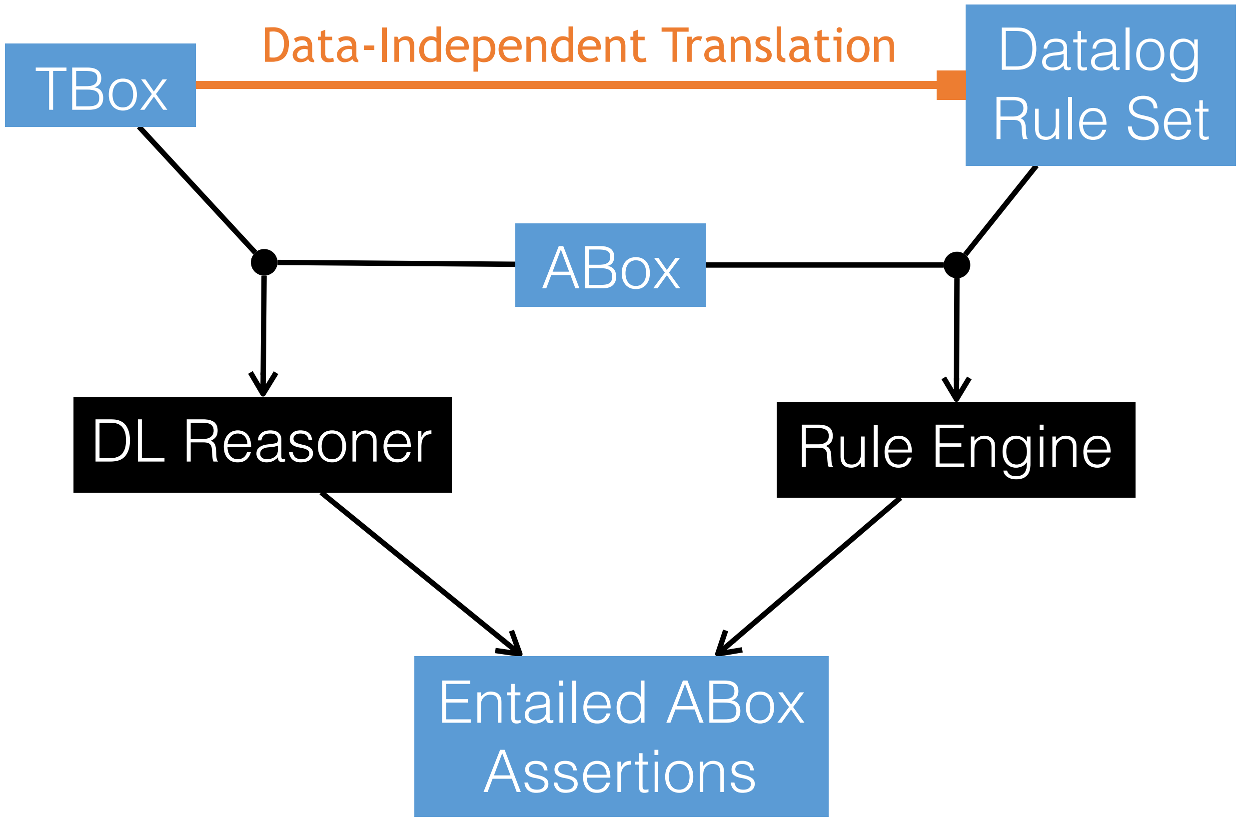
From Horn-$\mathcal{SRIQ}$ to Datalog: A Data-Independent Transformation that Preserves Assertion Entailment

David Carral, Larry González, Patrick Koopmann

TECHNISCHE UNIVERSITÄT DRESDEN

Introduction

Solving IQs with Datalog Rewritings



(i.e., facts over named individuals)

In **theory**:

* Correctness
* Complexity

In **practice**:

* Implement mapping
* Evaluate performance

The DL Horn-SRIQ

$$C_1 \sqcap \ldots \sqcap C_n \sqsubseteq D \qquad C \sqsubseteq \leq 1R.D \qquad A(a)$$

