# Chasing Sets: How to Use Existential Rules for Expressive Reasoning

David Carral, Irina Dragoste, Markus Krötzsch, Christian Lewe

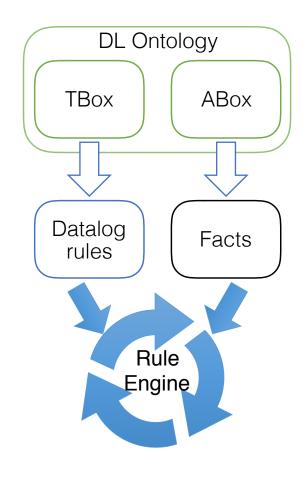


- ExpTime-complete combined complexity
- Fast and scalable reasoners available

Can we use Datalog to solve hard problems?

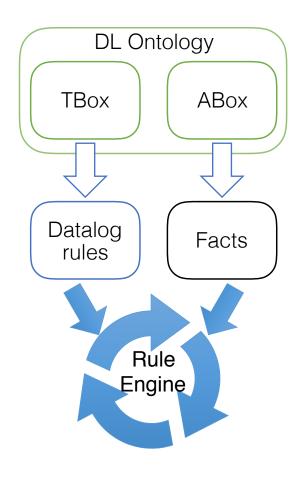
- ExpTime-complete combined complexity
- Fast and scalable reasoners available

✓ Carral et al. (KR 2018): Horn-ALCHOIQ



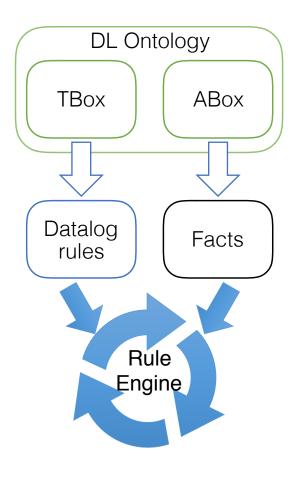
- ExpTime-complete combined complexity
- Fast and scalable reasoners available

- √Carral et al. (KR 2018): Horn-ALCHOIQ
  - exponentially many rules



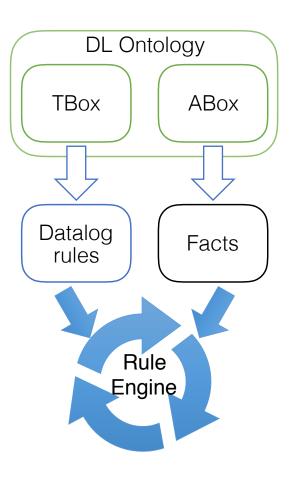
- ExpTime-complete combined complexity
- Fast and scalable reasoners available

- √Carral et al. (KR 2018): Horn-ALCHOIQ
  - exponentially many rules
- ✓Ahmetaj et al. (ICDT 2018): guarded existential rules



- ExpTime-complete combined complexity
- Fast and scalable reasoners available

- √Carral et al. (KR 2018): Horn-ALCHOIQ
  - exponentially many rules
- ✓Ahmetaj et al. (ICDT 2018): guarded existential rules
  - predicates with linearly large arities



Is there an efficient way to solve hard problems with rule engines, nonetheless?

By moving from **Datalog** to **existential rules** we can

solve hard (ExpTime-complete) real-world problems

- solve hard (ExpTime-complete) real-world problems
- using existing rule engines

- solve hard (ExpTime-complete) real-world problems
- using existing rule engines
- with a fixed set of rules

- solve hard (ExpTime-complete) real-world problems
- using existing rule engines
- with a fixed set of rules

- solve hard (ExpTime-complete) real-world problems
- using existing rule engines
- with a fixed set of rules
- Available reasoners use the chase algorithm

- solve hard (ExpTime-complete) real-world problems
- using existing rule engines
- with a fixed set of rules
- Available reasoners use the chase algorithm
- Chase algorithm may not terminate

- solve hard (ExpTime-complete) real-world problems
- using existing rule engines
- with a fixed set of rules
- Available reasoners use the chase algorithm
- Chase algorithm may not terminate
- Sufficient conditions for chase termination

- solve hard (ExpTime-complete) real-world problems
- using existing rule engines
- with a fixed set of rules
- Available reasoners use the chase algorithm
- Chase algorithm may not terminate
- Sufficient conditions for chase termination
  - characterise rule sets of PTime data complexity (like Datalog)

## How can we get the required expressivity?

- ExpTime-complete data complexity
- polynomial translation from Datalog(S) to existential rules

- ExpTime-complete data complexity
- polynomial translation from Datalog(S) to existential rules

$$person(x) \rightarrow likesAll(x, \emptyset)$$
 (1)

$$likesAll(x, S) \land likes(x, y) \rightarrow likesAll(x, S \cup \{y\})$$
 (2)

- ExpTime-complete data complexity
- polynomial translation from Datalog(S) to existential rules

$$person(x) \rightarrow likesAll(x, \emptyset)$$
 (1)

$$likesAll(x, S) \land likes(x, y) \rightarrow likesAll(x, S \cup \{y\})$$
 (2)

$$likesAll(x, S) \rightarrow allLikeAll(\{x\}, S)$$
 (3)

$$allLikeAll(S,T) \land likesAll(x,T) \rightarrow allLikeAll(S \cup \{x\},T)$$
 (4)

