

Stefan Borgwardt
Technische Universität Dresden

Explaining Description Logic Reasoning

DL'26, Lisboa, Portugal, 19th July, 2026

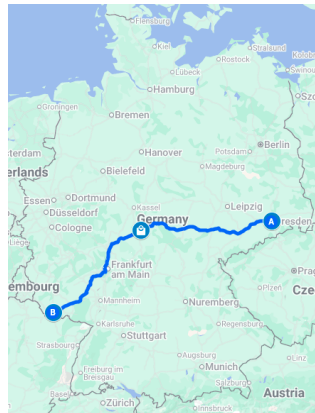


Team

Logical Reasoning

Human-Computer Interaction

Natural Language Generation



Map data ©2026 GeoBasis-DE/BKG (©2009),
Google, Inst. Geogr. Nacional

Assistants: Brisca Balthes, Steffen Breuer, Kathryn Chapman, Tom Frieze, Philipp Max Hermann, Mikhail Kotlov, Duy Nhu, Fedor Grigoryevich Pogulskiy, Alexej Popovič, Anna Milena Rothermel, Lukas Schmitz, Yifan Wang, Hui-Syuan Yeh

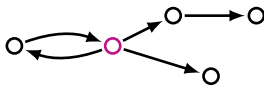
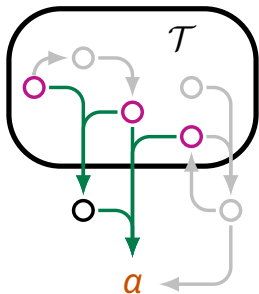


inspection time



Outline

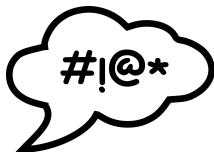
Proofs



Missing entailments

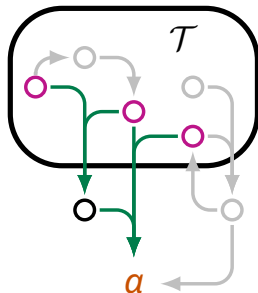


User studies



Verbalization

Justifications



TBox (ontology, knowledge base, logical theory) $\mathcal{T}$:

Collection of axioms (formulae, rules, facts)

Entailment (consequence, query) $\mathcal{T} \models a$:

Axiom (formula, rule, fact) that follows logically from the TBox

Justification (MUS, minimal conflict set, why-provenance):

Minimal subset of $\mathcal{T}$ from which a follows

(Horridge, Bail, Parsia, and Sattler 2013; Peñaloza and Ricca 2022)

Justifications are sometimes not enough for ...

complex reasoning

novice users

LLM-generated explanations

Explanation 135

☒ Display laconic explanation

Explanation for: 'positive regulation of unsaturated fatty acid biosynthetic process by positive regulation of transcription from RNA polymerase II promoter' SubClassOf 'regulation of cellular ketone metabolic process by positive

- 1) 'positive regulation of unsaturated fatty acid biosynthetic process by positive regulation of transcription from RNA polymerase II promoter' SubClassOf 'regulation of fatty acid biosynthetic process by regulation of transcription from RNA polymerase II promoter' In 27 other justifications ?
- 2) 'regulation of fatty acid biosynthetic process by regulation of transcription from RNA polymerase II promoter' SubClassOf regulates some 'regulation of fatty acid biosynthetic process by transcription from RNA polymerase II promoter' In 9 other justifications ?
- 3) 'regulation of fatty acid biosynthetic process by transcription from RNA polymerase II promoter' SubClassOf 'regulation of fatty acid biosynthetic process' In 10 other justifications ?
- 4) 'regulation of fatty acid biosynthetic process' SubClassOf regulates some 'fatty acid biosynthetic process' In 16 other justifications ?
- 5) 'fatty acid biosynthetic process' SubClassOf 'fatty acid metabolic process' In 65 other justifications ?
- 6) Transitive: regulates In 18 other justifications ?
- 7) 'positive regulation of unsaturated fatty acid biosynthetic process by positive regulation of transcription from RNA polymerase II promoter' SubClassOf biological_process and ('positively regulates' some 'transcription by RNA polymerase II') In NO other justifications ?
- 8) 'regulation of cellular ketone metabolic process' SubClassOf regulates some 'cellular ketone metabolic process' In 67 other justifications ?
- 9) 'regulation of fatty acid metabolic process' SubClassOf 'regulation of cellular ketone metabolic process' In 135 other justifications ?
- 10) biological_process and (regulates some 'fatty acid metabolic process') SubClassOf 'regulation of fatty acid metabolic process' In NO other justifications ?
- 11) biological_process and (regulates some 'cellular ketone metabolic process') and ('positively regulates' some 'transcription by RNA polymerase II') SubClassOf 'regulation of cellular ketone metabolic process by positive regulation of transcription from RNA polymerase II promoter' In NO other justifications ?

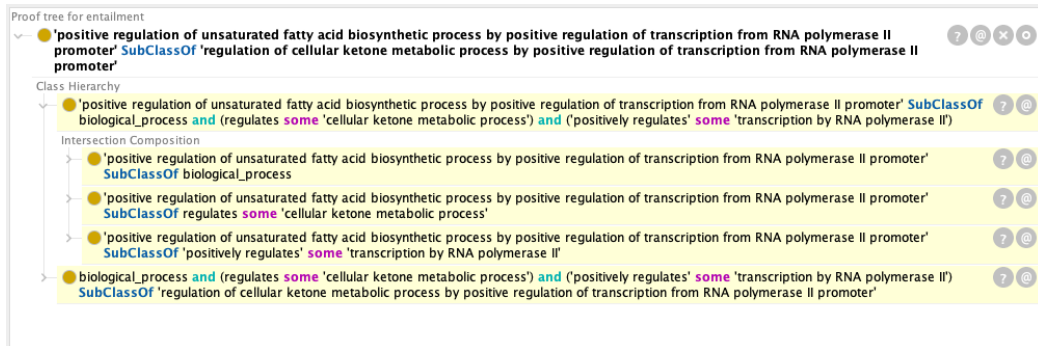
Justification in the ontology editor PROTÉGÉ