$$\exists R.C \sqsubseteq D \qquad C \sqsubseteq \exists R.D \qquad R(a,b)$$

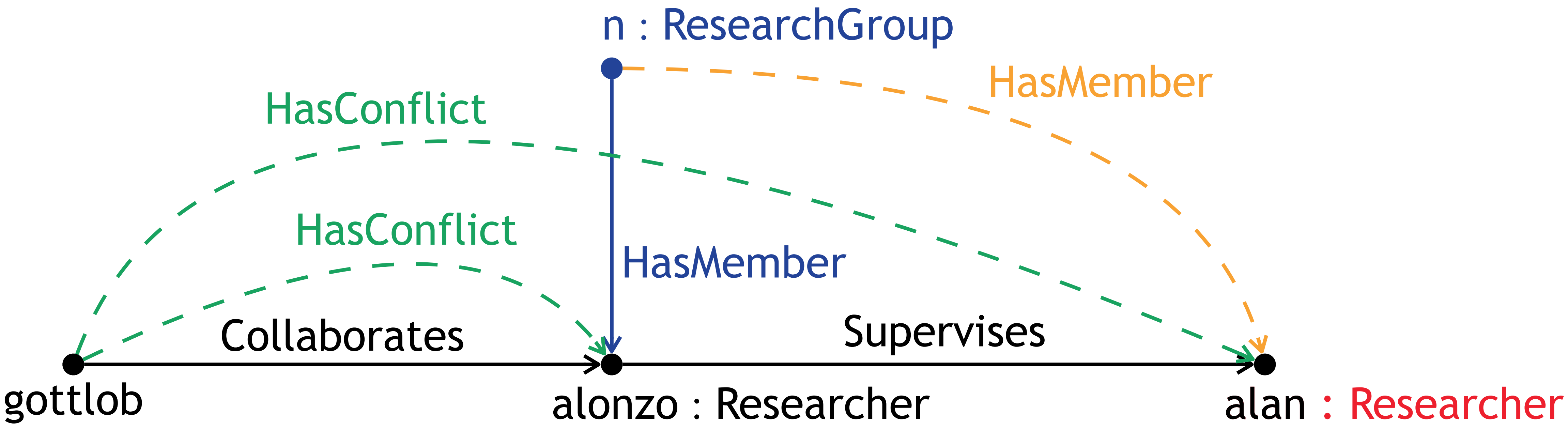
$$\boxed{R_1 \circ \ldots \circ R_n \sqsubseteq S}$$

$\text{ResearchGroup} \sqsubseteq \forall \text{HasMember} . \text{Researcher}$

$\text{Researcher} \sqsubseteq \exists \text{HasMember}^{-} . \text{ResearchGroup}$

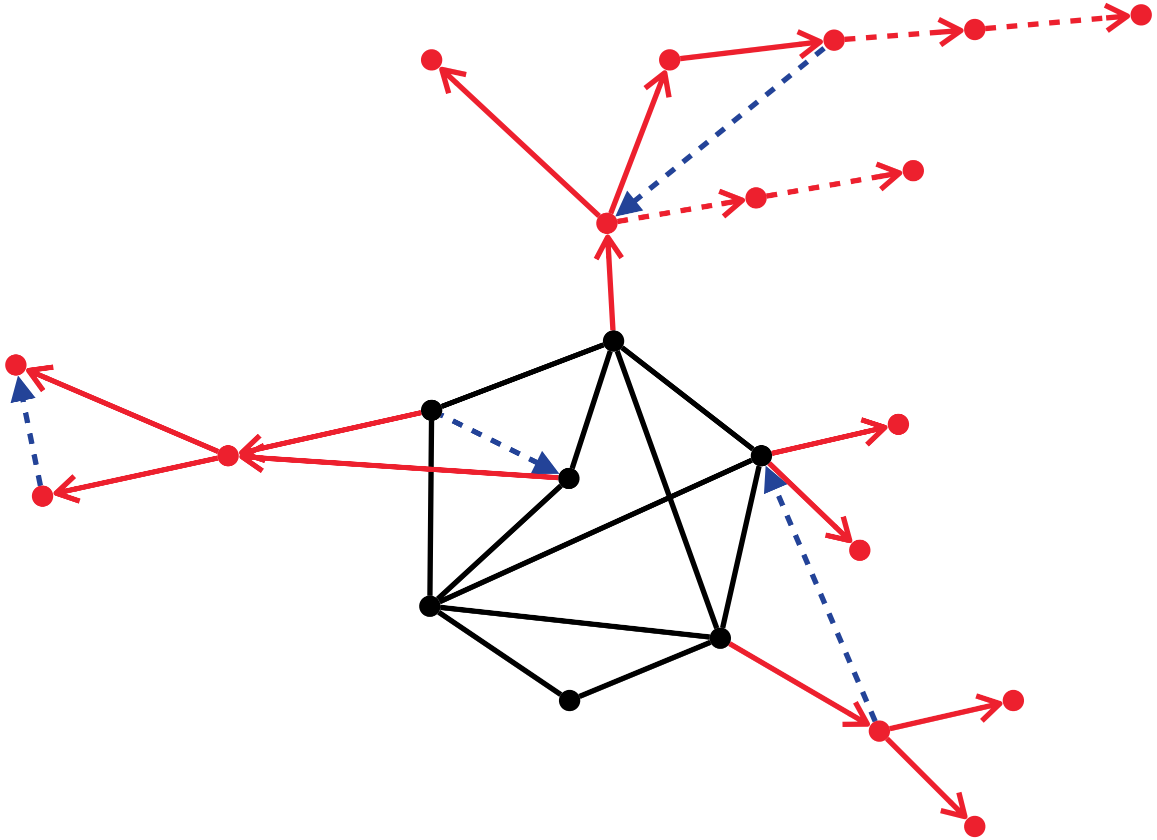
$\text{Collaborates} \circ \text{HasMember}^{-} \circ \text{HasMember} \sqsubseteq \text{HasConflict}$

