



PRACTICAL USES OF EXISTENTIAL RULES IN KNOWLEDGE REPRESENTATION

Part 2: Solving Horn-ALC Classification with Existential Rules

David Carral,¹ Markus Krötzsch,¹ and Jacopo Urbani² 1. TU Dresden 2. VU University Amsterdam

Special thanks to Irina Dragoste,¹ Ceriel Jacobs,² and Maximilian Marx¹ for their invaluable contributions to the software used in this tutorial

KR 2020, September 13, 2020

Outline

Goal

Show some example where either **rules** or **related ideas** were crucial to achieve the state of the art

- Horn-*ALC* reasoning
- PLP
- Data integration
- Stream reasoning

1st Scenario: Horn-*ALC* reasoning

The Description Logic Horn-*ALC*: Syntax

Definition. A Horn-*ALC* ontology is a set of Horn-*ALC* axioms:

 $A \sqsubseteq \bot \quad \top \sqsubseteq B \quad A \sqsubseteq B \quad A \sqcap E \sqsubseteq B \quad \exists R.A \sqsubseteq B \quad A \sqsubseteq \forall R.B \quad A \sqsubseteq \exists R.B$

In the above; *A*, *B*, and *E* are concept names; and *R* is a role name.

Remark. Note the axioms of the form $A \sqsubseteq \forall R.B$, which are not \mathcal{EL} , such as:

CheesePizza ⊑ ∀HasTopping.Cheese

The axiom states that "all toppings in a cheese pizza are cheese toppings".

Even though Horn- \mathcal{ALC} is not much more expressive than \mathcal{EL} , (Krötzsch, Rudolph, and Hitzler 2013) have showed that:

Theorem. Solving classification over Horn-*ALC* is ExpTime-complete.

A Consequence-Based Calculus to Solve Classification

$$\begin{array}{ll} R_A^C \\ R_A^{\Box} \\ \hline A \sqsubseteq A \\ \hline C \sqsubseteq B \\ \hline C \sqsubseteq A \\ \hline C \sqsubseteq B \\ \hline C \sqsubseteq A \\ \hline C \sqsubseteq B \\ \hline C \sqsubseteq A \\ \hline C \sqsubseteq B \\ \hline C \sqsubseteq A \\ \hline C \sqsubseteq B \\ \hline C \sqsubseteq A \\ \hline C \sqsubseteq B \\ \hline C \sqsubseteq B \\ \hline C \sqsubseteq A \\ \hline C \sqsubseteq B \\ \hline C \sqsubseteq B \\ \hline C \sqsubseteq A \\ \hline C \sqsubseteq B \\ \hline C \blacksquare C \\ \hline C \blacksquare B \\ \hline C \blacksquare C \\$$

Figure: Classification Calculus for Horn- \mathcal{ALC} . Where *A*, *B*, and *E* are concept names; *R* is a role name; and \mathbb{C} and \mathbb{D} are conjunctions of concept names

Remark. The above procedure is based on the work of (Kazakov 2009).

Consequence-Based Calculus: Complexity

Theorem. The Horn- \mathcal{ALC} classification calculus runs in exponential time in the size of the input ontology *O*.

Remark. Note that this calculus produces inferences of the form (1) $\mathbb{C} \sqsubseteq B$ and (2) $\mathbb{C} \sqsubseteq \exists R.\mathbb{D}$ where *B* is a concept name, *R* is a role name, and \mathbb{C} and \mathbb{D} are conjunctions of concept names. Therefore, on input *O*, the calculus may produce at most $2^{|\text{Concepts}(O)|} \times |\text{Concepts}(O)|$ and $2 \times 2^{|\text{Concepts}(O)|} \times |\text{Roles}(O)|$ inferences of type (1) and (2), respectively.

Remark. Since classification over Horn- \mathcal{RLC} is an ExpTime-complete problem, the calculus is worst-case optimal.

Because of the following result, we can not implement the Horn- \mathcal{RLC} classification calculus using a fixed Datalog rule set:

Theorem. The data complexity of fact entailment over Datalog is in P.

Assume that we can implement the Horn- \mathcal{ALC} classification calculus with a fixed Datalog rule set (as we did for the \mathcal{EL} classification calculus). Then:

- 1. By the above theorem, we could solve Horn-ALC classification in polynomial time.
- 2. By (1), we could solve an ExpTime-hard problem in polynomial time.
- 3. By (2), P = ExpTime(4)

Remark. To implement the Horn- \mathcal{ALC} classification calculus (or any other procedure that solves Horn- \mathcal{ALC} classification), we need a rule-based language with ExpTime-hard data complexity!

We study Datalog(S), an extension of Datalog that can model exponential computations.

