Start Axiom, or follows from Σ_n by Inclusion Introduction, then it suffices to find a $G : X' \to \mathcal{P}(M^{|\boldsymbol{z}|}) \setminus \{\emptyset\}$, such that $\mathcal{M} \models_{X'[G/\boldsymbol{z}]} \phi_{n+1}$ (note that in the first case this is due to Lemma 1). For this note that no variable of $\boldsymbol{z}$ is in $\mathrm{Var}(\Sigma_n)$, and hence by Proposition 2.2 $\mathcal{M} \models_{X'[G/\boldsymbol{z}]} \Sigma_n$ follows from (4). Otherwise, if $\boldsymbol{z}$ is empty, then it suffices to show that $\mathcal{M} \models_{X'} \phi_{n+1}$.

The cases where ϕ_{n+1} is an instance of Reflexivity, or follows from Σ_n by a conjunction rule, Projection and Permutation, Transitivity or Identity are straightforward. We prove the claim in the cases where one of the last four rules is applied.

- Inclusion Introduction: Then ϕ_{n+1} is of the form $\boldsymbol{a}x \subseteq \boldsymbol{b}c$ where $\boldsymbol{a} \subseteq \boldsymbol{b}$ is in Σ_n. Let $s \in X'$. Since $\mathcal{M} \models_{X'} \boldsymbol{a} \subseteq \boldsymbol{b}$ there is a $s' \in X'$ such that $s(\boldsymbol{a}) = s'(\boldsymbol{b})$. We let $G(s) = \{s'(c)\}$. Since $x \notin \mathrm{Dom}(X')$ we conclude that $\mathcal{M} \models_{X'[G/x]} \boldsymbol{a}x \subseteq \boldsymbol{b}c$.
- Start Axiom: Then ϕ_{n+1} is of the form $\boldsymbol{ac} \subseteq \boldsymbol{ax} \wedge \boldsymbol{b} \perp_{\boldsymbol{a}} \boldsymbol{x} \wedge \boldsymbol{ax} \subseteq \boldsymbol{ac}$. We define $G : X' \to \mathcal{P}(M^{|\boldsymbol{x}|}) \setminus \{\emptyset\}$ as follows:

$$G(s) = \{s'(\boldsymbol{c}) \mid s' \in X', s'(\boldsymbol{a}) = s(\boldsymbol{a})\}.$$

  Again, since $\boldsymbol{x}$ does not list any of the variables in $\mathrm{Dom}(X')$, it is straightforward to show that

$$\mathcal{M} \models_{X'[G/\boldsymbol{x}]} \boldsymbol{ac} \subseteq \boldsymbol{ax} \wedge \boldsymbol{b} \perp_{\boldsymbol{a}} \boldsymbol{x} \wedge \boldsymbol{ax} \subseteq \boldsymbol{ac}.$$

- Chase Rule: Then ϕ_{n+1} is of the form $\boldsymbol{abc} \subseteq \boldsymbol{xyz}$ where

$$\boldsymbol{y} \perp_{\boldsymbol{x}} \boldsymbol{z} \wedge \boldsymbol{ab} \subseteq \boldsymbol{xy} \wedge \boldsymbol{ac} \subseteq \boldsymbol{xz} \in \Sigma_n.$$

  Let $s \in X'$. Since $\mathcal{M} \models_{X'} \boldsymbol{ab} \subseteq \boldsymbol{xy} \wedge \boldsymbol{ac} \subseteq \boldsymbol{xz}$ there are $s', s'' \in X'$ such that $s'(\boldsymbol{xy}) = s(\boldsymbol{ab})$ and $s''(\boldsymbol{xz}) = s(\boldsymbol{ac})$. Since $s'(\boldsymbol{x}) = s''(\boldsymbol{x})$ and $\mathcal{M} \models_{X'} \boldsymbol{y} \perp_{\boldsymbol{x}} \boldsymbol{z}$, there is a $s_0 \in X'$ such that $s_0(\boldsymbol{xyz}) = s(\boldsymbol{abc})$ which shows the claim.
- Final Rule: Then ϕ_{n+1} is of the form $\boldsymbol{b} \perp_{\boldsymbol{a}} \boldsymbol{c}$ where

$$\boldsymbol{ac} \subseteq \boldsymbol{ax} \wedge \boldsymbol{b} \perp_{\boldsymbol{a}} \boldsymbol{x} \wedge \boldsymbol{abx} \subseteq \boldsymbol{abc} \in \Sigma_n.$$

  Let $s, s' \in X'$ be such that $s(\boldsymbol{a}) = s'(\boldsymbol{a})$. Since $\mathcal{M} \models_{X'} \boldsymbol{ac} \subseteq \boldsymbol{ax}$ there is a $s_0 \in X'$ such that $s'(\boldsymbol{ac}) = s_0(\boldsymbol{ax})$. Since $\mathcal{M} \models_{X'} \boldsymbol{b} \perp_{\boldsymbol{a}} \boldsymbol{x}$ and $s(\boldsymbol{a}) = s_0(\boldsymbol{a})$ there is a $s_1 \in X'$ such that $s_1(\boldsymbol{abx}) = s(\boldsymbol{ab})s_0(\boldsymbol{x})$. And since $\mathcal{M} \models_{X'} \boldsymbol{abx} \subseteq \boldsymbol{abc}$ there is a $s'' \in X'$ such that $s''(\boldsymbol{abc}) = s_1(\boldsymbol{abx})$. Then $s''(\boldsymbol{abc}) = s(\boldsymbol{ab})s'(\boldsymbol{c})$ which shows the claim and concludes the proof. □

This gives us the following soundness theorem.

Theorem 1. *Let $\Sigma \cup \{\phi\}$ be a finite set of conditional independence and inclusion atoms. Then $\Sigma \models \phi$ if $\Sigma \vdash \phi$.*

Proof. Assume that $\Sigma \vdash \phi$. Then there is a deduction $(\phi_1, \ldots, \phi_n)$ from Σ such that $\phi = \phi_n$ and no variables in ϕ are new in $\phi_1, \ldots, \phi_n$. Let $\mathcal{M}$ and X be such that $\mathrm{Var}(\Sigma \cup \{\phi\}) \subseteq \mathrm{Dom}(X)$ and $\mathcal{M} \models_X \Sigma$. We need to show that $\mathcal{M} \models_X \phi$. Let $\boldsymbol{y}$ list all the new variables in $\phi_1, \ldots, \phi_n$, and let $\boldsymbol{z}$ list all the variables in $\mathrm{Var}(\Sigma_n) \setminus \boldsymbol{y}$ which are not in $\mathrm{Dom}(X)$. We first let $X' := X[\boldsymbol{0}/z]$ for some dummy sequence $\boldsymbol{0}$ thus by Proposition 2.2, $\mathcal{M} \models_{X'} \Sigma$. Then by Lemma 2, $\mathcal{M} \models_{X'} \exists \boldsymbol{y} \bigwedge \Sigma_n$ implying there exists a $F : X' \to \mathcal{P}(M^{|\boldsymbol{y}|}) \setminus \{\emptyset\}$ such that $\mathcal{M} \models_{X''} \phi$, for $X'' := X'[F/\boldsymbol{y}]$. Since $X'' = X[\boldsymbol{0}/\boldsymbol{z}][F/\boldsymbol{y}]$ and no variables of $\boldsymbol{y}$ or $\boldsymbol{z}$ appear in ϕ, we conclude by Proposition 2.2 that $\mathcal{M} \models_X \phi$. □

5. Completeness

In this section we will prove that the set of axioms and rules presented in Definition 3 is complete with respect to the implication problem for conditional independence and inclusion atoms. For this purpose we introduce a graph characterization for the implication problem in Sect. 5.1. This characterization is based on the classical characterization of the implication problem for various database dependencies using the chase procedure [22]. The completeness proof is presented in Sect. 5.2. Also, in this section we will write $X \models \phi$ instead of $\mathcal{M} \models_X \phi$, since we will only deal with atoms, and the satisfaction of an atom depends only on the team X.

5.1. Graph characterization

We will consider graphs consisting of vertices and edges labeled by (possibly multiple) pairs of variables. The informal meaning is that a vertice will correspond to an assignment of a team, and an edge between s and s', labeled by uw, will express that $s(u) = s'(w)$. The graphical representation of the chase procedure is adapted from [24].

Definition 4. Let $G = (V, E)$ be a graph where E consists of directed labeled edges $(u, w)_{ab}$ where ab is a pair of variables, and for every pair (u, w) of vertices there can be several ab such that $(u, w)_{ab} \in E$. Then we say that u and w are ab-connected, written $u \sim_{ab} w$, if $u = w$ and $a = b$, or if there are vertices $v_0, \ldots, v_n$ and variables $x_0, \ldots, x_n$ such that

$$(u, v_0)_{ax_0}, (v_0, v_1)_{x_0x_1}, \ldots, (v_{n-1}, v_n)_{x_{n-1}x_n}, (v_n, w)_{x_nb} \in E^*$$

where $E^* := E \cup \{(w, u)_{ba} \mid (u, w)_{ab} \in E\}$.

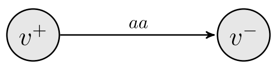
Next we define a graph $G_{\Sigma,\phi}$ in the style of Definition 4 for a set $\Sigma \cup \{\phi\}$ of conditional independence and inclusion atoms.

