The Journal of Symbolic Logic
Volume 84, Number 3, September 2019

CHARACTERIZING DOWNWARDS CLOSED, STRONGLY FIRST-ORDER, RELATIVIZABLE DEPENDENCIES

PIETRO GALLIANI

Abstract. In Team Semantics, a dependency notion is strongly first order if every sentence of the logic obtained by adding the corresponding atoms to First-Order Logic is equivalent to some first-order sentence. In this work it is shown that all nontrivial dependency atoms that are strongly first order, downwards closed, and relativizable (in the sense that the relativizations of the corresponding atoms with respect to some unary predicate are expressible in terms of them) are definable in terms of constancy atoms.

Additionally, it is shown that any strongly first-order dependency is safe for any family of downwards closed dependencies, in the sense that every sentence of the logic obtained by adding to First-Order Logic both the strongly first-order dependency and the downwards closed dependencies is equivalent to some sentence of the logic obtained by adding only the downwards closed dependencies.

§1. Introduction. Team Semantics [26] generalizes Tarskian semantics for First-Order Logic by defining the satisfaction relation with respect to sets of assignments, called *teams*, rather with respect to single assignments. Team Semantics was originally developed in order to provide a compositional semantics equivalent to the imperfect-information, game-theoretic semantics of *Independence-Friendly Logic* [23, 24, 37]; however, with the work of Väänänen [39] it became clear that this semantics is a natural generalization of Tarskian semantics, one which—even putting aside its connections with the theory of databases (see for instance [30])—greatly extends its expressive capabilities by allowing the study of a far greater range of atoms and operators.

One of the most peculiar aspects of Team Semantics is the way in which it straddles the boundary between first order and second order: while the syntaxes of the logics based on Team Semantics studied so far have for the most part a solidly first-order flavour, in the sense that they involve no *explicit* higher order quantification,[1] their expressive power often rises well above the one of First-Order Logic, all the way up to Existential Second Order Logic (as in the cases of Dependence Logic [39] and Independence Logic [20]) or even to full Second Order Logic (as in the case of Team Logic [31, 40]).

This is the case even for logics, such as the above mentioned Dependence Logic and Independence Logic or for Inclusion Logic [13], which only add to the language of First-Order Logic *dependency atoms* whose satisfaction conditions are first-order

Received September 1, 2018.
2010 *Mathematics Subject Classification.* 03C80, 03C85.
Key words and phrases. Team Semantics. Dependence Logic, Second Order Logic.
[1] The *majority quantifier* of [7] could arguably be considered an exception to this.

© 2019, Association for Symbolic Logic
0022-4812/19/8403-0012
DOI:10.1017/jsl.2019.12

https://doi.org/10.1017/jsl.2019.12 Published online by Cambridge University Press

definable as properties of relations. In brief, this is due to the higher order quantification hidden in Team Semantics rules for disjunction and existential quantification (see Rules **TS-**$\vee$ and **TS-**$\exists$ of Definition 2.6). For First-Order Logic proper there exists a strict equivalence between Team Semantics and the usual Tarskian semantics on the level of sentences (see Proposition 2.8), but this equivalence fails badly on the level of formulas: there exist properties of relations that are first-order definable (in the sense that they are defined by a first-order sentence $\phi(R)$), but that do not correspond to the satisfaction conditions of any first order *formula* in Team Semantics (that is, there is no first-order formula $\psi(\mathbf{x})$ which is satisfied by a set of assignments X if and only if the corresponding relation satisfies $\phi(R)$). Adding new atoms with these (first-order definable) satisfaction conditions, therefore, will increase the expressive power of the logic, in the sense that there will be formulas of this new logic which are not equivalent to any first-order formula.

Does it follow that this new logic will be more expressive also with respect to *sentences*, in the sense that there will exist sentences of this new logic which are not equivalent to any first-order sentence? Not necessarily. While this is true for e.g., Dependence Logic and Independence Logic (that is, for the logics obtained by adding the *functional dependence atoms* or the *independence atoms* of Definition 2.11, respectively), it is possible to find dependencies for which this is not the case: while adding them to the language of First-Order Logic makes it more expressive *with respect to formulas*, every sentence of this new logic is still equivalent to some first-order sentence. Dependency atoms, or families of dependency atoms, for which this is the case are called *strongly first order* [15, 16].

This asymmetry between expressivity on the level of formulas and expressivity on the level of sentences is one of the most intriguing aspects of Team Semantics; and, in particular, the fact that when using Team Semantics it is possible to generate logics with expressive powers between that of First Order-Logic FO and that of Existential Second Order Logic ESO (included) simply by adding to FO atoms with first-order definable satisfaction conditions makes Team Semantics an eminently suitable tool for the study and classification of fragments of ESO, a topic rich with open questions and with deep connections with important complexity-theoretic conjectures.

The study of strongly first-order dependencies, in particular, can be thought of as an attempt to investigate the border between first order and second order "from below" by seeking to characterize precisely which choices of dependency atoms (or families of dependency atoms) breach or fail to breach it. The conjecture according to which a dependency is strongly first order if and only if it is definable in terms of upwards closed dependencies and constancy dependencies[2] is, however, still unproven. In this work, a proof for a special case of it will be found: if a nontrivially-false dependency is strongly first order, is downwards closed, and is furthermore *relativizable* (a new, natural property of dependencies which will be introduced in this work)[3] then it is definable in terms of constancy atoms alone.

A property related to strong first orderness and introduced in the recent article [17] is *safety*. In brief, a dependency (or a set of dependencies) is safe for some logic

[2]As proved in [15] these families of dependencies are both strongly first order, as is their union.
[3]This property is related to the *relativization operator* of [38].

https://doi.org/10.1017/jsl.2019.12 Published online by Cambridge University Press

if it can be added to it without increasing the expressive power (wrt sentences) of the resulting formalism. In particular, a dependency is strongly first order if and only if it is safe for First-Order Logic. This notion of safety, thus, generalizes the notion of strong first orderness to logics more expressive than pure First-Order Logic, and a complete characterization of safety would provide a full classification of the expressive powers of the logics obtained by adding dependencies to logics with Team Semantics. Safety, however, is a surprisingly delicate notion: for instance, as shown in [17], constancy—despite being perhaps the simplest example of strongly first-order dependency—is *not* safe for First-Order Logic plus certain classes of dependencies (e.g., unary inclusion dependencies). In this work it will be shown that, in the case of downwards closed dependencies, safety is more robust: any strongly first-order dependency is safe for First-Order Logic plus any family containing only downwards closed dependencies. This result, aside from being interesting in its own right, will be essential for the characterization of strongly first order downwards closed relativizable dependencies mentioned above.

Much of the research in the area of Team Semantics thus far has focused on the study of specific logics obtained by adding specific atoms (or specific operators) to First-Order Logic with Team Semantics. This is an important direction of research, and many of the resulting logics are of independent interest (and, often, have intriguing connections with the theory of databases). But an equivalently important, if so far understandably[4] less studied, one consists in the *classification* of general families of such logics in terms of their relationships and of their meta-logical properties.[5] The study of strongly first-order and safe dependencies, to which this article contributes, is a promising subtopic of this intriguing and largely unexplored field of research.

§2. Preliminaries.

2.1. Team semantics and dependencies. As mentioned in the Introduction, Team Semantics generalizes Tarskian semantics for First-Order Logic by letting formulas be satisfied or not satisfied by sets of assignments, which are called *teams* for historical reasons:

DEFINITION 2.1 (Team). Let $\mathfrak{M}$ be a first-order model (over any signature Σ) with domain M and let V be a finite set of variables. Then a team X over $\mathfrak{M}$ with domain $\mathtt{Dom}(X) = V$ is a set of variable assignments $s : V \to M$ over $\mathfrak{M}$.

There exists an obvious correspondence between teams and relations:

DEFINITION 2.2 (Teams to relations). Let X be a team over some model $\mathfrak{M}$, and let $\mathbf{t} = t_1 \dots t_n$ be a finite tuple of terms in the signature of $\mathfrak{M}$ with variables in $\mathtt{Dom}(X)$. Then we write $X(\mathbf{t})$ for the n-ary relation

$$X(\mathbf{t}) = \{\langle t_1^{\mathfrak{M}}(s), t_2^{\mathfrak{M}}(s), \dots, t_n^{\mathfrak{M}}(s)\rangle : s \in X\}$$

where each $t_i^{\mathfrak{M}}(s)$ is the interpretation of t_i in $\mathfrak{M}$ for the assignment s.

[4] Indeed, when studying the properties of a new type of semantics, it is a good strategy to begin by identifying and investigating logics that make use of it and are of independent interest.

[5] For examples of works along these lines, see for instance [15] or [29].

https://doi.org/10.1017/jsl.2019.12 Published online by Cambridge University Press

Teams can be restricted to a subset of the variables in their domain in the obvious way:

DEFINITION 2.3 (Team restriction). Let X be a team over some model $\mathfrak{M}$ and let $V \subseteq \mathtt{Dom}(X)$ be a subset of the variables in its domain. Then we write $X_{|V}$ for the restriction of X to the variables in V, that is for

$$X_{|V} = \{s_{|V} : s \in X\}$$

where, for all assignments $s \in X$, $s_{|V}$ is the unique assignment with domain V such that $s_{|V}(v) = s(v)$ for all variables $v \in V$.

It will also be useful, in several places of this work, to consider the subteam of a given team which contains only the assignments satisfying (in the sense of the usual Tarskian Semantics) some first-order formula θ. This is defined in the obvious way:

DEFINITION 2.4 (Team selection). Let X be a team over some model $\mathfrak{M}$ and let $\theta(\mathbf{x})$ be a first-order formula over the signature of $\mathfrak{M}$ whose free variables $\mathrm{Free}(\theta)$ are contained in the domain $\mathtt{Dom}(X)$ of X. Then

$$X \upharpoonright \theta(\mathbf{x}) = \{s \in X : \mathfrak{M} \models_s \theta(\mathbf{x})\}$$

where the expression $\mathfrak{M} \models_s \theta(\mathbf{x})$ means that the assignment s satisfies $\theta(\mathbf{x})$ in $\mathfrak{M}$ according to the usual Tarskian Semantics.

Finally, in order to define Team Semantics we will also need the two following operations (see Figure 1):

$X =$

	v_0
s_0	0
s_1	1

$$H(s) = \begin{cases} \{\langle 1, 0\rangle\} & \text{if } s = s_0, \\ \{\langle 0, 0\rangle, \langle 0, 1\rangle\} & \text{if } s = s_1 \end{cases}$$

$X[H/v_1v_2] =$

	v_0	v_1	v_2
s'_0	0	1	0
s'_1	1	0	0
s'_2	1	0	1

$X[M/v_1v_2] =$

	v_0	v_1	v_2
s''_0	0	0	0
s''_1	0	0	1
s''_2	0	1	0
s''_3	0	1	1
s''_4	1	0	0
s''_5	1	0	1
s''_6	1	1	0
s''_7	1	1	1

FIGURE 1. Supplementation and duplication examples in a model with only two elements 0 and 1.

DEFINITION 2.5 ((Lax) supplementation and duplication). Let X be a team over some first-order model $\mathfrak{M}$ and let $H : X \to \mathcal{P}(M^k)\backslash\{\emptyset\}$ be a function associating to each assignment $s \in X$ some nonempty set $H(s) \subseteq M^k$ of k-tuples of possible

https://doi.org/10.1017/jsl.2019.12 Published online by Cambridge University Press

values, and let $\mathbf{v} = v_1 \dots v_k$ be any k-tuple of pairwise distinct variable symbols (which may or may not occur already in $\mathtt{Dom}(X)$). Then we write $X[H/\mathbf{v}]$ for the team, with domain $\mathtt{Dom}(X) \cup \{v_i : i = 1 \dots k\}$, defined as

$$X[H/\mathbf{v}] = \{s[\mathbf{m}/\mathbf{v}] : s \in X, \mathbf{m} \in H(s)\}$$

where, as usual, $s[\mathbf{m}/\mathbf{v}]$ is the assignment obtained starting from s and fixing the values of the variables $\mathbf{v} = v_1 \dots v_k$ to $\mathbf{m} = m_1 \dots m_k$. This team is called the *supplementation* of X along H.

For every $\mathbf{m} \in M^k$, we also write $X[\mathbf{m}/\mathbf{v}]$ for the set $\{s[\mathbf{m}/\mathbf{v}] : s \in X\}$, that is for the supplementation of X along $\mathbf{v}$ via some H with $H(s) = \{\mathbf{m}\}$ for all $s \in X$.

Finally, we write $X[M/\mathbf{v}]$ for the largest possible supplementation of X, which is also called the *duplication* of X and which extends X by assigning *all* possible values to the variables in $\mathbf{v}$:

$$X[M/\mathbf{v}] = \{s[\mathbf{m}/\mathbf{v}] : s \in X, \mathbf{m} \in M^k\}.$$

The reason why the above supplementation operation is described as "lax" is because there also exists in the literature a "strict" version, in which $H(s)$ is a singleton for every $s \in X$. As discussed in [13], the choice between these two operations (as well as between two possible semantics for disjunction) corresponds precisely to the choice between allowing or disallowing nondeterministic strategies in the equivalent imperfect-information game-theoretic semantics; however, when considering non downwards-closed dependencies, the strict variant of Team Semantics can fail to satisfy locality (Proposition 2.13 in this work) in the sense that the satisfiability of a formula ϕ in a team X may depend on the values taken in X by variables which do *not* appear free in ϕ. Because of this, in this work (as in most of the recent literature in the area of Team Semantics) we will focus only on the lax version of the semantics.

DEFINITION 2.6 ((Lax) Team Semantics for First-Order Logic). Let $\mathfrak{M}$ be a first-order model, let X be a team over it, and let $\phi(\mathbf{x})$ be a first-order formula in Negation Normal Form[6] over the signature of $\mathfrak{M}$ such that the free variables of ϕ are contained in the domain of X. Then we say that X satisfies ϕ in $\mathfrak{M}$, and write $\mathfrak{M} \models_X \phi$, if and only if this follows from the following rules:

[6]It is possible to define Team Semantics for expressions not in Negation Normal Form, as it was done for instance in [39]. However, doing so requires taking track of positive and negative satisfaction conditions, which makes the semantics more cumbersome; and, as discussed in [33], the usual ("dual") negation operator has little semantic meaning in certain Team Semantics-based extensions of First-Order Logic, as the satisfaction conditions of a formula and of its negation are essentially unrelated. Furthermore, it is often unclear what the interpretation of the dual negation of a dependency atom should be: for instance, in [39] it was decided that the negations of functional dependence atoms $= (\mathbf{x}; \mathbf{y})$ are only satisfied by empty teams, that is they are all equivalent to $\bot$. For this reason, in this work we will assume that all expressions are in Negation Normal Form. Another possible negation operator, of clearer interpretation in Team Semantics, is the contradictory negation $\mathfrak{M} \models_X \sim \phi \Leftrightarrow \mathfrak{M} \not\models_X \phi$; but adding it to First-Order Logic together with even very simple dependencies (e.g., constancy atoms) brings the expressive power of the resulting formalism all the way up to Second Order Logic. The logic $\mathtt{FO}(\sim)$ obtained by taking First-Order Logic (with Team Semantics) and adding to it the contradictory negation (but no dependencies) is however equivalent to First-Order Logic wrt sentences, as shown in [16], and in [36] an axiomatization for it is found. This operator will not be further discussed in this work.

https://doi.org/10.1017/jsl.2019.12 Published online by Cambridge University Press

TS-lit: If α is a first order literal then $\mathfrak{M} \models_X \alpha$ if and only if, for all $s \in X$, $\mathfrak{M} \models_s \alpha$ according to the usual rules of Tarskian Semantics;

TS-$\vee$: For all formulas ψ_1 and ψ_2, $\mathfrak{M} \models_X \psi_1 \vee \psi_2$ if and only if there exist teams Y and Z such that $X = Y \cup Z$, $\mathfrak{M} \models_Y \psi_1$ and $\mathfrak{M} \models_Z \psi_2$;[7]

TS-$\wedge$: For all formulas ψ_1 and ψ_2, $\mathfrak{M} \models_X \psi_1 \wedge \psi_2$ if and only if $\mathfrak{M} \models_X \psi_1$ and $\mathfrak{M} \models_X \psi_2$;

TS-$\exists$: For all formulas ψ and all variables v, $\mathfrak{M} \models_X \exists v\psi$ if and only if there exists some $H : X \to \mathcal{P}(M)\backslash\{\emptyset\}$ such that $\mathfrak{M} \models_{X[H/v]} \psi$;

TS-$\forall$: For all formulas ψ and all variables v, $\mathfrak{M} \models_X \forall v\psi$ if and only if $\mathfrak{M} \models_{X[M/v]} \psi$.

Given a first-order model $\mathfrak{M}$ and a negation normal form, first-order sentence ϕ over its signature, we say that ϕ is true in $\mathfrak{M}$ in Team Semantics, and write $\mathfrak{M} \models \phi$, if and only if $\mathfrak{M} \models_{\{\epsilon\}} \phi$, where $\{\epsilon\}$ is the team containing the unique assignment ϵ over the empty domain.