Justifications are sometimes not enough for ...

complex reasoning

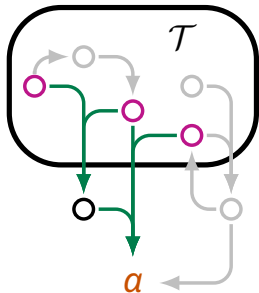
novice users

LLM-generated explanations



Proof in the ontology editor PROTÉGÉ

Proofs are hypergraphs



Inference step (hyperedge) (S, d) : List of premises S + conclusion d

Proof: **acyclic, non-redundant** hypergraph with sink a

Derivation structure $\mathfrak{D}(\mathcal{T}, a)$: Hypergraph of **all admissible inference steps** for $\mathcal{T} \models a$

Goal: Find optimal proofs in $\mathfrak{D}(\mathcal{T}, a)$

Measures: Size (5), tree size (6), depth (2), justification size (3), ...

Complexity of finding proofs

“Is there a proof in $\mathfrak{D}(\mathcal{T}, a)$ with measure $\leq p$?”

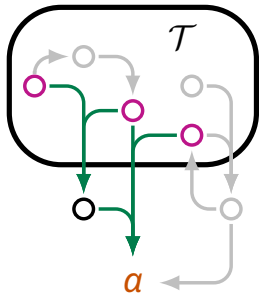
(Alrabbaa, Baader, B, Koopmann, and Kovtunova 2020, 2021)

	polynomial unary	polynomial binary	exponential unary	exponential binary
Φ -measures	$\leq \text{NP}$	$\leq \text{NP}$	$\leq \text{NExpTime}$	$\leq \text{NExpTime}$
Justification size	NP	NP	ExpTime	ExpTime
Size	NP	NP	NP	NExpTime
Monotone recursive Φ -measures	$\leq \text{P}$	$\leq \text{P}$	$\leq \text{ExpTime}$	$\leq \text{ExpTime}$
Tree size	P	P	NP	PSpace
Depth	P	P	PSpace	ExpTime

guess proof in $\mathfrak{D}(\mathcal{T}, a)$

polynomial Dijkstra-like
algorithm for directed
hypergraphs

Computing proofs



Consequence-based reasoners (ELK, CLIPPER, SEQUOIA)

(Kazakov, Krötzsch, and Simancik 2014; Eiter, Ortiz, Simkus, Tran, and Xiao 2012; Tena Cucala, Cuenca Grau, and Horrocks 2021)

Justification-oriented proofs

(Horridge, Parsia, and Sattler 2010)

Concept interpolation: $A \sqsubseteq ? \sqsubseteq B$

(Schlobach 2004; Peuter 2024)

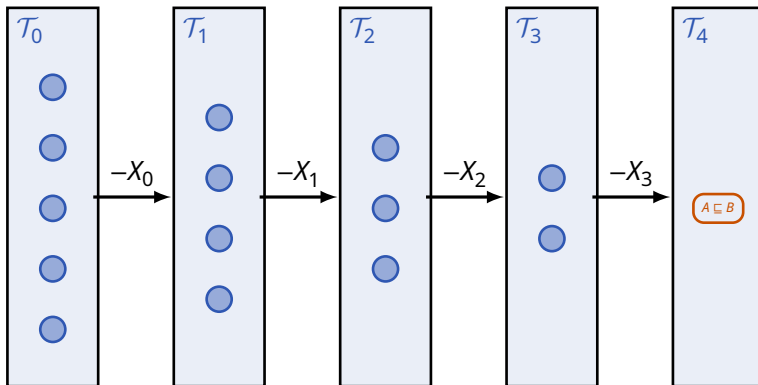
Elimination proofs based on forgetting (LETHE, FAME)

(Alrabbaa, Baader, B, Koopmann, and Kovtunova 2020)

Computing proofs

Elimination proofs based on forgetting (LETHE, FAME)

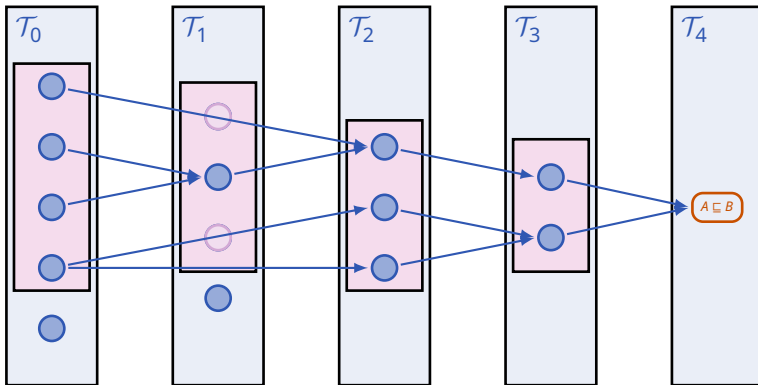
(Alrabbaa, Baader, B, Koopmann, and Kovtunova 2020)



Computing proofs

Elimination proofs based on forgetting (LETHE, FAME)

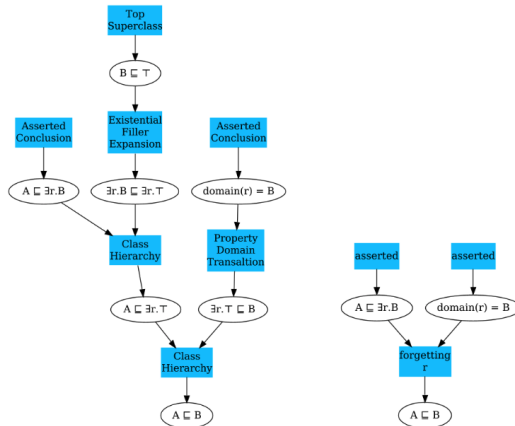
(Alrabbaa, Baader, B, Koopmann, and Kovtunova 2020)



Computing proofs

Elimination proofs based on forgetting (LETHE, FAME)