Definition 5. Let $\Sigma \cup \{\phi\}$ be a finite set of conditional independence and inclusion atoms. We let $G_{\Sigma,\phi} := (\bigcup_{n\in\mathbb{N}} V_n, \bigcup_{n\in\mathbb{N}} E_n)$ where $G_n = (V_n, E_n)$ is defined as follows:

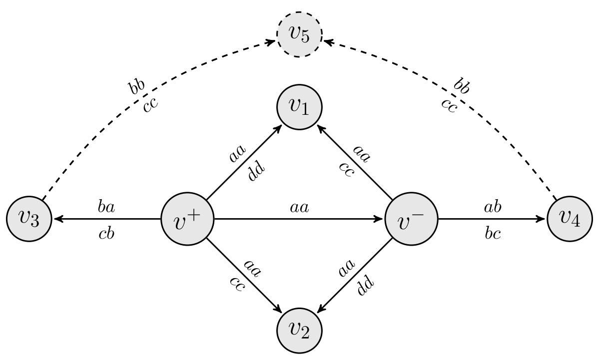
- If ϕ is $\boldsymbol{b} \perp_{\boldsymbol{a}} \boldsymbol{c}$, then $V_0 := \{v^+, v^-\}$ and $E_0 := \{(v^+, v^-)_{aa} \mid a \in \boldsymbol{a}\}$. If ϕ is $\boldsymbol{a} \subseteq \boldsymbol{b}$, then $V_0 := \{v\}$ and $E_0 := \emptyset$.
- Assume that G_n is defined. Then for every $v \in V_n$ and $x_1 \ldots x_k \subseteq y_1 \ldots y_k \in \Sigma$ we introduce a new vertex v_{new} and new edges $(v, v_{\text{new}})_{x_iy_i}$, for $1 \leq i \leq k$. Also for every $u, w \in V_n$, $u \neq w$, and $\boldsymbol{y} \perp_{\boldsymbol{x}} \boldsymbol{z} \in \Sigma$ where $u \sim_{xx} w$, for $x \in \boldsymbol{x}$, we introduce a new vertex v_{new} and new edges $(u, v_{\text{new}})_{yy}$, $(w, v_{\text{new}})_{zz}$, for $y \in \boldsymbol{xy}$ and $z \in \boldsymbol{xz}$. We let V_{n+1} and E_{n+1} be obtained by adding these new vertices and edges to the sets V_n and E_n.

Note that $G_{\Sigma,\phi} = G_0$ if $\Sigma = \emptyset$.

The construction of $G_{\Sigma,\phi}$ can be illustrated through an example. Suppose $\phi = b \perp_a c$ and $\Sigma = \{c \perp_a d, c \perp_b c, ab \subseteq bc\}$. Then, at level 0 of the construction of $G_{\Sigma,\phi}$, we have two nodes v^+ and v^- and an edge between them labeled by the pair aa.



At level 1, four new nodes $v_1, \ldots, v_4$ and the corresponding edges are introduced: v_1 and v_2 for $c \perp_a d$, and v_3 and v_4 for $ab \subseteq bc$. The dashed node v_5 is an example of a new node introduced at level 2, due to $c \perp_b c \in \Sigma$ and $v_3 \sim_{bb} v_4$.



We will next show in detail how $G_{\Sigma,\phi}$ yields a characterization of the implication problem $\Sigma \models \phi$.

Theorem 2. *Let $\Sigma \cup \{\phi\}$ be a finite set of conditional independence and inclusion atoms.*

1. *If ϕ is $a_1 \dots a_k \subseteq b_1 \dots b_k$, then $\Sigma \models \phi \Leftrightarrow \exists w \in V_{\Sigma,\phi}(v \sim_{a_i b_i} w$ for all $1 \le i \le k)$.*
2. *If ϕ is $\boldsymbol{b} \perp_{\boldsymbol{a}} \boldsymbol{c}$, then $\Sigma \models \phi \Leftrightarrow \exists v \in V_{\Sigma,\phi}(v^+ \sim_{bb} v$ and $v^- \sim_{cc} v$ for all $b \in \boldsymbol{ab}$ and $c \in \boldsymbol{ac})$.*

Proof. We deal with cases 1 and 2 simultaneously. First we will show the direction from right to left. So assume that the right-hand side of assumption holds. We show that $\Sigma \models \phi$. Let X be a team such that $X \models \Sigma$. We show that $X \models \phi$. For this, let $s, s' \in X$ be such that $s(\boldsymbol{a}) = s'(\boldsymbol{a})$. If ϕ is $\boldsymbol{b} \perp_{\boldsymbol{a}} \boldsymbol{c}$, then we need to find a s'' such that $s''(\boldsymbol{abc}) = s(\boldsymbol{ab})s'(\boldsymbol{c})$. If ϕ is $a_1 \dots a_k \subseteq b_1 \dots b_k$, then we need to find a s'' such that $s(a_1 \dots a_k) = s''(b_1 \dots b_k)$. We will now define inductively, for each natural number n, a function $f_n : V_n \to X$ such that $f_n(u)(x) = f_n(w)(y)$ if $(u, w)_{xy} \in E_n$. This will suffice for the claim as we will later show.

- Assume that $n = 0$.
  1. If ϕ is $a_1 \dots a_k \subseteq b_1 \dots b_k$, then $V_0 = \{v\}$ and $E_0 = \emptyset$, and we let $f_0(v) := s$.
  2. If ϕ is $\boldsymbol{b} \perp_{\boldsymbol{a}} \boldsymbol{c}$, then $V_0 = \{v^+, v^-\}$ and $E_0 = \{(v^+, v^-)_{aa} \mid a \in \boldsymbol{a}\}$. We let $f_0(v^+) := s$ and $f_0(v^-) := s'$. Then $f(v^+)(a) = f(v^-)(a)$, for $a \in \boldsymbol{a}$, as wanted.
- Assume that $n = m + 1$, and that f_m is defined so that $f_m(u)(x) = f_m(w)(y)$ if $(u, w)_{xy} \in E_m$. We let $f_{m+1}(u) = f_m(u)$, for $u \in V_m$. Assume that $v_{\text{new}} \in V_{m+1} \setminus V_m$ and that there are $u \in V_m$ and $x_1 \dots x_l \subseteq y_1 \dots y_l \in \Sigma$ such that $(u, v_{\text{new}})_{x_i y_i} \in E_{m+1} \setminus E_m$, for $1 \le i \le l$. Since $X \models x_1 \dots x_l \subseteq y_1 \dots y_l$, there is a $s_0 \in X$ such that $f_{m+1}(u)(x_i) = s_0(y_i)$, for $1 \le i \le l$. Define $f_{m+1}(v_{\text{new}}) := s_0$, hence $f_{m+1}(u)(x_i) = f_{m+1}(v_{\text{new}})(y_i)$, for $1 \le i \le l$, as wanted.
  Assume then that $v_{\text{new}} \in V_{m+1} \setminus V_m$ and that there are $u, w \in V_m$, $u \ne w$, and $\boldsymbol{y} \perp_{\boldsymbol{x}} \boldsymbol{z} \in \Sigma$ such that $(u, v_{\text{new}})_{yy}$, $(w, v_{\text{new}})_{zz} \in E_{m+1} \setminus E_m$, for $y \in \boldsymbol{xy}$ and $z \in \boldsymbol{xz}$. Then $u \sim_{xx} w$ in G_m, for $x \in \boldsymbol{x}$. This means that there are vertices $v_0, \dots, v_n$ and variables $x_0, \dots, x_n$, for $x \in \boldsymbol{x}$, such that
  $$(u, v_0)_{xx_0}, (v_0, v_1)_{x_0 x_1}, \dots, (v_{n-1}, v_n)_{x_{n-1} x_n}, (v_n, w)_{x_n x} \in E_m^*,$$
  where $E_m^* := E_m \cup \{(w, u)_{ba} \mid (u, w)_{ab} \in E_m\}$. By the induction assumption then
  $$f_m(u)(x) = f_m(v_0)(x_0) = \dots = f_m(v_n)(x_n) = f_m(w)(x).$$
  Hence, since $X \models \boldsymbol{y} \perp_{\boldsymbol{x}} \boldsymbol{z}$, there is a s_0 such that $s_0(\boldsymbol{xyz}) = f_m(u)(\boldsymbol{xy}) f_m(w)(\boldsymbol{z})$. We let $f_{m+1}(v_{\text{new}}) := s_0$ and conclude that $f_{m+1}(u)(y) = f_{m+1}(v_{\text{new}})(y)$ and $f_{m+1}(w)(z) = f_{m+1}(v_{\text{new}})(z)$, for $y \in \boldsymbol{xy}$ and $z \in \boldsymbol{xz}$. This concludes the construction.

Now, in case 2 there is a $v \in V_{\Sigma,\phi}$ such that $v^+ \sim_{bb} v$ and $v^- \sim_{cc} v$ for all $b \in \boldsymbol{ab}$ and $c \in \boldsymbol{ac}$. Let n be such that each path witnessing this is in G_n. We want to show that choosing s'' as $f_n(v)$, $s''(\boldsymbol{abc}) = s(\boldsymbol{ab})s'(\boldsymbol{c})$. Recall that $s = f_n(v^+)$ and $s' = f_n(v^-)$. First, let $b \in \boldsymbol{ab}$. The case where $v = v^+$ is trivial, so assume that $v \ne v^+$ in which case there are vertices $v_0, \dots, v_n$ and variables $x_0, \dots, x_n$ such that

$$(v^+, v_0)_{bx_0}, (v_0, v_1)_{x_0 x_1}, \dots, (v_{n-1}, v_n)_{x_{n-1} x_n}, (v_n, v)_{x_n b} \in E_n^*.$$

Therefore by the construction, $f_n(v^+)(b) = f_n(v)(b)$. Analogously $f_n(v^-)(c) = f_n(v)(c)$, for $c \in \boldsymbol{c}$, which concludes this case.

In case 1, s'' is found analogously. This concludes the proof of the direction from right to left.

For the other direction, assume that the right-hand side assumption fails in $G_{\Sigma,\phi}$. Again, we deal with both cases simultaneously. We will now construct a team X such that $X \models \Sigma$ and $X \not\models \phi$. We let $X := \{s_u \mid u \in V_{\Sigma,\phi}\}$ where each $s_u : \mathrm{Var}(\Sigma \cup \{\phi\}) \to \mathcal{P}(V_{\Sigma,\phi})^{|\mathrm{Var}(\Sigma \cup \{\phi\})|}$ is defined as follows:

$$s_u(x) := \prod_{y \in \mathrm{Var}(\Sigma \cup \{\phi\})} \{w \in V_{\Sigma,\phi} \mid u \sim_{xy} w\}.$$

We claim that $s_u(x) = s_w(y) \Leftrightarrow u \sim_{xy} w$. Indeed, assume that $u \sim_{xy} w$. If now v is in the set with the index z of the product $s_u(x)$, then $u \sim_{xz} v$. Since $w \sim_{yx} u$, we have that $w \sim_{yz} v$. Thus v is in the set with the index z of the product $s_w(y)$. Hence by symmetry we conclude that $s_u(x) = s_w(y)$. For the other direction assume that $s_u(x) = s_w(y)$. Then consider the set with the index y of the product $s_w(y)$. Since $w \sim_{yy} w$ by the definition, the vertex w is in this set, and thus by the assumption it is in the set with the index y of the product $s_u(x)$. It follows by the definition that $u \sim_{xy} w$ which shows the claim.