The following two widely known facts describe completely the relationship between Team Semantics and Tarskian Semantics for First-Order Logic:

PROPOSITION 2.7. *Let $\mathfrak{M}$ be a first-order model, let X be a team over it, and let $\phi(\mathbf{x}) \in \mathtt{FO}$ be a first-order formula over the signature of $\mathfrak{M}$ and with free variables in $\mathtt{Dom}(X)$. Then $\mathfrak{M} \models_X \phi(\mathbf{x})$ if and only if, for all assignments $s \in X$, $\mathfrak{M} \models_s \phi$ according to the usual Tarskian Semantics.*

PROOF. Straightforward induction. ⊣

PROPOSITION 2.8. *Let $\mathfrak{M}$ be a first-order model and let $\phi \in \mathtt{FO}$ be a first-order sentence over the signature of $\mathfrak{M}$. Then $\mathfrak{M} \models \phi$ in the sense of Team Semantics if and only if $\mathfrak{M} \models \phi$ in the sense of Tarskian Semantics.*

PROOF. By definition, $\mathfrak{M} \models \phi$ in the sense of Team Semantics if and only if $\mathfrak{M} \models_{\{\epsilon\}} \phi$. By Proposition 2.7, this is the case if and only if $\mathfrak{M} \models_\epsilon \phi$ in the sense of Tarskian Semantics. This holds if and only if ϕ is true in $\mathfrak{M}$ according to Tarskian Semantics. ⊣

Do these two results show that Team Semantics is a pointlessly overwrought, but practically equivalent, variant of Tarskian semantics? Well, no: as mentioned in the Introduction, the richer structure of the satisfaction relation in Team Semantics makes it possible to extend First-Order Logic in new ways. In Team Semantics, the satisfaction conditions of a first-order formula $\phi(\mathbf{x})$ always correspond to some first-order definable property of relations, in the sense that there exists some first-order sentence $\phi'(R)$, where R is a relation symbol not in the signature of $\mathfrak{M}$, such that $\mathfrak{M} \models_X \phi(\mathbf{x})$ if and only if $\mathfrak{M}[X(\mathbf{x})/R] \models \phi'(R)$:[8] indeed, by Proposition 2.7, it suffices to take $\phi'(R) := \forall\mathbf{x}(R\mathbf{x} \to \phi(\mathbf{x}))$. However, a moment's thought shows that there exist first-order definable properties of relations that do not correspond

[7] In this rule we do not require that $Y \cap Z = \emptyset$. Doing that would give us the strict semantics for disjunction, which—as in the case of supplementation and existential quantification—corresponds to allowing only deterministic strategies in the game theoretic semantics and would result in the failure of Proposition 2.13 for certain nondownwards closed dependencies.

[8] Here we write $\mathfrak{M}[X(\mathbf{t})/R]$ for the first-order model obtained by adding the relation symbol R to $\mathfrak{M}$—if it is not present already—and fixing the relation $X(\mathbf{t}) = \{s(\mathbf{t}) : s \in X\}$ as its interpretation.

https://doi.org/10.1017/jsl.2019.12 Published online by Cambridge University Press

to the satisfaction conditions of any first-order formula in Team Semantics. For instance,

COROLLARY 2.9. *There is no first-order formula $\phi(x)$ such that $\mathfrak{M} \models_X \phi(x)$ if and only if $|X(x)| \leq 1$, that is, if and only if the variable x takes at most one value in X.*

Thus, it is possible to extend the Team Semantics of First-Order Logic by introducing new types of atoms, such as the following *constancy atoms*:

DEFINITION 2.10 (Constancy atoms, Constancy Logic). Constancy Logic $\mathtt{FO}(=(\cdot))$ is the logic obtained by adding to the language of First-Order Logic (in Negation Normal Form) *constancy atoms* $=(x)$ for all variables x, with satisfaction conditions

TS-const: $\mathfrak{M} \models_X =(x)$ if and only if $|X(x)| \leq 1$.

It is then possible to inquire about the properties of this Constancy Logic. For example, by Corollary 2.9 we know already that there exist formulas in constancy logic which are not equivalent to any first-order formula; but what about sentences? Can Constancy Logic be used to define classes of models which are not definable in First-Order Logic? As shown in [13], the answer is no. However, this *is* the case for the following atoms and for the logics that they characterize:

DEFINITION 2.11 (Functional dependence, independence, and inclusion). Dependence Logic, Independence Logic, and Inclusion Logic are the logics obtained by adding to the language of First-Order Logic (in Negation Normal Form) *functional dependence atoms* $=(\mathbf{x};\mathbf{y})$, *independence atoms* $\mathbf{x}\perp_{\mathbf{y}}\mathbf{z}$, and *inclusion atoms* $\mathbf{x} \subseteq \mathbf{y}$, respectively, with the following semantics:

TS-dep: $\mathfrak{M} \models_X =(\mathbf{x};\mathbf{y})$ iff for any two $s, s' \in X$, if $s(\mathbf{x}) = s'(\mathbf{x})$ then $s(\mathbf{y}) = s'(\mathbf{y})$;

TS-indep: $\mathfrak{M} \models_X \mathbf{x}\perp_{\mathbf{y}}\mathbf{z}$ iff for any two $s, s' \in X$ with $s(\mathbf{y}) = s'(\mathbf{y})$ there exists some $s'' \in X$ such that $s''(\mathbf{xy}) = s(\mathbf{xy})$ and $s''(\mathbf{yz}) = s'(\mathbf{yz})$;

TS-inc: $\mathfrak{M} \models_X \mathbf{x} \subseteq \mathbf{y}$ iff for any $s \in X$ there exists some $s' \in X$ such that $s(\mathbf{x}) = s'(\mathbf{y})$.

As shown in [39], [20], and [19] ,respectively, Dependence Logic and Independence Logic are equivalent to Existential Second Order Logic $\mathtt{ESO}$ over sentences, while Inclusion Logic is equivalent to the positive fragment of Greatest Fixed Point Logic.[9] This is the case despite the fact that the satisfaction conditions of the above atoms are easily definable as first-order properties of relations; and, as mentioned in the Introduction, the fact that Team Semantics allows to greatly increase the expressive power of First-Order Logic via first-order definable properties is precisely what makes it an eminently suitable tool for the study of the boundary between first-order and higher order logics. Additionally, it is worth remarking here that these dependencies have clear connections with database theory: the relationship between functional dependence and inclusion atoms and the functional [1, 6] and inclusion [3] dependencies of database theory is obvious, and as discussed in [10] there likewise exists a correspondence between independence atoms and database-theoretic *multivalued dependencies* [12]. Additionally, these atoms have a strong

[9] This implies, in particular, that Inclusion Logic captures PTIME over finite ordered models.

https://doi.org/10.1017/jsl.2019.12 Published online by Cambridge University Press

doxastic flavour: if a team X represents the set of the states of things that an agent believes possible then $=(\mathbf{x}, \mathbf{y})$ can be read as "if I learned the true value of $\mathbf{x}$, I could infer the value of $\mathbf{y}$", while $\mathbf{x}\perp_{\mathbf{y}}\mathbf{z}$ can be read as "if I learned the true value of $\mathbf{y}$, the value of $\mathbf{x}$ would give me no information whatsoever regarding the value of $\mathbf{z}$" and $\mathbf{x} \subseteq \mathbf{y}$ can be read as "every possible value for $\mathbf{x}$ is also a possible value for $\mathbf{y}$". As discussed in [14], all the connectives and operators of Team Semantics also admit natural interpretations in terms of the dynamics of belief states; but we will not pursue this line of thought any further in this work.

A very fruitful research direction in Team Semantics research so far has consisted in the study of the properties of these logics and of fragments thereof (see e.g., [8, 9, 13, 18, 19, 21, 22, 27, 34]). This is a valuable topic of investigation, with rich connections with open problems in the classification of better known fragments of Second Order Logic and in descriptive complexity theory. Furthermore, as briefly mentioned above, these logics are of independent interest, and hence their properties and those of their fragments are deserving of study for their own sake.

The present work is a contribution to a different—if obviously related—research programme, one in which the central topic of investigation is not Dependence Logic or Independence Logic or any other specific Team Semantics-based logic per se but rather *Team Semantics itself*; and, under this perspective, rather than selecting "interesting" additional atoms and studying the logics obtained by adding them to First-Order Logic we want to select "interesting" properties for Team Semantics-based logics and investigate which atoms (or more in general which operators, although as we will see there are plenty of open questions even in the more limited case of atoms) or collections of atoms satisfy them if added to the language of First-Order Logic.

In order to do this properly, we must first clarify exactly what a dependency atom *is* in Team Semantics. The following definition, from [35], is a natural starting point:

DEFINITION 2.12 (Dependency). For any $k \in \mathbb{N}$, a k-ary *dependency* $\mathbf{D}$ is a class of models over the signature $\{R\}$ which is closed under isomorphisms, where R is a k-ary relation symbol.

Given a family of dependencies $\mathcal{D} = \{\mathbf{D}_i : i \in I\}$, we write $\mathtt{FO}(\mathcal{D})$ for the logic obtained by adding to the language of First-Order Logic (in Negation Normal Form) all dependency atoms $\mathbf{Dt}$, where $\mathbf{D} \in \mathcal{D}$ and $\mathbf{t}$ is a tuple of terms of length equal to the arity of $\mathbf{D}$. The Team Semantics for $\mathtt{FO}(\mathcal{D})$ is defined precisely as in Definition 2.6, with the additional condition.

TS-**D**: For all models $\mathfrak{M}$ with domain M, all teams X over $\mathfrak{M}$, all $\mathbf{D} \in \mathcal{D}$ and all tuples $\mathbf{t}$ of terms in the signature of $\mathfrak{M}$ with variables in the domain of X and of length equal to the arity of $\mathbf{D}$, $\mathfrak{M} \models_X \mathbf{Dt}$ if and only if $\langle M, X(\mathbf{t})\rangle \in \mathbf{D}$.[10]

The following result, that can be found in [29] together with a number of results concerning the complexity of the satisfiability problem for fragments of logics with generalized dependencies, shows that—regardless of the choice of $\mathcal{D}$—$\mathtt{FO}(\mathcal{D})$ is *local*

[10] In this work we identify the tuple $\langle M, R_1 \dots R_n\rangle$ with the model $\mathfrak{M}$ with domain M and relations $R_1 \dots R_n$. When no ambiguity regarding the choice of model is possible, as in this case, we also use the same symbol R_i for the relation symbol and for its interpretation $R_i^{\mathfrak{M}}$.

https://doi.org/10.1017/jsl.2019.12 Published online by Cambridge University Press

in the sense that the values of variables which do not appear free in a formula are irrelevant to its satisfaction or lack thereof:

PROPOSITION 2.13 (Locality). *Let $\mathcal{D}$ be any family of dependencies, let ϕ be a formula of $\mathtt{FO}(\mathcal{D})$, and let $Free(\phi)$ be the set of its free variables. Then, for all models $\mathfrak{M}$ whose signature contains that of ϕ and for all teams X with domain containing $Free(\phi)$,*

$$\mathfrak{M} \models_X \phi \Leftrightarrow \mathfrak{M} \models_{X_{|Free(\phi)}} \phi.$$

PROOF. By structural induction. ⊣

It is easy to verify that all the dependency atoms discussed so far are special cases of Definition 2.12. Note, furthermore, that nothing in Definition 2.12 requires the class $\mathbf{D}$ to be first-order definable. For instance, $\mathbf{U} = \{\langle M, R\rangle : |R| \text{ is uncountable}\}$[11] is a perfectly acceptable unary dependency, and the corresponding satisfaction condition in Team Semantics is: $\mathfrak{M} \models_X \mathbf{U}t$ if and only if $X(t)$ is uncountable.

If, as in the case of $\mathbf{U}$, the satisfaction conditions of a dependency atom $\mathbf{D}$ are not first-order definable then obviously not all sentences of $\mathtt{FO}(\mathbf{D})$ are equivalent to first-order sentences: indeed, it is easy to see that the $\mathtt{FO}(\mathbf{D})$ sentence $\forall\mathbf{x}(\neg R\mathbf{x} \vee (R\mathbf{x} \wedge \mathbf{D}\mathbf{x}))$ characterizes precisely the class $\mathbf{D}$, which is not first-order definable by assumption. A more interesting case is the one in which $\mathbf{D}$—understood as a class of models—*is* first-order definable:

DEFINITION 2.14 (First order dependencies). A k-ary dependency $\mathbf{D}$ is *first order* if there exists some sentence $\phi(R)$ in the signature $\{R\}$, where R is k-ary, such that $\mathbf{D} = \{\langle M, R\rangle : \langle M, R\rangle \models \phi(R)\}$.

PROPOSITION 2.15. *Constancy atoms, functional dependence atoms, independence atoms and inclusion atoms are all first order.*

PROOF. Choose for $\phi(R)$ the expressions $\forall\mathbf{x}\mathbf{x}'(R\mathbf{x} \wedge R\mathbf{x}' \to \mathbf{x} = \mathbf{x}')$, $\forall\mathbf{x}\mathbf{x}'\mathbf{y}(R\mathbf{x}\mathbf{y} \wedge R\mathbf{x}'\mathbf{y} \to \mathbf{x} = \mathbf{x}')$, $\forall\mathbf{x}\mathbf{x}'\mathbf{y}\mathbf{z}\mathbf{z}'(R\mathbf{x}\mathbf{y}\mathbf{z} \wedge R\mathbf{x}'\mathbf{y}\mathbf{z}' \to R\mathbf{x}\mathbf{y}\mathbf{z}')$ and $\forall\mathbf{x}\mathbf{y}(R\mathbf{x}\mathbf{y} \to \exists\mathbf{z}R\mathbf{z}\mathbf{x})$, respectively. ⊣

The following properties, well studied in the literature, provide a useful starting point for the classification of dependencies:

DEFINITION 2.16 (Closure properties for dependencies). Let $\mathbf{D}$ be any dependency. Then

- $\mathbf{D}$ has the *empty team property* if, for all sets of elements M, $\langle M, \emptyset\rangle \in \mathbf{D}$;
- $\mathbf{D}$ is *downwards closed* if for all relations R over a domain M, $\langle M, R\rangle \in \mathbf{D}$ and $Q \subseteq R$ imply $\langle M, Q\rangle \in \mathbf{D}$;
- $\mathbf{D}$ is *upwards closed* if for all relations R over some domain M, $\langle M, R\rangle \in \mathbf{D}$ and $Q \supseteq R$ imply $\langle M, Q\rangle \in \mathbf{D}$;
- $\mathbf{D}$ is *union closed* if for all families $\{R_i : i \in I\}$ of relations of the same arity over the same domain M such that $\langle M, R_i\rangle \in \mathbf{D}$ for all $i \in I$ we have that $\langle M, \bigcup_i R_i\rangle \in \mathbf{D}$ as well.

[11] Strictly speaking, this is not a set but a proper class since M ranges over all possible (uncountable) sets of elements. This is not a problem for the purposes of this work—see Definition 2.12—and we will use expressions such as the above for defining the interpretations of dependencies.

https://doi.org/10.1017/jsl.2019.12 Published online by Cambridge University Press

Aside from upwards closure, these properties are preserved by Team Semantics in the following sense:

PROPOSITION 2.17. *Let $\mathcal{D}$ be a family of dependencies and let $\phi(\mathbf{x})$ be any formula of $\mathtt{FO}(\mathcal{D})$. Then, for all models $\mathfrak{M}$ with signature containing the signature of $\phi(\mathbf{x})$,*

1. *If all $\mathbf{D} \in \mathcal{D}$ have the empty team property then $\mathfrak{M} \models_\emptyset \phi$;*
2. *If all $\mathbf{D} \in \mathcal{D}$ are downwards closed and X is a team such that $\mathfrak{M} \models_X \phi$ then $\mathfrak{M} \models_Y \phi$ for all $Y \subseteq X$;*
3. *If all $\mathbf{D} \in \mathcal{D}$ are union closed and $(X_i : i \in I)$ is a family of teams such that $\mathfrak{M} \models_{X_i} \phi$ for all $i \in I$ then $\mathfrak{M} \models_{\bigcup_i X_i} \phi$.*

PROOF. Straightforward induction. ⊣

Closure properties such as these are useful to establish nondefinability relations between dependencies.

DEFINITION 2.18 (Definability of dependencies). Let $\mathcal{D}$ be a family of dependencies and let $\mathbf{E}$ be another dependency. Then $\mathbf{E}$ is said to be *definable* in $\mathtt{FO}(\mathcal{D})$ if for all tuples $\mathbf{t}$ of terms there exists some formula $\phi(\mathbf{t})$ such that $\mathfrak{M} \models_X \mathbf{Et}$ if and only if $\mathfrak{M} \models_X \phi(\mathbf{t})$.[12]

It is easy to verify that the dependencies discussed thus far all have the empty team property; that none of them is upwards closed; that constancy and functional dependency are downwards closed but not union closed; that inclusion is union closed but not downwards closed; and that independence is neither union closed nor downwards closed. By Proposition 2.17, this allows us to infer at once for example that neither inclusion nor functional dependence are definable in terms of the other, and that neither of them alone suffices to define independence.[13]

It follows from known results in the literature that functional dependence atoms, inclusion atoms ,and independence atoms are "maximal" among the corresponding classes of first-order dependencies. More precisely, for all first-order dependencies $\mathbf{D}$,

1. If $\mathbf{D}$ has the empty team property and is downwards closed then it is definable in terms of functional dependence atoms ([32]);
2. If $\mathbf{D}$ has the empty team property then it is definable in terms of independence atoms ([13]);
3. If $\mathbf{D}$ has the empty team property and is union closed then it is definable in terms of inclusion atoms ([19]).

In fact, the results concerning functional dependence atoms and independence atoms are stronger, in that they also hold for *existential second order* (rather than first-order) dependencies $\mathbf{D}$.[14] This is not, however, the case for Inclusion Logic: while all first-order union-closed dependencies with the empty team property are definable in it, as discussed in [19] it is possible to find existential second order

[12]The choice of $\mathbf{t}$ may force us to rename bound variables in ϕ; but aside from that, it is not difficult to see that if such a ϕ exists then it works for any $\mathbf{t}$.

[13]However, as shown in [13], independence atoms are definable if we have *both* inclusion atoms and functional dependence atoms.

[14]Analogously to Definition 2.14, a dependency $\mathbf{D}$ is existential second order if there exists some $\mathtt{ESO}$ sentence $\phi(R)$ such that $\mathbf{D} = \{\langle M, R\rangle : \langle M, R\rangle \models \phi(R)\}$.

https://doi.org/10.1017/jsl.2019.12 Published online by Cambridge University Press

union-closed dependencies with the empty team property which are not definable in Inclusion Logic.

We conclude this section by providing a few definitions and elementary results that will be used in this work.

DEFINITION 2.19 ($\exists\mathbf{v}$ and $\forall\mathbf{v}$). Let $\mathbf{v} = v_1 \dots v_k$ be a tuple of pairwise distinct variables. Then, for all formulas ϕ, we write $\exists\mathbf{v}\phi$ as a shorthand for $\exists v_1 \exists v_2 \dots \exists v_k \phi$, and we write $\forall\mathbf{v}\phi$ as a shorthand for $\forall v_1 \forall v_2 \dots \forall v_k \phi$.

