Proceedings of the Thirty-Fifth International Joint Conference on Artificial Intelligence (IJCAI-26)
Sister Conferences Best Papers Track

Finite Axiomatizability by Disjunctive Existential Rules (Extended Abstract)*

Marco Calautti[1] , **Marco Console**[2] , **Andreas Pieris**[3,4]
[1]University of Milan
[2]Sapienza University of Rome
[3]University of Cyprus
[4]University of Edinburgh
marco.calautti@unimi.it, console@diag.uniroma.it, apieris@inf.ed.ac.uk

Abstract

In this short paper, we discuss a model-theoretic characterization of when a collection of relational structures is axiomatizable via a finite set of disjunctive existential rules. The presented characterization uses the property of criticality, a refined version of closure under direct products, and a new property that relies on the method of diagrams.

1 Introduction

Model theory is the study of the interaction between sentences in some logical formalism and their models, that is, structures that satisfy the sentences. There are two directions in this interaction: from syntax to semantics and from semantics to syntax. The first direction aims to identify properties that are satisfied by all models of sentences having common syntactic features, i.e., the syntax of a logical formalism is introduced first and then the properties of the mathematical structures that satisfy sentences of that formalism are explored. The second direction aims to characterize sentences in terms of their model-theoretic properties, i.e., given a set of properties, the goal is to determine whether or not these properties characterize some logical formalism. In general, establishing results in the second direction is much harder than establishing results in the first. In other words, obtaining model-theoretic characterizations of logical formalisms is a far greater challenge than identifying properties enjoyed by all models of sentences from a certain formalism.

In this work, we are concerned with the second direction of the interaction between logical sentences and their models. More precisely, we are interested in model-theoretic results of the following form. For a logical formalism $\mathcal{L}$, given a collection $\mathcal{C}$ of structures, the following are equivalent:

1. There exists a finite set Φ of sentences from $\mathcal{L}$ such that $\mathcal{C}$ is precisely the set of models of Φ, in which case we say that $\mathcal{C}$ is *finitely axiomatizable by* $\mathcal{L}$.[1]
2. $\mathcal{C}$ enjoys certain model-theoretic properties.

Such results are welcome as they pinpoint the absolute expressive power of a logical formalism in terms of insightful model-theoretic properties. In other words, they equip us with model-theoretic tools that allow us to prove or disprove that a set of sentences Φ from some logical formalism can be equivalently rewritten as a finite set of sentences from $\mathcal{L}$. Indeed, by showing that the collection of models of Φ enjoy the properties listed in item (2), then Φ is logically equivalent to a finite set Φ' of sentences from $\mathcal{L}$; otherwise, we can safely conclude that such Φ' does not exist.

Characterizations for Rule-based Formalisms. Interesting results of the above form have been established for rule-based formalisms, expressed in suitable fragments of first-order logic, that have been used in numerous areas of central importance to databases and knowledge representation and reasoning. A prominent such formalism is that of existential rules (a.k.a. tuple-generating dependencies in the database literature) that are first-order sentences

$$\forall \bar{x} \forall \bar{y} \left(\phi(\bar{x}, \bar{y}) \ \rightarrow \ \exists \bar{z} \, \psi(\bar{x}, \bar{z}) \right),$$

where ϕ and ψ are conjunctions of atoms. Existential rules have been originally introduced as a framework for database constraints [Abiteboul *et al.*, 1995]. Later on, they have been deployed in the study of data exchange [Fagin *et al.*, 2005] and data integration [Lenzerini, 2002]. More recently, they have been used for modeling ontologies intended for data-intensive tasks such as ontology-based data access [Calì *et al.*, 2010; Mugnier and Thomazo, 2014].

Early characterizations of axiomatizability and finite axiomatizability by full existential rules (i.e., existential rules without existentially quantified variables) have been established by Makowsky and Vardi in the 1980s [Makowsky and Vardi, 1986]. Moreover, ten Cate and Kolaitis obtained analogous characterizations for source-to-target existential rules, which have been used to formalize data exchange between a source schema and a target schema [ten Cate and Kolaitis, 2009]. More recently, Console, Kolaitis and Pieris characterized finite axiomatizability by arbitrary existential rules [Console *et al.*, 2021].

To the best of our knowledge, the above results are the only known results in the literature concerning characterizations of axiomatizability and finite axiomatizability by rule-based

*This short paper is based on a conference paper appeared in the 22nd International Conference on Principles of Knowledge Representation and Reasoning (KR 2025).