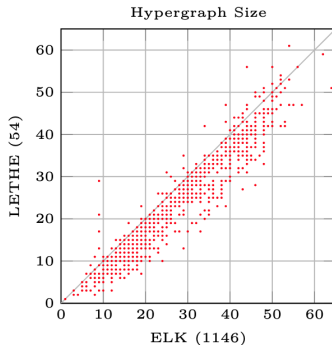
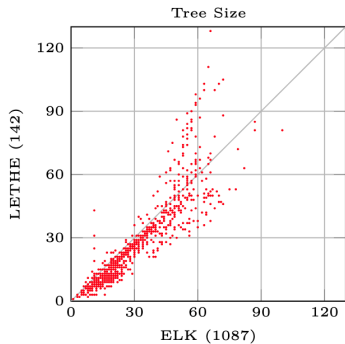
(Alrabbaa, Baader, B, Koopmann, and Kovtunova 2020)



Size of proofs

(Alrabbaa, Baader, B, Koopmann, and Kovtunova 2020)

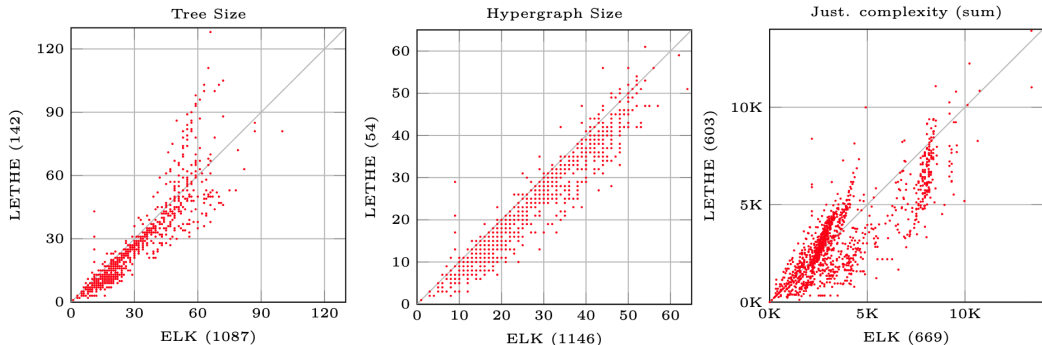
- Prototypical proofs from ORE 2015 ontologies ($\mathcal{ELH}$)
- Proof of minimal size = proof of minimal tree size



Size of proofs

(Alrabbaa, Baader, B, Koopmann, and Kovtunova 2020)

- Prototypical proofs from ORE 2015 ontologies ($\mathcal{ELH}$)
- Proof of minimal size = proof of minimal tree size

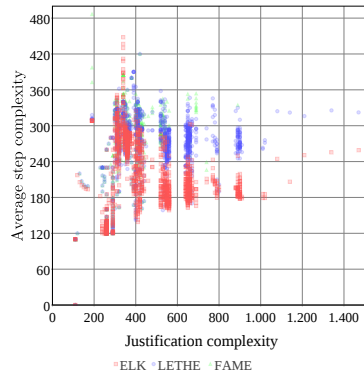
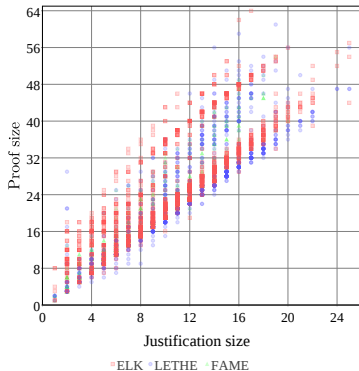


(Horridge, Bail, Parsia, and Sattler 2013)

Proofs vs. justifications

(B 2021)

- justifications of average size 10.7 (max. 25), average complexity 495 (max. 1470)



Proofs for ontology-mediated queries

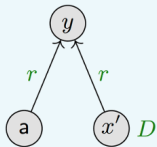
(Borgida, Calvanese, and Rodriguez-Muro 2008; Alrabbaa, B, Koopmann, and Kovtunova 2022)

ABox: $A(a)$

TBox: $A \sqsubseteq \exists r.B, B \sqsubseteq \exists s.A, \exists s \sqcap \exists r^- \sqsubseteq E, \exists r.E \sqsubseteq D$

Query:

$\exists x', y. r(x, y) \wedge r(x', y) \wedge D(x')$



Proofs for ontology-mediated queries

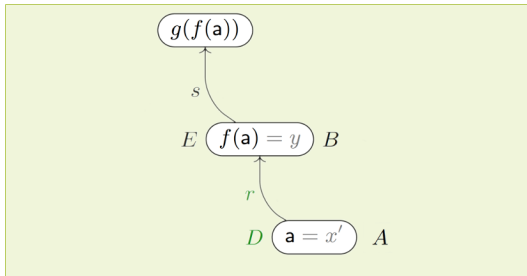
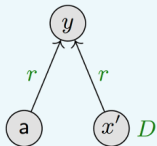
(Borgida, Calvanese, and Rodriguez-Muro 2008; Alrabbaa, B, Koopmann, and Kovtunova 2022)

ABox: $A(a)$

TBox: $A \sqsubseteq \exists r.B$, $B \sqsubseteq \exists s.A$, $\exists s \sqcap \exists r^- \sqsubseteq E$, $\exists r.E \sqsubseteq D$

Query:

$\exists x', y. r(x, y) \wedge r(x', y) \wedge D(x')$



Proofs for ontology-mediated queries

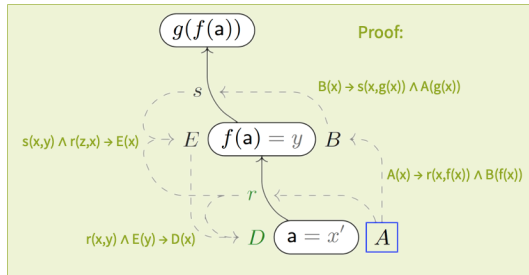
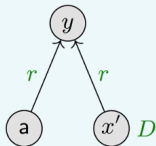
(Borgida, Calvanese, and Rodriguez-Muro 2008; Alrabbaa, B, Koopmann, and Kovtunova 2022)