PROPOSITION 2.20. *For all models $\mathfrak{M}$, all families of dependencies $\mathcal{D}$, all formulas $\phi \in \mathtt{FO}(\mathcal{D})$, all tuples $\mathbf{v} = v_1 \dots v_k$ of variables, all models $\mathfrak{M}$ whose domain contains the signature of ϕ and all teams X whose domain contains $Free(\phi)\backslash\{v_1 \dots v_k\}$*

$$\mathfrak{M} \models_X \exists\mathbf{v}\phi \Leftrightarrow \text{there is a } H : X \to \mathcal{P}(M^k)\backslash\{\emptyset\} \text{ s.t. } \mathfrak{M} \models_{X[H/\mathbf{v}]} \phi$$

and

$$\mathfrak{M} \models_X \forall\mathbf{v}\phi \Leftrightarrow \mathfrak{M} \models_{X[M/\mathbf{v}]} \phi.$$

PROOF. Trivial induction on the length k of the tuple $\mathbf{v}$. ⊣

DEFINITION 2.21 (Selective implication). For all families of dependencies $\mathcal{D}$, all formulas $\phi \in \mathtt{FO}(\mathcal{D})$ and all first-order formulas $\theta \in \mathtt{FO}$, we write $\theta \hookrightarrow \phi$ as a shorthand for $(\neg\theta) \vee (\theta \wedge \phi)$ where $\neg\theta$ is the first-order negation normal form expression equivalent to the negation of θ.

PROPOSITION 2.22. *For all models $\mathfrak{M}$, all teams X, all families of dependencies $\mathcal{D}$, all formulas $\phi \in \mathtt{FO}(\mathcal{D})$ and all formulas $\theta \in \mathtt{FO}$,*

$$\mathfrak{M} \models_X \theta \hookrightarrow \phi \Leftrightarrow \mathfrak{M} \models_{X\upharpoonright\theta} \phi.$$

PROOF. Follows easily from Proposition 2.7 and from the rule for disjunction in Team Semantics. ⊣

DEFINITION 2.23 (Boolean disjunction). Let $\phi, \psi \in \mathtt{FO}(\mathcal{D})$ for some choice of $\mathcal{D}$. Then we write $\phi \sqcup \psi$ for the $\mathtt{FO}(\mathcal{D} \cup \{=(\cdot)\})$-formula

$$\exists z_1 z_2 (=(z_1) \wedge =(z_2) \wedge ((z_1 = z_2 \wedge \phi) \vee (z_1 \neq z_2 \wedge \psi)))$$

where z_1 and z_2 are two new variables not occurring in ϕ or ψ and $=(z_i)$ is the constancy atom of Definition 2.10, which holds in a team X if and only if z_i takes at most one value in it.

PROPOSITION 2.24. *For all models $\mathfrak{M}$, all families of dependencies with the empty team property $\mathcal{D}$, all formulas $\phi, \psi \in \mathtt{FO}(\mathcal{D})$ and all teams X,*

$$\mathfrak{M} \models_X \phi \sqcup \psi \Leftrightarrow \mathfrak{M} \models_X \phi \text{ or } \mathfrak{M} \models_X \psi.$$

PROOF. Straightforward from definitions (observe, however, that the requirement that all $\mathbf{D} \in \mathcal{D}$ have the empty team property cannot be removed). ⊣

https://doi.org/10.1017/jsl.2019.12 Published online by Cambridge University Press

PROPOSITION 2.25 (Positive occurrences of relations). *Let $\mathcal{D}$ be any family of dependencies and let ϕ be any formula of $\mathtt{FO}(\mathcal{D})$ in which some relation symbol R occurs only positively.*[15] *Then ϕ is* upwards closed *in R, in the sense that*

$$\mathfrak{M} \models_X \phi, R^{\mathfrak{M}} \subseteq Q \Rightarrow \mathfrak{M}[Q/R] \models_X \phi$$

for all suitable models $\mathfrak{M}$, teams X and relations Q.

PROOF. Trivial by structural induction. ⊣

2.2. Strongly first order and safe dependencies. As already mentioned, in [13] it was shown that the constancy atoms of Definition 2.10 do not increase the expressive power of First-Order Logic with respect to sentences, in the sense that every sentence of the logic $\mathtt{FO}(=(\cdot))$ obtained by adding these atoms to First-Order Logic is equivalent to some first-order sentence. In other words, First-Order Logic is not a natural "stopping point" in the family of the logics based on Team Semantics, as it is possible to find expressions (e.g., constancy atoms) that as per Corollary 2.9 cannot be defined in terms of it but however do not add to the expressive power of its sentences.

Thus, we say that constancy atoms are *strongly first order* according to the following definition:

DEFINITION 2.26 (Strongly first order dependencies and families). Let $\mathcal{D}$ be a family of first-order dependencies. Then $\mathcal{D}$ is *strongly first order* if every sentence of $\mathtt{FO}(\mathcal{D})$ is equivalent to some first-order sentence.

A single dependency $\mathbf{D}$ is said to be strongly first order if the singleton $\{\mathbf{D}\}$ is strongly first order in the above sense.

Constancy atoms are not the only strongly first-order dependencies, and the logic $\mathtt{FO}(=(\cdot))$ is not a natural "stopping point" in the above sense[16] either. This follows at once from the following result from [15]:

THEOREM 2.27. *Let $\mathcal{D}^{\uparrow}$ be the family of all first order upwards closed dependencies. Then the family $\mathcal{D}^{\uparrow} \cup \{=(\cdot)\}$, which contains all those dependencies and also constancy atoms, is strongly first order.*

Using this result, it is not difficult to find additional strongly first-order dependencies. For instance, as mentioned in [15], the contradictory negations of inclusion dependencies

$$\mathfrak{M} \models_X \mathbf{x} \not\subseteq \mathbf{y} \Leftrightarrow \exists s \in X \forall s' \in Xs(\mathbf{x}) \neq s'(\mathbf{y})$$

are all strongly first order, since they are definable (in the sense of Definition 2.18) in terms of constancy atoms and upwards closed dependencies. To the knowledge of the author all strongly first-order dependencies known so far are also definable in terms of $\mathtt{FO}(\mathcal{D}^{\uparrow}, =(\cdot))$.[17] This justifies the following

[15] Since all our expressions are in Negation Normal Form, this is the same as saying that no negated literal $\neg R\mathbf{t}$ appears in ϕ.

[16] This notion will be made more precise in a moment.

[17] We commit a minor abuse of notation and write $\mathtt{FO}(\mathcal{D}^{\uparrow}, =(\cdot))$ rather than $\mathtt{FO}(\mathcal{D}^{\uparrow} \cup \{=(\cdot)\})$.

https://doi.org/10.1017/jsl.2019.12 Published online by Cambridge University Press

CONJECTURE 2.28. *A dependency* **D** *is strongly first order if and only if it is definable in* $\mathtt{FO}(\mathcal{D}^{\uparrow}, =(\cdot))$.

As already mentioned, the main result of this work will be the proof of a special case of this conjecture.

In [17], the notion of strong first orderness was generalized to the following notion of *safety*:

DEFINITION 2.29 (Safe dependencies and families). Let $\mathcal{L}$ be a logic based on Team Semantics[18] and let $\mathcal{E}$ be a family of dependencies. Then $\mathcal{E}$ is *safe* for $\mathcal{L}$ if every sentence of $\mathcal{L}(\mathcal{E})$ is equivalent to some sentence of $\mathcal{L}$.

A single dependency **E** is said to be safe for $\mathcal{L}$ if the singleton $\{\mathbf{E}\}$ is safe for $\mathcal{L}$ in the above sense.

A dependency **E** or a dependency family $\mathcal{E}$ is said to be safe for a dependency **D** or for a dependency family $\mathcal{D}$ if it is safe for the logic $\mathtt{FO}(\mathbf{D})$ (resp. $\mathtt{FO}(\mathcal{D})$).

By definition, a dependency (or a family of dependencies) is strongly first order if and only if it is safe for FO. The notion of safety can be used to make precise the informal concept of "natural stopping point" mentioned above:

DEFINITION 2.30 (Definitionally closed logic). A logic $\mathcal{L}$, based on Team Semantics, is *definitionally closed* if and only if the following two properties are equivalent for all dependencies **D**:

1. **D** is safe for $\mathcal{L}$;
2. **D** is definable in $\mathcal{L}$.

Thus, Corollary 2.9 and the fact that constancy is strongly first order imply at once that First-Order Logic is not definitionally closed, while *Conjecture 1* would have as a direct consequence that $\mathtt{FO}(\mathcal{D}^{\uparrow}, =(\cdot))$ is definitionally closed (and is, in fact, the only "definitional closure" of FO). More in general, a full characterization of the definitionally closed extensions of First-Order Logic would be a very valuable contribution to the classification of logics based on Team Semantics.

Safety, however, is a delicate property: in particular, as shown in [17], constancy dependencies are not safe for $\mathtt{FO}(\subseteq_1)$, where $\subseteq_1$ is the family of only *unary* inclusion dependencies,[19] despite being as already mentioned strongly first order (that is, safe for FO). Thus, a dependency **D** may be safe for a logic $\mathcal{L}$ but unsafe for some other logic $\mathcal{L}'$ which strictly contains $\mathcal{L}$. This is reminiscent of certain phenomena in the theory of second-order generalized quantifiers [28],[20] and it suggests that the problem of fully characterizing safety relations and definitionally closed logics will not be of easy solution.

Nonetheless, studying the properties of safety is a fruitful endeavour. In particular, in this work we will prove that strongly first-order dependencies are safe for $\mathtt{FO}(\mathcal{D})$

[18] For the purposes of this work, we can always assume that $\mathcal{L}$ is of the form $\mathtt{FO}(\mathcal{D})$ for some choice of $\mathcal{D}$; but nothing prevents applying this definition to another logic—for instance, to some fragment of First-Order Logic or to some extension of it by additional connectives and operators.

[19] More precisely, $\mathtt{FO}(\subseteq_1)$ adds to First-Order Logic all inclusion atoms $v_1 \subseteq v_2$, where v_1 and v_2 are single variables (not tuples of variables).

[20] The author thanks an anonymous reviewer for pointing out this connection.

https://doi.org/10.1017/jsl.2019.12 Published online by Cambridge University Press

whenever $\mathcal{D}$ contains only downwards closed dependencies, and use this result to prove—as already mentioned—a special case of *Conjecture 1*.

2.3. Relativizable dependencies. Sometimes it may be necessary to check whether a dependency holds not with respect to the current model, but with respect to some submodel thereof. This justifies the following definitions:[21]

DEFINITION 2.31 (Relativized dependencies). Let P be any unary relation symbol, let $\mathbf{D}$ be any k-ary dependency, and let $\mathbf{t}$ be a tuple of terms of length k. Then for all first-order models $\mathfrak{M}$ and all teams X over $\mathfrak{M}$ with the variables that appear in $\mathbf{t}$ in their domain, $\mathfrak{M} \models_X \mathbf{D}^P\mathbf{t}$ if and only if

1. For all $t \in \mathbf{t}$, $X(t) \subseteq P^{\mathfrak{M}}$;[22]
2. $\langle P^{\mathfrak{M}}, X(\mathbf{t})\rangle \in \mathbf{D}$.

DEFINITION 2.32 (Relativizable dependencies). A dependency $\mathbf{D}$ is relativizable if every sentence of $\mathtt{FO}(\mathbf{D}^P)$ is equivalent to some sentence of $\mathtt{FO}(\mathbf{D})$ over the same signature.

It follows from the above definitions that if $\mathbf{D}$ is relativizable and strongly first order then every sentence constructed out of relativized dependency atoms $\mathbf{D}^P\mathbf{t}$ and first-order connectives and literals is equivalent to some first-order sentence.

All the dependency atoms widely studied in the literature—functional dependence, inclusion, independence and so on—are relativizable. This can be verified by observing that they have the following, stronger property:

DEFINITION 2.33 (Closed-world dependencies). A dependency $\mathbf{D}$ is *closed world* if, for all sets of elements M and all relations R, if $M' = \{m \in M : \exists \mathbf{a} \in R \text{ s.t. } m \in \mathbf{a}\}$ is the set of all elements occurring in any tuple of R then

$$\langle M, R\rangle \in \mathbf{D} \Leftrightarrow \langle M', R\rangle \in \mathbf{D}.$$

This notion is related to the *closed-world assumption* employed in database theory and knowledge representation, and it is also similar (albeit not identical) to the notion of *dependencies closed under substructures* of [29]: in short, if $\mathbf{D}$ is closed-world the validity of some atom $\mathbf{Dv}$ in a team X cannot be affected by the existence (or nonexistence) in the model of elements that do not appear as possible values of $\mathbf{v}$ in X.

It is easy to see that the dependencies examined so far are all closed-world. A first-order, unary dependency notion that is not closed-world is *nontotality* $\mathtt{nt} = \{\langle M, P\rangle : P \neq M\}$, corresponding to the atoms $\mathfrak{M} \models_X \mathtt{nt}(t) \Leftrightarrow X(t) \neq M$. Indeed, if $\langle M, P\rangle \in \mathtt{nt}$ and we restrict M to the set M' of the elements which occur in P then $(M', P) \notin \mathtt{nt}$. However, nontotality is definable in terms of constancy, which *is* closed-world:

$$\mathfrak{M} \models_X \mathtt{nt}(t) \Leftrightarrow \mathfrak{M} \models_X \exists z(=\!(z) \wedge z \neq t).$$

[21] This notion of relativization of dependencies is closely related to the notion of relativization of formulas in Inclusion-Exclusion Logic discussed by Rönnholm in [38].

[22] This condition is actually a consequence of the next one, but we state it explicitly for clarity.

https://doi.org/10.1017/jsl.2019.12 Published online by Cambridge University Press

Furthermore, despite not being closed-world, nontotality is still relativizable: indeed, it is straightforward to check that $\mathtt{nt}^P(t)$ is logically equivalent to $Pt \wedge \forall x((x = t \vee \neg Px) \hookrightarrow \mathtt{nt}(x))$.[23]

PROPOSITION 2.34. *If a dependency* $\mathbf{D}$ *is closed world then it is relativizable.*

PROOF. If $\mathbf{D}$ is closed world and all elements that appear in $X(\mathbf{v})$ are in P,

$$\begin{aligned}\mathfrak{M} \models_X \mathbf{D}^P\mathbf{v} &\Leftrightarrow \langle P, X(\mathbf{v})\rangle \in \mathbf{D}\\ &\Leftrightarrow \langle M', X(\mathbf{v})\rangle \in \mathbf{D}\\ &\Leftrightarrow \langle M, X(\mathbf{v})\rangle \in \mathbf{D} \Leftrightarrow \mathfrak{M} \models_X \mathbf{D}\mathbf{v}\end{aligned}$$

where $M' = \{s(v_i) : s \in X, v_i \in \mathbf{v}\}$ is the set of all values that variables in $\mathbf{v}$ take in X. Thus, $\mathbf{D}^P\mathbf{v}$ is definable in $\mathtt{FO}(\mathbf{D})$ simply as $\left(\bigwedge_{v_i \in \mathbf{v}} Pv_i\right) \wedge \mathbf{D}\mathbf{v}$, and hence $\mathbf{D}$ is relativizable. ⊣

As we saw, it is easy to find dependencies which are not closed-world, like nontotality. However, it is less obvious to find examples of dependencies that are not relativizable, which led the author to pose the following

PROBLEM (SOLVED BY FAUSTO BARBERO). Find a dependency $\mathbf{D}$ that is not relativizable, or prove that all dependencies are relativizable.

Fausto Barbero, in a personal communication, pointed out that a counting argument shows that there exist dependencies (not necessarily first order) that are not relativizable. A concrete example is given by the unary dependency $\mathbf{I}_\infty = \{\langle M, P\rangle : M \text{ is infinite}\}$, for which $\mathfrak{M} \models_X \mathbf{I}_\infty(v)$ iff the model $\mathfrak{M}$ is infinite regardless of X. Indeed, the $\mathtt{FO}(\mathbf{I}^P_\infty)$ sentence $\exists v\mathbf{I}^P_\infty(v)$ is true if and only if P is infinite. However, there is no $\mathtt{FO}(\mathbf{I}_\infty)$ sentence over the signature $\{P\}$ that satisfies this property, because in an infinite model every occurrence $\mathbf{I}_\infty(v)$ of $\mathbf{I}_\infty$ is equivalent to $\top$ and a standard back-and-forth argument shows that there is no first-order sentence which is true in an infinite model if and only if P is infinite. The following variant of the above problem is, to the knowledge of the author, still open, and if solved it would suffice to remove the requirement of relativizability from the results of this article:

OPEN PROBLEM. Find a strongly first order dependency $\mathbf{D}$ which is not relativizable, or prove that all such dependencies are relativizable.