- ExpTime-complete data complexity
- polynomial translation from Datalog(S) to existential rules

$$person(x) \rightarrow likesAll(x, \emptyset)$$
 (1)

$$likesAll(x, S) \land likes(x, y) \rightarrow likesAll(x, S \cup \{y\})$$
 (2)

$$likesAll(x, S) \rightarrow allLikeAll(\{x\}, S)$$
 (3)

$$allLikeAll(S,T) \land likesAll(x,T) \rightarrow allLikeAll(S \cup \{x\},T)$$
 (4)

$$allLikeAll(S, S) \land alice \in S \rightarrow cliqueOfAlice(S)$$
 (5)

## Datalog(S): Definition

Logic with two sorts: objects and sets of objects

- Each predicate position has a sort
- Object and set variables are distinct
- Set terms:  $\emptyset$  { object}  $Set_1 \cup Set_2$
- Built-in predicates (only in body):  $object \in Set$   $Set_1 \subseteq Set_2$

## Datalog(S): Definition

Logic with two sorts: objects and sets of objects

- Each predicate position has a sort
- Object and set variables are distinct
- Set terms:  $\emptyset$  { object}  $Set_1 \cup Set_2$
- Built-in predicates (only in body):  $object \in Set$   $Set_1 \subseteq Set_2$

All set variables must occur in a regular body atom (not built-in)

## Datalog(S): Definition

Logic with two sorts: objects and sets of objects

- Each predicate position has a sort
- Object and set variables are distinct
- Set terms:  $\emptyset$  { object}  $Set_1 \cup Set_2$
- Built-in predicates (only in body):  $object \in Set$   $Set_1 \subseteq Set_2$

All set variables must occur in a regular body atom (not built-in)

**Theorem:** Datalog(S) has **ExpTime-complete** combined and **data** complexity.

$$person(x) \rightarrow likesAll(x, \emptyset)$$
 (1)

$$likesAll(x, S) \land likes(x, y) \rightarrow likesAll(x, S \cup \{y\})$$
 (2)

$$person(x) \rightarrow likesAll(x, \emptyset) \qquad (1)$$

$$likesAll(x, S) \wedge likes(x, y) \rightarrow likesAll(x, S \cup \{y\}) \qquad (2)$$

$$\rightarrow \exists V. empty(V) \qquad (1')$$

$$person(x) \wedge empty(Y) \rightarrow likesAll(x, Y) \qquad (1'')$$

$$person(x) \rightarrow likesAll(x, \emptyset) \qquad (1)$$

$$likesAll(x, S) \wedge likes(x, y) \rightarrow likesAll(x, S \cup \{y\}) \qquad (2)$$

$$\rightarrow \exists V. empty(V) \qquad (1')$$

$$person(x) \wedge empty(Y) \rightarrow likesAll(x, Y) \qquad (1'')$$

$$likesAll(x, S) \wedge likes(x, y) \rightarrow \exists V. likesAll(x, V) \wedge SU(S, y, V) \qquad (2')$$

$$person(x) \rightarrow likesAll(x, \emptyset) \qquad (1)$$

$$likesAll(x, S) \wedge likes(x, y) \rightarrow likesAll(x, S \cup \{y\}) \qquad (2)$$

$$\rightarrow \exists V. empty(V) \qquad (1')$$

$$person(x) \wedge empty(Y) \rightarrow likesAll(x, Y) \qquad (1'')$$

$$likesAll(x, S) \wedge likes(x, y) \rightarrow \exists V. likesAll(x, V) \wedge SU(S, y, V) \qquad (2')$$

person(Eve)
likes(Eve, a)
likes(Eve, b)

```
person(x) \rightarrow likesAll(x, \emptyset) \qquad (1)
likesAll(x, S) \wedge likes(x, y) \rightarrow likesAll(x, S \cup \{y\}) \qquad (2)
\rightarrow \exists V. empty(V) \qquad (1')
person(x) \wedge empty(Y) \rightarrow likesAll(x, Y) \qquad (1'')
likesAll(x, S) \wedge likes(x, y) \rightarrow \exists V. likesAll(x, V) \wedge SU(S, y, V) \qquad (2')
```

#### Eve

```
person(Eve)
likes(Eve, a)
likes(Eve, b)
```

$$person(x) \rightarrow likesAll(x, \emptyset) \qquad (1)$$

$$likesAll(x, S) \wedge likes(x, y) \rightarrow likesAll(x, S \cup \{y\}) \qquad (2)$$

$$\rightarrow \exists V. empty(V) \qquad (1')$$

$$person(x) \wedge empty(Y) \rightarrow likesAll(x, Y) \qquad (1'')$$

$$likesAll(x, S) \wedge likes(x, y) \rightarrow \exists V. likesAll(x, V) \wedge SU(S, y, V) \qquad (2')$$