Example. Consider the following Datalog(S) rule set: $Person(x) \rightarrow LikesAll(x, \emptyset)$ $LikesAll(x, S) \wedge Likes(x, y) \rightarrow LikesAll(x, S \cup \{y\})$ $LikesAll(x, S) \rightarrow AllLikeAll(\{x\}, S)$ $AllLikeAll(S, T) \wedge LikesAll(x, T) \rightarrow AllLikeAll(X \cup \{x\}, T)$ $AllLikeAll(S, S) \wedge alice \in S \rightarrow CliqueOfAlice(S)$

Theorem. Checking fact entailment for Datalog(S) is ExpTime-complete for both data and combined complexity.

See (Carral et al. 2019) for a complete proof of the above result.

Using a function to encode the axioms and entities in an input ontology as facts and a fixed Datalog(S) rule set, we can implement the Horn- \mathcal{ALC} classification calculus.

Example. For an ontology *O*, let Facts(*O*) be the fact set such that: $A \sqsubseteq \bot \mapsto nf:axiom_{\sqsubseteq}(c_A, c_{\bot})$ $\exists R.A \sqsubseteq B \mapsto nf:axiom_{\exists \sqsubseteq}(c_A, c_R, c_B)$ $\top \sqsubseteq B \mapsto nf:axiom_{\sqsubseteq}(c_{\top}, c_B)$ $A \sqsubseteq \forall R.B \mapsto nf:axiom_{\sqsubseteq \forall}(c_A, c_R, c_B)$ $A \sqsubseteq B \mapsto nf:axiom_{\sqsubseteq}(c_A, c_B)$ $A \sqsubseteq \exists R.B \mapsto nf:axiom_{\sqsubseteq \exists}(c_A, c_R, c_B)$ $A \sqcap E \sqsubseteq B \mapsto nf:axiom_{\sqcap \sqsubseteq}(c_A, c_E, c_B)$ $A \in Concepts(O) \mapsto nf:concept(c_E)$ In the above; c_A, c_B, c_E, c_{\top} , and c_{\bot} are fresh constants unique for A, B, E, \top , and \bot , respectively; and c_R is a fresh constant unique R.

We translate the production rules in the Horn- \mathcal{ALC} classification calculus (left) into analogous Datalog(S) rules (right):

$\frac{1}{A \sqsubseteq A} : A \in Concepts(O)$	$nf:concept(a) \\ \rightarrow SC(\{a\}, a)$
$\frac{\mathbb{C} \sqsubseteq \exists R.\mathbb{D}}{\mathbb{D} \sqsubseteq D} : D \in \mathbb{D}$	$Ex(C, r, D) \land d \in D$ $\rightarrow SC(D, d)$
$\frac{\mathbb{C} \sqsubseteq A}{\mathbb{C} \sqsubseteq B} : \top \sqsubseteq B \in O$	$SC(C, a) \land nf:axiom_{\square}(c_{\top}, b)$ $\rightarrow SC(C, b)$
$\frac{\mathbb{C} \sqsubseteq A}{\mathbb{C} \sqsubseteq B} : A \sqsubseteq B \in O$	$SC(C, a) \land nf:axiom_{\Box}(a, b)$ $\rightarrow SC(C, b)$
$\frac{\mathbb{C} \sqsubseteq A \mathbb{C} \sqsubseteq E}{\mathbb{C} \sqsubseteq B} : A \sqcap E \sqsubseteq B \in O$	$\begin{split} SC(C,a) \wedge SC(C,e) \wedge \texttt{nf:axiom}_{\sqcap \sqsubseteq}(a,e,b) \\ \to SC(C,b) \end{split}$

We translate the production rules in the Horn- \mathcal{ALC} classification calculus (left) into analogous Datalog(S) rules (right):

$\frac{\mathbb{C} \sqsubseteq A}{\mathbb{C} \sqsubseteq \exists R.B} : A \sqsubseteq \exists R.B \in O$	$SC(C, a) \land nf:axiom_{\exists}(a, r, b) \\ \rightarrow Ex(C, r, \{b\})$
$\frac{\mathbb{C} \sqsubseteq \exists R.\mathbb{D} \mathbb{D} \sqsubseteq A}{\mathbb{C} \sqsubseteq B} : \exists R.A \sqsubseteq B \in O$	$\begin{split} \mathtt{Ex}(C,r,D)\wedge \mathtt{SC}(D,a)\wedge \mathtt{nf:axiom}_{\exists\sqsubseteq}(r,a,b)\\ \to \mathtt{SC}(C,b) \end{split}$
$\frac{\mathbb{C} \sqsubseteq \exists R. \mathbb{D} \mathbb{D} \sqsubseteq \bot}{\mathbb{C} \sqsubseteq \bot}$	$\begin{split} \mathtt{Ex}(C,r,D) \wedge \mathtt{SC}(D,c_{\perp}) \\ \to \mathtt{SC}(C,c_{\perp}) \end{split}$
$\frac{\mathbb{C} \sqsubseteq \exists R.\mathbb{D} \mathbb{C} \sqsubseteq A}{\mathbb{C} \sqsubseteq \exists R.(\mathbb{D} \sqcap B)} : A \sqsubseteq \forall R.B \in O$	$\begin{split} \mathtt{Ex}(C,r,D) &\wedge \mathtt{SC}(C,a) \wedge \mathtt{nf:axiom}_{\sqsubseteq \forall}(a,r,b) \\ &\rightarrow \mathtt{Ex}(C,r,D \cup \{b\}) \end{split}$