Next we will show that $X \models \Sigma$. So assume that $\boldsymbol{y} \perp_{\boldsymbol{x}} \boldsymbol{z} \in \Sigma$ and that $s_u, s_w \in X$ are such that $s_u(\boldsymbol{x}) = s_w(\boldsymbol{x})$. We need to find a $s_v \in X$ such that $s_v(\boldsymbol{xyz}) = s_u(\boldsymbol{xy}) s_w(\boldsymbol{z})$. Since $u \sim_{xx} w$, for $x \in \boldsymbol{x}$, there is a $v \in G_{\Sigma,\phi}$ such that $(u, v)_{yy}, (w, v)_{zz} \in E_{\Sigma,\phi}$, for $y \in \boldsymbol{xy}$ and $z \in \boldsymbol{xz}$. Then $s_u(\boldsymbol{xy}) = s_v(\boldsymbol{xy})$ and $s_w(\boldsymbol{xz}) = s_v(\boldsymbol{xz})$, as wanted. In case $x_1 \dots x_l \subseteq y_1 \dots y_l \in \Sigma$, $X \models x_1 \dots x_l \subseteq y_1 \dots y_l$ is shown analogously.

It suffices to show that $X \not\models \phi$. Assume first that ϕ is $\boldsymbol{b} \perp_{\boldsymbol{a}} \boldsymbol{c}$. Then $s_{v^+}(\boldsymbol{a}) = s_{v^-}(\boldsymbol{a})$, but by the assumption there is no $v \in V_{\Sigma,\phi}$ such that $v^+ \sim_{bb} v$ and $v^- \sim_{cc} v$ for all $b \in \boldsymbol{ab}$ and $c \in \boldsymbol{ac}$. Hence there is no $s_v \in X$ such that $s_v(\boldsymbol{ab}) = s_{v^+}(\boldsymbol{ab})$ and $s_v(\boldsymbol{ac}) = s_{v^-}(\boldsymbol{ac})$, and therefore $X \not\models \boldsymbol{b} \perp_{\boldsymbol{a}} \boldsymbol{c}$. In case ϕ is $a_1 \ldots a_k \subseteq b_1 \ldots b_k$, $X \not\models \phi$ is shown analogously. □

Let us now see how to use this theorem with our concrete example (see the paragraph after Definition 5). First we notice that v_5 witnesses $v^+ \sim_{bb} v^-$. Also $v^+ \sim_{aa} v^-$ since $(v^+, v^-)_{aa} \in E_{\Sigma,\phi}$, and $v^- \sim_{xx} v^-$ for any x by the definition. Therefore, choosing v as v^-, we obtain $\Sigma \models b \perp_a c$ by the previous theorem.

5.2. Completeness proof

We are now ready to prove the completeness. Let us first define some notation needed in the proof. We will write $x = y$ for syntactical identity, $x \equiv y$ for an atom of the form $xy \subseteq zz$ implying the identity of x and y, and $\boldsymbol{x} \equiv \boldsymbol{y}$ for a conjunction the form $\bigwedge_{i \leq |\boldsymbol{x}|} \mathrm{pr}_i(\boldsymbol{x}) \equiv \mathrm{pr}_i(\boldsymbol{y})$. Let $\boldsymbol{x}$ be a sequence listing $\mathrm{Var}(\Sigma \cup \{\phi\})$. If $\boldsymbol{x}_v$ is a vector of length $|\boldsymbol{x}|$ (representing vertex v of the graph $G_{\Sigma,\phi}$), and $\boldsymbol{a} = (x_{i_1}, \ldots, x_{i_l})$ is a sequence of variables from $\boldsymbol{x}$, then we write $\boldsymbol{a}_v$ for

$$(\mathrm{pr}_{i_1}(\boldsymbol{x}_v), \ldots, \mathrm{pr}_{i_l}(\boldsymbol{x}_v)).$$

Also, for a deduction d from Σ, we write $\Sigma \vdash^d \psi$ if ψ appears as a proof step in d. Note that then new variables of the proof steps are allowed to appear in ψ.

We will next prove the completeness by using the following lemma (which will be proved later). Recall that (V_n, E_n) refers to the nth level of the construction of $G_{\Sigma,\phi}$.

Lemma 3. *Let n be a natural number, $\Sigma \cup \{\phi\}$ a finite set of conditional independence and inclusion atoms, and $\boldsymbol{x}$ a sequence listing $\mathrm{Var}(\Sigma \cup \{\phi\})$. Then there is a deduction $d = (\phi_1, \ldots, \phi_N)$ from Σ such that for each $u \in V_n$, there is a sequence $\boldsymbol{x}_u$ of length $|\boldsymbol{x}|$ (and possibly with repetitions) such that $\Sigma \vdash^d \boldsymbol{x}_u \subseteq \boldsymbol{x}$, and for each $(u, w)_{x_i x_j} \in E_n^*$, $\Sigma \vdash^d \mathrm{pr}_i(\boldsymbol{x}_u) \equiv \mathrm{pr}_j(\boldsymbol{x}_w)$. Moreover,*

- *if ϕ is of the form $\boldsymbol{a} \subseteq \boldsymbol{b}$, then $\phi_1 = \boldsymbol{x}_v \subseteq \boldsymbol{x}$ (obtained by Reflexivity), for $\boldsymbol{x}_v$ defined as $\boldsymbol{x}$,*
- *if ϕ is of the form $\boldsymbol{b} \perp_{\boldsymbol{a}} \boldsymbol{c}$, then $\phi_1 = \boldsymbol{ac} \subseteq \boldsymbol{ac}^* \wedge \boldsymbol{b} \perp_{\boldsymbol{a}} \boldsymbol{c}^* \wedge \boldsymbol{ac}^* \subseteq \boldsymbol{ac}$ (obtained by Start Axiom), for $\boldsymbol{a}_{v^+}\boldsymbol{b}_{v^+}\boldsymbol{c}_{v^-} = \boldsymbol{abc}^*$.*

Theorem 3. *Let $\Sigma \cup \{\phi\}$ be a finite set of conditional independence and inclusion atoms. Then $\Sigma \vdash \phi$ if $\Sigma \models \phi$.*

Proof. Let Σ and ϕ be such that $\Sigma \models \phi$. We will show that $\Sigma \vdash \phi$.

We have two cases: either

1. ϕ is $x_{i_1} \ldots x_{i_m} \subseteq x_{j_1} \ldots x_{j_m}$ and, by Theorem 2, there is a $w \in V_{\Sigma,\phi}$ such that $v \sim_{x_{i_k} x_{j_k}} w$ for all $1 \leq k \leq m$, or
2. ϕ is $\boldsymbol{b} \perp_{\boldsymbol{a}} \boldsymbol{c}$ and, by Theorem 2, there is a $v \in V_{\Sigma,\phi}$ such that $v^+ \sim_{x_i x_i} v$ and $v^- \sim_{x_j x_j} v$ for all $x_i \in \boldsymbol{ab}$ and $x_j \in \boldsymbol{ac}$.

Assume now first that ϕ is $\boldsymbol{a} \subseteq \boldsymbol{b}$ where $\boldsymbol{a} := x_{i_1} \ldots x_{i_m}$ and $\boldsymbol{b} := x_{j_1} \ldots x_{j_m}$. Then there is a $w \in V_{\Sigma,\phi}$ such that $v \sim_{x_{i_k} x_{j_k}} w$, for $1 \leq k \leq m$. Let n be such that all the witnessing paths are in G_n, and let $d = (\phi_1, \ldots, \phi_N)$ be a deduction from Σ obtained by Lemma 3, for $\Sigma \cup \{\phi\}$, n and $\boldsymbol{x}$ listing $\mathrm{Var}(\Sigma \cup \{\phi\})$. For $\Sigma \vdash \phi$, it now suffices to show that $\Sigma \cup \{\phi_1, \ldots, \phi_N\} \vdash \phi$ since, by Lemma 3, the variables that appear in ϕ appear already in ϕ_1 (as not new) and therefore cannot appear as new in any step of $(\phi_1, \ldots, \phi_N)$.

Let first $1 \leq k \leq m$. We show that from $\Sigma \cup \{\phi_1, \ldots, \phi_N\}$ we may derive

$$\mathrm{pr}_{i_k}(\boldsymbol{x}_v) \equiv \mathrm{pr}_{j_k}(\boldsymbol{x}_w). \tag{5}$$

If $w = v$ and $i_k = j_k$, then (5) is obtained by Reflexivity. If $w \neq v$ or $i_k \neq j_k$, then there are vertices $v_0, \ldots, v_p \in V_n$ and variables $x_{l_0}, \ldots, x_{l_p}$ such that

$$(v, v_0)_{x_{i_k} x_{l_0}}, (v_0, v_1)_{x_{l_0} x_{l_1}}, \ldots, (v_{p-1}, v_p)_{x_{l_{p-1}} x_{l_p}}, (v_p, w)_{x_{l_p} x_{j_k}} \in E_n^*.$$

Then by Lemma 3,

$$\Sigma \vdash^d \mathrm{pr}_{i_k}(\boldsymbol{x}_v) \equiv \mathrm{pr}_{l_0}(\boldsymbol{x}_{v_0}) \wedge \ldots \wedge \mathrm{pr}_{l_p}(\boldsymbol{x}_{v_p}) \equiv \mathrm{pr}_{j_k}(\boldsymbol{x}_w) \tag{6}$$

from which we obtain $\mathrm{pr}_{i_k}(\boldsymbol{x}_v) \equiv \mathrm{pr}_{j_k}(\boldsymbol{x}_w)$ by Identity Rule. Hence, we may now derive

$$\boldsymbol{a}_v \equiv \boldsymbol{b}_w. \tag{7}$$

Since $\Sigma \vdash^d \boldsymbol{x}_w \subseteq \boldsymbol{x}$ by Lemma 3, then by Permutation and Projection we obtain

$$\boldsymbol{b}_w \subseteq \boldsymbol{b}. \tag{8}$$

Note that by Lemma 3, $\boldsymbol{x}_v = \boldsymbol{x}$ hence $\boldsymbol{a}_v = \boldsymbol{a}$. Thus we obtain $\boldsymbol{a} \subseteq \boldsymbol{b}$ from (7) and (8) using repeatedly Identity Rule. Since none of the steps above introduce any new variables, we get $\Sigma \cup \{\phi_1, \ldots, \phi_N\} \vdash \phi$ which concludes case 1.