$\text{HasMember} \circ \text{Supervises} \sqsubseteq \text{HasMember}$



From Horn-$\mathcal{ALCHIQ}$ to Datalog

Forest Model Property



$C_1 \sqcap \ldots \sqcap C_n \sqsubseteq D$

$C \sqsubseteq \exists R . D$

$\exists R . C \sqsubseteq D$

$C \sqsubseteq \leq 1 R . D$

$R \sqsubseteq S$

"Unnamed-to-Named" Consequences

Successor-to-predecessor

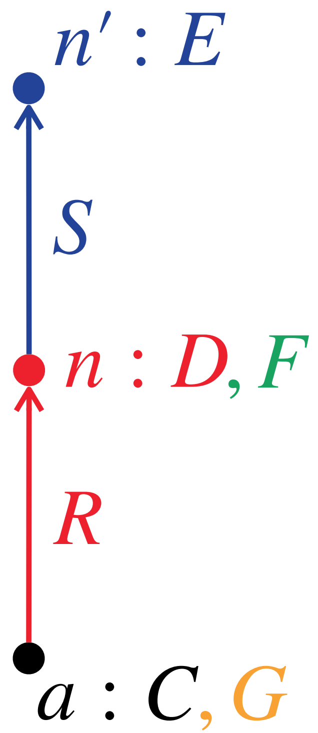
$C \sqsubseteq \exists R \,.\, D$

$D \sqsubseteq \exists S \,.\, E$

$\exists S \,.\, E \sqsubseteq F$

$\exists R \,.\, F \sqsubseteq G$

$C(x) \rightarrow G(x)$



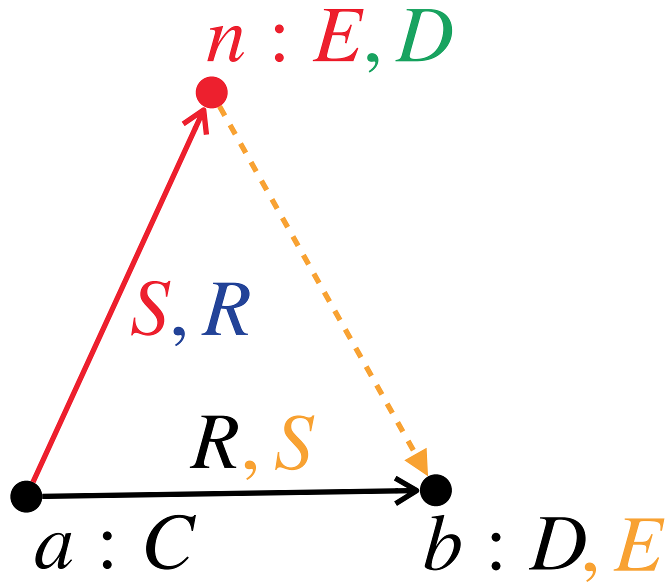
Folding

$C \sqsubseteq \exists S \,.\, E$

$S \sqsubseteq R$

$E \sqsubseteq D$

$C \sqsubseteq\, \leq 1 R \,.\, D$



$C(x) \wedge R(x, y) \wedge D(y) \rightarrow S(x, y)$

$C(x) \wedge R(x, y) \wedge D(y) \rightarrow E(y)$

Computing IQ Rewritings for Horn-$\mathcal{ALCHIQ}$

Consider some Horn-ALCHIQ TBox $\mathcal{T}$.
Then, the rule set $\mathcal{R}_{\mathcal{T}}$ defined as follows is an IQ-preserving rewriting for $\mathcal{T}$.