§3. Strongly first-order dependencies are safe for downwards closed dependencies. As already mentioned, constancy dependencies are strongly first order—that is, safe for $\mathtt{FO}$—but they are not safe for unary inclusion dependencies. However, in this section, we will see that strongly first-order dependencies are safe for *downwards closed* dependencies. Let us begin by recalling an easy variant of Lyndon's Theorem:

PROPOSITION 3.1. *Let* $\theta(S_1 \dots S_n)$ *be a first-order sentence in a signature containing the symbols* $S_1 \dots S_n$ *(and possibly others) which is* upwards closed *in all* S_i*, in the sense that, for any model* $\mathfrak{M}$*, if* $\mathfrak{M} \models \theta(S_1 \dots S_n)$ *and* $S_i' \supseteq S_i$ *for all* i *then*

[23] See Definition 2.21 and Proposition 2.22 for the interpretation of the selective implication $\hookrightarrow$. The intuition here is that we force x to take all values outside P as well as all possible values of t: so if x does not take all possible values, there must be some value *inside* P that t does not take.

https://doi.org/10.1017/jsl.2019.12 Published online by Cambridge University Press

$\mathfrak{M} \models \theta(S'_1 \dots S'_n)$ *Then* $\theta(S_1 \dots S_n)$ *is equivalent to some first order* $\theta'(S_1 \dots S_n)$ *in which all* S_i *occur only positively.*

PROOF. Recall (see e.g., Theorem 8.3.3 of [25]) that if θ is not equivalent to any formula in which the S_i occur only positively[24] then it is not preserved under bijective homomorphisms that fix all relations other than $S_1 \dots S_n$: therefore, there exist two models $\mathfrak{A}$ and $\mathfrak{B}$ such that $\mathfrak{A} \models \phi$, $\mathfrak{B} \not\models \phi$, and there is a bijective homomorphism $\mathfrak{h} : \mathfrak{A} \to \mathfrak{B}$ that fixes all relations other than $S_1 \dots S_n$. For all $i = 1 \dots n$, let $T_i = \mathfrak{h}[S_i^{\mathfrak{A}}]$ be the image of $S_i^{\mathfrak{A}}$ under $\mathfrak{h}$. Then since $\mathfrak{h}$ is a homomorphism, $T_i \subseteq S_i^{\mathfrak{B}}$; since $\mathfrak{h}$ is bijective and fixes all other relations, $\mathfrak{B}[T_1 \dots T_n / S_1 \dots S_n]$ is isomorphic to $\mathfrak{A}$, and hence $\mathfrak{B} \models \phi(T_1 \dots T_n)$; and thus, since by hypothesis $\mathfrak{B} \not\models \phi(S_1 \dots S_n)$, ϕ is not upwards closed in $S_1 \dots S_n$. ⊣

The main idea of this section will be to "extract" the downwards closed dependency atoms $\mathbf{Dt}$ ($\mathbf{D} \in \mathcal{D}$) from a sentence $\phi \in \mathtt{FO}(\mathcal{D}, \mathcal{S})$, leaving only a sentence of $\mathtt{FO}(\mathcal{S})$; then use the fact that $\mathcal{S}$ is strongly first order to find an equivalent first-order sentence; and finally add back the dependencies previously removed.

The following lemmas show how to do the first step of this procedure:

LEMMA 3.2. *Let* ϕ *be a* $\mathtt{FO}(\mathcal{D})$ *formula, where* $\mathcal{D}$ *is a set of dependencies* $\{\mathbf{D}_0, \mathbf{D}_1, \dots\}$ *such that* $\mathbf{D}_0$ *is downwards closed. Suppose furthermore that* S *is a new relation symbol with the same arity* k *as* $\mathbf{D}_0$*, and let* $\phi^*(S)$ *be the formula obtained by replacing a* single *instance* $\mathbf{D}_0\mathbf{t}$ *of* $\mathbf{D}_0$ *with* $S\mathbf{t}$*. Then for all models* $\mathfrak{M}$ *with signature containing the signature of* ϕ *and for all teams* X *with domain containing the free variables of* ϕ*,* $\mathfrak{M} \models_X \phi$ *if and only if there exists a relation* $S \subseteq M^k$ *such that* $\langle M, S\rangle \in \mathbf{D}_0$ *and* $\mathfrak{M} \models_X \phi^*(S)$.

PROOF. By structural induction on ϕ:

- If ϕ is of the form $\mathbf{D}_0\mathbf{t}$ then $\phi^*(S)$ is simply $S\mathbf{t}$. Suppose that $\mathfrak{M} \models_X \mathbf{D}_0\mathbf{t}$; then for $S = X(\mathbf{t})$ we have that $\langle M, S\rangle \in \mathbf{D}_0$ and $\mathfrak{M} \models_X S\mathbf{t}$, as required. Conversely, suppose that such a S exists. Then since $\mathfrak{M} \models_X S\mathbf{t}$ we have that $X(\mathbf{t}) \subseteq S$; but then, since $\mathbf{D}_0$ is downwards closed and $\langle M, S\rangle \in \mathbf{D}_0$ we have that $\langle M, X(\mathbf{t})\rangle \in \mathbf{D}_0$ as well, and hence $\mathfrak{M} \models_X \mathbf{D}_0\mathbf{t}$ as required.
- If ϕ is of the form $\psi_1 \vee \psi_2$ then—assuming without loss of generality that the instance of $\mathbf{D}_0$ which we are replacing is in ψ_1—$\phi^*(S)$ is $\psi_1^*(S) \vee \psi_2$. Then $\mathfrak{M} \models_X \psi_1 \vee \psi_2$ iff $X = Y \cup Z$ for two Y, Z such that $\mathfrak{M} \models_Y \psi_1$ and $\mathfrak{M} \models_Z \psi_2$ iff (by induction hypothesis) there exist a relation S and two Y, Z such that $\langle M, S\rangle \in \mathbf{D}_0$, $X = Y \cup Z$, $\mathfrak{M} \models_Y \psi_1^*(S)$ and $\mathfrak{M} \models_Z \psi_2$ iff there exists a relation S such that $\langle M, S\rangle \in \mathbf{D}_0$ and $\mathfrak{M} \models_X \phi^*(S)$, as required.
- If ϕ is of the form $\psi_1 \wedge \psi_2$ then—assuming without loss of generality that the instance of $\mathbf{D}_0$ which we are replacing is in ψ_1—$\phi^*(S)$ is $\psi_1^*(S) \wedge \psi_2$. Then $\mathfrak{M} \models_X \phi$ iff $\mathfrak{M} \models_X \psi_1$ and $\mathfrak{M} \models_X \psi_2$ iff (by induction hypothesis) there is some S such that $\langle M, S\rangle \in \mathbf{D}_0$, $\mathfrak{M} \models_X \psi_1^*(S)$ and $\mathfrak{M} \models_X \psi_2$ iff there is some S such that $\langle M, S\rangle \in \mathbf{D}_0$ and $\mathfrak{M} \models_X \phi^*(S)$.
- If ϕ is of the form $\exists v\psi$ then $\phi^*(S)$ is $\exists v\psi^*(S)$. Then $\mathfrak{M} \models_X \phi$ iff there exists some H such that $\mathfrak{M} \models_{X[H/v]} \psi$ iff there exist some S and H such

[24] But in which the identity symbol may occur positively or negatively - this is necessary to guarantee bijectivity.

https://doi.org/10.1017/jsl.2019.12 Published online by Cambridge University Press

that $\langle M, S\rangle \in \mathbf{D}_0$ and $\mathfrak{M} \models_{X[H/v]} \psi^*(S)$ iff there exists some S such that $\langle M, S\rangle \in \mathbf{D}_0$ and $\mathfrak{M} \models \phi^*(S)$.

- If ϕ is of the form $\forall v\psi$ then $\phi^*(S)$ is $\forall v\psi^*(S)$. Then $\mathfrak{M} \models_X \phi$ iff $\mathfrak{M} \models_{X[M/v]} \psi$ iff there exists some S such that $\langle M, S\rangle \in \mathbf{D}_0$ and $\mathfrak{M} \models_{X[M/v]} \psi^*(S)$ iff there exists some S such that $\langle M, S\rangle \in \mathbf{D}_0$ and $\mathfrak{M} \models \phi^*(S)$. ⊣

LEMMA 3.3. *Let $\mathcal{S}$ be a strongly first-order family of dependencies, let $\mathcal{D}$ be a set of downwards closed dependencies, and let ϕ be a sentence of $\mathtt{FO}(\mathcal{S}, \mathcal{D})$. Then there exists a first-order sentence $\chi(S_1 \dots S_n)$, whose signature contains the signature of ϕ as well as new relation symbols $S_1 \dots S_n$, and dependencies $\mathbf{D}_{j(1)} \dots \mathbf{D}_{j(n)} \in \mathcal{D}$ such that, for all models $\mathfrak{M}$, $\mathfrak{M} \models \phi$ if and only if there exist relations $S_1 \dots S_n$ such that*

- *For all $i = 1 \dots n$, $\langle M, S_i\rangle \in \mathbf{D}_{j(i)}$;*
- *$\mathfrak{M} \models \chi(S_1 \dots S_n)$.*

Moreover, all S_i occur only positively in χ.

PROOF. Starting from ϕ, we can apply iteratively the previous lemma to remove all instances of dependencies $\mathbf{D} \in \mathcal{D}$ from ϕ. In this way, we obtain a sentence $\theta(S_1 \dots S_n)$ in $\mathtt{FO}(\mathcal{S})$ which—aside from being in $\mathtt{FO}(\mathcal{S})$ rather than in $\mathtt{FO}$—would satisfy our requirements.[25] But since $\mathcal{S}$ is a strongly first-order family of dependencies, $\theta(S_1 \dots S_n)$ is logically equivalent to some first order $\chi(S_1 \dots S_n)$. Furthermore, since the S_i occur only positively in θ by Proposition 2.25 it is the case that θ is upwards closed in all S_i, and hence so is χ, and hence—by Proposition 3.1—we can also require that all S_i occur only positively in χ. ⊣

Note that there is no guarantee that the first-order sentence $\chi(S_1 \dots S_n)$ obtained via Lemma 3.3 would contain only one occurrence of each symbol S_i, as the translation from $\mathtt{FO}(\mathcal{S})$ to $\mathtt{FO}$ may introduce additional occurrences. However, as we will now see, as long as the S_i occur only positively this can be dealt with:

LEMMA 3.4. *Let $\mathbf{D}$ be a downwards closed dependency, let $\chi(S)$ be a first-order formula where S is a relational symbol that occurs only positively in χ, and let $\chi'(W_1 \dots W_n)$ be the formula obtained by replacing each occurrence of S in χ with a different new symbol W_i. Then the following are equivalent for all suitable models $\mathfrak{M}$ and assignments s:*

1. *There exists some S such that $\langle M, S\rangle \in \mathbf{D}$ and $\mathfrak{M} \models \chi(S)$;*
2. *There exist $W_1 \dots W_n$ such that $\langle M, (\bigcup_{i=1}^n W_i)\rangle \in \mathbf{D}$ and $\mathfrak{M} \models_s \chi'(W_1 \dots W_n)$.*

PROOF. $1 \Rightarrow 2$ If such a S exists, let $W_1 = W_2 = \cdots W_n = S$. Then since $\bigcup_i W_i = S$ we have that $\langle M, (\bigcup_i W_i)\rangle \in \mathbf{D}$, and since $\chi'(S \dots S)$ is the same as $\chi(S)$ we have that $\mathfrak{M} \models \chi'(W_1 \dots W_n)$.

$2 \Rightarrow 1$ Suppose that such $W_1 \dots W_n$ exist, and let $S = \bigcup_i W_i$. Then by assumption $\langle M, S\rangle \in \mathbf{D}$; and furthermore, since the W_i occur only positively in χ' and $W_i \subseteq S$ for all i we have that $\mathfrak{M} \models \chi'(S \dots S)$, that is, $\mathfrak{M} \models \chi(S)$. ⊣

[25] As per Definition 2.12, in $\mathtt{FO}(\mathcal{D})$, the atoms $\mathbf{Dt}$ for $\mathbf{D} \in \mathcal{D}$ cannot occur negated. Indeed, $\mathtt{FO}(\mathcal{D})$ contains only expressions in Negation Normal Form with first order literals, first-order connectives and atoms $\mathbf{Dt}$.

https://doi.org/10.1017/jsl.2019.12 Published online by Cambridge University Press

$$X = \begin{array}{|c|cc|cccc|cc|}\hline & v_1 & w_1 & v_2 & v_3 & w_2 & w_3 & v_4 & w_4 \\ \hline s_0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 \\ s_1 & 1 & 1 & 1 & 2 & 1 & 2 & 0 & 1 \\ s_2 & 2 & 2 & 0 & 0 & 1 & 1 & 0 & 1 \\ \hline \end{array}$$

$$X \upharpoonright (v_1 = w_1)(v_1) = \{s_0, s_1, s_2\}(v_1) = \{0, 1, 2\}.$$
$$X \upharpoonright (v_2 v_3 = w_2 w_3)(v_2 v_3) = \{s_0, s_1\}(v_2 v_3) = \{\langle 0, 1\rangle, \langle 1, 2\rangle\}.$$
$$X \upharpoonright (v_4 = w_4)(v_4) = \emptyset(v_4) = \emptyset.$$

FIGURE 2. A possible representation of the set of elements $A = \{0, 1, 2\}$, of the set of tuples $B = \{\langle 0, 1\rangle, \langle 1, 2\rangle\}$, and of the empty set $C = \emptyset$ in a single team. The variables v_1 and w_1 encode A, the variables $v_2 v_3$ and $w_2 w_3$ encode B, and the variables v_4 and w_4 encode C. In this way, it is possible to represent empty and nonempty relations within a single team.

In this way, we managed to "bring out" all the dependencies of $\mathcal{D}$ and convert the remaining expression to First-Order Logic. Now we need to show that it is possible to "put them back in".

To do so, we will have to encode multiple relations into a single team. The obvious way to do so would be to fix tuples of variables $\mathbf{v}_1 \dots \mathbf{v}_n$ and let a team X correspond to the sets of tuples $X(\mathbf{v}_i) = \{s(\mathbf{v}_i) : s \in X\}$ for $i = 1 \dots n$; but a problem with this is that, as long as the team X is nonempty, it would not be possible to encode empty sets of tuples. So we will instead use *two* tuples of variables[26] $\mathbf{v}_i$ and $\mathbf{w}_i$ to encode the relation $R_i = X \upharpoonright (\mathbf{v}_i = \mathbf{w}_i)(\mathbf{v}_i) = \{s(\mathbf{v}_i) : s \in X, s(\mathbf{v}_i) = s(\mathbf{w}_i)\}$: see Figure 2 for an illustration.

Using this representation, it is first of all possible to state that the union of the sets of tuples of elements encoded by certain variables satisfy a certain dependency:

LEMMA 3.5. *Let* $\mathbf{D}$ *be any dependency, let* k *be its arity, let* $\mathbf{v}_1 \dots \mathbf{v}_n$, $\mathbf{w}_1 \dots \mathbf{w}_n$ *be tuples of variables such that all* $\mathbf{v}_i$ *and* $\mathbf{w}_i$ *have length equal to the arity of* $\mathbf{D}$, *and let* $\mathbf{D}_\cup(\mathbf{v}_1 \dots \mathbf{v}_n; \mathbf{w}_1 \dots \mathbf{w}_n)$ *be the* $\mathtt{FO}(\mathbf{D})$ *formula*

$$\forall p_1 \dots p_n \forall q \exists \mathbf{z}_0 \mathbf{z}_1 \left(\left(\bigvee_i q = p_i \right) \hookrightarrow \left(\bigwedge_{i \le n} \left(\left(\bigwedge_{j<i} q \neq p_j \wedge q = p_i \right) \hookrightarrow \mathbf{z}_0 \mathbf{z}_1 = \mathbf{v}_i \mathbf{w}_i \right) \wedge (\mathbf{z}_0 = \mathbf{z}_1 \hookrightarrow \mathbf{D}\mathbf{z}_0) \right) \right)$$

where $p_1 \dots p_n$ *and* q *are new, distinct variables and* $\mathbf{z}_0$ *and* $\mathbf{z}_1$ *are tuples of new variables of the same length as the arity of* $\mathbf{D}$.

[26]This is more than what is strictly necessary—a tuple of variables and one extra variable to act as a flag would be enough—but it makes the formulas somewhat simpler, and in this work the number of existential quantifiers is not a concern.

https://doi.org/10.1017/jsl.2019.12 Published online by Cambridge University Press

Then for all suitable models $\mathfrak{M}$ (with at least two elements) and teams X, $\mathfrak{M} \models_X \mathbf{D}_\cup(\mathbf{v}_1 \dots \mathbf{v}_n; \mathbf{w}_1 \dots \mathbf{w}_n)$ if and only if $\langle M, \left(\bigcup_{i=1}^n W_i\right)\rangle \in \mathbf{D}$, where $W_i = X \upharpoonright (\mathbf{v}_i = \mathbf{w}_i)(\mathbf{v}_i) = \{s(\mathbf{v}_i) : s \in X, s(\mathbf{v}_i) = s(\mathbf{w}_i)\}$.

Proof. Suppose that $\mathfrak{M} \models_X \mathbf{D}_\cup(\mathbf{v}_1 \dots \mathbf{v}_n; \mathbf{w}_1 \dots \mathbf{w}_n)$. Then there exist some H and some

$$Y = X[M/p_1 \dots p_n q][H/\mathbf{z}_0\mathbf{z}_1] \upharpoonright \left(\bigvee_i q = p_i\right)$$

such that

1. For all $i = 1 \dots n$ and for all $s \in X[M/p_1 \dots p_n q]$, if i is the smallest index such that $s(q) = s(p_i)$ then $H(s) = \{s(\mathbf{v}_i\mathbf{w}_i)\}$;
2. For $Z = Y \upharpoonright (\mathbf{z}_0 = \mathbf{z}_1)$, $\langle M, Z(\mathbf{z}_0)\rangle \in \mathbf{D}$.

Now, Y is precisely the set of all tuples in $X[M/p_1 \dots p_n q][H/\mathbf{z}_0\mathbf{z}_1]$ such that the value of q is equal to the value of *some* p_i. Therefore, the first condition implies that $Y \upharpoonright (\mathbf{z}_0 = \mathbf{z}_1)(\mathbf{z}_0) = \bigcup_i (Y \upharpoonright (\mathbf{v}_i = \mathbf{w}_i)(\mathbf{v}_i)) = \bigcup_i (X \upharpoonright (\mathbf{v}_i = \mathbf{w}_i)(\mathbf{v}_i))$; and then the second condition implies that $\left\langle M, \left(\bigcup_i X \upharpoonright (\mathbf{v}_i = \mathbf{w}_i)(\mathbf{v}_i)\right)\right\rangle \in \mathbf{D}$, as required.