Assume next that ϕ is $\boldsymbol{b} \perp_{\boldsymbol{a}} \boldsymbol{c}$. Then there is a $v \in V_{\Sigma,\phi}$ such that $v^+ \sim_{x_i x_i} v$ and $v^- \sim_{x_j x_j} v$ for all $x_i \in \boldsymbol{ab}$ and $x_j \in \boldsymbol{ac}$. Analogously to the previous case, by Lemma 3, we obtain a deduction $d = (\phi_1, \ldots, \phi_N)$ from Σ for which

$$\Sigma \vdash^d \boldsymbol{x}_v \subseteq \boldsymbol{x} \tag{9}$$

and

$$\Sigma \vdash^d \boldsymbol{a}_v \boldsymbol{b}_v \equiv \boldsymbol{a}_{v^+} \boldsymbol{b}_{v^+} \wedge \boldsymbol{a}_v \boldsymbol{c}_v \equiv \boldsymbol{a}_{v^-} \boldsymbol{c}_{v^-}. \tag{10}$$

Again, for $\Sigma \vdash \phi$, it suffices to show that $\Sigma \cup \{\phi_1, \ldots, \phi_N\} \vdash \phi$. By Projection and Permutation we first deduce

$$\boldsymbol{a}_v \boldsymbol{b}_v \boldsymbol{c}_v \subseteq \boldsymbol{abc} \tag{11}$$

from (9), and using repeatedly Projection and Permutation and Identity Rule we get

$$\boldsymbol{a}_{v^+} \boldsymbol{b}_{v^+} \boldsymbol{c}_{v^-} \subseteq \boldsymbol{abc} \tag{12}$$

from (10) and (11). Note that by Lemma 3, $\boldsymbol{a}_{v^+} \boldsymbol{b}_{v^+} \boldsymbol{c}_{v^-} = \boldsymbol{abc}^*$ and $\Sigma \vdash^d \boldsymbol{ac} \subseteq \boldsymbol{ac}^* \wedge \boldsymbol{b} \perp_{\boldsymbol{a}} \boldsymbol{c}^*$. Therefore we can derive $\boldsymbol{b} \perp_{\boldsymbol{a}} \boldsymbol{c}$ with one application of Final Rule. Since none of the steps above introduce any new variables, we have $\Sigma \cup \{\phi_1, \ldots, \phi_N\} \vdash \phi$ which concludes case 2 and the proof. □

We are left to prove Lemma 3.

Proof of Lemma 3. Let n be a natural number, $\Sigma \cup \{\phi\}$ a finite set of conditional independence and inclusion atoms, and $\boldsymbol{x}$ a sequence listing $\mathrm{Var}(\Sigma \cup \{\phi\})$. We show the claim by induction on n. Note that at each step n it suffices to consider only edges $(u, w)_{x_i x_j} \in E_n$, since for $(w, u)_{x_j x_i} \in E_n^*$, $\mathrm{pr}_j(\boldsymbol{x}_w) \equiv \mathrm{pr}_i(\boldsymbol{x}_u)$ can be deduced from $\mathrm{pr}_i(\boldsymbol{x}_u) \equiv \mathrm{pr}_j(\boldsymbol{x}_w)$ (using Reflexivity for $\mathrm{pr}_i(\boldsymbol{x}_u)\mathrm{pr}_i(\boldsymbol{x}_u)$ and then Identity Rule).

- Assume that $n = 0$. We show in two cases how to construct a deduction d from Σ such that it meets the requirements of Lemma 3.
  1. Assume that ϕ is $\boldsymbol{a} \subseteq \boldsymbol{b}$. Recall that $V_0 = \{v\}$ and $E_0 = \emptyset$. Then defining $\boldsymbol{x}_v := \boldsymbol{x}$ we derive $\boldsymbol{x}_v \subseteq \boldsymbol{x}$ as a first step by Reflexivity.
  2. Assume that ϕ is $\boldsymbol{b} \perp_{\boldsymbol{a}} \boldsymbol{c}$. Recall that $V_0 = \{v^+, v^-\}$ and $E_0 = \{(v^+, v^-)_{x_i x_i} \mid x_i \in \boldsymbol{a}\}$. As a first step we use Start Axiom to obtain

     $$\boldsymbol{ac} \subseteq \boldsymbol{ac}^* \wedge \boldsymbol{b} \perp_{\boldsymbol{a}} \boldsymbol{c}^* \wedge \boldsymbol{ac}^* \subseteq \boldsymbol{ac} \tag{13}$$

     where $\boldsymbol{c}^*$ is a sequence of pairwise distinct new variables. Then using Inclusion Introduction and Projection and Permutation we may deduce

     $$\boldsymbol{ab}^* \boldsymbol{c}^* \boldsymbol{d}^* \subseteq \boldsymbol{abcd} \tag{14}$$

     from $\boldsymbol{ac}^* \subseteq \boldsymbol{ac}$ where $\boldsymbol{d}$ lists $\boldsymbol{x} \setminus \boldsymbol{abc}$ and $\boldsymbol{b}^* \boldsymbol{c}^* \boldsymbol{d}^*$ is a sequence of pairwise distinct new variables. By Projection and Permutation and Identity Rule we may assume that $\boldsymbol{ab}^* \boldsymbol{c}^* \boldsymbol{d}^*$ has repetitions exactly where $\boldsymbol{abcd}$ has. Therefore we can list the variables of $\boldsymbol{ab}^* \boldsymbol{c}^* \boldsymbol{d}^*$ in a sequence $\boldsymbol{x}_{v^-}$ of length $|\boldsymbol{x}|$ where

     $$\boldsymbol{ab}^* \boldsymbol{c}^* \boldsymbol{d}^* = (\mathrm{pr}_{i_1}(\boldsymbol{x}_{v^-}), \ldots, \mathrm{pr}_{i_l}(\boldsymbol{x}_{v^-})),$$

     for $\boldsymbol{abcd} = (x_{i_1}, \ldots, x_{i_l})$. Then $\boldsymbol{a}_{v^-} \boldsymbol{b}_{v^-} \boldsymbol{c}_{v^-} \boldsymbol{d}_{v^-} = \boldsymbol{ab}^* \boldsymbol{c}^* \boldsymbol{d}^*$, and we can derive $\boldsymbol{x}_{v^-} \subseteq \boldsymbol{x}$ from (14) by Projection and Permutation. We also let $\boldsymbol{x}_{v^+} := \boldsymbol{x}$ in which case $\boldsymbol{x}_{v^+} \subseteq \boldsymbol{x}$ is derivable by Reflexivity and $\boldsymbol{a}_{v^+} \boldsymbol{b}_{v^+} \boldsymbol{c}_{v^-} = \boldsymbol{abc}^*$. Moreover, $\boldsymbol{a}_{v^+} \equiv \boldsymbol{a}_{v^-}$ is derivable by Reflexivity because $\boldsymbol{a}_{v^+} = \boldsymbol{a}_{v^-}$. This concludes the case $n = 0$.
- Assume that $n = m + 1$. Then by the induction assumption, there is a deduction d such that for each $u \in V_m$ there is a sequence $\boldsymbol{x}_u$ such that $\Sigma \vdash^d \boldsymbol{x}_u \subseteq \boldsymbol{x}$, and for each $(u, w)_{x_i x_j} \in E_m$ also $\Sigma \vdash^d \mathrm{pr}_i(\boldsymbol{x}_u) \equiv \mathrm{pr}_j(\boldsymbol{x}_w)$. Assume that $v_{\mathrm{new}} \in V_{m+1} \setminus V_m$ is such that there are $u \in V_m$ and $x_{i_1} \ldots x_{i_l} \subseteq x_{j_i} \ldots x_{j_l} \in \Sigma$ for which we have added new edges $(u, v_{\mathrm{new}})_{x_{i_k} x_{j_k}}$ to V_{m+1}, for $1 \le k \le l$. We will introduce a sequence $\boldsymbol{x}_{v_{\mathrm{new}}}$ and show how to extend d to a deduction d^* such that $\Sigma \vdash^{d^*} \boldsymbol{x}_{v_{\mathrm{new}}} \subseteq \boldsymbol{x}$ and $\Sigma \vdash^{d^*} \mathrm{pr}_{i_k}(\boldsymbol{x}_u) \equiv \mathrm{pr}_{j_k}(\boldsymbol{x}_{v_{\mathrm{new}}})$, for $1 \le k \le l$.
  By Projection and Permutation we deduce first

  $$\mathrm{pr}_{i_1}(\boldsymbol{x}_u) \ldots \mathrm{pr}_{i_l}(\boldsymbol{x}_u) \subseteq x_{i_1} \ldots x_{i_l} \tag{15}$$

  from $\boldsymbol{x}_u \subseteq \boldsymbol{x}$. Then we obtain

  $$\mathrm{pr}_{i_1}(\boldsymbol{x}_u) \ldots \mathrm{pr}_{i_l}(\boldsymbol{x}_u) \subseteq x_{j_i} \ldots x_{j_l} \tag{16}$$

  from (15) and the assumption $x_{i_1} \ldots x_{i_l} \subseteq x_{j_i} \ldots x_{j_l}$ by Transitivity.