[1]Note that one can also talk about the notion of axiomatizability by $\mathcal{L}$, where the set of sentences Φ from $\mathcal{L}$ might be infinite.

formalisms. Having said that, let us stress that several preservation results can be found in the literature. When the target is a preservation result, one considers two formalisms $\mathcal{L}$ and $\mathcal{L}'$, where $\mathcal{L}'$ is typically a proper fragment of $\mathcal{L}$, and the goal is to obtain characterizations of the following form: a set Φ of sentences from $\mathcal{L}$ is logically equivalent to a set of sentences from $\mathcal{L}'$ iff the models of Φ enjoy certain model-theoretic properties. A prototypical example of such a result is the Łos-Tarski Theorem [Chang and Keisler, 1992], which states that a set Φ of first-order sentences is equivalent to a set of universal first-order sentences iff the collection of models of Φ is closed under substructures. Such results for rule-based formalisms have been obtained in the context of description logics by Lutz et al. [Lutz *et al.*, 2011] and of existential rules by Zhang et al. [Zhang *et al.*, 2020].

Disjunctive Existential Rules. The extension of existential rules with the feature of disjunction in the right-hand side of the implication, originally proposed in [Deutsch and Tannen, 2003], leads to the central rule-based formalism of disjunctive existential rules that are first-order sentences of the form

$$\forall \bar{x} \forall \bar{y} \left(\phi(\bar{x}, \bar{y}) \ \rightarrow \ \bigvee_{i=1}^{k} \exists \bar{z}_i \, \psi_i(\bar{x}_i, \bar{z}_i) \right).$$

Disjunctive existential rules have also found numerous applications in several areas such as data exchange [Fagin *et al.*, 2008], expressive database query languages [Eiter *et al.*, 1997], and knowledge representation and reasoning [Alviano *et al.*, 2012], to name a few. Despite the characterizations concerning existential rules discussed above, similar characterizations for disjunctive existential rules have remained largely unexplored so far. The present work changes this state of affairs by focusing on finite axiomatizability.

We establish that a collection of structures $\mathcal{C}$ is finitely axiomatizable by disjunctive existential rules with at most n universally quantified variables, at most m existentially quantified variables, and at most ℓ disjuncts iff $\mathcal{C}$ is critical, closed under repairable direct products, and diagrammatically (n, m, ℓ)-compatible; in fact, this has been originally shown in [Calautti *et al.*, 2025]. Criticality is a standard property, which has been used in several works (see, e.g., [Console *et al.*, 2021]), and states that, for each $\kappa > 0$, $\mathcal{C}$ contains a structure I_κ with κ domain elements, and each relation of I_κ contains all the tuples that can be formed using those κ elements. Closure under repairable direct products is obtained by carefully refining the standard property of closure under direct products since it is easy to show that disjunctive existential rules violate closure under direct products. Finally, diagrammatic (n, m, ℓ)-compatibility, which is the main innovation of this work, relies on the method of diagrams.

2 Preliminaries

Let $\mathbf{C}$ and $\mathbf{V}$ be disjoint countably infinite sets of constants and variables, respectively. For an integer $n > 0$, we may write $[n]$ for the set of integers $\{1, \ldots, n\}$.