For all $C \sqsubseteq \forall R\,.\,D \in \mathcal{T}$,
$C(x) \wedge R(x,y) \to D(y) \in \mathcal{R}_{\mathcal{T}}$

For all $R \sqsubseteq S \in \mathcal{T}$,
$R(x,y) \to S(x,y) \in \mathcal{R}_{\mathcal{T}}$

For all $C_1 \sqcap \ldots \sqcap C_n \sqsubseteq D \in \Omega(\mathcal{T})$
$C_1(x) \wedge \ldots \wedge C_n(x) \to D(x) \in \mathcal{R}_{\mathcal{T}}$

For all $C \sqsubseteq \leq 1R\,.\,D \in \mathcal{T}$,
$C(x) \wedge R(x,y) \wedge D(y) \wedge R(x,z) \wedge D(z) \to y \approx z \in \mathcal{R}_{\mathcal{T}}$,
$C(x) \wedge C_1(x) \wedge \ldots \wedge C_n(x) \wedge R(x,y) \wedge D(y) \to E(y) \in \mathcal{R}_{\mathcal{T}}$ if $C_1 \sqcap \ldots \sqcap C_n \sqsubseteq \exists R\,.\,(D \sqcap E) \in \Omega(\mathcal{T})$, and
$C(x) \wedge C_1(x) \wedge \ldots \wedge C_n(x) \wedge R(x,y) \wedge D(y) \to S(x,y) \in \mathcal{R}_{\mathcal{T}}$ if $C_1 \sqcap \ldots \sqcap C_n \sqsubseteq \exists (R \sqcap S)\,.\,D \in \Omega(\mathcal{T})$

$\Omega(\mathcal{T})$ is the set of all axioms of one of the following forms entailed by $\mathcal{T}$.
$C_1 \sqcap \ldots \sqcap C_n \sqsubseteq D$
$C_1 \sqcap \ldots \sqcap C_n \sqsubseteq \exists (R_1 \sqcap \ldots \sqcap R_m)\,.\,(D_1 \sqcap \ldots \sqcap D_k)$

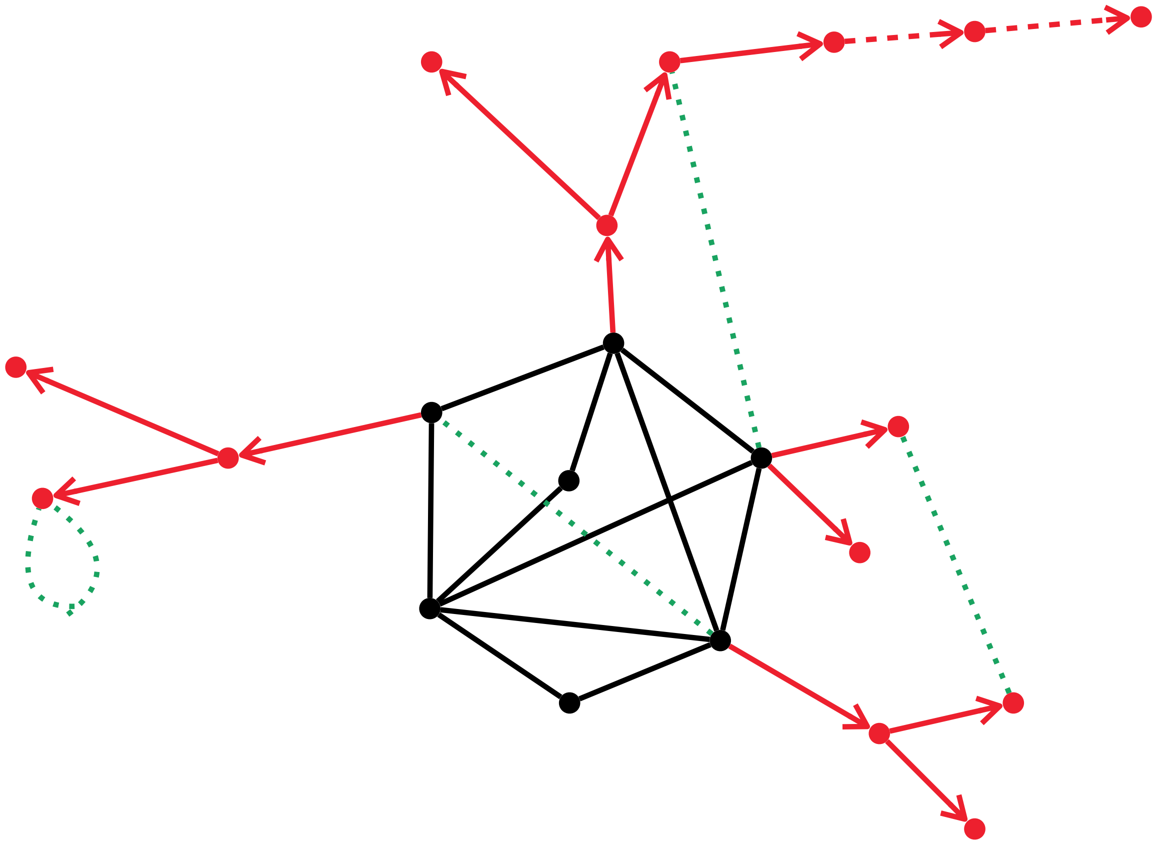
Consequence-based Reasoning Calculi
[IJCAI 2017] Yevgeny
[AAAI 2012] Eiter et al.

