A Journey to the Frontiers of Query Rewritability

Piotr Ostropolski-Nalewaja, Jerzy Marcinkowski, David Carral, and Sebastian Rudolph June 7, 2022

The setting: existential rules

An \exists -*rule* or a *Tuple Generating Dependency* is an FO sentence of the following form:

An \exists -*rule* or a *Tuple Generating Dependency* is an FO sentence of the following form:

$\forall x \; Human(x) \rightarrow \exists y \; Mother(x, y)$

An \exists -*rule* or a *Tuple Generating Dependency* is an FO sentence of the following form:

$$orall x \ Human(x)
ightarrow \exists y \ Mother(x,y) \ orall x,y \ Mother(x,y)
ightarrow Human(y)$$

An \exists -*rule* or a *Tuple Generating Dependency* is an FO sentence of the following form:

The setting: knowledge representation using ∃-rules

"Everyone has a mother"

 $Human(x) \rightarrow \exists y \; Mother(x, y)$ $Mother(x, y) \rightarrow Human(y)$

The setting: knowledge representation using \exists -rules

"Everyone has a mother"

 $Human(x) \rightarrow \exists y \; Mother(x, y)$ $Mother(x, y) \rightarrow Human(y)$

"E is transitive"

$$E(x,y), E(y,z) \rightarrow E(x,z)$$

The setting: knowledge representation using \exists -rules

"Everyone has a mother"

"E is transitive"

 $Human(x) \rightarrow \exists y \; Mother(x, y)$ $Mother(x, y) \rightarrow Human(y)$

 $E(x,y), E(y,z) \to E(x,z)$

Query entailment problem

The setting: knowledge representation using ∃-rules

"Everyone has a mother"

"E is transitive"

 $Human(x) \rightarrow \exists y \; Mother(x, y)$ $Mother(x, y) \rightarrow Human(y)$

 $E(x,y), E(y,z) \to E(x,z)$

Query entailment problem

Given a ruleset \mathcal{R} and a database \mathcal{D} we ask if some query \mathcal{Q} holds in every model of \mathcal{R} and \mathcal{D} ?

The setting: knowledge representation using ∃-rules

"Everyone has a mother"

"E is transitive"

 $Human(x) \rightarrow \exists y \; Mother(x, y)$ $Mother(x, y) \rightarrow Human(y)$ $E(x,y), E(y,z) \to E(x,z)$

Query entailment problem

Given a ruleset \mathcal{R} and a database \mathcal{D} we ask if some query \mathcal{Q} holds in every model of \mathcal{R} and \mathcal{D} ?

To denote that a query holds in every model of ${\mathcal R}$ and ${\mathcal D}$ we write

$$\mathcal{R}, \mathcal{D} \models \mathcal{Q}.$$

The main character: BDD class of ∃-rules

$$\mathcal{D}, \mathcal{R} \models \mathcal{Q} \iff \mathcal{D} \models \mathcal{Q}'$$

$$\mathcal{D}, \mathcal{R} \models \mathcal{Q} \iff \mathcal{D} \models \mathcal{Q}'$$

entail

$$\begin{array}{c} \mathcal{D}, \mathcal{R} \models \mathcal{Q} \iff \mathcal{D} \models \mathcal{Q}' \\ entail \\ holds \end{array}$$

The BDD class

BDD is an **undecidable property**.

BDD is an **undecidable property**. Thus, a lot of **decidable subclasses** of BDD were invented, such as: *Linear, Sticky, Sticky-Join, Backward Shy...*

BDD is an **undecidable property**. Thus, a lot of **decidable subclasses** of BDD were invented, such as: *Linear, Sticky, Sticky-Join, Backward Shy...*

It is thought that **BDD** is well understood.

BDD is an **undecidable property**. Thus, a lot of **decidable subclasses** of BDD were invented, such as: *Linear, Sticky, Sticky-Join, Backward Shy...*

It is thought that **BDD** is well understood. Soon, however, we will see that there is a lot more to learn about it.