Definition. Let \mathcal{R}_{HALC} be the rule set containing all of the above rules: $nf:concept(a) \rightarrow SC(\{a\}, a)$ $SC(C, a) \land nf:axiom_{\square}(c_{\top}, b) \rightarrow SC(C, b)$ $Ex(C, r, D) \land d \in D \rightarrow SC(D, d)$ $SC(C, a) \land nf:axiom_{\square}(a, b) \rightarrow SC(C, b)$ $SC(C, a) \land SC(C, e) \land nf:axiom_{\square\square}(a, e, b) \rightarrow SC(C, b)$ $SC(C, a) \land nf:axiom_{\square}(a, r, b) \rightarrow Ex(C, r, \{b\})$ $Ex(C, r, D) \land SC(D, a) \land nf:axiom_{\square\square}(r, a, b) \rightarrow SC(C, b)$ $Ex(C, r, D) \land SC(D, c_{\bot}) \rightarrow SC(C, c_{\bot})$ $Ex(C, r, D) \land SC(C, a) \land nf:axiom_{\square}(a, r, b) \rightarrow Ex(C, r, D \cup \{b\})$

Theorem. Consider a Horn- \mathcal{ALC} ontology O and an axiom of the form $A \sqsubseteq B$. Then, $O \models A \sqsubseteq B$ if and only if $\mathcal{R}_{\mathsf{HALC}} \cup \mathsf{Facts}(O) \models \mathsf{SC}(c_A, c_B)$.

David Carral, September 13, 2020

Implementing the Classification Calculus: 3-Rules

Alas, VLog does not support Datalog(S) reasoning. But there maybe some other rule-based language with ExpTime-hard data complexity that we can use!

The following result is a recent finding by (Krötzsch, Marx, and Rudolph 2019):

Theorem. The data complexity of fact entailment over rule sets that terminate with respect to the restricted chase is ExpTime-hard.

Remark. Note that the data complexity of fact entailment over existential rule sets that terminate with respect to the Skolem chase is in P.

(Carral et al. 2019) have proposed a translation from Datalog(S) into ∃-rules such that:

- The resulting rule sets terminate w.r.t. the Datalog-first restricted chase.
- Fact entailment is "preserved".

$$\rightarrow \exists V. empty(V)$$

$$person(x) \land empty(Y) \rightarrow likesAll(x, Y)$$

$$(1.1)$$

$$(1.2)$$

$\rightarrow \exists V. empty(V)$	(1.1)
$person(x) \land empty(Y) \rightarrow likesAll(x, Y)$	(1.2)
$likesAll(x, S) \land likes(x, y) \rightarrow \exists V. likesAll(x, V) \land SU(S, y, V)$	(2.1)

 $\mathsf{Person}(x) \to \mathsf{LikesAll}(x, \emptyset) \qquad \mathsf{LikesAll}(x, S) \land \mathsf{Likes}(x, y) \to \mathsf{LikesAll}(x, S \cup \{y\})$

$\rightarrow \exists V. empty(V)$	(1.1)
$person(x) \land empty(Y) \rightarrow likesAll(x, Y)$	(1.2)
$likesAll(x, S) \land likes(x, y) \rightarrow \exists V. likesAll(x, V) \land SU(S, y, V)$	(2.1)

person(eve)
likes(eve, a)
likes(eve, b)

 $\mathsf{Person}(x) \to \mathsf{LikesAll}(x, \emptyset) \qquad \mathsf{LikesAll}(x, S) \land \mathsf{Likes}(x, y) \to \mathsf{LikesAll}(x, S \cup \{y\})$

$\rightarrow \exists V. empty(V)$	(1.1)
$person(x) \land empty(Y) \rightarrow likesAll(x, Y)$	(1.2)
$likesAll(x, S) \land likes(x, y) \rightarrow \exists V. likesAll(x, V) \land SU(S, y, V)$	(2.1)

eve

person(eve)
likes(eve, a)
likes(eve, b)