#### Eve

```
person(Eve)
likes(Eve, a)
likes(Eve, b)
```



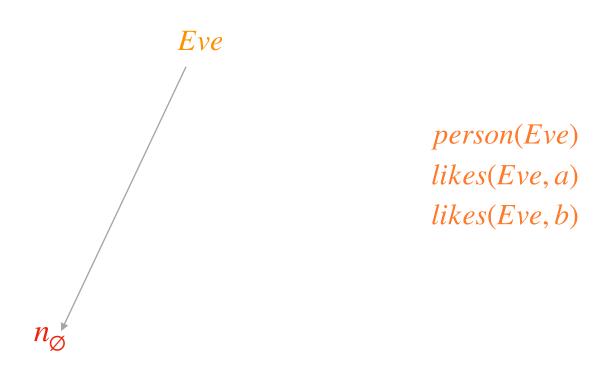
$$person(x) \rightarrow likesAll(x, \emptyset) \qquad (1)$$

$$likesAll(x, S) \wedge likes(x, y) \rightarrow likesAll(x, S \cup \{y\}) \qquad (2)$$

$$\rightarrow \exists V. empty(V) \qquad (1')$$

$$person(x) \wedge empty(Y) \rightarrow likesAll(x, Y) \qquad (1'')$$

$$likesAll(x, S) \wedge likes(x, y) \rightarrow \exists V. likesAll(x, V) \wedge SU(S, y, V) \qquad (2')$$



→ likesAll

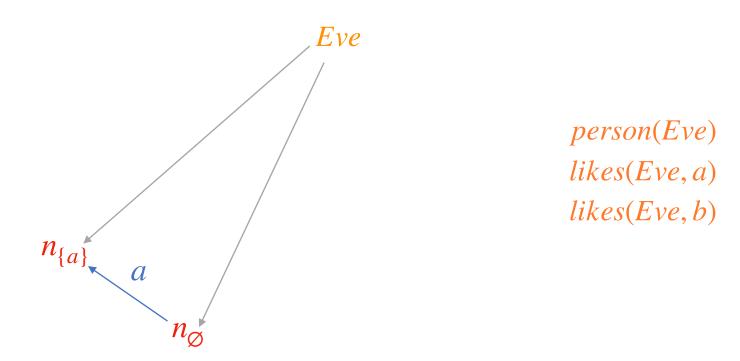
$$person(x) \rightarrow likesAll(x, \emptyset) \qquad (1)$$

$$likesAll(x, S) \wedge likes(x, y) \rightarrow likesAll(x, S \cup \{y\}) \qquad (2)$$

$$\rightarrow \exists V. empty(V) \qquad (1')$$

$$person(x) \wedge empty(Y) \rightarrow likesAll(x, Y) \qquad (1'')$$

$$likesAll(x, S) \wedge likes(x, y) \rightarrow \exists V. likesAll(x, V) \wedge SU(S, y, V) \qquad (2')$$





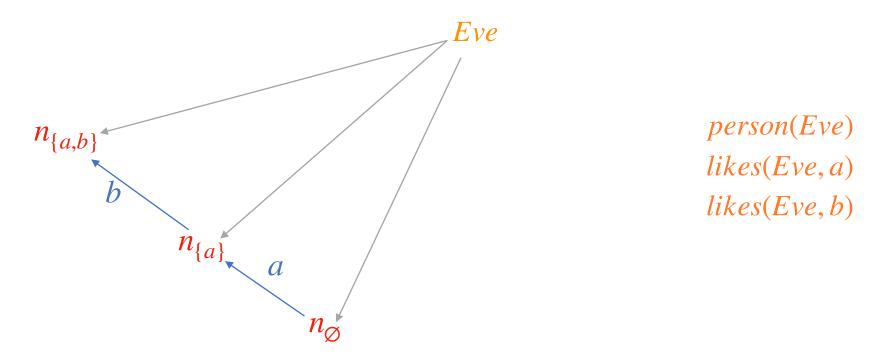
$$person(x) \rightarrow likesAll(x, \emptyset) \qquad (1)$$

$$likesAll(x, S) \wedge likes(x, y) \rightarrow likesAll(x, S \cup \{y\}) \qquad (2)$$

$$\rightarrow \exists V. empty(V) \qquad (1')$$

$$person(x) \wedge empty(Y) \rightarrow likesAll(x, Y) \qquad (1'')$$

$$likesAll(x, S) \wedge likes(x, y) \rightarrow \exists V. likesAll(x, V) \wedge SU(S, y, V) \qquad (2')$$



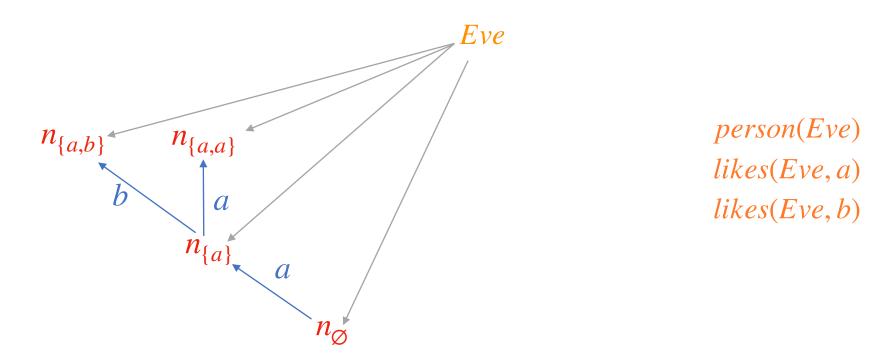
$$person(x) \rightarrow likesAll(x, \emptyset) \qquad (1)$$

$$likesAll(x, S) \wedge likes(x, y) \rightarrow likesAll(x, S \cup \{y\}) \qquad (2)$$

$$\rightarrow \exists V. empty(V) \qquad (1')$$

$$person(x) \wedge empty(Y) \rightarrow likesAll(x, Y) \qquad (1'')$$

$$likesAll(x, S) \wedge likes(x, y) \rightarrow \exists V. likesAll(x, V) \wedge SU(S, y, V) \qquad (2')$$