ABox: $A(a)$

TBox: $A \sqsubseteq \exists r.B, B \sqsubseteq \exists s.A, \exists s \sqcap \exists r^- \sqsubseteq E, \exists r.E \sqsubseteq D$

Query:

$\exists x', y. r(x, y) \wedge r(x', y) \wedge D(x')$



Proofs for ontology-mediated queries

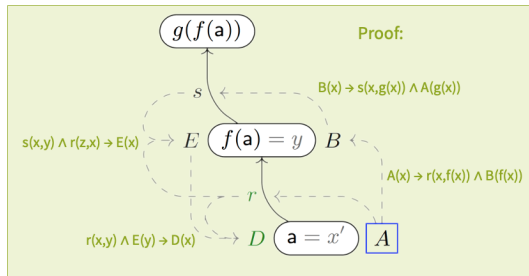
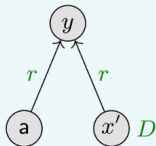
(Borgida, Calvanese, and Rodriguez-Muro 2008; Alrabbaa, B, Koopmann, and Kovtunova 2022)

ABox: $A(a)$

TBox: $A \sqsubseteq \exists r.B, B \sqsubseteq \exists s.A, \exists s \sqcap \exists r^- \sqsubseteq E, \exists r.E \sqsubseteq D$

Query:

$\exists x', y. r(x, y) \wedge r(x', y) \wedge D(x')$



	DL-Lite		$\mathcal{EL}$		Horn- $\mathcal{ALCHOI}$
	ELQ	CQ	IQ	CQ	CQ
Domain size	NP-c		in ExpTime		in NExpTime
Tree size	in P	NP-c	P-c	NP-c	in PSpace
Proof size	NP-c		in ExpTime		
Proof size bound	polynomial		exponential		double exponential

Proofs with concrete domains

(Alrabbaa, Baader, B, Koopmann, and Kovtunova 2023)

$[\text{fuel_level} = 0.1 \cdot \text{fuel_capacity}] \sqsubseteq \text{LowFuel}$

Proofs with concrete domains

(Alrabbaa, Baader, B, Koopmann, and Kovtunova 2023)

$$[\text{fuel_level} = 0.1 \cdot \text{fuel_capacity}] \sqsubseteq \text{LowFuel} \quad \rightsquigarrow \quad A' \sqsubseteq \text{LowFuel}$$

Proofs with concrete domains

(Alrabbaa, Baader, B, Koopmann, and Kovtunova 2023)

$[\text{fuel_level} = 0.1 \cdot \text{fuel_capacity}] \sqsubseteq \text{LowFuel} \quad \rightsquigarrow \quad A' \sqsubseteq \text{LowFuel}$

$$\boxed{\begin{array}{c} C \sqsubseteq A_\alpha \quad C \sqsubseteq A_\beta \\ \hline C \sqsubseteq A_\alpha \sqcap A_\beta \quad A_\alpha \sqcap A_\beta \sqsubseteq A_\gamma \\ \hline C \sqsubseteq A_\gamma \end{array}}$$

Proofs with concrete domains

(Alrabbaa, Baader, B, Koopmann, and Kovtunova 2023)

$$[\text{fuel_level} = 0.1 \cdot \text{fuel_capacity}] \sqsubseteq \text{LowFuel} \quad \rightsquigarrow \quad A' \sqsubseteq \text{LowFuel}$$

$$\frac{\frac{C \sqsubseteq A_\alpha \quad C \sqsubseteq A_\beta}{C \sqsubseteq A_\alpha \sqcap A_\beta} \quad A_\alpha \sqcap A_\beta \sqsubseteq A_\gamma}{C \sqsubseteq A_\gamma}$$

$$\frac{2x + 3y = 5 \quad \frac{4y = 3}{-12y = -9} \quad [-3]}{4x - 6y = 1} \quad [2, 1]$$

Proofs with concrete domains

(Alrabbaa, Baader, B, Koopmann, and Kovtunova 2023)

$[\text{fuel_level} = 0.1 \cdot \text{fuel_capacity}] \sqsubseteq \text{LowFuel} \quad \rightsquigarrow \quad A' \sqsubseteq \text{LowFuel}$

$$\frac{\frac{C \sqsubseteq A_\alpha \quad C \sqsubseteq A_\beta}{C \sqsubseteq A_\alpha \sqcap A_\beta} \quad A_\alpha \sqcap A_\beta \sqsubseteq A_\gamma}{C \sqsubseteq A_\gamma}$$

$$\frac{2x + 3y = 5 \quad \frac{4y = 3}{-12y = -9} \text{ [-3]}}{4x - 6y = 1} \text{ [2, 1]}$$

$\Rightarrow$

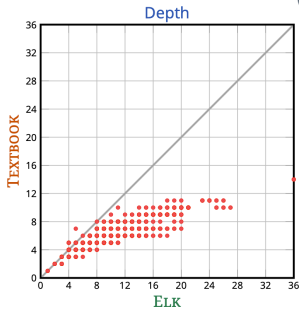
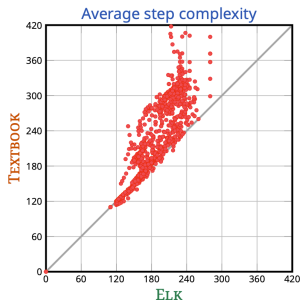
$$\frac{C \sqsubseteq [2x + 3y = 5] \quad \frac{C \sqsubseteq [4y = 3]}{C \sqsubseteq [-12y = -9]} \text{ [-3]}}{C \sqsubseteq [4x - 6y = 1]} \text{ [2, 1]}$$

Proofs for rules

- Proofs for fact entailment from Datalog and existential rules (NEMO)
(Ivliev, Gerlach, Meusel, Steinberg, and Krötzsch 2024)
- Even with (stratified) negation and aggregation
(Dachse, Gerlach, Hanisch, Ivliev, Krötzsch, Marx, and Méndez 2026)