Then by Reflexivity we may deduce $\mathrm{pr}_{i_1}(\boldsymbol{x}_u) \subseteq \mathrm{pr}_{i_1}(\boldsymbol{x}_u)$ from which we derive by Inclusion Introduction

$$\mathrm{pr}_{i_1}(\boldsymbol{x}_u)y_1 \subseteq \mathrm{pr}_{i_1}(\boldsymbol{x}_u)\mathrm{pr}_{i_1}(\boldsymbol{x}_u) \tag{17}$$

where y_1 is a new variable. Then from (16) and (17) we derive by Identity Rule

$$y_1\mathrm{pr}_{i_2}(\boldsymbol{x}_u)\ldots\mathrm{pr}_{i_l}(\boldsymbol{x}_u) \subseteq x_{j_1}\ldots x_{j_l}. \tag{18}$$

Iterating this procedure l times leads us to a formula

$$\bigwedge_{1\le k\le l} \mathrm{pr}_{i_k}(\boldsymbol{x}_u) \equiv y_k \wedge y_1\ldots y_l \subseteq x_{j_1}\ldots x_{j_l} \tag{19}$$

where $y_1,\ldots,y_l$ are pairwise distinct new variables. Let $x_{j_{l+1}},\ldots,x_{j_{l'}}$ list $\boldsymbol{x}\setminus\{x_{j_1},\ldots,x_{j_l}\}$. Repeating Inclusion Introduction for the inclusion atom in (19) gives us a formula

$$y_1\ldots y_{l'} \subseteq x_{j_1}\ldots x_{j_{l'}} \tag{20}$$

where $y_{l+1},\ldots,y_{l'}$ are pairwise distinct new variables. Let $\boldsymbol{y}$ now denote the sequence $y_1\ldots y_{l'}$. Then

$$\bigwedge_{1\le k\le l} \mathrm{pr}_{i_k}(\boldsymbol{x}_u) \equiv \mathrm{pr}_k(\boldsymbol{y}) \wedge \boldsymbol{y} \subseteq x_{j_1}\ldots x_{j_{l'}} \tag{21}$$

is the formula obtained from (19) by replacing its inclusion atom with (20). By Projection and Permutation and Identity Rule we may assume that $\mathrm{pr}_k(\boldsymbol{y}) = \mathrm{pr}_{k'}(\boldsymbol{y})$ if and only if $j_k = j_{k'}$, for $1 \le k \le l'$. Analogously to the case $n = 0$, we can then order the variables of $\boldsymbol{y}$ as a sequence $\boldsymbol{x}_{v_{\mathrm{new}}}$ of length $|\boldsymbol{x}|$ such that $\mathrm{pr}_{j_k}(\boldsymbol{x}_{v_{\mathrm{new}}}) = \mathrm{pr}_k(\boldsymbol{y})$, for $1 \le k \le l'$. Then

$$\bigwedge_{1\le k\le l} \mathrm{pr}_{i_k}(\boldsymbol{x}_u) \equiv \mathrm{pr}_{j_k}(\boldsymbol{x}_{v_{\mathrm{new}}}) \wedge \mathrm{pr}_{j_1}(\boldsymbol{x}_{v_{\mathrm{new}}})\ldots\mathrm{pr}_{j_{l'}}(\boldsymbol{x}_{v_{\mathrm{new}}}) \subseteq x_{j_1}\ldots x_{j_{l'}} \tag{22}$$

is the formula (21). By Projection and Permutation we can now deduce $\boldsymbol{x}_{v_{\mathrm{new}}} \subseteq \boldsymbol{x}$ from the inclusion atom in (22). Hence $\boldsymbol{x}_{v_{\mathrm{new}}}$ is such that $\boldsymbol{x}_{v_{\mathrm{new}}} \subseteq \boldsymbol{x}$ and $\mathrm{pr}_{i_k}(\boldsymbol{x}_u) \equiv \mathrm{pr}_{j_k}(\boldsymbol{x}_{v_{\mathrm{new}}})$ can be derived, for $1 \le k \le l$. This concludes the case for inclusion. Assume then that $v_{\mathrm{new}} \in V_{m+1} \setminus V_m$ is such that there are $u, w \in V_m$, $u \ne w$, and $\boldsymbol{q} \perp_{\boldsymbol{p}} \boldsymbol{r} \in \Sigma$ for which we have added new edges $(u, v_{\mathrm{new}})_{x_i x_i}, (w, v_{\mathrm{new}})_{x_j x_j}$ to V_{m+1}, for $x_i \in \boldsymbol{pq}$ and $x_j \in \boldsymbol{pr}$. We will introduce a sequence $\boldsymbol{x}_{v_{\mathrm{new}}}$ and show how to extend d to a deduction d^* such that $\Sigma \vdash^{d^*} \boldsymbol{x}_{v_{\mathrm{new}}} \subseteq \boldsymbol{x}$, and $\Sigma \vdash^{d^*} \mathrm{pr}_i(\boldsymbol{x}_u) \equiv \mathrm{pr}_i(x_{v_{\mathrm{new}}})$ and $\Sigma \vdash^{d^*} \mathrm{pr}_j(\boldsymbol{x}_w) \equiv \mathrm{pr}_j(x_{v_{\mathrm{new}}})$, for $x_i \in \boldsymbol{pq}$ and $x_j \in \boldsymbol{pr}$. The latter means that

$$\Sigma \vdash^{d^*} \boldsymbol{p}_u\boldsymbol{q}_u \equiv \boldsymbol{p}_{v_{\mathrm{new}}}\boldsymbol{q}_{v_{\mathrm{new}}} \wedge \boldsymbol{p}_w\boldsymbol{r}_w \equiv \boldsymbol{p}_{v_{\mathrm{new}}}\boldsymbol{r}_{v_{\mathrm{new}}}.$$

First of all, we know that $u \sim_{x_k x_k} w$ in G_m for all $x_k \in \boldsymbol{p}$. Thus there are vertices $v_0,\ldots,v_n \in V_m$ and variables $x_{i_0},\ldots,x_{i_n}$ such that

$$(u, v_0)_{x_k x_{i_0}}, (v_0, v_1)_{x_{i_0} x_{i_1}}, \ldots, (v_{n-1}, v_n)_{x_{i_{n-1}} x_{i_n}}, (v_n, w)_{x_{i_n} x_k} \in E_m^*.$$

Hence by the induction assumption and Identity Rule, there are $\boldsymbol{x}_u$ and $\boldsymbol{x}_w$ such that $\Sigma \vdash^d \boldsymbol{x}_u \subseteq \boldsymbol{x}$ and $\Sigma \vdash^d \boldsymbol{x}_w \subseteq \boldsymbol{x}$, and $\Sigma \vdash^d \mathrm{pr}_k(\boldsymbol{x}_u) \equiv \mathrm{pr}_k(\boldsymbol{x}_w)$, for $x_k \in \boldsymbol{p}$. In other words,

$$\Sigma \vdash^d \boldsymbol{p}_u \equiv \boldsymbol{p}_w. \tag{23}$$

By Projection and Permutation we first derive

$$\boldsymbol{p}_u\boldsymbol{q}_u \subseteq \boldsymbol{pq} \tag{24}$$

and

$$\boldsymbol{p}_w\boldsymbol{r}_w \subseteq \boldsymbol{pr} \tag{25}$$

from $\boldsymbol{x}_u \subseteq \boldsymbol{x}$ and $\boldsymbol{x}_w \subseteq \boldsymbol{x}$, respectively. Then we derive

$$\boldsymbol{p}_u\boldsymbol{r}_w \subseteq \boldsymbol{pr} \tag{26}$$

from $\boldsymbol{p}_u \equiv \boldsymbol{p}_w$ and (25) by Identity Rule. By Chase Rule we then derive

$$\boldsymbol{p}_u\boldsymbol{q}_u\boldsymbol{r}_w \subseteq \boldsymbol{pqr} \tag{27}$$

from the assumption $\boldsymbol{q} \perp_{\boldsymbol{p}} \boldsymbol{r}$, (24) and (26). Now it can be the case that $x_i \in \boldsymbol{pq}$ and $x_i \in \boldsymbol{r}$, but $\mathrm{pr}_i(\boldsymbol{x}_u) \ne \mathrm{pr}_i(\boldsymbol{x}_w)$. Then we can derive

$$\mathrm{pr}_i(\boldsymbol{x}_u)\mathrm{pr}_i(\boldsymbol{x}_w) \subseteq x_i x_i \tag{28}$$

from (27) by Projection and Permutation, and

$$\boldsymbol{p}_u\boldsymbol{q}_u\boldsymbol{r}_w(\mathrm{pr}_i(\boldsymbol{x}_u)/\mathrm{pr}_i(\boldsymbol{x}_w)) \subseteq \boldsymbol{pqr} \tag{29}$$

from (28) and (27) by Identity Rule. Let now $\boldsymbol{r}^*$ be obtained from $\boldsymbol{r}_w$ by replacing, for each $x_i \in \boldsymbol{pq} \cap \boldsymbol{r}$, the variable $\mathrm{pr}_i(\boldsymbol{x}_w)$ with $\mathrm{pr}_i(\boldsymbol{x}_u)$. Iterating the previous derivation gives us then

$$\boldsymbol{r}^* \equiv \boldsymbol{r}_w \wedge \boldsymbol{p}_u\boldsymbol{q}_u\boldsymbol{r}^* \subseteq \boldsymbol{pqr}. \tag{30}$$

Let $\boldsymbol{s}$ list the variables in $\boldsymbol{x} \setminus \boldsymbol{pqr}$. From the inclusion atom in (30) we derive by Inclusion Introduction

$$\boldsymbol{p}_u\boldsymbol{q}_u\boldsymbol{r}^*\boldsymbol{s}^* \subseteq \boldsymbol{pqrs} \tag{31}$$

where $\boldsymbol{s}^*$ is a sequence of pairwise distinct new variables. Then $\boldsymbol{p}_u\boldsymbol{q}_u\boldsymbol{r}^*\boldsymbol{s}^*$ has repetitions at least where $\boldsymbol{pqrs}$ has, and hence we can define $\boldsymbol{x}_{v_{\text{new}}}$ as the sequence of length $|\boldsymbol{x}|$ where

$$\boldsymbol{p}_u\boldsymbol{q}_u\boldsymbol{r}^*\boldsymbol{s}^* = (\mathrm{pr}_{i_1}(\boldsymbol{x}_{v_{\text{new}}}), \dots, \mathrm{pr}_{i_l}(\boldsymbol{x}_{v_{\text{new}}})), \tag{32}$$

for $\boldsymbol{pqrs} = (x_{i_1}, \dots, x_{i_l})$. Then $\boldsymbol{p}_{v_{\text{new}}}\boldsymbol{q}_{v_{\text{new}}}\boldsymbol{r}_{v_{\text{new}}}\boldsymbol{s}_{v_{\text{new}}} = \boldsymbol{p}_u\boldsymbol{q}_u\boldsymbol{r}^*\boldsymbol{s}^*$, and we can thus derive