Conversely, suppose that $\left\langle M, \left(\bigcup_i X \upharpoonright (\mathbf{v}_i = \mathbf{w}_i)(\mathbf{v}_i)\right)\right\rangle \in \mathbf{D}$. Then let 0 and 1 be two distinct elements in M, let $\mathbf{0}$ and $\mathbf{1}$ represent tuples of k zeroes or ones, and let $H : X[M/p_1 \dots p_n q] \to \mathcal{P}(M^{2k})\backslash\{\emptyset\}$ be defined as

$$H(s) = \begin{cases} \{s(\mathbf{v}_i\mathbf{w}_i)\} & \text{if } i \text{ is the smallest index such that } s(q) = s(p_i); \\ \{\mathbf{01}\} & \text{if } s(q) \neq s(p_i)\ \forall i \in \{1 \dots n\}. \end{cases}$$

Then let $Y = X[M/p_1 \dots p_n q][H/\mathbf{z}_0\mathbf{z}_1] \upharpoonright \left(\bigvee_i q = p_i\right)$. By construction, it is clear that

$$\mathfrak{M} \models_Y \left(\bigwedge_{j<i} q \neq p_j \wedge q = p_i\right) \hookrightarrow \mathbf{z}_0\mathbf{z}_1 = \mathbf{v}_i\mathbf{w}_i$$

for all $i = 1 \dots n$. Moreover, $Y(\mathbf{z}_0\mathbf{z}_1) = \bigcup_i X(\mathbf{v}_i\mathbf{w}_i)$, because for each $s \in X$ and for each $i = 1 \dots n$ there exists an assignment $s^i = s[0 \dots 1 \dots 1/p_1 \dots p_n][1/q] \in X[M/p_1 \dots p_n q]$ such that $s^i(p_j)$ is 0 if $j < i$ and 1 otherwise and such that $s^i(q) = 1$, and for such an assignment we have that $H(s^i) = \{s^i(\mathbf{v}_i\mathbf{w}_i)\} = \{s(\mathbf{v}_i\mathbf{w}_i)\}$, as required. Thus we also have that $Y \upharpoonright (\mathbf{z}_0 = \mathbf{z}_1)(\mathbf{z}_0) = \bigcup_i X \upharpoonright (\mathbf{v}_i = \mathbf{w}_i)(\mathbf{v}_i)$. But by assumption we know that $\left\langle M, \bigcup_i X \upharpoonright (\mathbf{v}_i = \mathbf{w}_i)(\mathbf{v}_i)\right\rangle \in \mathbf{D}$; therefore $\langle M, Y \upharpoonright (\mathbf{z}_0 = \mathbf{z}_1)(\mathbf{z}_0)\rangle \in \mathbf{D}$, and thus $\mathfrak{M} \models_Y \mathbf{z}_0 = \mathbf{z}_1 \hookrightarrow \mathbf{D}\mathbf{z}_0$, and finally $\mathfrak{M} \models_X \mathbf{D}_\cup(\mathbf{v}_1 \dots \mathbf{v}_n; \mathbf{w}_1 \dots \mathbf{w}_n)$ as required. ⊣

Finally, we need a way to "put back" the dependencies of $\mathcal{D}$ into the formula. The next two lemmas take care of that:

Lemma 3.6. *Let $\psi(W)$ be a* quantifier-free $\mathtt{FO}(\mathcal{D})$ *formula in which the k-ary relation symbol W occurs at most once and only positively and let $\psi'(\mathbf{v}, \mathbf{w})$ be the formula obtained by replacing the (unique) occurrence $W\mathbf{t}$ of W in ψ with $\mathbf{v} = \mathbf{w} \wedge \mathbf{t} = \mathbf{w}$, where $\mathbf{v}$ and $\mathbf{w}$ are tuples of new variables of length equal to the arity of W. Then, for all suitable models $\mathfrak{M}$, relations W and teams X, $\mathfrak{M} \models_X \psi(W)$ if and only if there exists some H such that*

1. $\mathfrak{M} \models_{X[H/\mathbf{vw}]} \psi'(\mathbf{v}, \mathbf{w})$;
2. $X[H/\mathbf{vw}] \upharpoonright (\mathbf{v} = \mathbf{w})(\mathbf{v}) \subseteq W$.

https://doi.org/10.1017/jsl.2019.12 Published online by Cambridge University Press

PROOF. The proof is by structural induction on ψ.

- If ψ is of the form $W\mathbf{t}$, ψ' is simply $\mathbf{v} = \mathbf{w} \wedge \mathbf{t} = \mathbf{w}$. Now if $\mathfrak{M} \models_X W\mathbf{t}$ it must be that $s(\mathbf{t}) \in W$ for all $s \in X$. Thus, if we define H so that $H(s) = \{s(\mathbf{t})s(\mathbf{t})\}$ for all $s \in X$ we have at once that $X[H/\mathbf{vw}] \upharpoonright (\mathbf{v} = \mathbf{w})(\mathbf{v}) = \{s(\mathbf{t}) : s \in X\} \subseteq W$, and furthermore $\mathfrak{M} \models_{X[H/\mathbf{vw}]} \mathbf{v} = \mathbf{w} \wedge \mathbf{t} = \mathbf{w}$ as required.

  Conversely, suppose that there exists some H such that $X[H/\mathbf{vw}] \upharpoonright (\mathbf{v} = \mathbf{w})(\mathbf{v}) \subseteq W$ and $\mathfrak{M} \models_{X[H/\mathbf{vw}]} \mathbf{v} = \mathbf{w} \wedge \mathbf{t} = \mathbf{w}$. Now, for any $s \in X$, let $\mathbf{m} = s(\mathbf{t})$: then $s[\mathbf{mm}/\mathbf{vw}] \in X[H/\mathbf{vw}] \upharpoonright (\mathbf{v} = \mathbf{w})$, and therefore $\mathbf{m} \in W$. Thus $\mathfrak{M} \models_X W\mathbf{t}$ as required.
- If $\psi(W)$ is of the form $\psi_1 \vee \psi_2$, let us assume—without loss of generality—that W occurs only in ψ_1. Then $\psi'(\mathbf{v}, \mathbf{w})$ is $\psi_1'(\mathbf{v}, \mathbf{w}) \vee \psi_2$. Now suppose that $\mathfrak{M} \models_X \psi_1(W) \vee \psi_2$. Then $X = Y \cup Z$ for two Y, Z such that $\mathfrak{M} \models_Y \psi_1(W)$ and $\mathfrak{M} \models_Z \psi_2$. So by induction hypothesis we know that there exists some $H : Y \to \mathcal{P}(M^{2k})\backslash\{\emptyset\}$ such that $\mathfrak{M} \models_{Y[H/\mathbf{vw}]} \psi_1'(\mathbf{v}, \mathbf{w})$ and $Y[H/\mathbf{vw}] \upharpoonright (\mathbf{v} = \mathbf{w})(\mathbf{v}) \subseteq W$. Now let $\mathbf{m}_1$ and $\mathbf{m}_2$ be two arbitrary, distinct tuples of elements, and let $H' : X \to \mathcal{P}(M^{2k})\backslash\{\emptyset\}$ be defined as

$$H'(s) = \begin{cases} H(s) \cup \{\mathbf{m}_1\mathbf{m}_2\} & \text{if } s \in Y \cap Z; \\ H(s) & \text{if } s \in Y\backslash Z; \\ \{\mathbf{m}_1\mathbf{m}_2\} & \text{if } s \in Z\backslash Y. \end{cases}$$

  Then $X[H'/\mathbf{vw}] = Y[H/\mathbf{vw}] \cup Z[\mathbf{m}_1\mathbf{m}_2/\mathbf{vw}]$, $\mathfrak{M} \models_{Z[\mathbf{m}_1\mathbf{m}_2/\mathbf{vw}]} \psi_2$ by locality, and $X[H'/\mathbf{vw}] \upharpoonright (\mathbf{v} = \mathbf{w})(\mathbf{v}) = Y[H/\mathbf{vw}] \upharpoonright (\mathbf{v} = \mathbf{w})(\mathbf{v}) \subseteq W$. Moreover $\mathfrak{M} \models_{X[H'/\mathbf{vw}]} \psi_1'(\mathbf{v}, \mathbf{w}) \vee \psi_2$, as required.

  Conversely, suppose that there is some H such that $\mathfrak{M} \models_{X[H/\mathbf{vw}]} \psi_1'(\mathbf{v}, \mathbf{w}) \vee \psi_2$ and such that $X[H/\mathbf{vw}] \upharpoonright (\mathbf{v} = \mathbf{w})(\mathbf{v}) \subseteq W$. Then $X[H/\mathbf{vw}] = Y' \cup Z'$ for two Y', Z' such that $\mathfrak{M} \models_{Y'} \psi_1'(\mathbf{v}, \mathbf{w})$ and $\mathfrak{M} \models_{Z'} \psi_2$. Now let $Y = \{s \in X : \exists \mathbf{m}_1\mathbf{m}_2 \text{ s.t. } s[\mathbf{m}_1\mathbf{m}_2/\mathbf{vw}] \in Y'\}$ and let $H' : Y \to \mathcal{P}(M^{2k})\backslash\{\emptyset\}$ be defined as

$$H'(s) = \{\mathbf{m}_1\mathbf{m}_2 \in M^{2k} : s[\mathbf{m}_1\mathbf{m}_2/\mathbf{vw}] \in Y'\}.$$

  Then by construction $H'(s) \neq \emptyset$ for all $s \in Y$, and furthermore $Y[H'/\mathbf{vw}] = Y'$; therefore, $\mathfrak{M} \models_{Y[H'/\mathbf{vw}]} \psi_1'(\mathbf{v}, \mathbf{w})$, and since $Y' \subseteq X[H/\mathbf{vw}]$ we have that

$$\begin{aligned} Y[H'/\mathbf{vw}] \upharpoonright (\mathbf{v} = \mathbf{w})(\mathbf{v}) &= Y' \upharpoonright (\mathbf{v} = \mathbf{w})(\mathbf{v}) \\ &\subseteq X[H/\mathbf{vw}] \upharpoonright (\mathbf{v} = \mathbf{w})(\mathbf{v}) \subseteq W. \end{aligned}$$

  Then by induction hypothesis we have that $\mathfrak{M} \models_Y \psi_1(W)$. Now let $Z = \{s \in X : \exists \mathbf{m}_1\mathbf{m}_2 \text{ s.t. } s[\mathbf{m}_1\mathbf{m}_2/\mathbf{vw}] \in Z'\}$. By locality, we have that $\mathfrak{M} \models_Z \psi_2$, and moreover $X = Y \cup Z$ (since if $s \in X$ then there are some $\mathbf{m}_1\mathbf{m}_2$ such that $s[\mathbf{m}_1\mathbf{m}_2/\mathbf{vw}] \in X[H/\mathbf{vw}] = Y' \cup Z'$) and hence $\mathfrak{M} \models_X \psi_1(W) \vee \psi_2$ as required.
- If $\psi(W)$ is of the form $\psi_1 \wedge \psi_2$, let us again assume that W occurs only in ψ_1. Then $\psi'(\mathbf{v}, \mathbf{w}) = \psi_1'(\mathbf{v}, \mathbf{w}) \wedge \psi_2$. Suppose that $\mathfrak{M} \models_X \psi_1(W) \wedge \psi_2$: then $\mathfrak{M} \models_X \psi_1(W)$ and $\mathfrak{M} \models_X \psi_2$, and by induction hypothesis there exists some H such that $X[H/\mathbf{vw}] \upharpoonright (\mathbf{v} = \mathbf{w})(\mathbf{v}) \subseteq W$ and $\mathfrak{M} \models_{X[H/\mathbf{vw}]} \psi_1'(\mathbf{v}, \mathbf{w})$. Furthermore, by locality, we have that $\mathfrak{M} \models_{X[H/\mathbf{vw}]} \psi_2$, and therefore $\mathfrak{M} \models_{X[H/\mathbf{vw}]} \psi_1'(\mathbf{v}, \mathbf{w}) \wedge \psi_2$ as required.

https://doi.org/10.1017/jsl.2019.12 Published online by Cambridge University Press

Conversely, suppose that there is some H such that $\mathfrak{M} \models_{X[H/\mathbf{vw}]} \psi_1'(\mathbf{v},\mathbf{w}) \wedge \psi_2$ and $X[H/\mathbf{vw}] \upharpoonright (\mathbf{v}=\mathbf{w})(\mathbf{v}) \subseteq W$. Then by locality $\mathfrak{M} \models_X \psi_2$ and by induction hypothesis $\mathfrak{M} \models_X \psi_1(W)$, and hence $\mathfrak{M} \models_X \psi_1(W) \wedge \psi_2$ as required. ⊣

LEMMA 3.7. *Let $\chi(W_1 \dots W_q) = \forall \mathbf{x}_1 \exists \mathbf{y}_1 \dots \forall \mathbf{x}_n \exists \mathbf{y}_n \psi(W_1 \dots W_q)$ be a $\mathtt{FO}(\mathcal{D})$ sentence in Quantifier Normal Form in which the W_i—which all have the same arity k—occur only positively and at most once each, and let $\mathbf{D}$ be a downwards closed dependency of the same arity of the W_i. Then the following are equivalent*:

1. *There exist $W_1 \dots W_q$ such that $\langle M, \bigcup_i W_i \rangle \in \mathbf{D}$ and $\mathfrak{M} \models \chi(W_1 \dots W_q)$;*
2. *$\mathfrak{M} \models \forall \mathbf{x}_1 \dots \exists \mathbf{y}_n \exists \mathbf{v}_1 \mathbf{w}_1 \dots \mathbf{v}_q \mathbf{w}_q (\mathbf{D}_\cup(\mathbf{v}_1 \dots \mathbf{v}_q; \mathbf{w}_1 \dots \mathbf{w}_q) \wedge \psi'(\mathbf{v}_1 \mathbf{w}_1 \dots \mathbf{v}_q \mathbf{w}_q))$, where $\psi'(\mathbf{v}_1 \mathbf{w}_1 \dots \mathbf{v}_q \mathbf{w}_q)$ is obtained—as in the previous lemma—by replacing each $W_i \mathbf{t}_i$ with $\mathbf{v}_i = \mathbf{w}_i \wedge \mathbf{t}_i = \mathbf{w}_i$ for tuples of new variables $\mathbf{v}_i$ and $\mathbf{w}_i$.*

PROOF. Suppose that 1. holds. Then there exist $W_1 \dots W_q$ such that $\langle M, \bigcup_i W_i \rangle \in \mathbf{D}$ and $\mathfrak{M} \models \chi(W_1 \dots W_q)$. Then in particular there exist functions $H_1 \dots H_n$ such that, for $X = \{\epsilon\}[M/\mathbf{x}_1][H_1/\mathbf{y}_1] \dots [M/\mathbf{x}_n][H_n/\mathbf{y}_n]$, $\mathfrak{M} \models_X \psi(W_1 \dots W_q)$. Applying repeatedly Lemma 3.6, we can find $K_1 \dots K_q$ such that, for $Y = X[K_1/\mathbf{v}_1\mathbf{w}_1] \dots [K_q/\mathbf{v}_q\mathbf{w}_q]$,

a) $\mathfrak{M} \models_Y \psi'(\mathbf{v}_1 \mathbf{w}_1 \dots \mathbf{v}_q \mathbf{w}_q)$;
b) For all $i \in 1 \dots q$, $Y \upharpoonright (\mathbf{v}_i = \mathbf{w}_i)(\mathbf{v}_i) \subseteq W_i$.

Now, for all i, let $W_i' = Y \upharpoonright (\mathbf{v}_i = \mathbf{w}_i)(\mathbf{v}_i)$. Then since $\mathbf{D}$ is downwards closed and $\bigcup_i W_i' \subseteq \bigcup_i W_i$ we have that $\mathfrak{M} \models_Y \mathbf{D}(\bigcup_i W_i')$, and hence $\mathfrak{M} \models_Y \mathbf{D}_\cup(\mathbf{v}_1 \dots \mathbf{v}_q; \mathbf{w}_1 \dots \mathbf{w}_q)$ by Lemma 3.5. This proves that

$$\mathfrak{M} \models_X \exists \mathbf{v}_1 \mathbf{w}_1 \dots \mathbf{v}_q \mathbf{w}_q (\mathbf{D}_\cup(\mathbf{v}_1 \dots \mathbf{v}_q; \mathbf{w}_1 \dots \mathbf{w}_q) \wedge \psi'(\mathbf{v}_1 \mathbf{w}_1 \dots \mathbf{v}_q \mathbf{w}_q))$$

and hence that 2. holds.

Conversely, suppose that 2. holds. Then there exist functions $H_1 \dots H_n$, $K_1 \dots K_q$ such that, for $X = \{\epsilon\}[M/\mathbf{x}_1][H_1/\mathbf{y}_1] \dots [M/\mathbf{x}_n][H_n/\mathbf{y}_n]$ and for $Y = X[K_1/\mathbf{v}_1\mathbf{w}_1] \dots [K_q/\mathbf{v}_q\mathbf{w}_q]$, $\mathfrak{M} \models_Y \mathbf{D}_\cup(\mathbf{v}_1 \dots \mathbf{v}_q; \mathbf{w}_1 \dots \mathbf{w}_q) \wedge \psi'(\mathbf{v}_1 \mathbf{w}_1 \dots \mathbf{v}_q \mathbf{w}_q)$. Now, for each $i = 1 \dots q$, let $W_i = Y \upharpoonright (\mathbf{v}_i = \mathbf{w}_i)(\mathbf{v}_i)$. Then by Lemma 3.5 we have that $\langle M, \bigcup_i W_i \rangle \in \mathbf{D}$; and furthermore, by Lemma 3.6 $\mathfrak{M} \models_Y \psi(W_1 \dots W_n)$. But then by locality $\mathfrak{M} \models_X \psi(W_1 \dots W_n)$, and hence $\mathfrak{M} \models \chi(W_1 \dots W_n)$ and 1. holds. ⊣

We can now finally prove the main result of this section:

THEOREM 3.8. *Let $\mathcal{D}$ be any family of downwards closed dependencies and let $\mathcal{S}$ be any strongly first-order family of dependencies. Then $\mathcal{S}$ is safe for $\mathcal{D}$.*

PROOF. Let ϕ be a sentence of $\mathtt{FO}(\mathcal{S}, \mathcal{D})$. Then, by Lemmas 3.3 and 3.4, there exist a first-order formula $\theta(W_1^1 \dots W_1^{r(1)}, \dots, W_k^1 \dots W_k^{r(k)})$, in which each W_i^j occurs only once, and dependencies $\mathbf{D}_{j(1)} \dots \mathbf{D}_{j(k)} \in \mathcal{D}$ such that the following are equivalent for all models $\mathfrak{M}$:

A) $\mathfrak{M} \models \phi$;
B) There exist relations $W_1^1 \dots W_1^{r(1)}, \dots, W_k^1 \dots W_k^{r(k)}$ such that
   - For all $i = 1 \dots k$, $\left\langle M, \bigcup_{t=1}^{r(i)} W_i^t \right\rangle \in \mathbf{D}_{j(i)}$;
   - $\mathfrak{M} \models \theta(W_1^1 \dots W_1^{r(1)}, \dots, W_k^1 \dots W_k^{r(k)})$.