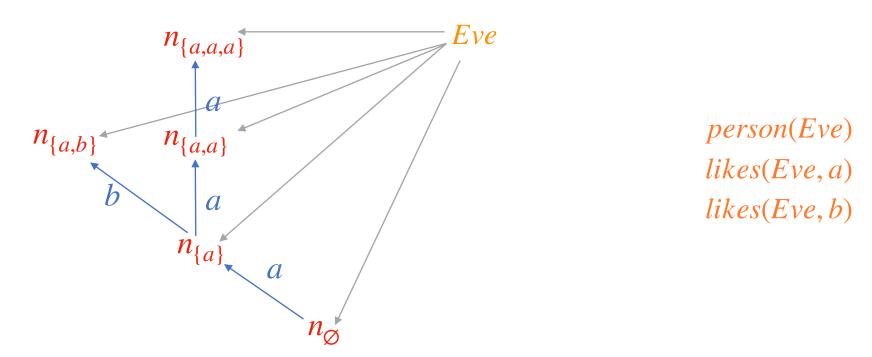
$$person(x) \rightarrow likesAll(x, \emptyset) \qquad (1)$$

$$likesAll(x, S) \wedge likes(x, y) \rightarrow likesAll(x, S \cup \{y\}) \qquad (2)$$

$$\rightarrow \exists V. empty(V) \qquad (1')$$

$$person(x) \wedge empty(Y) \rightarrow likesAll(x, Y) \qquad (1'')$$

$$likesAll(x, S) \wedge likes(x, y) \rightarrow \exists V. likesAll(x, V) \wedge SU(S, y, V) \qquad (2')$$



→ likesAll

→ SU

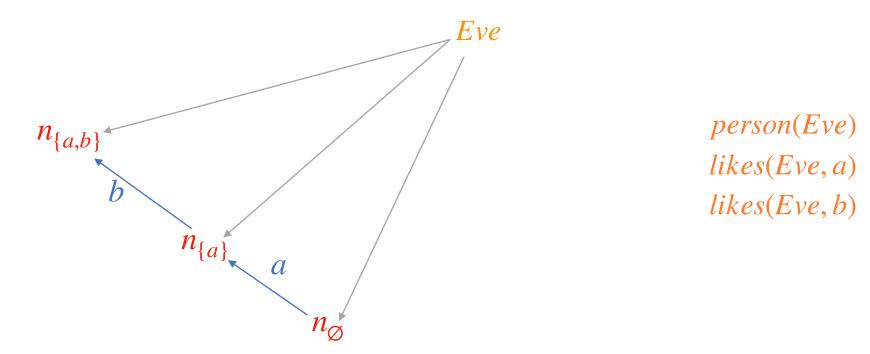
$$person(x) \rightarrow likesAll(x, \emptyset) \qquad (1)$$

$$likesAll(x, S) \wedge likes(x, y) \rightarrow likesAll(x, S \cup \{y\}) \qquad (2)$$

$$\rightarrow \exists V. empty(V) \qquad (1')$$

$$person(x) \wedge empty(Y) \rightarrow likesAll(x, Y) \qquad (1'')$$

$$likesAll(x, S) \wedge likes(x, y) \rightarrow \exists V. likesAll(x, V) \wedge SU(S, y, V) \qquad (2')$$



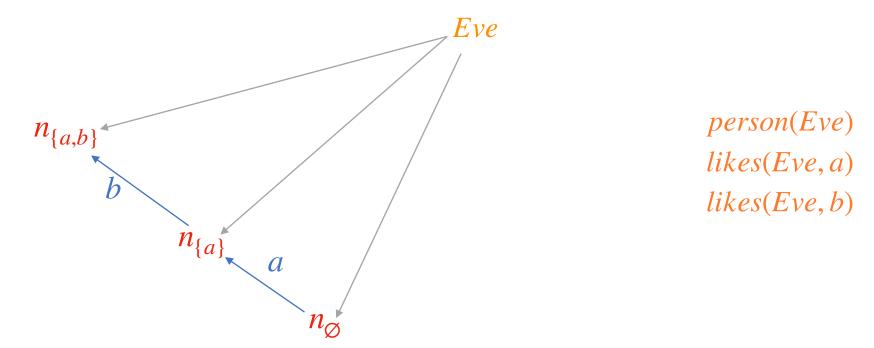
$$person(x) \rightarrow likesAll(x, \emptyset) \qquad (1)$$

$$likesAll(x, S) \wedge likes(x, y) \rightarrow likesAll(x, S \cup \{y\}) \qquad (2)$$

$$\rightarrow \exists V. empty(V) \qquad (1')$$

$$person(x) \wedge empty(Y) \rightarrow likesAll(x, Y) \qquad (1'')$$

$$likesAll(x, S) \wedge likes(x, y) \rightarrow \exists V. likesAll(x, V) \wedge SU(S, y, V) \wedge SU(V, y, V) \qquad (2')$$