$$\boldsymbol{x}_{v_{\text{new}}} \subseteq \boldsymbol{x} \tag{33}$$

from (31) by Projection and Permutation. Moreover,

$$\boldsymbol{p}_{v_{\text{new}}}\boldsymbol{q}_{v_{\text{new}}} \equiv \boldsymbol{p}_u\boldsymbol{q}_u \tag{34}$$

can be derived by Reflexivity, and

$$\boldsymbol{p}_{v_{\text{new}}}\boldsymbol{r}_{v_{\text{new}}} \equiv \boldsymbol{p}_w\boldsymbol{r}_w \tag{35}$$

is derivable since (35) is the conjunction of $\boldsymbol{p}_u \equiv \boldsymbol{p}_w$ in (23) and $\boldsymbol{r}^* \equiv \boldsymbol{r}_w$ in (30). Hence, for $\boldsymbol{x}_{v_{\text{new}}}$ we can derive

$$\boldsymbol{x}_{v_{\text{new}}} \subseteq \boldsymbol{x} \wedge \boldsymbol{p}_{v_{\text{new}}}\boldsymbol{q}_{v_{\text{new}}} \equiv \boldsymbol{p}_u\boldsymbol{q}_u \wedge \boldsymbol{p}_{v_{\text{new}}}\boldsymbol{r}_{v_{\text{new}}} \equiv \boldsymbol{p}_w\boldsymbol{r}_w$$

which concludes the case $n = m + 1$ and the proof. $\square$

By Theorem 1 and Theorem 3 we now have the following.

Corollary 1. *Let $\Sigma \cup \{\phi\}$ be a finite set of conditional independence and inclusion atoms. Then $\Sigma \vdash \phi$ if and only if $\Sigma \models \phi$.*

The following example shows how to deduce $b \perp_a c \vdash c \perp_a b$ and $b \perp_a cd \vdash b \perp_a c$.

Example 1.

- $b \perp_a c \vdash c \perp_a b$:
  1. $ab \subseteq ab' \wedge c \perp_a b' \wedge ab' \subseteq ab$ (Start Axiom)
  2. $ac \subseteq ac$ (Reflexivity)
  3. $b \perp_a c \wedge ab' \subseteq ab \wedge ac \subseteq ac \vdash ab'c \subseteq abc$ (Chase Rule)
  4. $ab'c \subseteq abc \vdash acb' \subseteq acb$ (Projection and Permutation)
  5. $ab \subseteq ab' \wedge c \perp_a b' \wedge acb' \subseteq acb \vdash c \perp_a b$ (Final Rule)
- $b \perp_a cd \vdash b \perp_a c$:
  1. $ac \subseteq ac' \wedge b \perp_a c' \wedge ac' \subseteq ac$ (Start Axiom)
  2. $ac'd' \subseteq acd$ (Inclusion Introduction)
  3. $ab \subseteq ab$ (Reflexivity)
  4. $b \perp_a cd \wedge ab \subseteq ab \wedge ac'd' \subseteq acd \vdash abc'd' \subseteq abcd$ (Chase Rule)
  5. $abc' \subseteq abc$ (Projection and Permutation)
  6. $ac \subseteq ac' \wedge b \perp_a c' \wedge abc' \subseteq abc \vdash b \perp_a c$ (Final Rule)

Our results show that for any consequence $\boldsymbol{b} \perp_{\boldsymbol{a}} \boldsymbol{c}$ of Σ there is a deduction starting with an application of Start Axiom and ending with an application of Final Rule.

6. Generalization to embedded join dependencies

The method of axiomatizing embedded multivalued dependencies using implicit existential quantification and inclusion dependencies can be generalized to the case where EMVDs are replaced with embedded join dependencies (EJDs). In the team semantics setting we will now define a *join atom* that corresponds to EJD. For this, we first formulate the concept of natural join for teams. For teams $X_1, \dots, X_n$ over $\mathcal{M}$ with domains $V_1, \dots, V_n$, the *(natural) join* of X_i's is defined as follows:

$$\bowtie_{i=1}^n \{X_i\} := \{s : \bigcup_{i=1}^n V_i \to M : s \upharpoonright V_i \in X_i \text{ for all } 1 \le i \le n\}.$$

The semantics of join atoms is defined as follows:

	a	b	c	x	y	z	w
s	1	2	3	1	2	3	4
s'	1	3	4	1	3	2	3

Fig. 1. Team X.

Definition 6. Let $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$ be tuples of variables listing $V_1, \ldots, V_n$, respectively, and let $V := \bigcup_{i=1}^n V_i$. Then $*(\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n)$ is a *join atom* with the semantic rule

– $\mathcal{M} \models_X *(\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n)$ if and only if $X \upharpoonright V = \bowtie_{i=1}^n \{X \upharpoonright V_i\}$.

Let FO($\bowtie$) denote first-order logic extended with join atoms.[2] Note that all FO($\bowtie$) formulas are ESO-expressible and by [8] all ESO-expressible properties of teams are FO($\perp_c$)-expressible. Also note that a conditional independence atom $\boldsymbol{b} \perp_{\boldsymbol{a}} \boldsymbol{c}$ can be expressed as $\exists \boldsymbol{d}(\boldsymbol{d} = \boldsymbol{c} \wedge *(\boldsymbol{ab}, \boldsymbol{ad}))$, where $\boldsymbol{d}$ is a tuple of new variables. Hence the following holds:

Proposition 3. *Any* FO($\perp_c$) *formula is logically equivalent to some* FO($\bowtie$) *formula, and vice versa.*

Axioms for join atoms and inclusion atoms can be defined as in Definition 3 except that the last three rules are replaced with analogues. In the following definition we use $\boldsymbol{x} \approx \boldsymbol{y}$ to denote $\boldsymbol{x} \subseteq \boldsymbol{y} \wedge \boldsymbol{y} \subseteq \boldsymbol{x}$.

Definition 7. In addition to the usual introduction and elimination rules for conjunction and the first 5 rules of Definition 3, we adopt the following rules for join atoms and inclusion atoms. The prerequisite for applying each of the following rules is that $\boldsymbol{a}_1 \ldots \boldsymbol{a}_n = (a_{i_1}, \ldots, a_{i_l})$ and $\boldsymbol{b}_1 \ldots \boldsymbol{b}_n = (b_{j_1}, \ldots, b_{j_l})$ where $j_k = j_{k'} \Rightarrow i_k = i_{k'}$, for $1 \leq k < k' \leq l$.

6. Start Axiom:

$$\bigwedge_{1 \leq i \leq n} \boldsymbol{a}_i \approx \boldsymbol{b}_i \wedge *(\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n)$$

where $\boldsymbol{b}_1 \ldots \boldsymbol{b}_n$ is a sequence of *new* variables.

7. Chase Rule:

$$\text{if } \bigwedge_{1 \leq i \leq n} \boldsymbol{a}_i \subseteq \boldsymbol{b}_i \wedge *(\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n), \text{ then } \boldsymbol{a}_1 \ldots \boldsymbol{a}_n \subseteq \boldsymbol{b}_1 \ldots \boldsymbol{b}_n.$$

8. Final Rule:

$$\text{if } \bigwedge_{1 \leq i \leq n} \boldsymbol{a}_i \subseteq \boldsymbol{b}_i \wedge *(\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n) \wedge \boldsymbol{b}_1 \ldots \boldsymbol{b}_n \subseteq \boldsymbol{a}_1 \ldots \boldsymbol{a}_n, \text{ then } *(\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n).$$

Example 2. The prerequisite that the tuple $\boldsymbol{a}_1 \ldots \boldsymbol{a}_n$ enjoy at least the same repetitions as the tuple $\boldsymbol{b}_1 \ldots \boldsymbol{b}_n$ is important. For example, the rule

$$\text{if } xy \subseteq ab \wedge yz \subseteq bc \wedge *(ab, bc), \text{ then } xyyz \subseteq abbc$$

is an instance of Chase Rule and clearly sound. However, the rule

$$\text{if } xy \subseteq ab \wedge zw \subseteq bc \wedge *(ab, bc), \text{ then } xyzw \subseteq abbc$$

is not an instance of Chase Rule and not sound since there exists a team X such that $X \models xy \subseteq ab \wedge zw \subseteq bc \wedge *(ab, bc)$ but $X \not\models xyzw \subseteq abbc$ (see Fig. 1).

Consider now the implication problem of join atoms and inclusion atoms with respect to the axioms given in Definition 7. In Subsect. 6.1 and 6.2 we will show that the axioms are sound and complete, respectively.

[2] The possibility of extending first-order logic with join atoms was first considered in [30].

6.1. Soundness

We will next show that the axioms are sound. This can be done analogously to Sect. 4 where new variables of the proof steps are interpreted as existentially quantified in the sense of Definition 1. For proving the variant of Lemma 2, we need the following locality property for $\mathrm{FO}(\bowtie)$. The proof is a straightforward structural induction.

Proposition 4 *(Locality). If $\phi \in \mathrm{FO}(\bowtie)$ is such that $\mathrm{Fr}(\phi) \subseteq V$, then for all models $\mathcal{M}$ and teams X,*

$$\mathcal{M} \models_X \phi \Leftrightarrow \mathcal{M} \models_{X \upharpoonright V} \phi.$$

Next we prove the variant of Lemma 2 for join atoms and inclusion atoms. For the proof note that Lemma 1 holds also for $\mathrm{FO}(\bowtie)$.