We identify a class of "local theories" subsuming every known decidable subclass of BDD!

We identify a class of "local theories" subsuming every known decidable subclass of BDD!

We show that this class is strictly contained in BDD.

We identify a class of "local theories" subsuming every known decidable subclass of BDD!

We show that this class is strictly contained in BDD.

And we show that the FUS/FES conjecture (BDD \cap Core Terminating = Uniform BDD) holds for this class.

We identify a class of "local theories" subsuming every known decidable subclass of BDD!

We show that this class is strictly contained in BDD.

And we show that the FUS/FES conjecture (BDD ∩ Core Terminating = Uniform BDD) holds for this class. (not a part of this talk)



$E(x,y) \rightarrow \exists z \ E(y,z)$



0

 $E(x, y) \rightarrow \exists z \ E(y, z)$ $A(x), E(x, y), E(y, z) \rightarrow E(x, z)$



•

 $E(x, y) \rightarrow \exists z \ E(y, z)$ $A(x), E(x, y), E(y, z) \rightarrow E(x, z)$



•

 $E(x,y) \rightarrow \exists z \ E(y,z)$ $A(x), E(x, y), E(y, z) \rightarrow E(x, z)$



•

 $E(x,y) \rightarrow \exists z \ E(y,z)$ $A(x), E(x, y), E(y, z) \rightarrow E(x, z)$


$E(x,y) \rightarrow \exists z \ E(y,z)$ $A(x), E(x, y), E(y, z) \rightarrow E(x, z)$



 $E(x, y) \rightarrow \exists z \ E(y, z)$ $A(x), E(x, y), E(y, z) \rightarrow E(x, z)$



 $E(x,y) \rightarrow \exists z \ E(y,z)$ $A(x), E(x, y), E(y, z) \rightarrow E(x, z)$



 $E(x, y) \rightarrow \exists z \ E(y, z)$ $A(x), E(x, y), E(y, z) \rightarrow E(x, z)$



 $E(x, y) \rightarrow \exists z \ E(y, z)$ $A(x), E(x, y), E(y, z) \rightarrow E(x, z)$



 $E(x, y) \rightarrow \exists z \ E(y, z)$ $A(x), E(x, y), E(y, z) \rightarrow E(x, z)$



The main character: BDD class of ∃-rules

A ruleset ${\mathcal R}$ is BDD if:

A ruleset \mathcal{R} is BDD if:

$\forall \mathcal{Q} \ \exists k \ \forall \mathcal{D} \quad \mathcal{D}, \mathcal{R} \models \mathcal{Q} \iff \mathit{Chase}_k(\mathcal{R}, \mathcal{D}) \models \mathcal{Q}$

A ruleset ${\mathcal R}$ is BDD if:

$$\forall \mathcal{Q} \ \exists k \ \forall \mathcal{D} \quad \mathcal{D}, \mathcal{R} \models \mathcal{Q} \iff \mathsf{Chase}_k(\mathcal{R}, \mathcal{D}) \models \mathcal{Q}$$

Query

A ruleset ${\mathcal R}$ is BDD if:

$$\begin{array}{c} \forall \mathcal{Q} \ \exists k \ \forall \mathcal{D} \quad \mathcal{D}, \mathcal{R} \models \mathcal{Q} \iff Chase_k(\mathcal{R}, \mathcal{D}) \models \mathcal{Q} \\ & & \\ & \\ & \\ & & \\ & \\ & &$$

A ruleset \mathcal{R} is BDD if:







 $E(x, y), E(y, z) \rightarrow E(x, z)$



 $E(x, y), E(y, z) \rightarrow E(x, z)$



 $E(x, y), E(y, z) \rightarrow E(x, z)$



 $E(x, y), E(y, z) \rightarrow E(x, z)$



 $E(x,y), E(y,z) \rightarrow E(x,z)$



 $E(x, y), E(y, z) \rightarrow E(x, z)$



$E(x,y), E(y,z) \to E(x,z)$



$E(x,y), E(y,z) \to E(x,z)$



 $(\neg \text{ BDD}) \exists \mathcal{Q} \forall k \exists \mathcal{D} \quad \mathcal{D}, \mathcal{R} \models \mathcal{Q} \land Chase_k(\mathcal{R}, \mathcal{D}) \not\models \mathcal{Q}$

$E(x,y), E(y,z) \to E(x,z)$



 $(\neg \text{ BDD}) \exists \mathcal{Q} \forall k \exists \mathcal{D} \quad \mathcal{D}, \mathcal{R} \models \mathcal{Q} \land Chase_k(\mathcal{R}, \mathcal{D}) \not\models \mathcal{Q}$ Q = E(a, b)

 $A(x), E(x, y) \rightarrow A(y)$



 $A(x), E(x, y) \rightarrow A(y)$



 $A(x), E(x, y) \rightarrow A(y)$



 $A(x), E(x, y) \rightarrow A(y)$



 $A(x), E(x, y) \rightarrow A(y)$



 $A(x), E(x, y) \rightarrow A(y)$



 $A(x), E(x, y) \rightarrow A(y)$



 $A(x), E(x, y) \rightarrow A(y)$



 $A(x), E(x, y) \rightarrow A(y)$



 $(\neg \mathsf{BDD}) \exists \mathcal{Q} \forall k \exists \mathcal{D} \quad \mathcal{D}, \mathcal{R} \models \mathcal{Q} \land \mathit{Chase}_k(\mathcal{R}, \mathcal{D}) \not\models \mathcal{Q}$

 $A(x), E(x, y) \rightarrow A(y)$





 $(\neg \mathsf{BDD}) \exists \mathcal{Q} \forall k \exists \mathcal{D} \quad \mathcal{D}, \mathcal{R} \models \mathcal{Q} \land \mathit{Chase}_k(\mathcal{R}, \mathcal{D}) \not\models \mathcal{Q}$
Class of local theories

We say that a ruleset \mathcal{R} is local if every atom of the chase can be derived only from a constant number of atoms of the database.

We say that a ruleset \mathcal{R} is local if every atom of the chase can be derived only from a constant number of atoms of the database.

Every local ruleset is BDD!

We say that a ruleset \mathcal{R} is local if every atom of the chase can be derived only from a constant number of atoms of the database.

Every local ruleset is BDD!

Moreover, it contains every known decidable subclass of BDD.

We say that a ruleset \mathcal{R} is local if every atom of the chase can be derived only from a constant number of atoms of the database.

Every local ruleset is BDD!

Moreover, it contains every known decidable subclass of BDD.

almost

$E(x,y) \rightarrow \exists x', y' Box(x, y, x', y')$





$E(x, y) \rightarrow \exists x', y' Box(x, y, x', y')$ $Box(x, y, x', y'), E(x, y), E(y, z) \rightarrow \exists z' Box(y, z, y', z')$







$E(x, y) \rightarrow \exists x', y' Box(x, y, x', y')$ Box(x, y, x', y'), E(x, y), E(y, z) \rightarrow \exists z' Box(y, z, y', z')





















$$\begin{split} & \textit{E}(x,y) \rightarrow \exists x', y' \textit{Box}(x,y,x',y') \\ & \textit{Box}(x,y,x',y'), \textit{E}(x,y), \textit{E}(y,z) \rightarrow \exists z' \textit{Box}(y,z,y',z') \end{split}$$



$$\begin{split} & \textit{E}(x,y) \rightarrow \exists x', y' \textit{Box}(x,y,x',y') \\ & \textit{Box}(x,y,x',y'), \textit{E}(x,y), \textit{E}(y,z) \rightarrow \exists z' \textit{Box}(y,z,y',z') \end{split}$$