Relational Structures. A *(relational) schema* $\mathbf{S}$ is a finite set of relation symbols (or predicates) with positive arity; we write $\mathsf{ar}(R)$ for the arity of the relation symbol R. A *(relational) structure* I over a schema $\mathbf{S} = \{R_1, \ldots, R_n\}$, or $\mathbf{S}$-*structure*, is a tuple $(\mathsf{dom}(I), R_1^I, \ldots, R_n^I)$, where $\mathsf{dom}(I) \subseteq \mathbf{C}$ is a (finite or infinite) domain and $R_1^I, \ldots, R_n^I$ are relations over $\mathsf{dom}(I)$, i.e., $R_i^I \subseteq \mathsf{dom}(I)^{\mathsf{ar}(R_i)}$ for each $i \in [n]$. We say that I is a *finite structure* if $\mathsf{dom}(I)$ is finite. A *fact* of I is an expression of the form $R_i(\bar{c})$, where $\bar{c} \in R_i^I$, and we denote by $\mathsf{facts}(I)$ the set of facts of I. The *active domain* of a structure I, denoted $\mathsf{adom}(I)$, is the set of elements of $\mathsf{dom}(I)$ that occur in at least one fact of I. For an $\mathbf{S}$-structure $J = (\mathsf{dom}(J), R_1^J, \ldots, R_n^J)$, we write $J \subseteq I$ if $\mathsf{facts}(J) \subseteq \mathsf{facts}(I)$. We say that J is a *substrucure* of I, denoted $J \preceq I$, if $\mathsf{dom}(J) \subseteq \mathsf{dom}(I)$ and $R^J = R^I_{|\mathsf{dom}(J)}$ for each $R \in \mathbf{S}$ with $R^I_{|\mathsf{dom}(J)}$ being the *restriction of* R^I *over* $\mathsf{dom}(J)$, i.e., the relation $\left\{\bar{c} \in R^I \mid \bar{c} \in \mathsf{dom}(J)^{\mathsf{ar}(R)}\right\}$. Note that $J \preceq I$ implies $J \subseteq I$, but the other direction does not necessarily hold. A *homomorphism* from I to J is a function $h : \mathsf{dom}(I) \to \mathsf{dom}(J)$ such that, for each $i \in [n]$, $\bar{c} = (c_1, \ldots, c_m) \in R_i^I$ implies $h(\bar{c}) = (h(c_1), \ldots, h(c_m)) \in R_i^J$. We write $h : I \to J$ for the fact that h is a homomorphism from I to J. Let $h(\mathsf{facts}(I))$ be the set $\{R(h(\bar{c})) \mid R(\bar{c}) \in \mathsf{facts}(I)\}$. Finally, we say that I and J are *isomorphic*, written $I \simeq J$, if there is a bijective homomorphism from I to J such that $h^{-1} : J \to I$.

Extended Structures. An *extended schema* is a pair $(\mathbf{S}, \mathbf{F})$, where $\mathbf{S} = \{R_1, \ldots, R_n\}$ is a relational schema and $\mathbf{F} = \{f_1, \ldots, f_m\}$ is a set of 0-ary function symbols. A *structure* I over $(\mathbf{S}, \mathbf{F})$ is a tuple $(\mathsf{dom}(I), R_1^I, \ldots, R_n^I, f_1^I, \ldots, f_m^I)$, where $\mathsf{dom}(I) \subseteq \mathbf{C}$, $R_1^I, \ldots, R_n^I$ are relations over $\mathsf{dom}(I)$, and $f_1^I, \ldots, f_m^I$ are constants of $\mathsf{dom}(I)$. Given a relational schema $\mathbf{S}$ and a finite non-empty set $C = \{c_1, \ldots, c_m\}$ of constants from $\mathbf{C}$, we denote by $\mathbf{S}[C]$ the extended schema $(\mathbf{S}, \{f_{c_1}, \ldots, f_{c_m}\})$. Furthermore, given an $\mathbf{S}$-structure $I = (\mathsf{dom}(I), R_1^I, \ldots, R_n^I)$, we denote by $I[C]$ the structure over $\mathbf{S}[C]$ defined as $(\mathsf{dom}(I) \cup C, R_1^I, \ldots, R_n^I, f_{c_1}^I, \ldots, f_{c_m}^I)$, where $f_{c_j}^I = c_j$ for each $j \in [m]$.[2]

By abuse of notation, we may write down first-order formulas that mention a constant $c \in \mathbf{C}$, which is actually the 0-ary function symbol f_c. For example, $\exists x(R(c, x, d) \wedge \neg(x = d))$, where $c, d \in \mathbf{C}$ and $x \in \mathbf{V}$, is essentially the sentence $\exists x(R(f_c, x, f_d) \wedge \neg(x = f_d))$ over the extended schema $(\{R\}, \{f_c, f_d\})$. Now, by abuse of terminology, given a structure I over a schema $\mathbf{S}$ and a first-order sentence Φ that mentions relation symbols of $\mathbf{S}$ and constants (not necessarily from $\mathsf{dom}(I)$) of $\mathbf{C}$, we may say that I satisfies Φ, denoted $I \models \Phi$. With $C \subseteq \mathbf{C}$ being the set of constants occurring in Φ, $I \models \Phi$ essentially denotes the fact that $I[C] \models \Phi$, that is, the structure $I[C]$ over the extended schema $\mathbf{S}[C]$ satisfies Φ under the standard first-order semantics.