Lemma 4. *Let $(\phi_1, \ldots, \phi_n)$ be a deduction from Σ, and let $\boldsymbol{y}$ list all the new variables of the deduction steps. Let $\mathcal{M}$ and X be such that $\mathcal{M} \models_X \Sigma$ and $Var(\Sigma_n) \setminus \boldsymbol{y} \subseteq Dom(X)$ where $\Sigma_n := \Sigma \cup \{\phi_1, \ldots, \phi_n\}$. Then*

$$\mathcal{M} \models_X \exists \boldsymbol{y} \bigwedge \Sigma_n.$$

Proof. We show the claim by induction on n. Assume that the claim holds for any deduction of length n. We prove that the claim holds for deductions of length $n+1$ also. Let $(\phi_1, \ldots, \phi_{n+1})$ be a deduction from Σ, and let $\boldsymbol{y}$ and $\boldsymbol{z}$ list all the new variables of the deduction steps $\phi_1, \ldots, \phi_n$ and ϕ_{n+1}, respectively. Note that ϕ_{n+1} might not contain any new variables in which case $\boldsymbol{z}$ is empty. Assume that $\mathcal{M} \models_X \Sigma$ for some $\mathcal{M}$ and X, where $\mathrm{Var}(\Sigma_{n+1}) \setminus \boldsymbol{yz} \subseteq \mathrm{Dom}(X)$. By Proposition 4 we may assume that $\mathrm{Var}(\Sigma_{n+1}) \setminus \boldsymbol{yz} = \mathrm{Dom}(X)$. We need to show that

$$\mathcal{M} \models_X \exists \boldsymbol{y} \exists \boldsymbol{z} \bigwedge \Sigma_{n+1}.$$

By the induction assumption

$$\mathcal{M} \models_X \exists \boldsymbol{y} \bigwedge \Sigma_n$$

therefore by Lemma 1 there is a function $F : X \to \mathcal{P}(M^{|\boldsymbol{y}|}) \setminus \{\emptyset\}$ such that

$$\mathcal{M} \models_{X'} \bigwedge \Sigma_n \tag{36}$$

where $X' := X[F/\boldsymbol{y}]$. Since no variable listed in $\boldsymbol{z}$ is in $\mathrm{Dom}(X')$, then by Proposition 4, it suffices to show that

$$\mathcal{M} \models_{X'} \exists \boldsymbol{z} \phi_{n+1}. \tag{37}$$

Since all the other cases are covered in the proof of Proposition 2, we will consider the cases where ϕ_{n+1} is an instance of Start Axiom or follows from Σ_n by Chase Rule or Final Rule. Note that in the last two cases $\boldsymbol{z}$ is empty hence it suffices to show that $\mathcal{M} \models_{X'} \phi_{n+1}$. In the following items B_i (or A_i) denotes the set of variables listed in $\boldsymbol{b}_i$ (or in $\boldsymbol{a}_i$), and $B := \bigcup_{i=1}^n B_i$ ($A := \bigcup_{i=1}^n A_i$).

- Start Axiom: Then ϕ_{n+1} is of the form $\bigwedge_{1 \le i \le n} \boldsymbol{a}_i \approx \boldsymbol{b}_i \wedge *(\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n)$ where $\boldsymbol{b}_1 \ldots \boldsymbol{b}_n$ is a sequence of *new* variables. We first let

  $$Y_i := \{s : B_i \to M \mid s(\boldsymbol{b}_i) = s'(\boldsymbol{a}_i) \text{ for some } s' \in X'\}.$$

  Then we let $Y := \bowtie_{i=1}^n \{Y_i\}$. Note that $Y \neq \emptyset$. For, taking $s' \in X'$ we find $s_1, \ldots, s_n$ in $Y_1, \ldots, Y_n$ such that $s_i(\boldsymbol{b}_i) = s'(\boldsymbol{a}_i)$. Then letting $s : B \to M$ be such that $s(\boldsymbol{b}_i) = s_i(\boldsymbol{b}_i)$, we obtain that $s \in Y$. Due to the prerequisite for applying Start Axiom, s is well-defined. Then letting $X'' := X' \bowtie Y$, we obtain that

  $$\mathcal{M} \models_{X''} \bigwedge_{1 \le i \le n} \boldsymbol{a}_i \approx \boldsymbol{b}_i \wedge *(\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n)$$

  from which (37) follows.
- Chase Rule: Then ϕ_{n+1} is of the form $\boldsymbol{a}_1 \ldots \boldsymbol{a}_n \subseteq \boldsymbol{b}_1 \ldots \boldsymbol{b}_n$ where

  $$\bigwedge_{1 \le i \le n} \boldsymbol{a}_i \subseteq \boldsymbol{b}_i \wedge *(\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n) \in \Sigma_n.$$

  Let $s \in X'$. Since $\mathcal{M} \models_{X'} \bigwedge_{1 \le i \le n} \boldsymbol{a}_i \subseteq \boldsymbol{b}_i$, there exist $s_i \in X' \upharpoonright B_i$ such that $s(\boldsymbol{a}_i) = s_i(\boldsymbol{b}_i)$ for $1 \le i \le n$. We let $s' : B \to M$ be such that $s'(\boldsymbol{b}_i) = s_i(\boldsymbol{b}_i)$ for $1 \le i \le n$. Then

  $$s' \in \bowtie_{i=1}^n \{X' \upharpoonright B_i\}.$$

Note that due to the prerequisite for applying Chase Rule, s' is well-defined. Since $\mathcal{M} \models_{X'} *(\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n)$, we obtain that $s' \in X' \upharpoonright B$. We conclude that $s(\boldsymbol{a}_i) = s'(\boldsymbol{b}_i)$ for $1 \le i \le n$.
- Final Rule: Then ϕ_{n+1} is of the form $*(\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n)$ where

$$\bigwedge_{1\le i\le n} \boldsymbol{a}_i \subseteq \boldsymbol{b}_i \wedge *(\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n) \wedge \boldsymbol{b}_1 \ldots \boldsymbol{b}_n \subseteq \boldsymbol{a}_1 \ldots \boldsymbol{a}_n \in \Sigma_n.$$

Let $s : A \to M$ be such that $s \upharpoonright A_i \in X' \upharpoonright A_i$ for $1 \le i \le n$. We will show that $s \in X' \upharpoonright A$. Analogously to the previous case, applying the first two conjuncts we find a $s' \in X' \upharpoonright B$ such that $s(\boldsymbol{a}_i) = s'(\boldsymbol{b}_i)$ for $1 \le i \le n$. Now due to $\mathcal{M} \models_{X'} \boldsymbol{b}_1 \ldots \boldsymbol{b}_n \subseteq \boldsymbol{a}_1 \ldots \boldsymbol{a}_n$ we obtain that $s \in X' \upharpoonright A$. Hence $\mathcal{M} \models_{X'} *(\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n)$. □

By Lemma 4 we obtain the following soundness theorem for join atoms and inclusion atoms. The proof is identical to the proof of Theorem 1.

Theorem 4. *Let $\Sigma \cup \{\phi\}$ be a finite set of join atoms and inclusion atoms. Then $\Sigma \models \phi$ if $\Sigma \vdash \phi$.*

6.2. Completeness

For showing completeness of the axioms, we can use the same technique as in Sect. 5. That is, we first present a chase characterization of the implication problem using directed labeled graphs, and then apply this in the deduction. For the chase characterization, let us first define a graph $G_{\Sigma,\phi}$ in the style of Definition 4 for a set $\Sigma \cup \{\phi\}$ of join atoms and inclusion atoms.

Definition 8. Let $\Sigma \cup \{\phi\}$ be a finite set of join atoms and inclusion atoms. We let $G_{\Sigma,\phi} := (\bigcup_{n\in\mathbb{N}} V_n, \bigcup_{n\in\mathbb{N}} E_n)$ where $G_n = (V_n, E_n)$ is defined as follows:

- If ϕ is $*(\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n)$, then $V_0 := \{v_1, \ldots, v_n\}$ and $E_0 := \{(v_i, v_j)_{aa} \mid 1 \le i < j \le n, a \in \boldsymbol{a}_i \cap \boldsymbol{a}_j\}$. If ϕ is $\boldsymbol{a} \subseteq \boldsymbol{b}$, then $V_0 := \{v\}$ and $E_0 := \emptyset$.
- Assume that G_n is defined. Then for every $u \in V_n$ and $x_1 \ldots x_k \subseteq y_1 \ldots y_k \in \Sigma$ we introduce a new vertex v_{new} and new edges $(u, v_{\text{new}})_{x_i y_i}$, for $1 \le i \le k$. Also for every $u_1, \ldots, u_n \in V_n$ and $*(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n) \in \Sigma$ where $u_i \sim_{xx} u_j$, for all $1 \le i < j \le n$ and $x \in \boldsymbol{x}_i \cap \boldsymbol{x}_j$, we introduce a new vertex v_{new} and new edges $(u_i, v_{\text{new}})_{xx}$, for $1 \le i \le n$ and $x \in \boldsymbol{x}_i$. We let V_{n+1} and E_{n+1} be obtained by adding these new vertices and edges to the sets V_n and E_n.

We then obtain the following characterization.

Theorem 5. *Let $\Sigma \cup \{\phi\}$ be a finite set of join atoms and inclusion atoms.*

1. *If ϕ is $a_1 \ldots a_k \subseteq b_1 \ldots b_k$, then $\Sigma \models \phi \Leftrightarrow \exists w \in V_{\Sigma,\phi}(v \sim_{a_i b_i} w$ for all $1 \le i \le k)$.*
2. *If ϕ is $*(\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n)$, then $\Sigma \models \phi \Leftrightarrow \exists w \in V_{\Sigma,\phi}(v_i \sim_{aa} w$ for all $1 \le i \le n$ and $a \in \boldsymbol{a}_i)$.*

Proof. Analogous to the proof of Theorem 2. □

Next we present a variant of Lemma 3. Recall that we write $x = y$ for syntactical identity, $x \equiv y$ for an atom of the form $xy \subseteq zz$ implying the identity of x and y, and $\boldsymbol{x} \equiv \boldsymbol{y}$ for a conjunction of the form $\bigwedge_{i\le|\boldsymbol{x}|} \mathrm{pr}_i(\boldsymbol{x}) \equiv \mathrm{pr}_i(\boldsymbol{y})$. Also if $\boldsymbol{x}$ is a sequence listing $\mathrm{Var}(\Sigma \cup \{\phi\})$, and $\boldsymbol{x}_v$ is a vector of length $|\boldsymbol{x}|$ (representing vertex v of the graph $G_{\Sigma,\phi}$), and $\boldsymbol{a} = (x_{i_1}, \ldots, x_{i_l})$ is a sequence of variables from $\boldsymbol{x}$, then we write $\boldsymbol{a}_v$ for

$$(\mathrm{pr}_{i_1}(\boldsymbol{x}_v), \ldots, \mathrm{pr}_{i_l}(\boldsymbol{x}_v)).$$

Also, for a deduction d from Σ, we write $\Sigma \vdash^d \psi$ if ψ appears as a proof step in d.