Proofs for rules

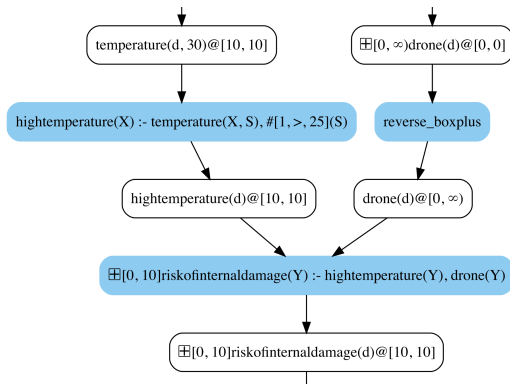
- Proofs for fact entailment from Datalog and existential rules (NEMO)
(Ivliev, Gerlach, Meusel, Steinberg, and Krötzsch 2024)
- Even with (stratified) negation and aggregation
(Dachselet, Gerlach, Hanisch, Ivliev, Krötzsch, Marx, and Méndez 2026)
- Can be used to implement consequence-based DL calculi
(Alrabbaa, B, Herrmann, and Krötzsch 2025)



Proofs for temporal rules

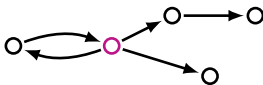
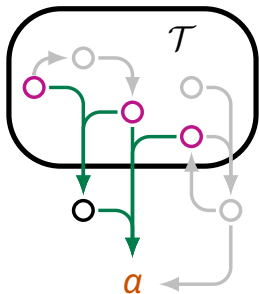
Proofs for DatalogMTL based on materialization in MeTeOR

(B, Demberg, Jobanputra, Kovtunova, and Nhu 2024)



Outline

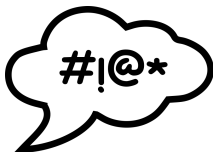
Proofs



Missing entailments



User studies



Verbalization

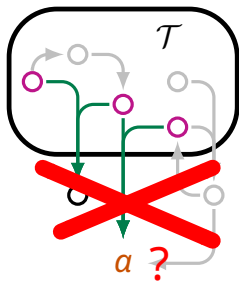
Explaining missing entailments

Missing entailment: $\mathcal{T} \not\models a$

Counterexample:

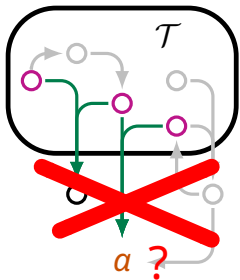
Model of $\mathcal{T}$ that does not satisfy a

(Bauer, Sattler, and Parsia 2009; Alrabbaa and Hieke 2022; Alrabbaa, B, Frieze, Koopmann, and Kotlov 2023)



Explaining missing entailments

Missing entailment: $\mathcal{T} \not\models a$



Counterexample:

Model of $\mathcal{T}$ that does not satisfy a

(Bauer, Sattler, and Parsia 2009; Alrabbaa and Hieke 2022; Alrabbaa, B, Friesse, Koopmann, and Kotlov 2023)

Abductive hypothesis:

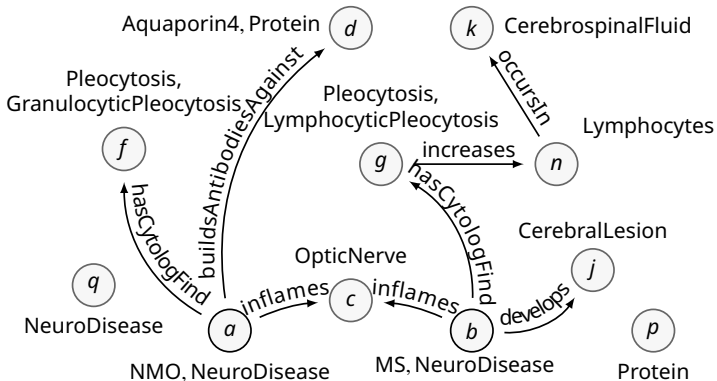
$\mathcal{H}$ such that $\mathcal{T} \cup \mathcal{H} \models a$ ($\mathcal{T} \cup \mathcal{H} \not\models \perp$ and $\mathcal{H} \neq \{a\}$)

(Calvanese, Ortiz, Simkus, and Stefanoni 2013; Wei-Kleiner, Dragisic, and Lambrix 2014; Koopmann, Del-Pinto, Tournet, and Schmidt 2020; Haifani, Koopmann, Tournet, and Weidenbach 2022)

Relevant counterexamples

Model of $\mathcal{T}$ that does not satisfy **NeuromyelitisOptica** $\sqsubseteq$ **MultipleSclerosis**

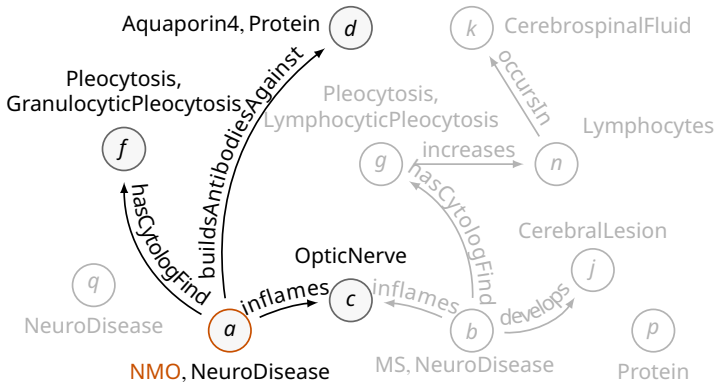
(Alrabbaa and Hieke 2022)



Relevant counterexamples

Part of a model of $\mathcal{T}$ that does not satisfy $\text{NeuromyelitisOptica} \sqsubseteq \text{MultipleSclerosis}$

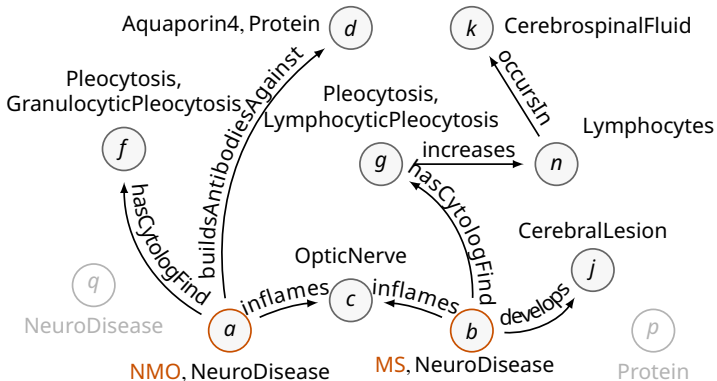
(Alrabbaa and Hieke 2022)