Disjunctive Existential Rules. An atom over a schema $\mathbf{S}$ is an expression of the form $R(\bar{v})$, where $R \in \mathbf{S}$ and $\bar{v}$ is an $\mathsf{ar}(R)$-tuple of variables from $\mathbf{V}$. A *disjunctive existential*

[2] The additional notions defined above for relational structures (such as subset, substructure, homomorphism, and isomorphism) are not needed in the paper for structures over extended schemas.

rule (dexr) σ over a schema $\mathbf{S}$ is a constant-free sentence

$$\forall\bar{x}\forall\bar{y}\left(\phi(\bar{x},\bar{y}) \;\rightarrow\; \bigvee_{i=1}^{k} \exists\bar{z}_i\,\psi_i(\bar{x}_i,\bar{z}_i)\right),$$

where $k > 0$, $\bar{x},\bar{y},\bar{x}_1,\ldots,\bar{x}_k,\bar{z}_1,\ldots,\bar{z}_k$ are tuples of variables of $\mathbf{V}$, the variables of $\bar{x}_i$ occur in $\bar{x}$ for $i \in [k]$, each variable of $\bar{x}$ occurs in $\bar{x}_i$ for some $i \in [k]$, $\phi(\bar{x},\bar{y})$ is a (possibly empty) conjunction of atoms over $\mathbf{S}$, and $\psi_i(\bar{x}_i,\bar{z}_i)$ is a non-empty conjunction of atoms over $\mathbf{S}$ for each $i \in [k]$. For brevity, we write σ as $\phi(\bar{x},\bar{y}) \rightarrow \bigvee_{i=1}^{k} \exists\bar{z}_i\,\psi_i(\bar{x}_i,\bar{z}_i)$ and use comma instead of $\wedge$ for joining atoms. When $\phi(\bar{x},\bar{y})$ is empty, σ is essentially the sentence $\bigvee_{i=1}^{k} \exists\bar{z}_i\,\psi_i(\bar{z}_i)$. We refer to $\phi(\bar{x},\bar{y})$ and $\bigvee_{i=1}^{k} \psi_i(\bar{x}_i,\bar{z}_i)$ as the *body* and *head* of σ, denoted $\mathsf{body}(\sigma)$ and $\mathsf{head}(\sigma)$, respectively. By abuse of notation, we may treat a tuple of variables as a set of variables and a conjunction of atoms as a set of atoms. Assuming that $\phi(\bar{x},\bar{y})$ is non-empty, an $\mathbf{S}$-structure I satisfies σ if the following holds: whenever there exists a function $h : \bar{x} \cup \bar{y} \rightarrow \mathsf{dom}(I)$ such that $h(\phi(\bar{x},\bar{y})) \subseteq \mathsf{facts}(I)$ (as usual, we write $h(\phi(\bar{x},\bar{y}))$ for the set $\{R(h(\bar{v})) \mid R(\bar{v}) \in \phi(\bar{x},\bar{y})\}$), then there exists an integer $i \in [k]$ and an extension h' of h such that $h'(\psi_i(\bar{x}_i,\bar{z}_i)) \subseteq \mathsf{facts}(I)$. When $\phi(\bar{x},\bar{y})$ is empty, an $\mathbf{S}$-structure I satisfies σ if there exists $i \in [k]$ and a function $h : \bar{z}_i \rightarrow \mathsf{dom}(I)$ such that $h(\psi_i(\bar{z}_i)) \subseteq \mathsf{facts}(I)$. We write $I \models \sigma$ for the fact that I satisfies σ. The $\mathbf{S}$-structure I satisfies a set Σ of dexrs over $\mathbf{S}$, written $I \models \Sigma$, in which case we say that I is a *model* of Σ, if $I \models \sigma$ for each $\sigma \in \Sigma$

Finite Axiomatizability. Let $\mathcal{C}$ be a collection of structures over a schema $\mathbf{S}$. We say that $\mathcal{C}$ is *finitely axiomatizable* by dexrs if there exists a finite set Σ of dexrs over $\mathbf{S}$ such that $I \in \mathcal{C}$ iff $I \models \Sigma$, that is, $\mathcal{C}$ is precisely the set of models of Σ. Henceforth, since dexrs cannot distinguish isomorphic structures, we implicitly assume that collections $\mathcal{C}$ of structures are *closed under isomorphisms*, i.e., if $I \in \mathcal{C}$ and J is a structure such that $I \simeq J$, then $J \in \mathcal{C}$.