$$\frac{M \sqsubseteq \exists S.(N \sqcap N') \quad N \sqsubseteq A}{M \sqsubseteq \exists S.(N \sqcap N' \sqcap A)}\ \mathbf{R}^{c}_{\sqsubseteq}$$

$$\frac{M \sqsubseteq \exists (S \sqcap S').N \quad S \sqsubseteq r}{M \sqsubseteq \exists (S \sqcap S' \sqcap r).N}\ \mathbf{R}^{r}_{\sqsubseteq}$$

From Horn-$\mathcal{SRIQ}$ to Datalog

Tree Model Property



$C_1 \sqcap \ldots \sqcap C_n \sqsubseteq D$

$C \sqsubseteq \exists R . D$

$\exists R . C \sqsubseteq D$

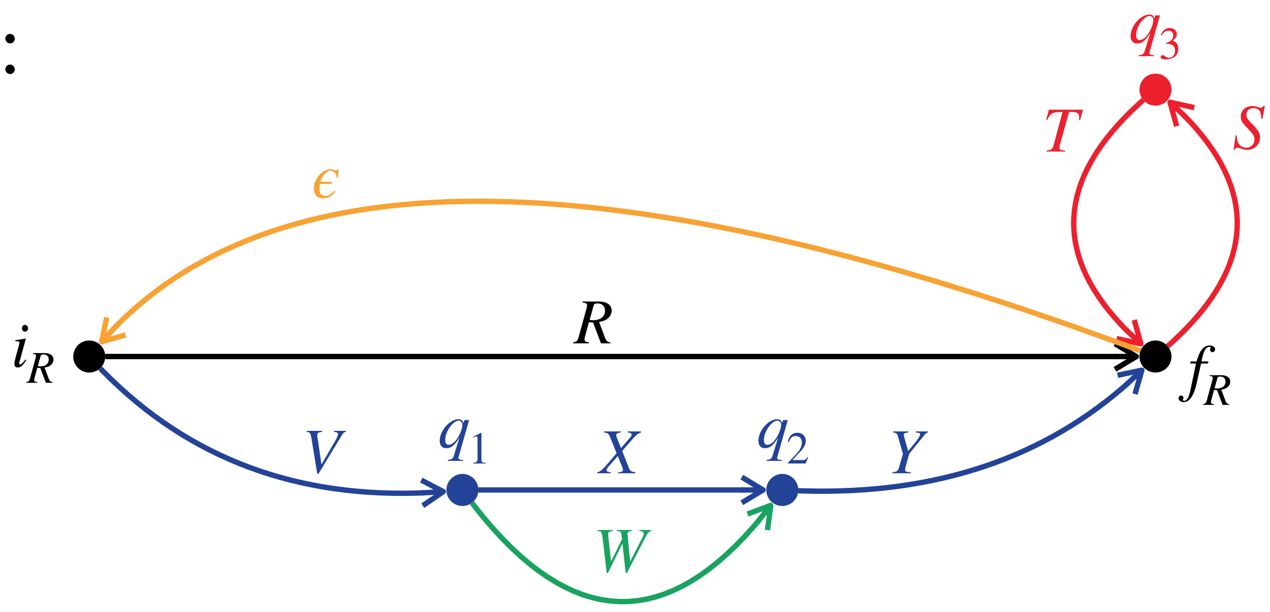
$C \sqsubseteq \leq 1 R . D$

$R_1 \circ \ldots \circ R_n \sqsubseteq S$

Complex Roles and NFA

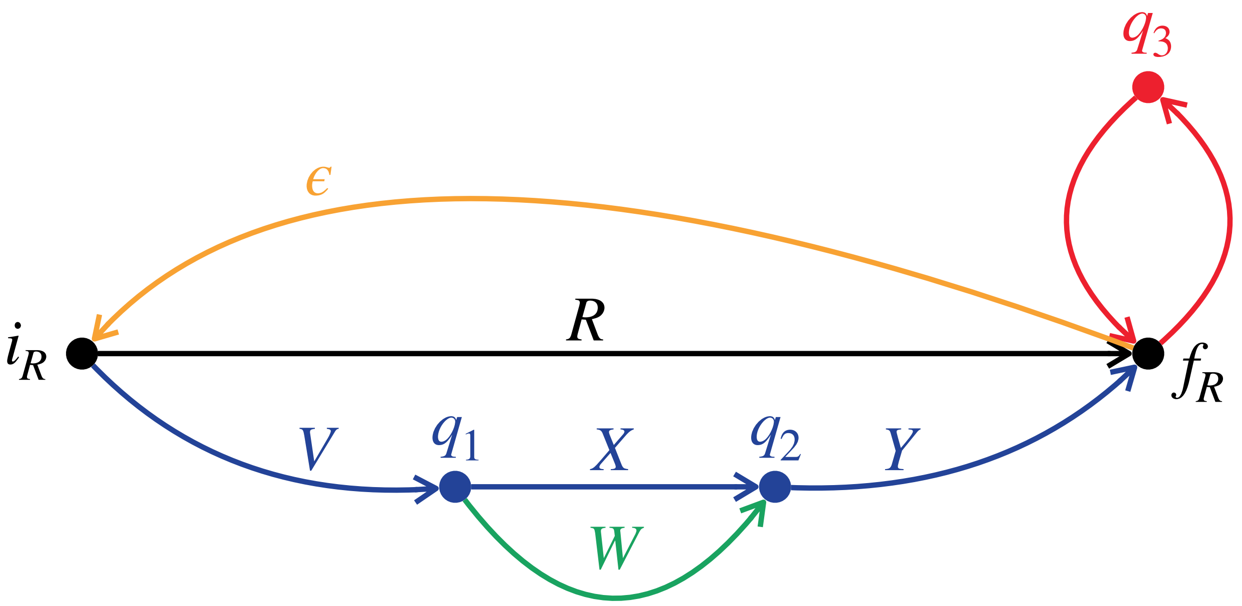