Relevant counterexamples

Part of a model of $\mathcal{T}$ that does not satisfy $\text{NeuromyelitisOptica} \sqsubseteq \text{MultipleSclerosis}$

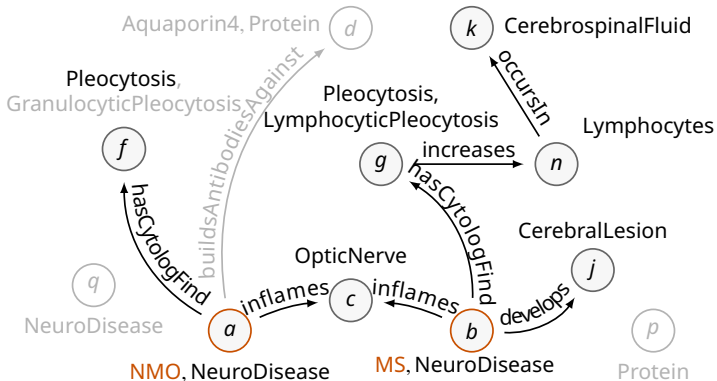
(Alrabbaa and Hieke 2022)



Relevant counterexamples

Part of a model of $\mathcal{T}$ that does not satisfy $\text{NeuromyelitisOptica} \sqsubseteq \text{MultipleSclerosis}$

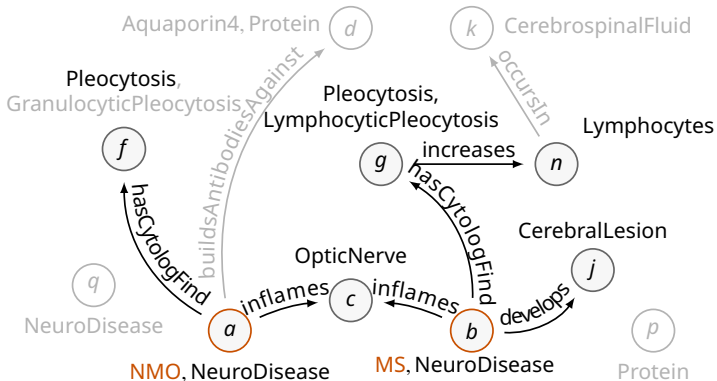
(Alrabbaa and Hieke 2022)



Relevant counterexamples

Part of a model of $\mathcal{T}$ that does not satisfy $\text{NeuromyelitisOptica} \sqsubseteq \text{MultipleSclerosis}$

(Alrabbaa and Hieke 2022)



Abduction

$\mathcal{H}$ such that $\mathcal{T} \cup \mathcal{H} \models a$ ($\mathcal{T} \cup \mathcal{H} \not\models \perp$ and $\mathcal{H} \not\equiv \{a\}$)

Complete signature-based abduction

(Koopmann, Del-Pinto, Tourret, and Schmidt 2020)

- Restrict $\mathcal{H}$ to a signature Σ
- High expressivity: $\mathcal{T}$ in $\mathcal{ALC}$, $\mathcal{H}$ with disjunction of axioms, inverse roles, nominals, fixpoint concepts
- Implementation based on LETHE

Abduction

$\mathcal{H}$ such that $\mathcal{T} \cup \mathcal{H} \models a$ ($\mathcal{T} \cup \mathcal{H} \not\models \perp$ and $\mathcal{H} \not\equiv \{a\}$)

Complete signature-based abduction

(Koopmann, Del-Pinto, Tournet, and Schmidt 2020)

- Restrict $\mathcal{H}$ to a signature Σ
- High expressivity: $\mathcal{T}$ in $\mathcal{ALC}$, $\mathcal{H}$ with disjunction of axioms, inverse roles, nominals, fixpoint concepts
- Implementation based on LETHE

Connection-minimal abduction

(Haifani, Koopmann, Tournet, and Weidenbach 2022)

- $\mathcal{H}$ is “connected” to a
- Low expressivity: $\mathcal{T}$ in $\mathcal{EL}$, $\mathcal{H}$ only conjunctions of concept names
- Implementation (CAPI) based on SPASS

(Non-)commercial break

youtu.be/DUDWp2YmhaY



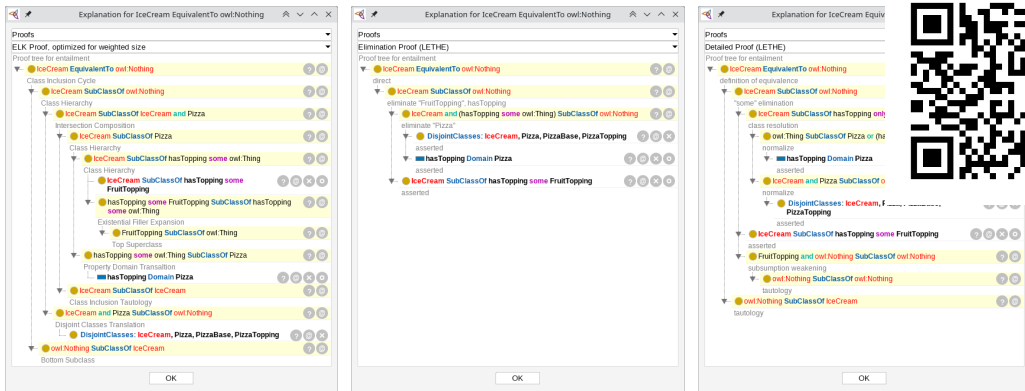
EEVEE

A tool for explaining the mysteries
of Pizza

and Description Logic reasoning!

EVEE – a collection of PROTÉGÉ plugins

github.com/de-tu-
dresden-inf-lat/evee



The image displays three screenshots of the EVEE plugin interface, each showing a different proof strategy for the goal "IceCream EquivalentTo owl:Nothing".