3 Finite Axiomatizability by Disjunctive Existential Rules

We proceed to characterize when a collection $\mathcal{C}$ of structures is finitely axiomatizable by dexrs. To this end, we introduce three model-theoretic properties of collections of structures, that is, criticality, closure under repairable direct products, and diagrammatic compatibility, and then discuss that they suffice for the desired characterization of finitely axiomatizable by dexrs. Fix an arbitrary schema $\mathbf{S} = \{R_1,\ldots,R_r\}$.

3.1 Criticality

An $\mathbf{S}$-structure $I = (\mathsf{dom}(I), R_1^I,\ldots,R_r^I)$ is κ-*critical*, for $\kappa > 0$, if $|\mathsf{dom}(I)| = \kappa$ and $R_i^I = \mathsf{dom}(I)^{\mathsf{ar}(R_i)}$ for each $i \in [r]$. We can now lift criticality to collections of structures.

Definition 1. *A collection $\mathcal{C}$ of $\mathbf{S}$-structures is κ-critical, for $\kappa > 0$, if it contains a κ-critical $\mathbf{S}$-structure. We further say that $\mathcal{C}$ is* critical *if it is κ-critical for each $\kappa > 0$.* ∎

It is easy to show the following lemma:

Lemma 2. *A collection $\mathcal{C}$ of structures that is finitely axiomatizable by dexrs is critical.*

3.2 Closure Under Repairable Direct Products

Consider two $\mathbf{S}$-structures $I = (\mathsf{dom}(I), R_1^I,\ldots,R_r^I)$ and $J = (\mathsf{dom}(J), R_1^J,\ldots,R_r^J)$. Their *direct product*, denoted $I \otimes J$, is the $\mathbf{S}$-structure $K = (\mathsf{dom}(K), R_1^K,\ldots,R_r^K)$, where $\mathsf{dom}(K) = \mathsf{dom}(I) \times \mathsf{dom}(J)$ and

$$R_i^K \;=\; \big\{\big((a_1,b_1),\ldots,(a_{\mathsf{ar}(R_i)},b_{\mathsf{ar}(R_i)})\big) \mid \big(a_1,\ldots,a_{\mathsf{ar}(R_i)}\big) \in R_i^I \text{ and } \big(b_1,\ldots,b_{\mathsf{ar}(R_i)}\big) \in R_i^J\big\},$$

for each $i \in [r]$. A collection of structures $\mathcal{C}$ is closed under direct products if, for every $I, J \in \mathcal{C}$, the direct product of I and J belongs to $\mathcal{C}$. Closure under direct products is a well-known property that has been extensively used in model theory. However, it is not appropriate towards a characterization of finite axiomatizability by dexrs.

Example 1. *Consider the set Σ consisting of the single dexr*

$$R(x) \;\rightarrow\; S(x) \vee T(x)$$

and the structures I_1 and I_2 with $\mathsf{facts}(I_1) = \{R(a), S(a)\}$ and $\mathsf{facts}(I_2) = \{R(a), T(a)\}$, and thus, $\mathsf{facts}(I_1 \otimes I_2) = \{R((aa))\}$. Clearly, $I_1 \models \Sigma$ and $I_2 \models \Sigma$, but $I_1 \otimes I_2 \not\models \Sigma$. ∎

The above example essentially tells us that a collection of structures that is finitely axiomatizable by dexrs is not necessarily closed under direct products, and thus, we need to revisit the standard notion of direct product. To this end, we define the function $\pi_K : \mathsf{dom}(K) \rightarrow \mathsf{dom}(I)$ as follows: for every $(a_I, a_J) \in \mathsf{dom}(K)$, $\pi_K((a_I,a_J)) = a_I$. It is clear that π_K is a homomorphism from K to I, which we call the *projective homomorphism for K*. Now, a *repairable direct product* of the $\mathbf{S}$-structures I and J is an $\mathbf{S}$-structure L such that (1) $I \otimes J \subseteq L$, and (2) there exists an extension h of π_K that is a homomorphism from L to I. We now define the following novel model-theoretic property:

Definition 3. *A collection $\mathcal{C}$ of $\mathbf{S}$-structures is* closed under repairable direct products *if, for every two $\mathbf{S}$-structures $I, J \in \mathcal{C}$, there exists an $\mathbf{S}$-structure in $\mathcal{C}$ that is a repairable direct product of I and J.* ∎