→ likesAll

→ SU

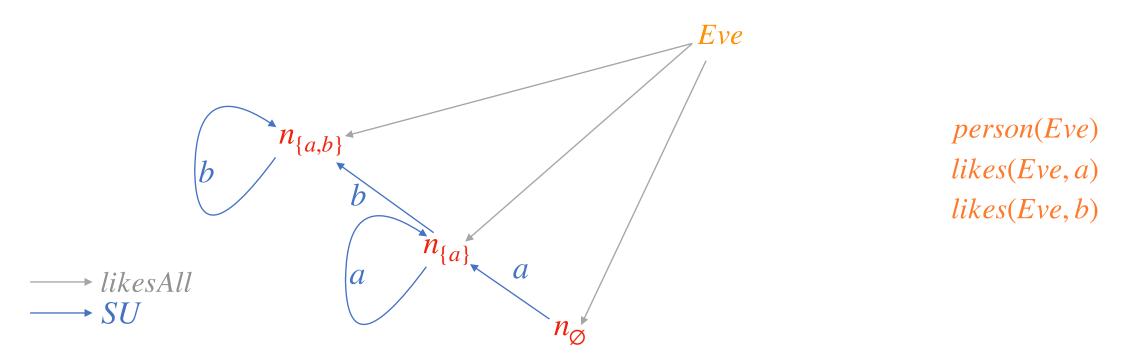
$$person(x) \rightarrow likesAll(x, \emptyset) \qquad (1)$$

$$likesAll(x, S) \wedge likes(x, y) \rightarrow likesAll(x, S \cup \{y\}) \qquad (2)$$

$$\rightarrow \exists V. empty(V) \qquad (1')$$

$$person(x) \wedge empty(Y) \rightarrow likesAll(x, Y) \qquad (1'')$$

$$likesAll(x, S) \wedge likes(x, y) \rightarrow \exists V. likesAll(x, V) \wedge SU(S, y, V) \wedge SU(V, y, V) \qquad (2')$$



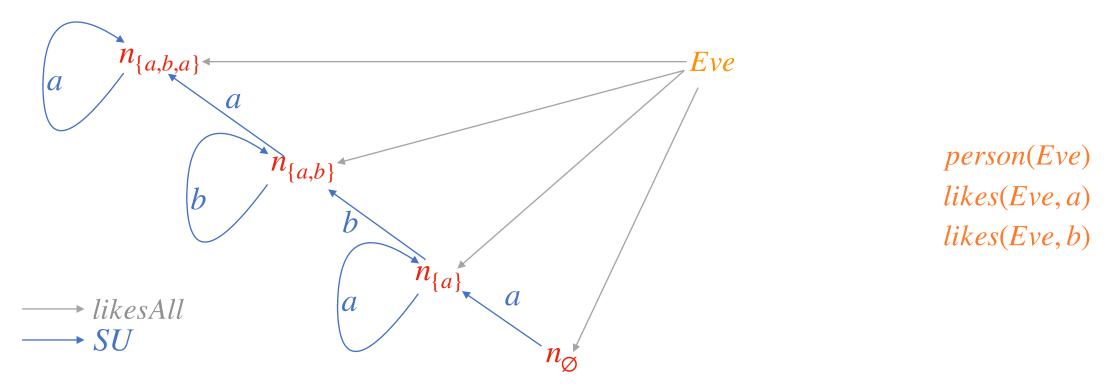
$$person(x) \rightarrow likesAll(x, \emptyset) \qquad (1)$$

$$likesAll(x, S) \wedge likes(x, y) \rightarrow likesAll(x, S \cup \{y\}) \qquad (2)$$

$$\rightarrow \exists V. empty(V) \qquad (1')$$

$$person(x) \wedge empty(Y) \rightarrow likesAll(x, Y) \qquad (1'')$$

$$likesAll(x, S) \wedge likes(x, y) \rightarrow \exists V. likesAll(x, V) \wedge SU(S, y, V) \wedge SU(V, y, V) \qquad (2')$$



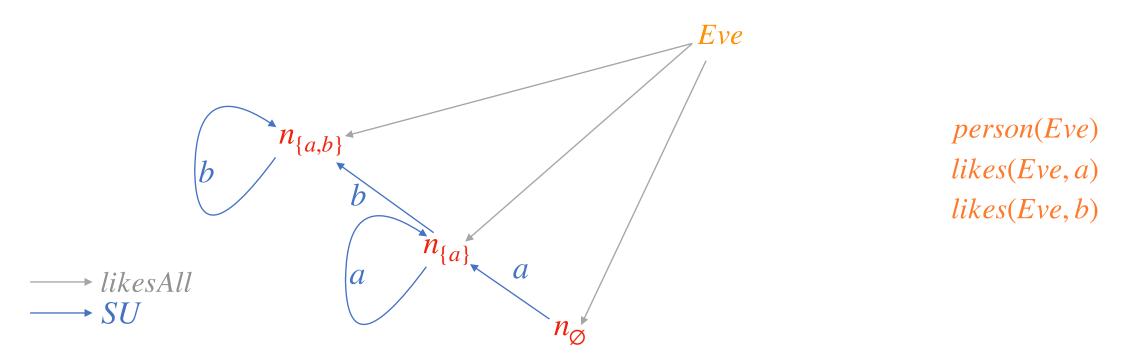
$$person(x) \rightarrow likesAll(x, \emptyset) \qquad (1)$$

$$likesAll(x, S) \wedge likes(x, y) \rightarrow likesAll(x, S \cup \{y\}) \qquad (2)$$