$$\mathcal{T} = \{R \circ S \circ T \sqsubseteq R, \quad V \circ X \circ Y \sqsubseteq R, \quad W \sqsubseteq X, \quad R \circ R \sqsubseteq R\}$$

$\mathcal{N}_{\mathcal{T}}(R)$:



Box Pushing

$$A \sqsubseteq \forall R . B \in \mathcal{T}$$



$$BP(\mathcal{T}) \supseteq \mathcal{T} \cup \{$$

$$A \sqsubseteq B_{i_R}, B_{f_R} \sqsubseteq B,$$

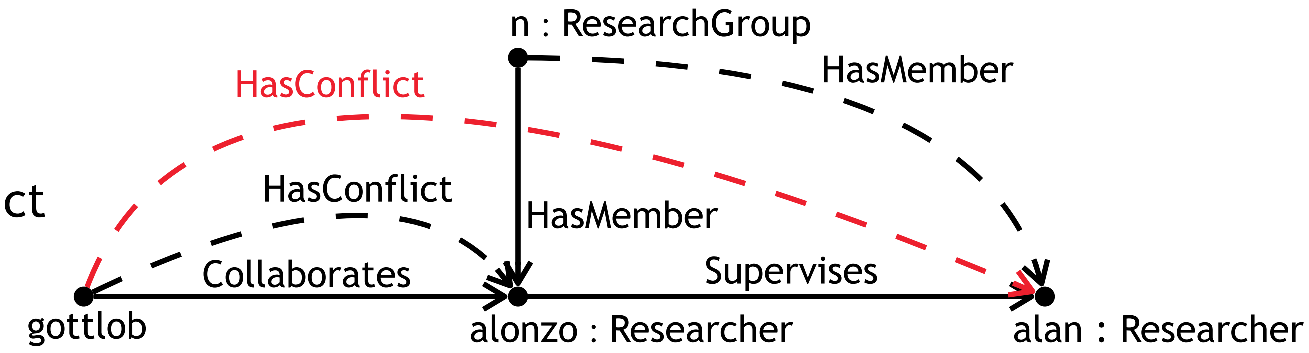
$$B_{i_R} \sqsubseteq \forall R . B_{f_R},$$

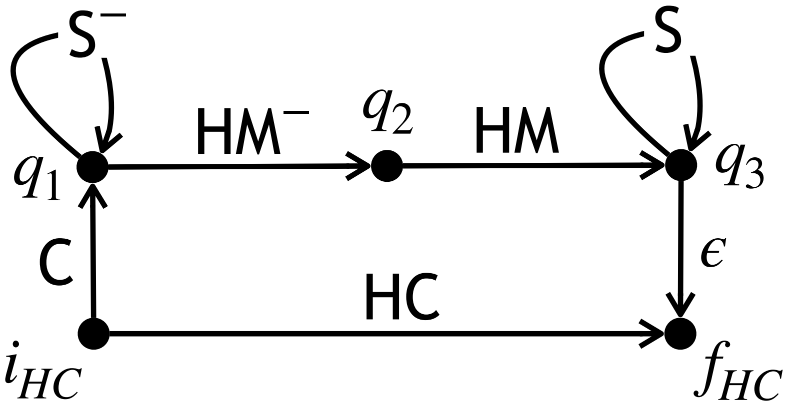
$$B_{f_R} \forall S . B_{q_3}, B_{q_3} \sqsubseteq \forall T . B_{f_R},$$

$$B_{i_R} \sqsubseteq \forall V . B_{q_1}, B_{q_1} \sqsubseteq \forall X . B_{q_2}, B_{q_2} \sqsubseteq \forall Y . B_{f_R},$$