Coming back to Example 1, it is easy to see that the structure L with $\mathsf{facts}(L) = \{R((aa)), S((aa))\}$ is a repairable direct product of I_1 and I_2; indeed, $I_1 \otimes I_2 \subseteq L$ and $\pi_{I_1 \otimes I_2}$ is a homomorphism from L to I_1. We can show the following:

Lemma 4. *A collection $\mathcal{C}$ of structures that is finitely axiomatizable by dexrs is closed under repairable direct products.*

3.3 Diagrammatic Compatibility

We now introduce our new property of collections of structures, which in turn relies on the notion of relative diagram of a finite structure. We first introduce relative diagrams and then define the new model-theoretic property of interest.

Relative Diagrams. Consider a structure I over a schema $\mathbf{S}$ and a finite structure $K \subseteq I$ with $\mathsf{dom}(K) = \mathsf{adom}(K)$. We are interested in the so-called (m,ℓ)-diagrams of K relative to I for integers $m, \ell \geq 0$, which we define below. To this end, let $A_{K,m}$ be the set of all atomic formulas of the form $R(\bar{u})$ that can be formed using predicates from $\mathbf{S}$, constants from $\mathsf{dom}(K)$, and m distinct variables $y_1,\ldots,y_m$ from $\mathbf{V}$, i.e., $R \in \mathbf{S}$ and $\bar{u} \in (\mathsf{dom}(K) \cup \{y_1, \ldots, y_m\})^{\mathsf{ar}(R)}$. Furthermore, let $C_{K,m}$ be the set of all conjunctions of atomic formulas from $A_{K,m}$. Observe that both $A_{K,m}$ and $C_{K,m}$ are finite sets since $\mathsf{dom}(K)$ is finite. Let

$$N^I_{K,m} \;=\; \{\gamma(\bar{y}) \in C_{K,m} \mid I \not\models \exists \bar{y}\, \gamma(\bar{y})\},$$

which is clearly finite since $C_{K,m}$ is finite. For a set of formulas $G \subseteq N^I_{K,m}$, the *G-diagram of K relative to I*, denoted $\Delta^I_{K,G}$, is defined as the first-order sentence

$$\underbrace{\bigwedge_{\alpha \in \mathsf{facts}(K)} \alpha}_{\Psi_1} \;\wedge\; \underbrace{\bigwedge_{\substack{c,d \in \mathsf{dom}(K),\\ c \neq d}} \neg(c = d) \;\wedge\; \bigwedge_{\gamma(\bar{y}) \in G} \neg\left(\exists \bar{y}\, \gamma(\bar{y})\right)}_{\Psi_2}.$$

Note that Ψ_1 and Ψ_2 might be empty (i.e., they have no conjuncts). In particular, Ψ_1 is empty if K is empty, whereas Ψ_2 is empty if I is 1-critical, which means that $|\mathsf{dom}(K)| = 1$ and $I \models \exists \bar{y}\, \gamma(\bar{y})$ for each $\gamma(\bar{y}) \in C_{K,m}$, and thus, $G = \emptyset$. When Ψ_1 and Ψ_2 are empty, $\Delta^I_{K,G}$ is the truth value $\mathtt{true}$, i.e., a tautology. Intuitively, $\Delta^I_{K,G}$ provides a witness for the fact that a dexr, whose head-disjuncts belong to G, is violated by the structure I due to the structure $K \subseteq I$.

Definition 5. *Consider an* $\mathbf{S}$*-structure* I *and a finite structure* $K \subseteq I$ *with* $\mathsf{dom}(K) = \mathsf{adom}(K)$*. An* (m, ℓ)-diagram of K relative to I*, for* $m, \ell \geq 0$*, is a G-diagram of* K *relative to* I*, where* $G \subseteq N^I_{K,m}$ *and* $|G| \leq \ell$. ∎

Let $\Phi^I_{K,G}(\bar{x})$ be the formula obtained from $\Delta^I_{K,G}$ by replacing each constant $c \in \mathsf{dom}(K)$ with a new variable $x_c \in \mathbf{V} \setminus \{y_1, \ldots, y_m\}$. If $\Delta^I_{K,G}$ is empty, then $\Phi^I_{K,G}$ is the truth value $\mathtt{true}$, i.e., a tautology.