$$\rightarrow \exists V. empty(V) \qquad (1')$$

$$person(x) \wedge empty(Y) \rightarrow likesAll(x, Y) \qquad (1'')$$

$$likesAll(x, S) \wedge likes(x, y) \rightarrow \exists V. likesAll(x, V) \wedge SU(S, y, V) \wedge SU(V, y, V) \qquad (2')$$



$$person(x) \rightarrow likesAll(x, \emptyset) \qquad (1)$$

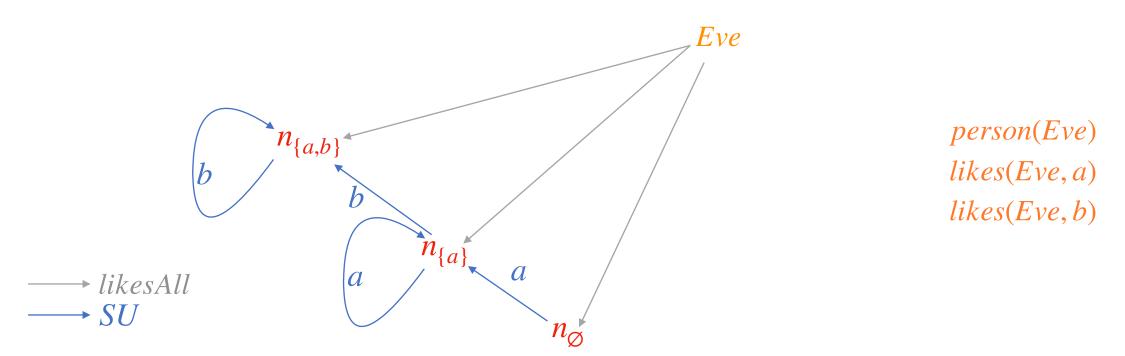
$$likesAll(x, S) \wedge likes(x, y) \rightarrow likesAll(x, S \cup \{y\}) \qquad (2)$$

$$\rightarrow \exists V. empty(V) \qquad (1')$$

$$person(x) \wedge empty(Y) \rightarrow likesAll(x, Y) \qquad (1'')$$

$$likesAll(x, S) \wedge likes(x, y) \rightarrow \exists V. likesAll(x, V) \wedge SU(S, y, V) \wedge SU(V, y, V) \qquad (2')$$

$$SU(U, x, V) \wedge SU(U, y, U) \rightarrow SU(V, y, V) \qquad (2'')$$



$$person(x) \rightarrow likesAll(x, \emptyset) \qquad (1)$$

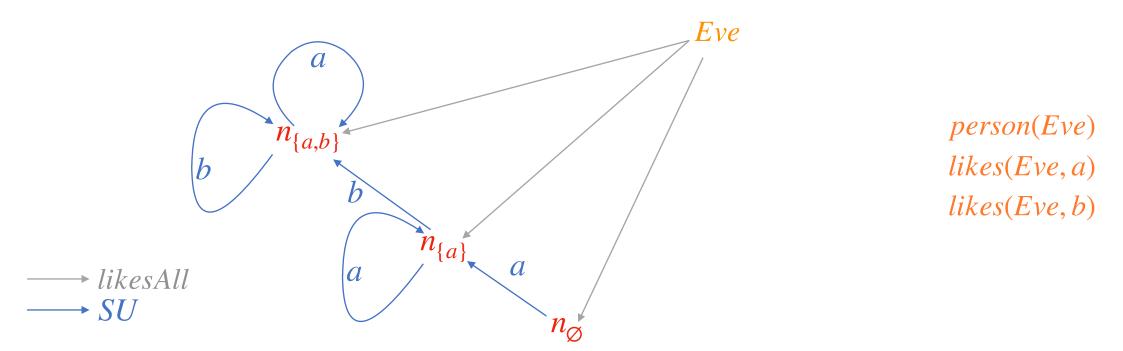
$$likesAll(x, S) \wedge likes(x, y) \rightarrow likesAll(x, S \cup \{y\}) \qquad (2)$$

$$\rightarrow \exists V. empty(V) \qquad (1')$$

$$person(x) \wedge empty(Y) \rightarrow likesAll(x, Y) \qquad (1'')$$

$$likesAll(x, S) \wedge likes(x, y) \rightarrow \exists V. likesAll(x, V) \wedge SU(S, y, V) \wedge SU(V, y, V) \qquad (2')$$

$$SU(U, x, V) \wedge SU(U, y, U) \rightarrow SU(V, y, V) \qquad (2'')$$



# Datalog(S) to existential rules

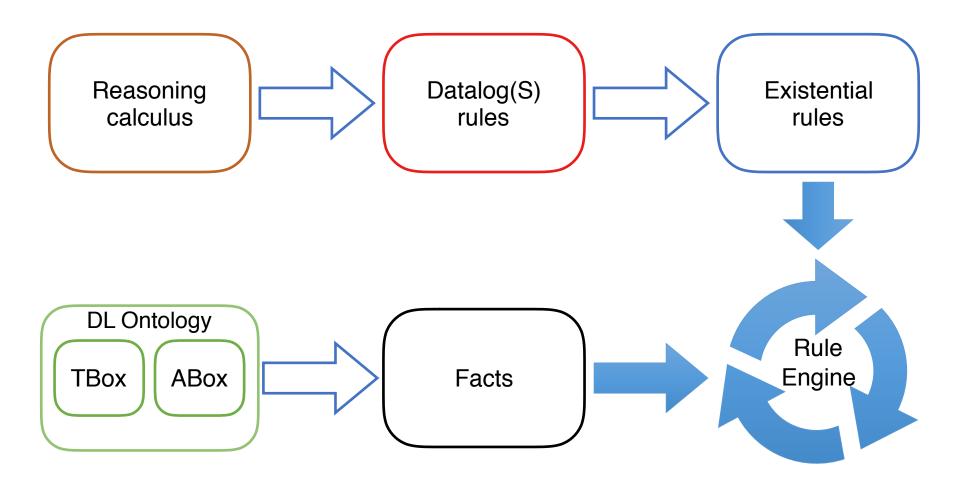
**Theorem:** Any Datalog(S) rule set can be

- polynomially translated
- into a consequence-preserving set of existential rules
- with a terminating Datalog-first standard chase.