Now let $\theta(W_1^1 \dots W_1^{r(1)}, \dots, W_k^1 \dots W_k^{r(k)})$ have Quantifier Normal Form

$$\forall \mathbf{x}_1 \exists \mathbf{y}_1 \dots \forall \mathbf{x}_l \exists \mathbf{y}_l \psi(W_1^1 \dots W_1^{r(1)}, \dots, W_k^1 \dots W_k^{r(n)}).$$

https://doi.org/10.1017/jsl.2019.12 Published online by Cambridge University Press

But applying repeatedly Lemma 3.7 and the fact that all $\mathbf{D}_{j(i)}$ are downwards closed[27] we can see that there exists some $\mathtt{FO}(\mathcal{D})$ sentence

$$\begin{aligned}\phi^* :=& \forall \mathbf{x}_1 \exists \mathbf{y}_1 \ldots \forall \mathbf{x}_l \exists \mathbf{y}_l \; \exists \mathbf{v}_1^1 \mathbf{w}_1^1 \ldots \mathbf{v}_1^{r(1)} \mathbf{w}_1^{r(1)} \ldots \mathbf{v}_k^1 \mathbf{w}_k^1 \ldots \mathbf{v}_k^{r(k)} \mathbf{w}_k^{r(k)} \\ & ((\bigwedge_i (\mathbf{D}_{j(i)})_\cup (\mathbf{v}_i^1 \ldots \mathbf{v}_i^{r(i)}, \mathbf{w}_i^1 \ldots \mathbf{w}_i^{r(i)})) \wedge \\ & \psi'(\mathbf{v}_1^1 \mathbf{w}_1^1 \ldots \mathbf{v}_1^{r(1)} \mathbf{w}_1^{r(1)} \ldots \mathbf{v}_k^1 \mathbf{w}_k^1 \ldots \mathbf{v}_k^{r(k)} \mathbf{w}_k^{r(k)}))\end{aligned}$$

which is true if and only if B) above holds. Thus, every $\mathtt{FO}(\mathcal{S}, \mathcal{D})$ sentence is equivalent to some $\mathtt{FO}(\mathcal{D})$ sentence - in other words, $\mathcal{S}$ is safe for $\mathcal{D}$. ⊣

§4. Characterizing strongly first-order, relativizable, downwards closed dependencies. We can now prove that all dependencies that are nontrivial, relativizable, downwards closed and strongly first order are definable in terms of constancy. To do so, we will need to use the Chang-Makkai Theorem, applying it not to unary relations (as is presented e.g., in Theorem 5.3.6 of [5]) but to k-ary relations (as mentioned in [4], the proof carries over to this case without problems. For reference, the Appendix contains the detailed proof of the case that interests us). On the other hand, it suffices to consider countable structures and vocabularies containing only a relation symbol R.

So this is the variant of the Chang-Makkai Theorem that we need:

THEOREM 4.1. *Let $\Phi(R)$ be a first-order sentence on the vocabulary $\{R\}$, where R is k-ary. The following are equivalent*:

1. *For every countable model $\mathfrak{M}$ over the signature $\{R\}$, $|\{R : R \subseteq M^k, \mathfrak{M} \models \Phi(R)\}| < \aleph_1$;*
2. *There are a finite number of formulas $\theta_1(\mathbf{x}, \mathbf{z}) \ldots \theta_n(\mathbf{x}, \mathbf{z})$, over the empty vocabulary, such that*

$$\phi(R) \models \bigvee_{i=1}^{n} \exists \mathbf{z} \forall \mathbf{x} (R\mathbf{x} \leftrightarrow \theta_i(\mathbf{x}, \mathbf{z})).$$

The idea of this section's main proof is to show that if $\mathbf{D}$ is strongly first order then the property $\mathbf{D}_{\mathbf{max}}$ of being maximal among the R that satisfy $\mathbf{D}$ is also strongly first order; that whenever $\langle M, R\rangle \in \mathbf{D}$ there is at least one $R' \supseteq R$ such that $\langle M, R'\rangle \in \mathbf{D}_{\mathbf{max}}$; and that for every countable model there must be a countable number of such maximal R'. Then we use the above version of the Chang-Makkai Theorem to show that these maximal sets are definable over the empty signature, and use the downwards closure property to define the $\mathbf{D}$ atoms via Boolean disjunctions (that is, unraveling Definition 2.23, via constancy atoms).

PROPOSITION 4.2. *Let $\mathbf{D}$ be a downwards closed, strongly first-order dependency. Then there exists a first-order dependency $\mathbf{D}_{max}$ such that $\langle M, R\rangle \in \mathbf{D}_{max}$ if and only if $\langle M, R\rangle \in \mathbf{D}$ and there is no $S \supsetneq R$ such that $\langle M, S\rangle \in \mathbf{D}$. Furthermore, $\mathbf{D}_{max}$ is strongly first order itself.*

[27] For the applications of Lemma 3.7 after the first one we can treat the $(\mathbf{D}_{j(i)})_\cup$ already introduced as ordinary dependency atoms, so that the expression is still in Quantifier Normal Form.

https://doi.org/10.1017/jsl.2019.12 Published online by Cambridge University Press

PROOF. Since $\mathbf{D}$ is strongly first order, the $\mathtt{FO}(\mathbf{D})$ sentence

$$\phi(R) := \exists \mathbf{x}(\neg R\mathbf{x} \wedge \forall \mathbf{y}((\neg R\mathbf{y} \wedge \mathbf{x} \neq \mathbf{y}) \vee \mathbf{D}\mathbf{y}))$$

is equivalent to some first-order sentence. Observe that $\mathfrak{M} \models \phi(R)$ if and only if there exists some S such that $R \subsetneq S$ and $\langle M, S\rangle \in \mathbf{D}$. Indeed, suppose that such a S exists and let $\mathbf{a} \in S \backslash R$. Then, choosing $\mathbf{a}$ as the only value for $\mathbf{x}$,[28] we have that $\mathfrak{M} \models_{\{(\mathbf{x}:\mathbf{a})\}} \neg R\mathbf{x}$. Now let $X = \{(\mathbf{x} : \mathbf{a})\}[M/\mathbf{y}]$, and split it into $Z = \{s \in X : s(\mathbf{y}) \in S\}$ and $Y = X \backslash Z = \{s \in X : s(\mathbf{y}) \notin S\}$. Clearly $\mathfrak{M} \models_Z \mathbf{D}\mathbf{y}$, because $Z(\mathbf{y}) = S$. Moreover, for all $s \in Y$ we have that $s(\mathbf{y}) \notin R$, because $s(\mathbf{y}) \notin S$ and $R \subseteq S$, and that $s(\mathbf{y}) \neq s(\mathbf{x})$, because $s(\mathbf{x}) = \mathbf{a} \in S$. Thus we have that $\mathfrak{M} \models_Y \neg R\mathbf{y} \wedge \mathbf{x} \neq \mathbf{y}$, as required.

Conversely, suppose that $\mathfrak{M} \models_{\{\epsilon\}} \phi(R)$, where ϵ is the empty assignment. Then there exists some $X = \{\epsilon\}[K/\mathbf{x}][M/\mathbf{y}]$ such that $X(\mathbf{x}) \cap R = \emptyset$ and $\mathfrak{M} \models_X (\neg R\mathbf{y} \wedge \mathbf{x} \neq \mathbf{y}) \vee \mathbf{D}\mathbf{y}$. Thus, $X = Y \cup Z$ for two Y, Z such that $\mathfrak{M} \models_Y (\neg R\mathbf{y} \wedge \mathbf{x} \neq \mathbf{y})$ and $\mathfrak{M} \models_Z \mathbf{D}\mathbf{y}$. Let $S = Z(\mathbf{y})$: then $\langle M, S\rangle \in \mathbf{D}$. Furthermore, R is contained in S: indeed, for all $\mathbf{m} \in R$ there is some $s \in X$ is such that $s(\mathbf{y}) = \mathbf{m}$, and then $s \in Z$, because $\mathfrak{M} \models_Y \neg R\mathbf{y}$ and $X = Y \cup Z$. Finally, S is not contained in R: indeed, if $\mathbf{a}$ is a possible value for $\mathbf{x}$ in X then $\mathbf{a} \notin R$ and $\mathbf{a} \in S$ (because the assignment $(\mathbf{xy} : \mathbf{aa})$ is in X but not in Y).

If $\phi(R)$ is equivalent to a first-order sentence, the same is true of its negation. Let $\psi(R)$ be the first-order sentence equivalent to the negation of $\phi(R)$: then $\mathfrak{M} \models \psi(R)$ if and only if there is no $S \supsetneq R$ such that $\langle M, S\rangle \in \mathbf{D}$. Thus $\mathbf{D}_{\mathbf{max}}(R)$ is equivalent to $\mathbf{D}(R) \wedge \psi(R)$, and therefore it is first order. It remains to show that it is *strongly* first order as a dependency.

Now $\psi(R)$ is upwards closed in R: indeed, if there is no $S \supsetneq R$ such that $\mathbf{D}(S)$ and $R \subseteq R'$ then clearly there is no $S \supsetneq R'$ such that $\mathbf{D}(S)$ either. By Theorem 2.27, upwards closed dependencies are strongly first order; therefore, the dependency $\mathbf{E}(R) = \{\langle M, R\rangle : \langle M, R\rangle \models \psi(R)\}$ is strongly first order.

But by Theorem 3.8, strongly first-order dependencies are safe for downwards closed dependencies. So in particular $\mathbf{E}$ is safe for $\mathbf{D}$, and the atom $\mathbf{D}_{\mathbf{max}}(\mathbf{t})$ is definable in $\mathtt{FO}(\mathbf{D}, \mathbf{E})$ as $\mathbf{D}_{\mathbf{max}}(\mathbf{t}) := \mathbf{D}\mathbf{t} \wedge \mathbf{E}\mathbf{t}$. Therefore every sentence $\phi \in \mathtt{FO}(\mathbf{D}_{\mathbf{max}})$ is equivalent to some sentence $\phi' \in \mathtt{FO}(\mathbf{D}, \mathbf{E})$, which by the safety of $\mathbf{E}$ for $\mathbf{D}$ is equivalent to some sentence $\phi'' \in \mathtt{FO}(\mathbf{D})$, which by the strongly first orderness of $\mathbf{D}$ is equivalent to some first-order sentence. Therefore $\mathbf{D}_{\mathbf{max}}$ is strongly first order, as required. ⊣

The following consequence of the above proposition will also be necessary for our proof:

COROLLARY 4.3. *Let $\mathbf{D}(R)$ be a downwards closed, strongly first-order dependency and let M, R be such that $\langle M, R\rangle \in \mathbf{D}$. Then there exists a $S \supseteq R$ such that $\langle M, S\rangle \in \mathbf{D}_{max}$.*

PROOF. Consider the $\mathtt{FO}(\mathbf{D}_{\mathbf{max}})$ sentence

$$\theta(T) := \forall \mathbf{x}(\neg T\mathbf{x} \vee \mathbf{D}_{\mathbf{max}}\mathbf{x}).$$

It is easy to check that $\mathfrak{M} \models \theta(T)$ if and only if there exists some $K \supseteq T$ such that $\langle M, K\rangle \in \mathbf{D}_{\mathbf{max}}$; and since, as we just saw, $\mathbf{D}_{\mathbf{max}}$ is strongly first order, $\theta(T)$

[28] More precisely: picking $\mathbf{x}$ according to the function $H : \{\epsilon\} \to \mathcal{P}(M) \backslash \{\emptyset\}$ such that $H(\epsilon) = \mathbf{a}$.

https://doi.org/10.1017/jsl.2019.12 Published online by Cambridge University Press

is equivalent to some first-order sentence. Let $\chi(T)$ be the first-order sentence equivalent to the negation of $\theta(T)$: then $\mathfrak{M} \models \chi(T)$ if and only if there is no $K \supseteq T$ such that $\langle M, K\rangle \in \mathbf{D}_{\mathbf{max}}$. Furthermore, $\chi(T)$ is clearly upwards closed in T, since if there are no $K \supseteq T$ with $\langle M, K\rangle \in \mathbf{D}_{\mathbf{max}}$ and $T \subseteq T'$ then there are no $K \supseteq T'$ with $\langle M, K\rangle \in \mathbf{D}_{\mathbf{max}}$ either.

Therefore by Theorem 2.27 the dependency $\mathbf{F} = \{\langle M, T\rangle : \langle M, T\rangle \models \chi(T)\}$ is strongly first order. Consider now then the $\mathtt{FO}(\mathbf{D}, \mathbf{F})$ sentence on the empty signature

$$\exists \mathbf{x}(\mathbf{Dx} \wedge \mathbf{Fx}).$$

By construction, this is true if and only if there exists a relation R which satisfies $\mathbf{D}(R)$ but is contained in no maximal superset; and since strongly first-order dependencies are safe for downwards closed dependencies by Theorem 3.8, this is equivalent to some $\mathtt{FO}(\mathbf{D})$ sentence and thus—since $\mathbf{D}$ is strongly first order itself—to some first-order sentence. Therefore, there exists a first-order sentence η over the empty vocabulary such that $\mathfrak{M} \models \eta$ if and only if there is some relation R with domain M which satisfies $\mathbf{D}(R)$ but which is not contained in any R' satisfying $\mathbf{D}_{\mathbf{max}}(R)$.

This is clearly not true if the model $\mathfrak{M}$ is finite, so for all finite models $\mathfrak{M}$ we have that $\mathfrak{M} \models \neg\eta$. But then, since $\neg\eta$ is a first-order sentence over the empty signature, by compactness the same is true for all models, finite or infinite.

By the semantics of η this implies that whenever $\langle M, R\rangle \in \mathbf{D}$ there exists some $R' \supseteq R$ such that $\langle M, R'\rangle \in \mathbf{D}_{\mathbf{max}}$, as required. ⊣

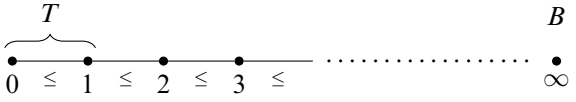
As a quick aside, it is perhaps worth briefly pointing out here that this corollary fails for arbitrary first-order dependencies: if $\langle M, R\rangle \in \mathbf{D}$, it is not necessarily the case that there exists some maximal $R' \supseteq R$ such that $\langle M, R'\rangle \in \mathbf{D}$. As a counterexample, let $\mathbf{D}(R)$ be the 4-ary dependency which is defined by the conjunction of the following axioms:

- If we write "$x \leq y$" for $\exists zuRxyzu$, the relation $\leq$ is a linear order with endpoints;
- If we write "Bz" for $\exists xyuRxyzu$, B is not true of the starting point of the linear order, it is true of some element, and whenever it is true of some element it is true of its immediate predecessor along $\leq$ (if any exists);
- If we write "Tu" for $\exists xyzRxyzu$, there exists some element a such that $\neg Ba$ and such that for all b, Tb if and only if $b \leq a$ in the above order.

Then consider a model $\mathfrak{M}$ with domain $\mathbb{N} \cup \{\infty\}$ and the team X with domain $\{x, y, z, w\}$ defined as

$$\{s : s(x) \leq s(y), s(z) = \infty \text{ and } s(u) \leq 1\}$$

corresponding to the graphical representation



Then it is easy to check that $\langle M, X(\mathbf{t})\rangle \in \mathbf{D}$, but no maximal $R' \supseteq X(\mathbf{t})$ satisfying $\mathbf{D}$ exists: the range of T (that is, the projection of the relation over the fourth element) can be extended to any set of the form $\{n \in \mathbb{N} : n \leq m\}$ for any $m \in \mathbb{N}$ but not to $\mathbb{N}$ or to $\mathbb{N} \cup \{\infty\}$, while nothing can be added to the linear order $\leq$ (that

https://doi.org/10.1017/jsl.2019.12 Published online by Cambridge University Press

is, to the projection over the first two elements) or to B (that is, to the projection over the third element) without violating our axioms.

The fact that whenever $\mathfrak{M} \models_X \mathbf{Dt}$ there exists some maximal $X' \supseteq X$ such that $\mathfrak{M} \models_{X'} \mathbf{Dt}$, therefore, is one that does not hold for all first-order dependencies (although it is easy to check that it holds for functional dependence, independence and inclusion, and although we just proved that it holds for all downwards closed strongly first-order dependencies).

PROPOSITION 4.4. *Let $A = \{a_1, a_2, \ldots\}$ be a countable set of elements and let* $\mathbf{D}$ *be a strongly first order, downwards closed, relativizable dependency of arity k. Then there exist only countably many $R \subseteq A^k$ such that $\langle A, R\rangle \in \mathbf{D}_{max}$.*

PROOF. Suppose that there exist uncountably many such R. Then let I be a $(k+1)$-ary relation symbol, let P be a unary symbol, and consider the first-order theory T containing the following axioms:

(a) $\exists x_1 \ldots x_n \left(\bigwedge_{i<j\leq n} x_i \neq x_j \wedge \bigwedge_{i=1}^{n} Px_i\right)$, for all $n \in \mathbb{N}$;
(b) $\forall \mathbf{z} \forall q (I\mathbf{z}q \rightarrow \bigwedge_{z_i \in \mathbf{z}} Pz_i)$;
(c) $\forall q \mathbf{D}^P_{\mathbf{max}}(I_q)$, where $\mathbf{D}^P_{\mathbf{max}}(I_q)$ is obtained from $\mathbf{D}^P_{\mathbf{max}}(R)$ (that is, from the relativization of $\mathbf{D}_{\mathbf{max}}(R)$ to P) by replacing every $R\mathbf{v}$ with $I\mathbf{v}q$;
(d) $\neg\chi(I, P)$, where $\chi(I, P)$ is the first-order sentence equivalent[29] to the $\mathtt{FO}(\mathbf{D}^P)$ sentence $\forall q \exists \mathbf{z}(\neg I\mathbf{z}q \wedge \mathbf{D}^P\mathbf{z})$.