The Property. Consider a collection $\mathcal{C}$ of structures over a schema $\mathbf{S}$ and an $\mathbf{S}$-structure I. For $n, m, \ell \geq 0$, we say that $\mathcal{C}$ is *diagrammatically* (n, m, ℓ)*-compatible* with I if, for every $K \preceq I$ with $\mathsf{dom}(K) = \mathsf{adom}(K)$ and $|\mathsf{dom}(K)| \leq n$, and every (m, ℓ)-diagram $\Delta^I_{K,G}$ of K relative to I, where $G \subseteq N^I_{K,m}$, there is $J \in \mathcal{C}$ such that $J \models \Delta^I_{K,G}$. Roughly, this states that no matter how a dexr σ with a bounded number of variables (at most n universally quantified and at most m existentially quantified variables) and a bounded number of disjuncts in its head (at most ℓ disjuncts) is violated by the structure I, there is always a structure $J \in \mathcal{C}$ that violates σ for the same reason. The new property of interest follows.

Definition 6. *A collection* $\mathcal{C}$ *of* $\mathbf{S}$*-structures is* diagrammatically (n, m, ℓ)-compatible*, for* $n, m, \ell \geq 0$*, if, for every* $\mathbf{S}$*-structure* I*, the following holds:* $\mathcal{C}$ *is diagrammatically* (n, m, ℓ)*-compatible with* I *implies* $I \in \mathcal{C}$. ∎

The next lemma establishes that collections of structures that are finitely axiomatizable by dexrs are diagrammatically compatible. Actually, it shows a stronger claim since it relates the integers n, m and ℓ that witness diagrammatic (n, m, ℓ)-compatibility with the number of universally quantified variables, existentially quantified variables, and head-disjuncts, respectively, occurring in the dexrs. A dexr is called (n, m, ℓ)*-dexr*, for $n, m \geq 0$, with $n + m > 0$, and $\ell > 0$, if it mentions at most n universally quantified variables in its body, at most m existentially quantified variables in each disjunct of its head, and at most ℓ disjuncts in its head. Note that we require $n + m > 0$ since, by definition, a dexr has at least one variable that is either universally or existentially quantified. Furthermore, we require $\ell > 0$ since, by definition, a dexr has at least one disjunct in its head.

Lemma 7. *For* $n, m \geq 0$*, with* $n + m > 0$*, and* $\ell > 0$*, every collection* $\mathcal{C}$ *of structures that is finitely axiomatizable by* (n, m, ℓ)*-dexrs is diagrammatically* (n, m, ℓ)*-compatible.*

3.4 The Characterization

We are now ready to present the characterization of finite axiomatizability by dexrs.

Theorem 8. *For a collection* $\mathcal{C}$ *of structures and* $n, m \geq 0$*, with* $n + m > 0$*, and* $\ell > 0$*, the following are equivalent:*

1. $\mathcal{C}$ *is finitely axiomatizable by* (n, m, ℓ)*-dexrs.*
2. $\mathcal{C}$ *is critical, closed under repairable direct products, and diagrammatically* (n, m, ℓ)*-compatible.*

The direction $(1) \Rightarrow (2)$ follows from Lemmas 2, 4 and 7. For $(2) \Rightarrow (1)$, consider a collection $\mathcal{C}$ of structures over a schema $\mathbf{S}$ that is critical, closed under repairable direct products, and diagrammatically (n, m, ℓ)-compatible for integers $n, m \geq 0$, with $n + m > 0$, and $\ell > 0$. We show that $\mathcal{C}$ is finitely axiomatizable by (n, m, ℓ)-dexrs in three steps:

1. We first define a finite set $\Sigma^{\vee}$ of disjunctive dependencies over $\mathbf{S}$, which extend dexrs by allowing equality formulas over the universally quantified variables as disjuncts, with at most n universally quantified variables, at most m existentially quantified variables, and at most ℓ disjuncts that are a conjunction of atoms (i.e., not an equality formula), such that, for every $\mathbf{S}$-structure I, $I \in \mathcal{C}$ iff $I \models \Sigma^{\vee}$. This exploits the fact that $\mathcal{C}$ is 1-critical (this follows from the fact that $\mathcal{C}$ is critical) and diagrammatically (n, m, ℓ)-compatible.
2. We then show that there is a finite set $\Sigma^{\exists,=}$ of (n, m, ℓ)-dexrs and disjunctive equality rules (i.e., disjunctive dependencies with only equality formulas as disjuncts) with at most n universally quantified variables over $\mathbf{S}$ such that $\Sigma^{\vee} \equiv \Sigma^{\exists,=}$; in fact, $\Sigma^{\exists,=}$ is the set of dexrs and disjunctive equality rules occurring in $\Sigma^{\vee}$. This exploits closure under repairable direct products.
3. We finally argue that $\Sigma^{\exists,=}$ consists only of dexrs, which implies that $\mathcal{C}$ is finitely axiomatizable by (n, m, ℓ)-dexrs. This exploits the fact that $\mathcal{C}$ is critical.