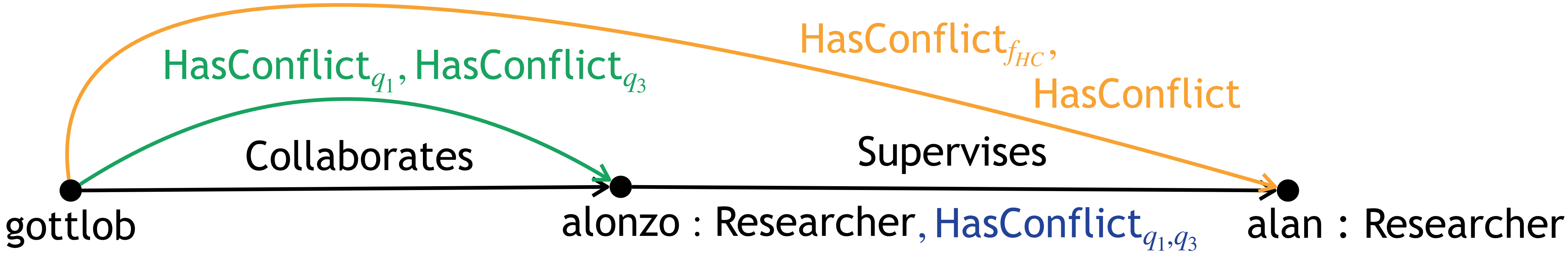
$$B_{q_1} \sqsubseteq \forall W . B_{q_2}, B_{f_R} \sqsubseteq B_{i_R} \}$$

$\mathsf{ResearchGroup} \sqsubseteq \forall \mathsf{HasMember} . \mathsf{Researcher}$
$\mathsf{Researcher} \sqsubseteq \exists \mathsf{HasMember}^{-} . \mathsf{ResearchGroup}$
$\mathsf{Collaborates} \circ \mathsf{HasMember}^{-} \circ \mathsf{HasMember} \sqsubseteq \mathsf{HasConflict}$
$\mathsf{HasMember} \circ \mathsf{Supervises} \sqsubseteq \mathsf{HasMember}$





$$\mathsf{Researcher}(x) \rightarrow \mathsf{HasConflict}_{q_1,q_3}(x)$$
$$\mathsf{Collaborates}(x, y) \rightarrow \mathsf{HasConflict}_{q_1}(x, y)$$
$$\mathsf{HasConflict}_{q_1}(x, y) \wedge \mathsf{HasConflict}_{q_1,q_3}(y) \rightarrow \mathsf{HasConflict}_{q_3}(x, y)$$
$$\mathsf{HasConflict}_{q_3}(x, y) \wedge \mathsf{Supervises}(y, z) \rightarrow \mathsf{HasConflict}_{f_{HC}}(x, y)$$
$$\mathsf{HasConflict}_{f_R}(x, y) \rightarrow \mathsf{HasConflict}(x, y)$$