 $\mathsf{Person}(x) \to \mathsf{LikesAll}(x, \emptyset) \qquad \mathsf{LikesAll}(x, S) \land \mathsf{Likes}(x, y) \to \mathsf{LikesAll}(x, S \cup \{y\})$

$\rightarrow \exists V. empty(V)$	(1.1)
$person(x) \land empty(Y) \rightarrow likesAll(x, Y)$	(1.2)
$likesAll(x, S) \land likes(x, y) \rightarrow \exists V. likesAll(x, V) \land SU(S, y, V)$	(2.1)

eve

person(eve)
likes(eve, a)
likes(eve, b)

 n_{\emptyset}

 $\mathsf{Person}(x) \to \mathsf{LikesAll}(x, \emptyset) \qquad \mathsf{LikesAll}(x, S) \land \mathsf{Likes}(x, y) \to \mathsf{LikesAll}(x, S \cup \{y\})$





→ likesAll

David Carral, September 13, 2020

Practical Uses of Existential Rules in Knowledge Representation







$$\rightarrow \exists V. empty(V)$$

$$person(x) \land empty(Y) \rightarrow likesAll(x, Y)$$

$$likesAll(x, S) \land likes(x, y) \rightarrow \exists V. likesAll(x, V) \land SU(S, y, V)$$

$$(2.1)$$

















Solving ExpTime-hard Problems

Step-by-step procedure to implement an ExpTime algorithm with VLog:

- 1. Encode input using a set of facts \mathcal{F} .
- 2. Encode ExpTime algorithm using a fixed Datalog(S) rule set \mathcal{R} .
- 3. Apply the translation by (Carral et al. 2019) to \mathcal{R} to obtain a set \mathcal{R}' of existential rules such that:
 - The rule set \mathcal{R}' "preserves" fact entailment over \mathcal{R} .
 - The rule set \mathcal{R}' terminates w.r.t. the Datalog-first restricted chase.
- 4. Use VLog to compute all of the consequences of $\mathcal{R}' \cup \mathcal{F}$.

Remark. For a detailed explanation of the above procedure, see (Carral, Dragoste, and Rudolph 2020).

Evaluation: Classification

Konclude	VLog	#SC	#Ax.	ID
5s	432s	1051K	223K	00040
3s	387s	718K	142K	00048
3s	1s	162K	318K	00477
2s	132s	965K	159K	00533
14s	549s	2283K	152K	00786

Figure: Ontologies and results for classification showing: axiom count, number of SC facts derived, and reasoning times for VLog and Konclude

Remark. Presented in (Carral et al. 2019).

Evaluation: Class Retrieval

Definition. A Horn- \mathcal{ALC} ontology is a set of Horn- \mathcal{ALC} axioms: $A \sqsubseteq \bot \qquad \top \sqsubseteq B \qquad A \sqsubseteq B \qquad A \sqcap E \sqsubseteq B$ $\exists R.A \sqsubseteq B \qquad A \sqsubseteq \forall R.B \qquad A \sqsubseteq \exists R.B \qquad A(a) \qquad R(a,b)$ Where *A*, *B*, and *E* are concepts; *R* is a role; and *a* and *b* are individuals.

Definition. Class retrieval is the reasoning task of computing all axioms of the form A(a) that are logically entailed by some input ontology *O*.

Remark. The Horn-*ALC* classification calculus can be extended with 3 rules (as done by (Carral et al. 2019)) to solve class retrieval.

Evaluation: Class Retrieval



Experimental results for class retrieval for VLog (pink/grey) and Konclude (black)

Remark. Presented in (Carral et al. 2019).

Conclusions and Future Work

Remark. We can use VLog to solve (ExpTime-)hard problems!

Future work:

- Rulewerk Extension: translate Datalog(S) to existential rules
- VLog Extension: native support for Datalog(S)
- Implement existing calculi using our approach

References I

Carral, David, Irina Dragoste, and Sebastian Rudolph (2020). "Reasoner = Logical Calculus + Rule Engine". In: Künstliche Intell.

- Carral, David et al. (2019). "Chasing Sets: How to Use Existential Rules for Expressive Reasoning". In: P. of the 28th Int. Joint Conf. on AI (IJCAI 2019).
- Kazakov, Yevgeny (2009). "Consequence-Driven Reasoning for Horn-*SHIQ* Ontologies". In: **P. of the 21st Int. Joint Conf. on AI (IJCAI 2009)**.
- Krötzsch, Markus, Maximilian Marx, and Sebastian Rudolph (2019). "The Power of the Terminating Chase (Invited Talk)". In: P. of the 22nd Int. Conf. on Database Theory (ICDT) 2019).
- Krötzsch, Markus, Sebastian Rudolph, and Pascal Hitzler (2013). "Complexities of Horn Description Logics". In: ACM Trans. Comput. Log.