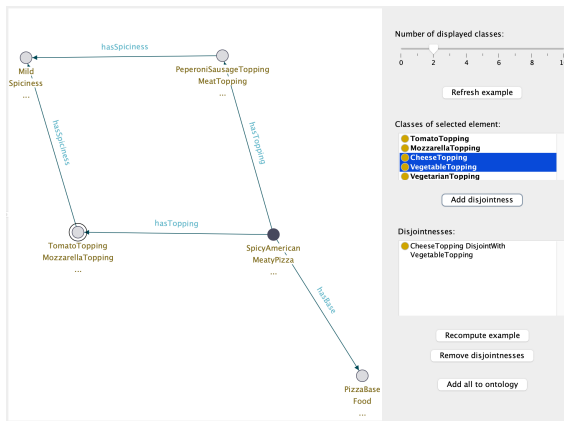
- Left Screenshot:** Shows the "ELK Proof, optimized for weighted size" strategy. The proof tree for entailment starts with "IceCream EquivalentTo owl:Nothing", followed by "IceCream SubClassOf owl:Nothing", "IceCream SubClassOf IceCream and Pizza", "Intersection Composition", "IceCream SubClassOf Pizza", "Class Hierarchy", "IceCream SubClassOf hasTopping some owl:Thing", "IceCream SubClassOf hasTopping some FruitTopping", "FruitTopping", "hasTopping some FruitTopping SubClassOf hasTopping some owl:Thing", "Existential Filler Expansion", "FruitTopping SubClassOf owl:Thing", "Top Superclass", "hasTopping some owl:Thing SubClassOf Pizza", "Property Domain Translation", "hasTopping Domain Pizza", "IceCream SubClassOf IceCream", "Class Inclusion Tautology", "IceCream and Pizza SubClassOf owl:Nothing", "Disjoint Classes Translation", "DisjointClasses: IceCream, Pizza, PizzaBase, PizzaTopping", and "owl:Nothing SubClassOf IceCream".
- Middle Screenshot:** Shows the "Elimination Proof (LETHE)" strategy. The proof tree for entailment starts with "IceCream EquivalentTo owl:Nothing", followed by "IceCream SubClassOf owl:Nothing", "eliminate 'FruitTopping', hasTopping", "IceCream and (hasTopping some owl:Thing) SubClassOf owl:Nothing", "eliminate 'Pizza'", "DisjointClasses: IceCream, Pizza, PizzaBase, PizzaTopping", "asserted", "hasTopping Domain Pizza", "asserted", and "IceCream SubClassOf hasTopping some FruitTopping".
- Right Screenshot:** Shows the "Detailed Proof (LETHE)" strategy. The proof tree for entailment starts with "IceCream EquivalentTo owl:Nothing", followed by "IceCream SubClassOf owl:Nothing", "definition of equivalence", "some 'elimination'", "IceCream SubClassOf hasTopping only", "class resolution", "owl:Thing SubClassOf Pizza or (hasTopping some owl:Thing)", "normalize", "hasTopping Domain Pizza", "asserted", "IceCream and Pizza SubClassOf owl:Nothing", "normalize", "DisjointClasses: IceCream, Pizza, PizzaBase, PizzaTopping", "asserted", "IceCream SubClassOf hasTopping some FruitTopping", "asserted", "FruitTopping and owl:Nothing SubClassOf owl:Nothing", "subsumption weakening", "owl:Nothing SubClassOf owl:Nothing", "tautology", and "owl:Nothing SubClassOf IceCream".



(Alrabbaa, B, Friese, Hirsch, Knieriemen, Koopmann, Kovtunova, Krüger, Popovic, and Siahaan 2024)

EVEE – a collection of PROTÉGÉ plugins

github.com/de-tu-dresden-inf-lat/evee



(Alrabbaa, B, Friese, Hirsch, Knieriemen, Koopmann, Kovtunova, Krüger, Popovic, and Siahaan 2024)

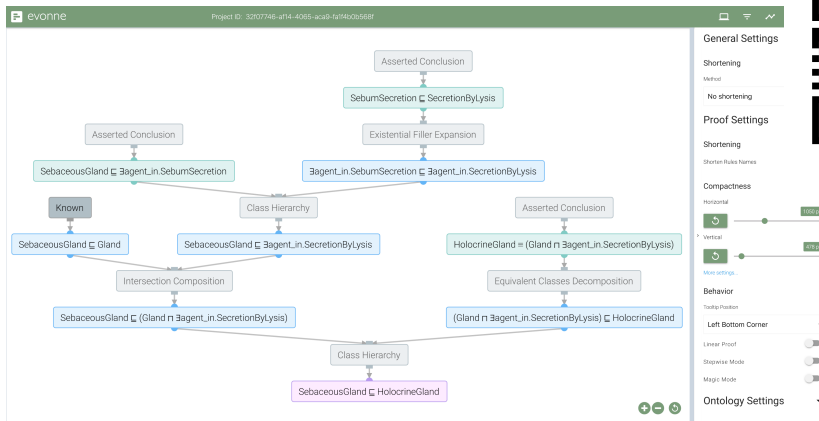
EVEE – a collection of PROTÉGÉ plugins

github.com/de-tu-
dresden-inf-lat/evee

The screenshot displays the EVEE interface, which is divided into several sections. On the left, there are two tabs: 'Missing entailment' and 'Vocabulary'. The 'Vocabulary' tab is active, showing a 'Permitted vocabulary' list with items like American, CheeseTopping, ChilliTopping, Country, DeepPanBase, DomainThing, Food, and HamTopping. Below this is a 'Forbidden vocabulary' section with a list containing SpicyPizza. The main area on the right is titled 'Hypotheses:' and contains two hypothesis boxes. Hypothesis 1 shows 'Mild SubClassOf Hot' with an 'Explain' button and a 'Forbid vocabulary' button. Hypothesis 2 shows a more complex statement: 'SpicyAmerican SubClassOf (hasTopping some ChilliTopping) or (hasTopping some SpicySalamiTopping)', with buttons for 'Explain', 'Forbid vocabulary', and 'Add to ontology'. A QR code is located to the right of the interface.