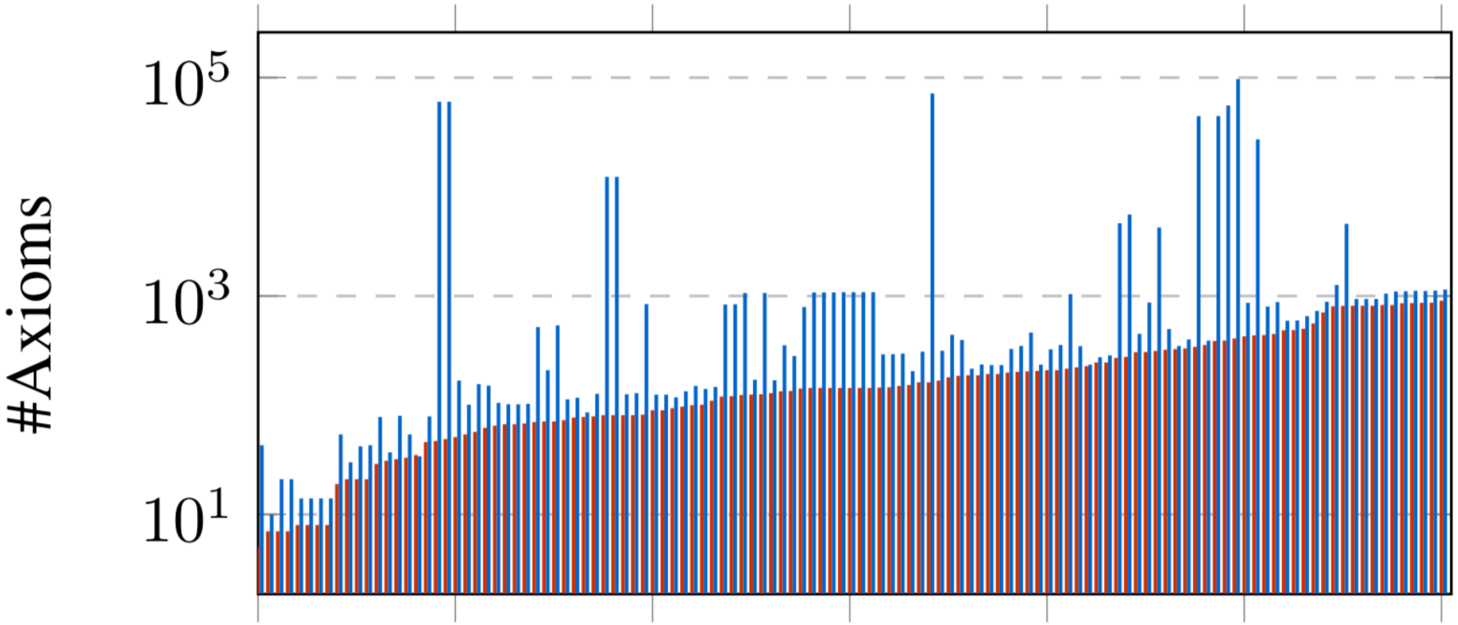


Evaluation

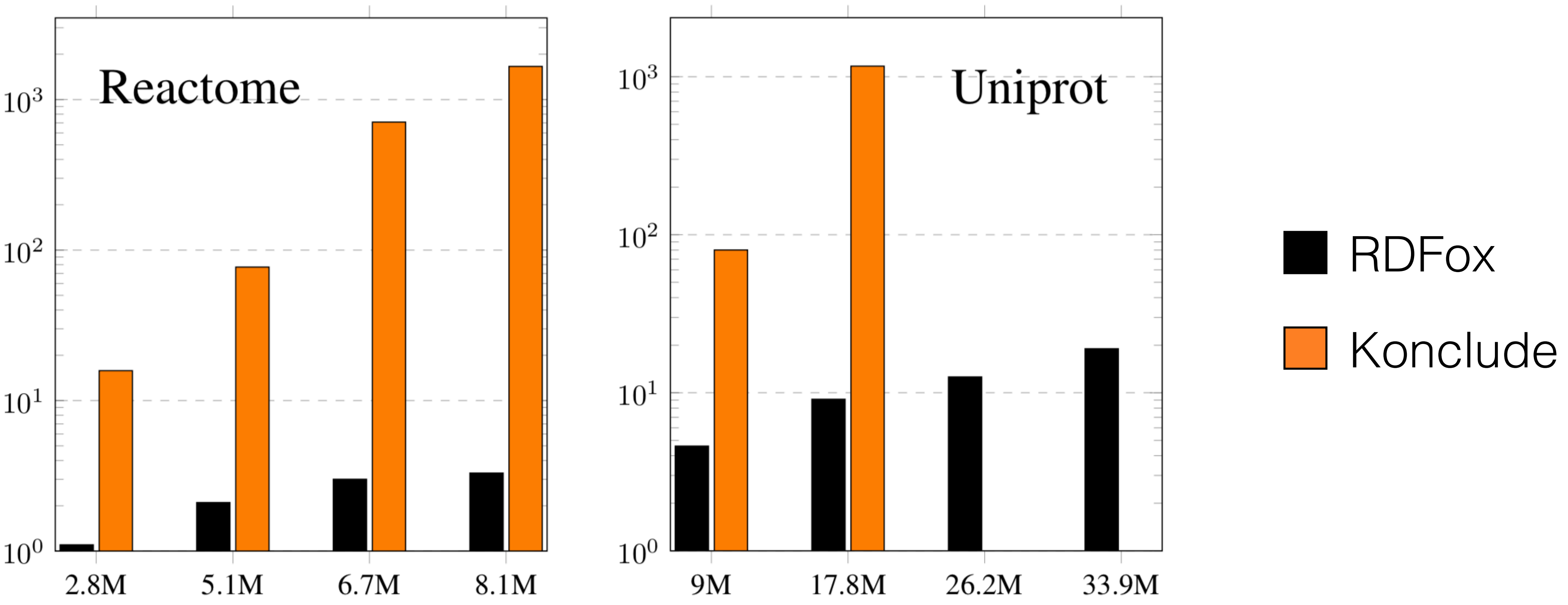
Size of Rewritings

- 187 ontologies from the MOWL Corpus

- 121 computed rewritings



Rewritings Performance



From Horn-$\mathcal{SRIQ}$ to Datalog: A Data-Independent Transformation that Preserves Assertion Entailment

David Carral, Larry González, Patrick Koopmann

TECHNISCHE UNIVERSITÄT DRESDEN