Lemma 5. *Let n be a natural number, $\Sigma \cup \{\phi\}$ a finite set of join atoms and inclusion atoms, and $\boldsymbol{x}$ a sequence listing $Var(\Sigma \cup \{\phi\})$. Then there is a deduction $d = (\phi_1, \ldots, \phi_N)$ from Σ such that for each $u \in V_n$, there is a sequence $\boldsymbol{x}_u$ of length $|\boldsymbol{x}|$ (and possibly with repetitions) such that $\Sigma \vdash^d \boldsymbol{x}_u \subseteq \boldsymbol{x}$, and for each $(u, w)_{x_i x_j} \in E_n^*$, $\Sigma \vdash^d \mathrm{pr}_i(\boldsymbol{x}_u) \equiv \mathrm{pr}_j(\boldsymbol{x}_w)$. Moreover,*

- *if ϕ is of the form $\boldsymbol{a} \subseteq \boldsymbol{b}$, then $\phi_1 = \boldsymbol{x}_v \subseteq \boldsymbol{x}$ (obtained by Reflexivity), for $\boldsymbol{x}_v$ defined as $\boldsymbol{x}$,*
- *if ϕ is of the form $*(\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n)$, then $\phi_1 = \bigwedge_{1\le i\le n} \boldsymbol{a}_i \approx \boldsymbol{b}_i \wedge *(\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n)$ (obtained by Start Axiom) where $(\boldsymbol{a}_i)_{v_i} = \boldsymbol{b}_i$ for $1 \le i \le n$.*

Please cite this article in press as: M. Hannula, J. Kontinen, A finite axiomatization of conditional independence and inclusion dependencies, Inf. Comput. (2016), http://dx.doi.org/10.1016/j.ic.2016.04.001

Proof. Analogous to the proof of Lemma 3. □

Using Lemma 5 we can prove the completeness.

Theorem 6. *Let $\Sigma \cup \{\phi\}$ be a finite set of join atoms and inclusion atoms. Then $\Sigma \vdash \phi$ if $\Sigma \models \phi$.*

Proof. Analogous to the proof of Theorem 3. □

By Theorem 4 and 6 we obtain the following corollary.

Corollary 2. *Let $\Sigma \cup \{\phi\}$ be a finite set of join atoms and inclusion atoms. Then $\Sigma \vdash \phi$ if and only if $\Sigma \models \phi$.*

7. Conclusion

We have given a finite axiomatization for the implication problem of conditional independence atoms and inclusion atoms in the team semantics setting. Using a chase characterization for the implication problem we have showed that the axiomatization is sound and complete. In database theory, this yields a sound and complete axiomatization for the implication problem of EMVDs, INDs and FDs in the unirelational case. Moreover, we have generalized the above setting by introducing a sound and complete axiomatization for the implication problem of so-called join atoms and inclusion atoms. This immediately gives us a sound and complete axiomatization for the implication problem of EJDs and INDs in the unirelational case.

More generally, this article introduces a method of finding axiomatizations for various collections of database dependencies using techniques that arise from dependence logic and team semantics. A usual deduction that applies these axiomatizations uses additional inclusion dependencies and implicit existential quantification in the intermediate steps. It remains a question of future work to study the range of application of the techniques introduced in this article.

Acknowledgments

The authors would like to thank Sebastian Link for useful comments and suggestions. The authors were supported by grants 264917 and 292767 of the Academy of Finland, and the first author was supported by University of Auckland research grant 3706751.

References

[1] S. Abiteboul, R. Hull, V. Vianu, Foundations of Databases, Addison–Wesley, 1995.
[2] W.W. Armstrong, Dependency structures of data base relationships, in: Proc. of IFIP World Computer Congress, 1974, pp. 580–583.
[3] C. Beeri, M.Y. Vardi, A proof procedure for data dependencies, J. ACM 31 (4) (1984) 718–741.
[4] C. Beeri, M.Y. Vardi, Formal systems for tuple and equality generating dependencies, SIAM J. Comput. 13 (1) (1984) 76–98.
[5] M.A. Casanova, R. Fagin, C.H. Papadimitriou, Inclusion dependencies and their interaction with functional dependencies, in: Proceedings of the 1st ACM SIGACT-SIGMOD Symposium on Principles of Database Systems. PODS '82, ACM, New York, NY, USA, 1982, pp. 171–176.
[6] A.K. Chandra, M.Y. Vardi, The implication problem for functional and inclusion dependencies is undecidable, SIAM J. Comput. 14 (3) (1985) 671–677.
[7] S.S. Cosmadakis, P.C. Kanellakis, Equational theories and database constraints, in: R. Sedgewick (Ed.), Proceedings of the 17th Annual ACM Symposium on Theory of Computing, May 6–8, 1985, ACM, Providence, Rhode Island, USA, 1985, pp. 273–284.
[8] P. Galliani, Inclusion and exclusion dependencies in team semantics: on some logics of imperfect information, Ann. Pure Appl. Log. 163 (1) (2012) 68–84.
[9] P. Galliani, General models and entailment semantics for independence logic, Notre Dame J. Form. Log. 54 (2) (2013) 253–275.
[10] P. Galliani, L. Hella, Inclusion logic and fixed point logic, in: S.R.D. Rocca (Ed.), Computer Science Logic 2013 (CSL 2013), Leibniz International Proceedings in Informatics (LIPIcs), vol. 23, Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, Dagstuhl, Germany, 2013, pp. 281–295.
[11] P. Galliani, J.A. Väänänen, On dependence logic, in: A. Baltag, S. Smets (Eds.), Johan van Benthem on Logical and Informational Dynamics, Springer Verlag, 2014, pp. 101–119.
[12] P. Galliani, M. Hannula, J. Kontinen, Hierarchies in independence logic, in: S.R.D. Rocca (Ed.), Computer Science Logic 2013 (CSL 2013), Leibniz International Proceedings in Informatics (LIPIcs), vol. 23, Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, Dagstuhl, Germany, 2013, pp. 263–280.
[13] D. Geiger, A. Paz, J. Pearl, Axioms and algorithms for inferences involving probabilistic independence, Inf. Comput. 91 (1) (1991) 128–141.
[14] E. Grädel, J. Väänänen, Dependence and independence, Stud. Log. 101 (2) (2013) 399–410.
[15] M. Hannula, Axiomatizing first-order consequences in independence logic, Ann. Pure Appl. Logic 166 (1) (2015) 61–91.
[16] M. Hannula, Reasoning about embedded dependencies using inclusion dependencies, in: M. Davis, A. Fehnker, A. McIver, A. Voronkov (Eds.), Logic for Programming, Artificial Intelligence, and Reasoning, Proceedings of 20th International Conference, LPAR-20 2015, Suva, Fiji, November 24–28, 2015, in: Lect. Notes Comput. Sci., vol. 9450, Springer, 2015, pp. 16–30.
[17] M. Hannula, J. Kontinen, A finite axiomatization of conditional independence and inclusion dependencies, in: C. Beierle, C. Meghini (Eds.), Foundations of Information and Knowledge Systems, Proceedings of 8th International Symposium, FoIKS 2014, Bordeaux, France, March 3–7, 2014, in: Lect. Notes Comput. Sci., vol. 8367, Springer, 2014, pp. 211–229.
[18] C. Herrmann, On the undecidability of implications between embedded multivalued database dependencies, Inf. Comput. 122 (2) (1995) 221–235.
[19] C. Herrmann, Corrigendum to "On the undecidability of implications between embedded multivalued database dependencies" [Inf. Comput. 122 (1995) 221–235], Inf. Comput. 204 (12) (2006) 1847–1851.
[20] J. Kontinen, J.A. Väänänen, Axiomatizing first-order consequences in dependence logic, Ann. Pure Appl. Logic 164 (11) (2013) 1101–1117.

[21] J. Kontinen, S. Link, J.A. Väänänen, Independence in database relations, in: L. Libkin, U. Kohlenbach, R.J.G.B. de Queiroz (Eds.), WoLLIC, in: Lect. Notes Comput. Sci., vol. 8071, Springer, 2013, pp. 179–193.

[22] D. Maier, A.O. Mendelzon, Y. Sagiv, Testing implications of data dependencies, ACM Trans. Database Syst. 4 (4) (December 1979) 455–469.

[23] J.C. Mitchell, The implication problem for functional and inclusion dependencies, Inf. Control 56 (3) (1983) 154–173.

[24] P. Naumov, B. Nicholls, R.E. axiomatization of conditional independence, in: Proceedings of the 14th Conference on Theoretical Aspects of Rationality and Knowledge (TARK '13), 2013, pp. 131–137.

[25] E. Pacuit, F. Yang, Dependence and independence in social choice: Arrow's theorem, in: S. Abramsky, J. Kontinen, H. Vollmer, J. Väänänen (Eds.), Dependence Logic: Theory and Applications, Springer, 2016, in press.

[26] G. Paolini, J. Väänänen, Dependence logic in pregeometries and ω-stable theories, J. Symb. Log. 81 (1) (2016) 32–55.

[27] J. Paredaens, The interaction of integrity constraints in an information system, J. Comput. Syst. Sci. 20 (3) (1980) 310–329.

[28] J. Rissanen, Theory of relations for databases — a tutorial survey, in: J. Winkowski (Ed.), Mathematical Foundations of Computer Science, in: Lect. Notes Comput. Sci., vol. 64, Springer, Berlin, Heidelberg, 1978, pp. 536–551.

[29] J. Väänänen, Dependence Logic, Cambridge University Press, 2007.

[30] J. Väänänen, Comments on Abramsky, in: Logic Across the University: Foundations and Applications, Proceedings of the Tsinghua Logic Conference, College Publications, London, 2013, pp. 152–156.

[31] Fan Yang, Jouko Väänänen, Propositional logics of dependence, Ann. Pure Appl. Log. 167 (7) (2016) 557–589, http://dx.doi.org/10.1016/j.apal.2016.03.003.

[32] M. Yannakakis, C.H. Papadimitriou, Algebraic dependencies, J. Comput. Syst. Sci. 25 (1) (1982) 2–41.

Please cite this article in press as: M. Hannula, J. Kontinen, A finite axiomatization of conditional independence and inclusion dependencies, Inf. Comput. (2016), http://dx.doi.org/10.1016/j.ic.2016.04.001