4 Further Results

Results analogous to Theorem 8 have been also established for the two main members of the guarded family of dexrs:

- A dexr is *linear* if it has at most one atom in its body.
- A dexr is *guarded* if its body is empty or has an atom that mentions all the universally quantified variables.

To achieve the desired characterizations for the above classes of dexrs, we simply need to replace the diagrammatic compatibility property in Theorem 8 with a refined version of it that takes into account the syntactic property underlying linear and guarded dexrs.

References

[Abiteboul *et al.*, 1995] Serge Abiteboul, Richard Hull, and Victor Vianu. *Foundations of Databases*. Addison-Wesley, 1995.

[Alviano *et al.*, 2012] Mario Alviano, Wolfgang Faber, Nicola Leone, and Marco Manna. Disjunctive datalog with existential quantifiers: Semantics, decidability, and complexity issues. *Theory Pract. Log. Program.*, 12(4-5):701–718, 2012.

[Calautti *et al.*, 2025] Marco Calautti, Marco Console, and Andreas Pieris. Finite axiomatizability by disjunctive existential rules. In *KR*, 2025.

[Calì *et al.*, 2010] Andrea Calì, Georg Gottlob, Thomas Lukasiewicz, Bruno Marnette, and Andreas Pieris. Datalog+/-: A family of logical knowledge representation and query languages for new applications. In *LICS*, pages 228–242, 2010.

[Chang and Keisler, 1992] Chen C. Chang and H. Jerome Keisler. *Model theory, Third Edition*. North-Holland, 1992.

[Console *et al.*, 2021] Marco Console, Phokion G. Kolaitis, and Andreas Pieris. Model-theoretic characterizations of rule-based ontologies. In *PODS*, pages 416–428, 2021.

[Deutsch and Tannen, 2003] Alin Deutsch and Val Tannen. Reformulation of XML queries and constraints. In *ICDT*, pages 225–241, 2003.

[Eiter *et al.*, 1997] Thomas Eiter, Georg Gottlob, and Heikki Mannila. Disjunctive datalog. *ACM Trans. Database Syst.*, 22(3):364–418, 1997.

[Fagin *et al.*, 2005] Ronald Fagin, Phokion G. Kolaitis, Renée J. Miller, and Lucian Popa. Data exchange: semantics and query answering. *Theor. Comput. Sci.*, 336(1):89–124, 2005.

[Fagin *et al.*, 2008] Ronald Fagin, Phokion G. Kolaitis, Lucian Popa, and Wang Chiew Tan. Quasi-inverses of schema mappings. *ACM Trans. Database Syst.*, 33(2):11:1–11:52, 2008.

[Lenzerini, 2002] Maurizio Lenzerini. Data integration: A theoretical perspective. In *PODS*, pages 233–246, 2002.

[Lutz *et al.*, 2011] Carsten Lutz, Robert Piro, and Frank Wolter. Description logic tboxes: Model-theoretic characterizations and rewritability. In *IJCAI*, pages 983–988, 2011.

[Makowsky and Vardi, 1986] Johann A. Makowsky and Moshe Y. Vardi. On the expressive power of data dependencies. *Acta Inf.*, 23(3):231–244, 1986.

[Mugnier and Thomazo, 2014] Marie-Laure Mugnier and Michaël Thomazo. An introduction to ontology-based query answering with existential rules. In *Reasoning Web*, pages 245–278, 2014.

[ten Cate and Kolaitis, 2009] Balder ten Cate and Phokion G. Kolaitis. Structural characterizations of schema-mapping languages. In *ICDT*, pages 63–72, 2009.

[Zhang *et al.*, 2020] Heng Zhang, Yan Zhang, and Guifei Jiang. Model-theoretic characterizations of existential rule languages. In *IJCAI*, 2020.