✓ Datalog-first is implemented by some rule engines

Datalog(S)
for
DL Reasoning?

# DL Reasoning using Datalog(S)



# Classification for Horn-SHIQ (Kazakov,IJCAI 2009)

*Proof.* By applying structural transformation to  $\mathcal{O}$ , we obtain an ontology  $\mathcal{O}'$  containing only concept inclusions of the form  $A_1 \sqsubseteq A_2$ ,  $A \sqsubseteq \mathsf{st}(C_+)$ , and  $\mathsf{st}(C_-) \sqsubseteq A$ , where  $C_+$  occurs positively in  $\mathcal{O}$  and  $C_-$  occurs negatively in  $\mathcal{O}$ . Since  $\mathcal{O}$  is a Horn  $\mathcal{SHIQ}$  ontology,  $C_+$  can only be of the form  $\top$ ,  $\bot$ , A,  $\neg C$ ,  $C \sqcap D$ ,  $\exists R.C$ ,  $\forall R.C$ ,  $\geqslant nS.C$ , or  $\leqslant 1S.C$ , and  $C_-$  only of the form  $\top$ ,  $\bot$ , A,  $C \sqcap D$ ,  $C \sqcup D$ ,  $\exists R.C$ , or  $\geqslant 1R.C$ .

Concept inclusions of the form  $A \sqsubseteq \operatorname{st}(C_+)$  that are not of form (n1), are transformed to form (n1) as follows:

- $A \sqsubseteq \mathsf{st}(\neg C) = \neg A_C \Rightarrow A \sqcap A_C \sqsubseteq \bot;$
- $A \sqsubseteq \operatorname{st}(\geqslant nS.C) = \geqslant nS.A_C \Rightarrow A \sqsubseteq \exists S.B_i, B_i \sqsubseteq A_C, 1 \le i \le n, B_i \sqcap B_j \sqsubseteq \bot, 1 \le i < j \le n,$  where  $B_i$  are fresh atomic concepts.

Concept inclusions of the form  $st(C_{-}) \sqsubseteq A$  that are not of form (n1) are transformed to form (n1) as follows:

- $\mathsf{st}(C \sqcup D) = A_C \sqcup A_D \sqsubseteq A \Rightarrow A_C \sqsubseteq A, A_D \sqsubseteq A;$
- $\operatorname{st}(\exists R.C) = \exists R.A_C \sqsubseteq A \Rightarrow A_C \sqsubseteq \forall R^-.A;$
- $\bullet \ \operatorname{st}(\geqslant\!1S.C)\!=\!\geqslant\!1S.A_C\sqsubseteq A \ \Rightarrow \ A_C\sqsubseteq \forall S^-.A.$

It is easy to show using Proposition 1, that  $\mathcal{O}' \models \alpha$  iff  $\mathcal{O} \models \alpha$  for every axiom  $\alpha$  containing no new symbols.  $\square$ 

#### 4.2 Elimination of Transitivity

After normalization, we apply a well-known technique, which allows the elimination of transitivity axioms. Transitivity axioms of form (n3) in Lemma 2 can interact only with axioms

$$\begin{array}{c} \mathbf{I}\mathbf{I} & \overline{M \, \square \, A \, \sqsubseteq \, A} & \mathbf{I}\mathbf{2} \, \overline{M \, \sqsubseteq \, \top} \\ \mathbf{R}\mathbf{1} & \overline{M \, \sqsubseteq \, A_i \quad \square \, A_i \, \sqsubseteq \, C \, \in \, \mathcal{O}} \\ \hline \mathbf{R}\mathbf{2} & \overline{M \, \sqsubseteq \, \exists R.N \quad N \, \sqsubseteq \, \bot} \\ \mathbf{R}\mathbf{2} & \overline{M \, \sqsubseteq \, \exists R.N \quad N \, \sqsubseteq \, \bot} \\ \hline \mathbf{R}\mathbf{3} & \overline{M \, \sqsubseteq \, \exists R_1.N \quad M \, \sqsubseteq \, \forall R_2.A \quad R_1 \, \sqsubseteq_{\mathcal{O}} \, R_2} \\ \hline \overline{M \, \sqsubseteq \, \exists R_1.(N \, \sqcap \, A)} \\ \mathbf{R}\mathbf{4} & \overline{M \, \sqsubseteq \, \exists R_1.N \quad N \, \sqsubseteq \, \forall R_2.A \quad R_1 \, \sqsubseteq_{\mathcal{O}} \, R_2} \\ \hline \overline{M \, \sqsubseteq \, A} & \overline{M \, \sqsubseteq \, \exists R_1.N_1 \quad N_1 \, \sqsubseteq \, B \quad R_1 \, \sqsubseteq_{\mathcal{O}} \, R_2} \\ \hline \overline{M \, \sqsubseteq \, \exists R_2.N_2 \quad N_2 \, \sqsubseteq \, B \quad R_2 \, \sqsubseteq_{\mathcal{O}} \, S} \\ \hline \mathbf{R}\mathbf{5} & \overline{M \, \sqsubseteq \, \exists R_1.N_1 \quad N_1 \, \sqsubseteq \, \exists R_2.(N_2 \, \sqcap \, A) \quad R_1 \, \sqsubseteq_{\mathcal{O}} \, S^-} \\ \overline{M \, \sqsubseteq \, \exists R_1.N_1 \quad N_1 \, \sqsubseteq \, \exists R_2.(N_2 \, \sqcap \, A) \quad R_1 \, \sqsubseteq_{\mathcal{O}} \, S^-} \\ \overline{M \, \sqsubseteq \, B \quad N_2 \, \sqcap \, A \, \sqsubseteq \, B} & \overline{R_2 \, \sqsubseteq_{\mathcal{O}} \, S} \\ \hline \mathbf{R}\mathbf{6} & \overline{M \, \sqsubseteq \, \exists S.B} \\ \hline \overline{M \, \sqsubseteq \, A \quad M \, \sqsubseteq \, \exists R_2^-.N_1} \\ \end{array}$$