I state that if our assumption is true, the above theory—that is first order and has a finite vocabulary—has only uncountable models. This is clearly impossible by the Löwenheim-Skolem Theorem, and therefore our assumption cannot hold.

Indeed, suppose that $\mathfrak{M} = \langle M, P, I\rangle$ is a countable model of T. Then by (a) P is infinite and countable, and we can assume that it is A up to isomorphism; by (b) and (c), for every $m \in M$ the relation $I_m = \{\mathbf{a} \in M^k : \langle \mathbf{a}, m\rangle \in I\}$ is in P^k and satisfies $\mathbf{D}^P_{\mathbf{max}}$ in M, that is to say satisfies $\mathbf{D}_{\mathbf{max}}$ in P; and by (d), as we will see, *every* relation that satisfies $\mathbf{D}_{\mathbf{max}}$ in P is equal to I_m for some $m \in M$. But we said that A—and, therefore, P—contains uncountably many distinct R which satisfy $\mathbf{D}_{\mathbf{max}}(R)$ in it, and therefore this is impossible.

On the other hand—again, postponing the verification that (d) holds if and only if I enumerates all R that satisfy $\mathbf{D}_{\mathbf{max}}(R)$ in P—there certainly exists an uncountable model $\mathfrak{M} = \langle M, P, I\rangle$ of the above theory: let M be a set of cardinality $2^{\aleph_0}$ containing A, let the interpretation of P be A itself, and let I_m range over all (uncountably many, but certainly no more than $|M| = 2^{|A|}$) $R \subseteq P^k$ such that $\langle P, R\rangle \in \mathbf{D}_{\mathbf{max}}$ as m ranges over M. Thus our theory has an uncountable model but no countable models, which is impossible.

We now verify that (d) holds if and only if I_m enumerates all relations that satisfy $\mathbf{D}_{\mathbf{max}}$ with respect to P, that is, that the $\mathtt{FO}(\mathbf{D}^P)$ sentence $\xi := \forall q \exists \mathbf{z}(\neg I\mathbf{z}q \wedge \mathbf{D}^P\mathbf{z})$ is true if and only if I_m does *not* enumerate all such subsets as m ranges over M.

- *Left to right:* Suppose that $\mathfrak{M} \models \xi$. Then there exists a team $X = \{\epsilon\}[M/q][H/\mathbf{z}]$, for some function H, such that $\mathfrak{M} \models_X \neg I\mathbf{z}q \wedge \mathbf{D}^P\mathbf{z}$. Now let $R = X(\mathbf{z})$. From the second conjunct, we have at once that $R \subseteq P^k$ and

[29] Such a first-order sentence exists because $\mathbf{D}$ is strongly first order *and* relativizable. Note that we do not need $\mathbf{D}_{\mathbf{max}}$ to be relativizable—for (c) we only need that $\mathbf{D}^P_{\mathbf{max}}(R)$ is a first-order sentence, which is certainly true since $\mathbf{D}_{\mathbf{max}}(R)$ is a first-order sentence—but just that $\mathbf{D}$ is relativizable.

https://doi.org/10.1017/jsl.2019.12 Published online by Cambridge University Press

that $\langle P, R\rangle \in \mathbf{D}$; and for every $q \in M$ there exists some assignment $s \in X$ for which $s(q) = m$, and thus from the first conjunct we have that $\langle s(\mathbf{z}), m\rangle \notin I$, that is $s(\mathbf{z}) \notin I_m$. Thus, $\langle P, R\rangle \in \mathbf{D}$ but R is not contained in any I_m; and since by Corollary 4.3 this R must be contained in some relation that satisfies $\mathbf{D_{max}}$ in P, it follows that I_m does not enumerate all maximal relations.

- *Right to left:* Suppose that I_m does not enumerate all subsets of P^k which satisfy $\mathbf{D_{max}}$ in P for m ranging in M. Then in particular there exists some $R \subseteq P^k$ such that $\langle P, R\rangle \in \mathbf{D_{max}}$ and $R \backslash I_m \neq \emptyset$ for all $m \in M$. Now consider the following function $H : \{\epsilon\}[M/q] \to \mathcal{P}(M^k)\backslash\{\emptyset\}$:

$$H(s) = R\backslash I_(s(q)) = \{\mathbf{a} \in R : \langle \mathbf{a}, s(q)\rangle \notin I\}.$$

As we just saw, $H(s) \neq \emptyset$ for all $s \in \{\epsilon\}[M/q]$. For $X = \{\epsilon\}[M/q][H/\mathbf{z}]$, we clearly have that $\mathfrak{M} \models_X \neg I\mathbf{z}q$. Furthermore, $\mathfrak{M} \models_X \mathbf{D}^P\mathbf{z}$: indeed, $X(\mathbf{z}) \subseteq R$ by construction, $\langle P, R\rangle \in \mathbf{D}$ since $\langle P, R\rangle \in \mathbf{D_{max}}$, thus $\langle P, X(\mathbf{z})\rangle \in \mathbf{D}$ by downwards closure and finally $\mathfrak{M} \models_X \mathbf{D}^P\mathbf{z}$. Therefore $\mathfrak{M} \models \xi$ as required. ⊣

At this point, our conclusion follows from the Chang-Makkai Theorem.

THEOREM 4.5. *Let* $\mathbf{D}(R)$ *be a downwards closed, strongly first-order, relativizable dependency that has the empty team property (that is, it is satisfied by the empty team).*[30] *Then it is definable in* $\mathtt{FO}(=(\cdot))$.

PROOF. By Proposition 4.4, $|\{R : R \subseteq M^k, \langle M, R\rangle \in \mathbf{D_{max}}\}|$ is countable for all countable models $\mathfrak{M}$. Thus, by Theorem 2 and Proposition 4.2, there exist a finite number of first-order formulas $\theta_1(\mathbf{x}, \mathbf{z}) \ldots \theta_n(\mathbf{x}, \mathbf{z})$ over the empty vocabulary (and in which we can assume—via variable renaming—that the variables of $\mathbf{xz}$ are not being quantified over), such that

$$\mathbf{D_{max}}(R) \models \bigvee_{i=1}^{n} \exists\mathbf{z}\forall\mathbf{x}(R\mathbf{x} \leftrightarrow \theta_i(\mathbf{x}, \mathbf{z})).$$

For each i, let $\chi_i(\mathbf{z}) := \mathbf{D}(\theta_i(_, \mathbf{z}))$ be the first-order formula obtained from $\mathbf{D}(R)$ by replacing every occurrence $R\mathbf{t}$ of R with $\theta_i(\mathbf{t}, \mathbf{z})$, so that $\mathfrak{M} \models_{(\mathbf{z}:\mathbf{m})} \chi_i(\mathbf{z})$ if and only if $\langle M, R\rangle \in \mathbf{D}$ for $R = \{\mathbf{a} : \mathfrak{M} \models \theta_i(\mathbf{a}, \mathbf{m})\}$.

Then I state that $\mathbf{Dv}$ is equivalent to the $\mathtt{FO}(=(\cdot), \sqcup)$ formula

$$\bigsqcup_{i=1}^{n} \exists\mathbf{z}(=(\mathbf{z}) \wedge \chi_i(\mathbf{z}) \wedge \theta_i(\mathbf{v}, \mathbf{z})). \tag{1}$$

Where $\sqcup$ represents the *Boolean disjunction* of Definition 2.23, which is definable in terms of constancy atoms.

Indeed, suppose that $\mathfrak{M} \models_X \mathbf{Dv}$. Then for $R = X(\mathbf{v})$ we have that $\langle M, R\rangle \in \mathbf{D}$, and thus by Corollary 4.3 there exists some $R' \supseteq R$ such that $\langle M, R'\rangle \in \mathbf{D_{max}}$. Thus, there exists some $i = 1 \ldots n$ and a fixed tuple $\mathbf{m}$ of elements of M such that $\mathfrak{M} \models_{(\mathbf{z}:\mathbf{m})} \forall\mathbf{x}(R'\mathbf{x} \leftrightarrow \theta_i(\mathbf{x}, \mathbf{z}))$. Since $\langle M, R'\rangle \in \mathbf{D}$, $\mathfrak{M} \models_{(\mathbf{z}:\mathbf{m})} \chi_i(\mathbf{z})$. Since $R \subseteq R'$, $\mathfrak{M} \models_{(\mathbf{z}:\mathbf{m})} \forall\mathbf{x}(R\mathbf{x} \to \theta_i(\mathbf{x}, \mathbf{z}))$ and therefore by Proposition 2.7 $\mathfrak{M} \models_{X[\mathbf{m}/\mathbf{z}]} \chi_i(\mathbf{z}) \wedge$

[30] The only downwards closed dependency that does not have this property is the trivial one that is false for all teams.

https://doi.org/10.1017/jsl.2019.12 Published online by Cambridge University Press

$\theta_i(\mathbf{v}, \mathbf{z})$. Moreover, clearly $\mathfrak{M} \models_{X[\mathbf{m}/\mathbf{z}]} =(\mathbf{z})$, and so $\mathfrak{M} \models_{X[\mathbf{m}/\mathbf{z}]} =(\mathbf{z}) \wedge \chi_i(\mathbf{z}) \wedge \theta_i(\mathbf{v}, \mathbf{z})$, and finally Equation (1) holds in X.

Conversely, suppose that (1) is true in X. Then there exists some i and some tuple $\mathbf{m}$ of elements such that $\mathfrak{M} \models_{X[\mathbf{m}/\mathbf{z}]} \chi_i(\mathbf{z})$ and $\mathfrak{M} \models_{X[\mathbf{m}/\mathbf{z}]} \theta_i(\mathbf{v}, \mathbf{z})$. From the first statement, we obtain at once that the set $R' = \{\mathbf{a} : \mathfrak{M} \models \theta_i(\mathbf{a}, \mathbf{m})\}$ is such that $\langle M, R' \rangle \in \mathbf{D}$. From the second one, we obtain that $R = X(\mathbf{v})$ is contained in R'. By the downwards closure of $\mathbf{D}$ we can then conclude that $\langle M, R \rangle \in \mathbf{D}$, that is $\mathfrak{M} \models_X \mathbf{Dv}$. ⊣

§5. Conclusions and further work. In this article, a characterization of downwards closed, relativizable, strongly first-order dependencies was found by showing that—aside from the trivially false one—these are precisely the ones that are definable in terms of constancy atoms. This is a special case of *Conjecture 1*, according to which all strongly first-order dependencies are definable in terms of constancy atoms *and upwards closed dependencies*; and it is the hope of the author that the techniques developed in this work may be adapted to generalize this result, ideally all the way to a full proof of *Conjecture 1*.

A reasonable enough starting point could be to attempt to get rid, or alternatively prove the necessity, of the requirement of relativizability, for instance by showing that all strongly first-order dependencies (or at least all downwards closed ones) are relativizable anyway.

Another possible direction in which the present work can be expanded could be to consider not only logics $\mathtt{FO}(\mathcal{D})$ obtained by adding dependency atoms to First-Order Logic, but more in general logics $\mathcal{L}(\mathcal{D})$ where $\mathcal{L}$ can be based on arbitrary choices of connectives and operators (interpreted in Team Semantics). Much like studying the more general notion of safety gave us in this work the tools necessary to prove the above mentioned result about strongly first-order dependencies, it is possible that studying the notions of safety and strong first orderness in the more general $\mathcal{L}(\mathcal{D})$ case may give us the tools for solving the $\mathtt{FO}(\mathcal{D})$ one; and, moreover, such an investigation would connect the research program to which this work belongs to the related area of the study of generalized quantifiers in Team Semantics [2, 10, 11, 35].

§6. Appendix: The Chang-Makkai theorem for non-unary relations and countable models. Here I report for convenience a proof of Theorem 2. It is no different from the usual proof (via recursive saturation) of the Chang-Makkai Theorem for countable models; but I recall it here anyway to verify that it works even for k-ary relations, since this theorem is typically presented only for unary relations.

We need to prove the equivalence of the following two statements:

1. For every countable model $\mathfrak{M}$ over the signature $\{R\}$, $|\{R : R \subseteq M^k, \mathfrak{M} \models \phi(R)\}| < \aleph_1$;
2. There are a finite number of formulas $\theta_1(\mathbf{x}, \mathbf{z}) \dots \theta_n(\mathbf{x}, \mathbf{z})$, over the empty vocabulary, such that

$$\phi(R) \models \bigvee_i \exists \mathbf{z} \forall \mathbf{x} (R\mathbf{x} \leftrightarrow \theta_i(\mathbf{x}, \mathbf{z})). \tag{2}$$

https://doi.org/10.1017/jsl.2019.12 Published online by Cambridge University Press

The direction from 2. to 1. is obvious. For the direction from 1. to 2., we reason as follows. Suppose that 2. fails. Then the theory

$$\{\phi(R)\} \cup \{\forall \mathbf{z} \neg \forall \mathbf{x}(R\mathbf{x} \leftrightarrow \theta(\mathbf{x}, \mathbf{z})) : \theta(\mathbf{x}, \mathbf{z}) \in \mathtt{FO}\}, \tag{3}$$

where $\mathbf{z}$ ranges over all tuples of all lengths and $\theta(\mathbf{x}, \mathbf{z})$ ranges over all first-order formulas over the empty signature, is satisfiable. Let $\mathfrak{M} = \langle M, R\rangle$ be a countable, recursively saturated model for it. Let k be the arity of R, and let $(\mathbf{t}(i) : i \in \mathbb{N})$ be an enumeration of all k-tuples of elements in M. We will show that there exist $2^{\aleph_0}$ distinct $R' \subseteq M^k$ such that $\langle M, R'\rangle$ is isomorphic to $\langle M, R\rangle$; and since $\mathfrak{M} \models \phi(R)$, this in particular will show that 1. fails.

The idea of the proof is to define two functions G and H, sending—for every $n \in \mathbb{N}$—every function $f : \{0 \dots n-1\} \to \{0, 1\}$ into some $G(f), H(f) \in \{0 \dots n-1\} \to M^k$ such that

(i) If $f \subseteq g$ then $G(f) \subseteq G(g)$ and $H(f) \subseteq H(g)$;

(ii) For all $f : \{0 \dots n-1\} \to \{0, 1\}$,

$$\langle M, G(f)\rangle \equiv \langle M, H(f)\rangle,$$

where we write $G(f)$ as a shorthand for the tuple of elements $(G(f)(i)_j : i \in 0 \dots n-1, j \in 1 \dots k)$ and similarly for $H(f)$;

(iii) If $\mathtt{Dom}(f) = \{0 \dots n-1\}$ for n of the form $3m$ for some $m \in \mathbb{N}$ then

$$G(f \cup \{\langle n, 0\rangle\}) = G(f \cup \{\langle n, 1\rangle\}) = G(f) \cup \{\langle n, \mathbf{t}(m)\rangle\};$$

(iv) If $\mathtt{Dom}(f) = \{0 \dots n-1\}$ for n of the form $3m+1$ then

$$H(f \cup \{\langle n, 0\rangle\}) = H(f \cup \{\langle n, 1\rangle\}) = H(f) \cup \{\langle n, \mathbf{t}(m)\rangle\};$$

(v) If $\mathtt{Dom}(f) = \{0, \dots n-1\}$ for n of the form $3m+2$ then there are two tuples $\mathbf{b}_0 \in R$, $\mathbf{b}_1 \notin R$ such that

$$G(f \cup \{\langle n, 0\rangle\}) = G(f) \cup \{\langle n, \mathbf{b}_0\rangle\};$$
$$G(f \cup \{\langle n, 1\rangle\}) = G(f) \cup \{\langle n, \mathbf{b}_1\rangle\};$$
$$H(f \cup \{\langle n, 0\rangle\}) = H(f \cup \{\langle n, 1\rangle\}).$$

Before proving that these G, H exist, let us verify that their existence would lead to the required conclusion of there being uncountably many R' isomorphic to R. Because of (i), we can apply G and H also to functions $f : \mathbb{N} \to \{0, 1\}$ in the obvious way by setting

$$G(f) = \bigcup_{n \in \mathbb{N}} \{G(f_{|0 \dots n}) : n \in \mathbb{N}\},\ H(f) = \bigcup_{n \in \mathbb{N}} \{H(f_{|0 \dots n}) : n \in \mathbb{N}\}$$

and still have that $\langle M, G(f)\rangle \equiv \langle M, H(f)\rangle$ (every first-order formula will involve a finite number of constant symbols, so this follows trivially from (ii)).

For any such $f : \mathbb{N} \to \{0, 1\}$, since M (without the relation R) is a countable model over the empty signature[31] and since by (iii) and (iv) every element of M appears eventually in both $G(f)$ and $H(f)$,[32] it follows that the function ι_f sending,

[31] We could use the fact that it is recursively saturated and countable, but it is overkill for our needs. In fact, much of this proof could probably be simplified for the case that interests us, but we report it fully anyway.

[32] In fact, the same element of M will occur multiple times, e.g., as $G(f)(i)_j$ and as $G(f)(i')_{j'}$. But then, by (ii), it is also true that $H(f)(i)_j = H(f)(i')_j$, and vice versa.

https://doi.org/10.1017/jsl.2019.12 Published online by Cambridge University Press

for each $i \in \mathbb{N}$ and each $j \in 1 \dots k$, $G(f)(i)_j$ into $H(f)(i)_j$ is an automorphism of M.