$$\begin{split} & \textit{E}(x,y) \rightarrow \exists x', y' \textit{Box}(x,y,x',y') \\ & \textit{Box}(x,y,x',y'), \textit{E}(x,y), \textit{E}(y,z) \rightarrow \exists z' \textit{Box}(y,z,y',z') \end{split}$$



$$\begin{split} & E(x,y) \rightarrow \exists x', y' \; \textit{Box}(x,y,x',y') \\ & \textit{Box}(x,y,x',y'), E(x,y), E(y,z) \rightarrow \exists z' \; \textit{Box}(y,z,y',z') \end{split}$$



$$\begin{split} & \textit{E}(x,y) \rightarrow \exists x', y' \textit{Box}(x,y,x',y') \\ & \textit{Box}(x,y,x',y'), \textit{E}(x,y), \textit{E}(y,z) \rightarrow \exists z' \textit{Box}(y,z,y',z') \end{split}$$



$$\begin{split} & \textit{E}(x,y) \rightarrow \exists x', y' \textit{Box}(x,y,x',y') \\ & \textit{Box}(x,y,x',y'), \textit{E}(x,y), \textit{E}(y,z) \rightarrow \exists z' \textit{Box}(y,z,y',z') \end{split}$$



Linear-width rewritings

A ruleset $\mathcal R$ is BDD (FO-rewritable) if for every CQ $\mathcal Q$ there exists a UCQ $\mathcal Q'$ such that for every database $\mathcal D$ the following holds

$$\mathcal{D}, \mathcal{R} \models \mathcal{Q} \iff \mathcal{D} \models \mathcal{Q}'$$

A ruleset \mathcal{R} is BDD (FO-rewritable) if for every CQ \mathcal{Q} there exists a UCQ \mathcal{Q}' such that for every database \mathcal{D} the following holds

$$\mathcal{D}, \mathcal{R} \models \mathcal{Q} \iff \mathcal{D} \models \mathcal{Q}'$$

The linear-width rewritings class

The width of a UCQ is the maximal size of its disjuncts.

A ruleset \mathcal{R} is BDD (FO-rewritable) if for every CQ \mathcal{Q} there exists a UCQ \mathcal{Q}' such that for every database \mathcal{D} the following holds

$$\mathcal{D}, \mathcal{R} \models \mathcal{Q} \iff \mathcal{D} \models \mathcal{Q}'$$

The linear-width rewritings class

The width of a UCQ is the maximal size of its disjuncts.

We say that a ruleset \mathcal{R} admits *linear-width rewritings* if it is BDD and for every CQ its UCQ rewriting is of linear size width.

A ruleset \mathcal{R} is BDD (FO-rewritable) if for every CQ \mathcal{Q} there exists a UCQ \mathcal{Q}' such that for every database \mathcal{D} the following holds

$$\mathcal{D}, \mathcal{R} \models \mathcal{Q} \iff \mathcal{D} \models \mathcal{Q}'$$

The linear-width rewritings class

The width of a UCQ is the maximal size of its disjuncts.

We say that a ruleset \mathcal{R} admits *linear-width rewritings* if it is BDD and for every CQ its UCQ rewriting is of linear size width.

Every known decidable subclass of BDD admits linear-width rewritings!

BDD is not about linear-width rewritings

BDD is not about linear-width rewritings

 $\rightarrow \exists x \ H(x,x), V(x,x)$

0
$$ightarrow \exists x \ H(x,x), V(x,x)$$

 $op (x)
ightarrow \exists y, z \ H(x,y), V(x,z)$

.



























Rewriting of $V^n H V^{-n}$ contains a disjunct H^{2^n} !

It shows deficiency in the understanding of an impactful class of existential rules.

It shows deficiency in the understanding of an impactful class of existential rules.

How can we use it?

It shows deficiency in the understanding of an impactful class of existential rules.

How can we use it?

It gives a concrete example of behaviour that can be formalized and used for more expressive knowledge representation. Thank You