Table 3: Saturation Rules for Horn SHIQ Ontologies

# Consequence-driven classification

$$\frac{H \sqsubseteq \exists R . K \quad H \sqsubseteq A}{H \sqsubseteq \exists R . (K \sqcap B)} : A \sqsubseteq \forall R . B \in \mathcal{O}$$

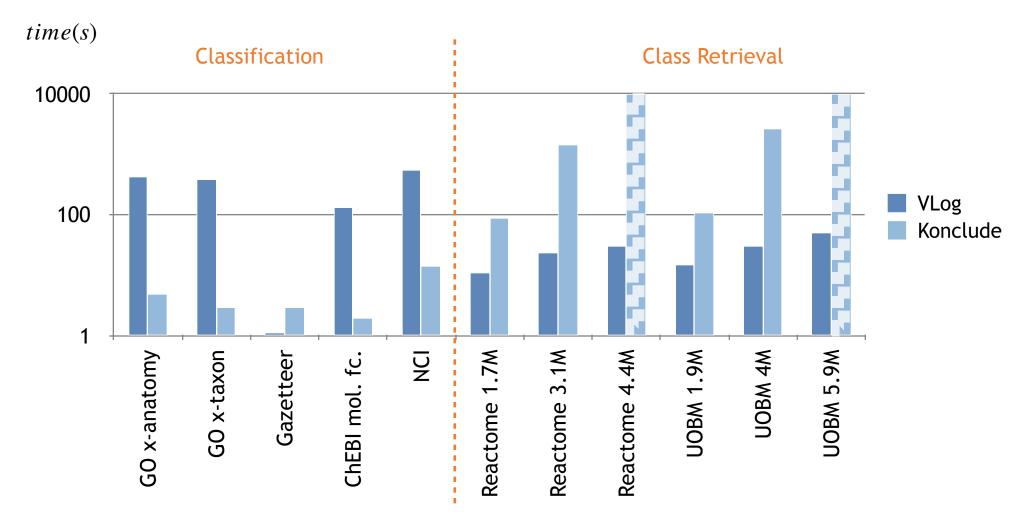
# Consequence-driven classification

$$\frac{H \sqsubseteq \exists R . K \quad H \sqsubseteq A}{H \sqsubseteq \exists R . (K \sqcap B)} : A \sqsubseteq \forall R . B \in \mathcal{O}$$

```
Exists(H, r, K) \land SubClass(H, a) \land ax_{\sqsubseteq \forall}(a, r, b)

\rightarrow Exists(H, r, K \cup \{b\})
```

### Evaluation



# What can we use Datalog(S) for?



- Consequence-based classification and class retrieval for **Horn-ALC**:
  - Kazakov (IJCAI 2011)



- Fact entailment for guarded existential rules:
  - Ahmetaj et al. (ICDT 2018)

# What can we use Datalog(S) for?



- Consequence-based classification and class retrieval for Horn-ALC:
  - Kazakov (IJCAI 2011)



- Fact entailment for guarded existential rules:
  - Ahmetaj et al. (ICDT 2018)
- Combined-approach materialisation for Horn-ALCHOIQ:
  - Carral et al. (KR 2018)
- Reasoning tasks for non-Horn DLs
  - Simančík et al. (IJCAI 2011): ALCH

# What can we use Datalog(S) for?



- Consequence-based classification and class retrieval for Horn-ALC:
  - Kazakov (IJCAI 2011)



- Fact entailment for guarded existential rules:
  - Ahmetaj et al. (ICDT 2018)
- Combined-approach materialisation for Horn-ALCHOIQ:
  - Carral et al. (KR 2018)
- Reasoning tasks for non-Horn DLs
  - Simančík et al. (IJCAI 2011): ALCH
- ?

# Summary

We provide a practical new way of solving

- ExpTime-complete problems
- using current existential rule engines

#### Next steps:

- Logical reasoning: solve new ExpTime-complete problems
- Rule engine development: optimise and benchmark
- Characterising chase termination: discover syntactic criteria