If the smallest index n such that $f(n) \neq g(n)$ is of the form $3m+2$ then the automorphisms ι_f and ι_g differ and map R into two different relations: indeed, ι_f maps some $\mathbf{b}_0 \in R$ into some tuple $\mathbf{c} = H(f)(n)$, so $\mathbf{c} \in \iota_f[R]$, while ι_g maps some $\mathbf{b}_1 \notin R$ into the same $\mathbf{c}$, so $\mathbf{c} \notin \iota_g[R]$. But there are $2^{\aleph_0}$ many functions every pair of which differs in the first place in an index of this form,[33] and therefore there exist $2^{\aleph_0}$ pairwise distinct automorphic images of R. Thus there are uncountably many R such that $\mathfrak{M} \models \phi(R)$, as required.

It remains to show that the F, G described above can be constructed. We proceed by induction on $n \in \mathbb{N}$, and there are three cases to consider:

- n is of the form $3m$:

  In this case, we extend G by setting—for all f—$G(f \cup \{\langle n, 0\rangle\}) = G(f \cup \{\langle n, 1\rangle\}) = G(f) \cup \{\langle n, \mathbf{t}(m)\rangle\}$, and we must show that there exists some tuple $\mathbf{t}' \in M^k$ such that

  $$\langle M, G(f), \mathbf{t}(m)\rangle \equiv \langle M, H(f), \mathbf{t}'\rangle$$

  where we know, by induction hypothesis, that $\langle M, G(f)\rangle \equiv \langle M, H(f)\rangle$. This is clearly the case, because recursively saturated models are countably homogeneous.[34]

  Then extend H by setting $H(f \cup \{\langle n, 0\rangle\}) = H(f \cup \{\langle n, 1\rangle\}) = H(f) \cup \{\langle n, \mathbf{t}'\rangle\}$: the conditions about F and G are still satisfied.
- n is of the form $3m+1$:

  In this case, we extend H as $H(f \cup \{\langle n, 0\rangle\}) = H(f \cup \{\langle n, 1\rangle\}) = H(f) \cup \{\langle n, \mathbf{t}(m)\rangle\}$, and we must show that there exists some tuple $\mathbf{t}' \in M^k$ such that

  $$\langle M, G(f), \mathbf{t}'\rangle \equiv \langle M, H(f), \mathbf{t}(m)\rangle.$$

  Again, this follows easily from countable homogeneity (or, more simply, from the fact that M has no relation symbols in the signature).

  Then extend G by setting $G(f \cup \{\langle n, 0\rangle\}) = G(f \cup \{\langle n, 1\rangle\}) = G(f) \cup \{\langle n, \mathbf{t}'\rangle\}$: the conditions about F and G are still satisfied.
- n is of the form $3m+2$:

  We need to find tuples $\mathbf{b}_0 \in R$, $\mathbf{b}_1 \notin R$, $\mathbf{c}$ such that

  $$\langle M, G(f), \mathbf{b}_0\rangle \equiv \langle M, H(f), \mathbf{c}\rangle$$

  and

  $$\langle M, G(f), \mathbf{b}_1\rangle \equiv \langle M, H(f), \mathbf{c}\rangle.$$

  To do this, we first find $\mathbf{b}_0 \in R$, $\mathbf{b}_1 \notin R$ such that $\langle M, G(f), \mathbf{b}_0\rangle \equiv \langle M, G(f), \mathbf{b}_1\rangle$. Suppose that no such $\mathbf{b}_0, \mathbf{b}_1$ exist: then for all $\mathbf{b}_0 \in R$, consider the recursive type

  $$\{\neg R\mathbf{x}\} \cup \{\theta(\mathbf{b}_0) \leftrightarrow \theta(\mathbf{x}) : \theta(\mathbf{x}) \in \mathtt{FO}\}$$

[33] For any $K \subseteq \mathbb{N}$ and for all $i \in \mathbb{N}$, let $f_K(i)$ be 1 if i is of the form $3m+2$ for some $m \in K$ and let it be 0 otherwise. Then for all $K_1, K_2 \subseteq \mathbb{N}$, if $K_1 \neq K_2$ then f_{K_1} and f_{K_2} differ in the first place in such an index.

[34] Or, more simply, because $\mathfrak{M}$ is infinite and has no relation symbol in the signature aside from R. Using the recursive saturation of $\langle M, R\rangle$ here is actually really overkill; but we need it anyway to deal with the $3m+2$ case.

https://doi.org/10.1017/jsl.2019.12 Published online by Cambridge University Press

where, in the above expression, θ ranges over all first-order formulas (over the empty signature) with parameters in $G(f)$. The above expression must be finitely unsatisfiable in $\langle M, R\rangle$; otherwise, by the recursive saturation of (M, R), we would have that a $\mathbf{b}_1$ exists as required. So, for each $\mathbf{b}_0 \in R$ there exists a $\sigma_{\mathbf{b}_0}(\mathbf{x})$ such that $\mathfrak{M} \models \sigma_{\mathbf{b}_0}(\mathbf{b}_0)$ and $\mathfrak{M} \models \forall \mathbf{x}(\sigma_{\mathbf{b}_0}(\mathbf{x}) \rightarrow R\mathbf{x})$.[35] Then consider the theory

$$\{R\mathbf{x}\} \cup \{\forall \mathbf{y}(\sigma(\mathbf{y}) \rightarrow R\mathbf{y}) \rightarrow \neg\sigma(\mathbf{x}) : \sigma(\mathbf{x}) \in \mathtt{FO}\}$$

where, again, $\sigma(\mathbf{x})$ ranges over all first-order formulas over the empty signature with parameters in $G(f)$. This must also be finitely unsatisfiable given the theory of M, since otherwise by recursive saturation there would exist some $\mathbf{b} \in R$ such that for no formula σ we have $\mathfrak{M} \models \forall \mathbf{y}(\sigma(\mathbf{y}) \rightarrow R\mathbf{y})$ and $\mathfrak{M} \models \sigma(\mathbf{b})$, and we just saw that $\sigma_{\mathbf{b}}$ is such a formula.

Therefore, there exists some finite number of σs, which we can write as $\sigma_1(\mathbf{x}, \mathbf{a}) \ldots \sigma_q(\mathbf{x}, \mathbf{a})$ for some tuple $\mathbf{a}$ of parameters in $G(f)$, such that $\mathfrak{M} \models \forall \mathbf{y}(\sigma_i(\mathbf{y}, \mathbf{a}) \rightarrow R\mathbf{y})$ for all i and such that $\mathfrak{M} \models \forall \mathbf{x}(R\mathbf{x} \rightarrow \bigvee_i \sigma_i(\mathbf{x}, \mathbf{a}))$. But then $\rho(\mathbf{x}, \mathbf{a}) := \bigvee_i \sigma_i(\mathbf{x}, \mathbf{a})$ defines R in M in the sense that $\mathfrak{M} \models \forall \mathbf{x}(R\mathbf{x} \leftrightarrow \rho(\mathbf{x}, \mathbf{a}))$, which is impossible since $\langle M, R\rangle$ is a model of (3).

Thus, it is possible to find $\mathbf{b}_0, \mathbf{b}_1$ such that $\langle M, G(f), \mathbf{b}_0\rangle \equiv \langle M, G(f), \mathbf{b}_1\rangle$.

Now let us find—using, once more, the countable homogeneity of M—some $\mathbf{c}$ such that $\langle M, G(f), \mathbf{b}_0\rangle \equiv \langle M, H(f), \mathbf{c}\rangle$. Since $\langle M, G(f), \mathbf{b}_0\rangle \equiv \langle M, G(f), \mathbf{b}_1\rangle$, we then also have that $\langle M, G(f), \mathbf{b}_1\rangle \equiv \langle M, H(f), \mathbf{c}\rangle$. Finally, extend G by setting $G(f \cup \{\langle n, 0\rangle\}) = G(f) \cup \{\langle n, \mathbf{b}_0\rangle\}$ and $G(f \cup \{\langle n, 1\rangle\}) = G(f) \cup \{\langle n, \mathbf{b}_1\rangle\}$, and extend H by setting $H(f \cup \{\langle n, 0\rangle\}) = H(f \cup \{\langle n, 1\rangle\}) = H(f) \cup \{\langle n, \mathbf{c}\rangle\}$. The conditions are all still satisfied, and this concludes the proof.

Acknowledgments. I thank Fausto Barbero for a great number of insightful comments, suggestions, and corrections (including, among others, his observation about the existence of nonrelativizable dependencies mentioned at the end of Section 2.3). I also thank the anonymous referee for their many helpful comments and corrections.

REFERENCES

[1] W. W. Armstrong, *Dependency structures of data base relationships*, ***Proceedings of IFIP World Computer Congress*** (J. L. Rosenfeld, editor), North-Holland, Amsterdam, 1974, pp. 580–583.

[2] F. Barbero, *Some observations about generalized quantifiers in logics of imperfect information*, arXiv preprint, 2017, arXiv:1709.07301.

[3] M. A. Casanova, R. Fagin, and C. H. Papadimitriou, *Inclusion dependencies and their interaction with functional dependencies*, ***Proceedings of the 1st ACM Sigact-Sigmod Symposium on Principles of Database Systems***, PODS '82, ACM, 1982, New York, NY, USA, pp. 171–176.

[4] C. C. Chang, *Some new results in definability*. ***Bulletin of the American Mathematical Society***, vol. 70 (1964), no. 6, pp. 808–813.

[35] Indeed, let $\Gamma_0 = \{\neg R\mathbf{x}\} \cup \{\theta_i(\mathbf{b}_0) \leftrightarrow \theta_i(\mathbf{x}) : i = 1 \ldots l\}$ be unsatisfiable in $\langle M, R\rangle$. Then define $\sigma_{\mathbf{b}_0}(\mathbf{x})$ as $\bigwedge\{\theta_i(\mathbf{x}) : i = 1 \ldots l, \mathfrak{M} \models \theta_i(\mathbf{b}_0)\} \wedge \bigwedge\{\neg\theta_i(\mathbf{x}) : i = 1 \ldots l, \mathfrak{M} \not\models \theta_i(\mathbf{b}_0)\}$. Clearly $\mathfrak{M} \models \sigma_{\mathbf{b}_0}(\mathbf{b}_0)$; and moreover, if $\mathfrak{M} \models \sigma_{\mathbf{b}_0}(\mathbf{c})$ we have that $\mathfrak{M} \models \theta_i(\mathbf{b}_0) \leftrightarrow \theta_i(\mathbf{c})$ for all $i = 1 \ldots l$, and therefore by the unsatisfiability of Γ_0 we have that $\mathfrak{M} \models R\mathbf{c}$.

https://doi.org/10.1017/jsl.2019.12 Published online by Cambridge University Press

[5] C. C. Chang and H. J. Keisler, ***Model Theory***, Studies in Logic and the Foundations of Mathematics, vol. 73, Elsevier, Amsterdam, 1990.

[6] E. F. Codd, *Further normalization of the data base relational model*, ***Data Base Systems*** (R. Rustin, editor), Prentice-Hall, New York, 1972, pp. 33–64.

[7] A. Durand, J. Ebbing, J. Kontinen, and H. Vollmer, *Dependence logic with a majority quantifier*. ***Journal of Logic, Language and Information***, vol. 24 (2015), no. 3, pp. 289–305.

[8] A. Durand and J. Kontinen, *Hierarchies in dependence logic*. ***ACM Transactions on Computational Logic***, vol. 13 (2012), no. 4, pp. 31:1–31:21.

[9] A. Durand, J. Kontinen, and H. Vollmer, *Expressivity and complexity of dependence logic*, ***Dependence Logic*** (S. Abramsky, J. Kontinen, J. Väänänen, and H. Vollmer, editors), Springer, Cham, 2016, pp. 5–32.

[10] F. Engström, *Generalized quantifiers in dependence logic*. ***Journal of Logic, Language and Information***, vol. 21 (2012), no. 3, pp. 299–324.

[11] F. Engström, J. Kontinen, and J. Väänänen, *Dependence logic with generalized quantifiers*: *Axiomatizations*. ***Journal of Computer and System Sciences***, vol. 88 (2017), pp. 90–102.

[12] R. Fagin, *Multivalued dependencies and a new normal form for relational databases*. ***ACM Transactions on Database Systems***, vol. 2 (1977), pp. 262–278.

[13] P. Galliani, *Inclusion and exclusion dependencies in team semantics*: *On some logics of imperfect information*. ***Annals of Pure and Applied Logic***, vol. 163 (2012), no. 1, pp. 68–84.

[14] ———, *The doxastic interpretation of team semantics*, ***Logic Without Borders*: *Essays on Set Theory, Model Theory, Philosophical Logic and Philosophy of Mathematics*** (Ä. Hirvonen, J. Kontinen, R. Kossak, and A. Villaveces, editors), Ontos Mathematical Logic, vol. 5, Walter de Gruyter, Berlin, 2015, pp. 167–191.

[15] ———, *Upwards closed dependencies in team semantics*. ***Information and Computation***, vol. 245 (2015), pp. 124–135.

[16] ———, *On strongly first-order dependencies*, ***Dependence Logic*** (S. Abramsky, J. Kontinen, J. Väänänen, and H. Vollmer, editors), Springer, Cham, 2016, pp. 53–71.

[17] ———, *Safe Dependency Atoms and Possibility Operators in Team Semantics*, ***Proceedings Ninth International Symposium on Games, Automata, Logics, and Formal Verification, GandALF 2018*** (A. Orlandini and M. Zimmermann, editors), Open Publishing Association, 2018, pp. 58–72.

[18] P. Galliani, M. Hannula, and J. Kontinen, *Hierarchies in independence logic*, ***Computer Science Logic 2013 (CSL 2013)*** (S. R. D. Rocca, editor), Leibniz International Proceedings in Informatics (LIPIcs), vol. 23, Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, Dagstuhl, Germany, 2013, pp. 263–280.

[19] P. Galliani and L. Hella, *Inclusion logic and fixed point logic*, ***Computer Science Logic 2013 (CSL 2013)*** (S. R. D. Rocca, editor), Leibniz International Proceedings in Informatics (LIPIcs), vol. 23, Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, Dagstuhl, Germany, 2013, pp. 281–295.

[20] E. Grädel and J. Väänänen, *Dependence and independence*. ***Studia Logica***, vol. 101 (2013), no. 2, pp. 399–410.

[21] M. Hannula, *Axiomatizing first-order consequences in independence logic*. ***Annals of Pure and Applied Logic***, vol. 166 (2015), no. 1, pp. 61–91.

[22] ———, *Hierarchies in inclusion logic with lax semantics*. ***ACM Transactions on Computational Logic (TOCL)***, vol. 19 (2018), no. 3, p. 16.

[23] J. Hintikka, ***The Principles of Mathematics Revisited***, Cambridge University Press, Cambridge, 1996.

[24] J. Hintikka and G. Sandu, *Informational independence as a semantic phenomenon*, ***Logic, Methodology and Philosophy of Science*** (J. E. Fenstad, I. T. Frolov, and R. Hilpinen, editors), Elsevier, Amsterdam, 1989, pp. 571–589.

[25] W. Hodges, ***A Shorter Model Theory***, Cambridge University Press, Cambridge, 1997.

[26] ———, *Compositional semantics for a language of imperfect information*. ***Journal of the Interest Group in Pure and Applied Logics***, vol. 5 (1997), no. 4, pp. 539–563.

[27] J. Kontinen, *Coherence and computational complexity of quantifier-free dependence logic formulas*. ***Studia Logica***, vol. 101 (2013), no. 2, pp. 267–291.

[28] J. Kontinen, *Definability of second order generalized quantifiers*. ***Archive for Mathematical Logic***, vol. 49 (2010), no. 3, pp. 379–398.

[29] J. Kontinen, A. Kuusisto, and J. Virtema, *Decidability of predicate logics with team semantics*, ***41st International Symposium on Mathematical Foundations of Computer Science (MFCS 2016)*** (P.

https://doi.org/10.1017/jsl.2019.12 Published online by Cambridge University Press

Faliszewski, A. Muscholl, and R. Niedermeier, editors), Leibniz International Proceedings in Informatics (LIPIcs), vol. 58, Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, Dagstuhl, Germany, 2016, pp. 60:1–60:14.

[30] J. Kontinen, S. Link, and J. Väänänen, *Independence in database relations*, ***Logic, Language, Information, and Computation*** (L. Libkin, U. Kohlenbach, and R. de Queiroz, editors), Springer, Cham, 2013, pp. 179–193.

[31] J. Kontinen and V. Nurmi, *Team logic and second-order logic*. ***Fundamenta Informaticae***, vol. 106 (2011), no. 2–4, pp. 259–272.

[32] J. Kontinen and J. Väänänen, *On definability in dependence logic*. ***Journal of Logic, Language and Information***, vol. 3 (2009), no. 18, pp. 317–332.

[33] ———, *A remark on negation of dependence logic*. ***Notre Dame Journal of Formal Logic***, vol. 52 (2011), no. 1, pp. 55–65.

[34] ———, *Axiomatizing first-order consequences in dependence logic*. ***Annals of Pure and Applied Logic***, vol. 164 (2013), no. 11, pp. 1101–1117.

[35] A. Kuusisto, *A double team semantics for generalized quantifiers*. ***Journal of Logic, Language and Information***, vol. 24 (2015), no. 2, pp. 149–191.

[36] M. Lück, *Axiomatizations of team logics*. ***Annals of Pure and Applied Logic***, vol. 169 (2018), no. 9, pp. 928–969.

[37] A. L. Mann, G. Sandu, and M. Sevenster, ***Independence-Friendly Logic: A Game-Theoretic Approach***, Cambridge University Press, Cambridge, 2011.

[38] R. Rönnholm, ***Arity fragments of logics with team semantics***, Ph.D. thesis, Tampere University, 2018.

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

[40] ———, *Team logic*, ***Interactive Logic. Selected Papers from the 7th Augustus de Morgan Workshop*** (J. van Benthem, D. Gabbay, and B. Löwe, editors), Amsterdam University Press, Amsterdam, 2007, pp. 281–302.

FREE UNIVERSITY OF BOZEN-BOLZANO
FACOLTÀ DI SCIENZE E TECNOLOGIE INFORMATICHE
PIAZZA DOMENICANI 3, 39100 BOLZANO, ITALY
E-mail: pietro.galliani@unibz.it

https://doi.org/10.1017/jsl.2019.12 Published online by Cambridge University Press