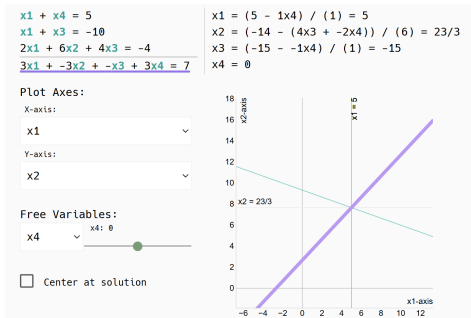
(Alrabbaa, B, Friese, Hirsch, Knieriemen, Koopmann, Kovtunova, Krüger, Popovic, and Siahaan 2024)

EVONNE – EEE's bigger sister



(Alrabbaa, Baader, Dachzelt, Kovtunova, and Méndez 2025)

EVONNE – EVEE's bigger sister



Premises:

$$x4 + 0 = x5$$

$$x3 - 2 = x2$$

$$x2 - 1 = x1$$

$$x4 + 5 = x3$$

Conclusion:

$$x5 + 2 = x1$$

Negated:

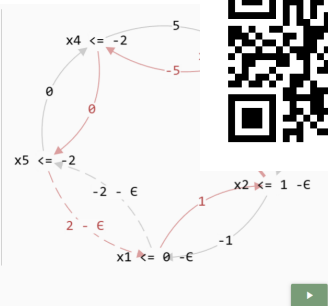
$$x5 - x1 \leq -2.0 - \epsilon$$

$$x1 - x5 \leq 2.0 - \epsilon$$

Cycle Value: $-\epsilon$

Propagate a value for "x3"!

3



(Alrabbaa, Baader, Dachzelt, Kovtunova, and Méndez 2025)

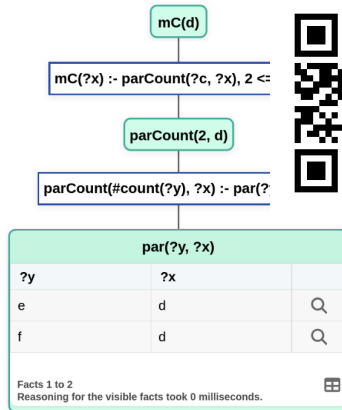
NEMO and NEV

github.com/
knowsys/nemo



Examples Help

```
1 par(a, b). par(b, c). sp(c, d).
2 par(e, d). par(f, d).
3
4 kin(?x, ?y) :- par(?x, ?y).
5 kin(?x, ?y) :- par(?x, ?z),
6   || sp(?z, ?z2), par(?y, ?z2).
7 kin(?x, ?y) :- kin(?y, ?x).
8 kin(?x, ?y) :- kin(?x, ?z), kin(?z, ?y).
9 parCount(#count(?y), ?x) :- par(?y, ?x).
10 mC(?x) :- parCount(?c, ?x), 2 <= ?c.
```



(Ivliev, Gerlach, Meusel, Steinberg, and Krötzsch 2024; Dachzelt, Gerlach, Hanisch, Ivliev, Krötzsch, Marx, and Méndez 2026)

Proofs for DatalogMTL and DLs with concrete domains

`github.com/de-
tu-dresden-inf-
lat/meteor-proofs`

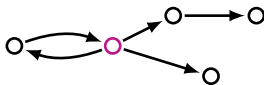
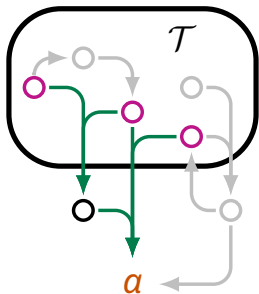


`doi.org/10.5281/
zenodo.8208780`



Outline

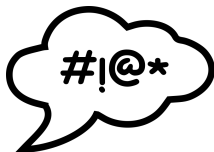
Proofs



Missing entailments

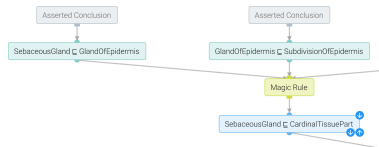


User studies



Verbalization

User studies

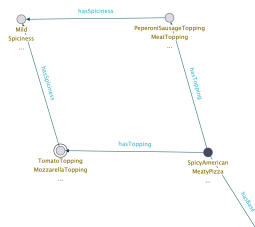
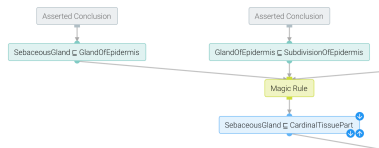


Proof format

(Alrabbaa, B, Hirsch, Knieriemen, Kovtunova, Rothermel, and Wiehr 2022)

- Prefer **shorter**, **tree-shaped**, **interactive** proofs
- Preferences are **subjective**
- IQ test (ICAR16) predicts performance on proofs

User studies



Proof format

(Alrabbaa, B, Hirsch, Knieriemen, Kovtunova, Rothermel, and Wiehr 2022)

- Prefer **shorter**, **tree-shaped**, **interactive** proofs
- Preferences are **subjective**
- IQ test (ICAR16) predicts performance on proofs

Usability of EVEE

(Alrabbaa, B, Friese, Hirsch, Knieriemen, Koopmann, Kovtunova, Krüger, Popovic, and Siahaan 2024)

- Preferred explanation methods depend on **users** and **use cases**

Designing user studies

domain-specific / abstract performance / preferences experts / students / laypeople

Since every cell culture is a material object and every material object contains an atom, every cell culture contains an atom. From the facts that every cell culture contains an atom and that every cell culture contains a cell and contains only cells, it follows that every cell culture contains something which is both an atom and a cell.

Every cell is a compound. Thus, any object which is an atom and a cell at the same time is also an atom and a compound. There is no object which is an atom and a compound at the same time. Therefore, there is no object which is both an atom and a cell.

Furthermore, since every cell culture contains something which is both an atom and a cell and there is no object which is both an atom and a cell, there is no cell culture.

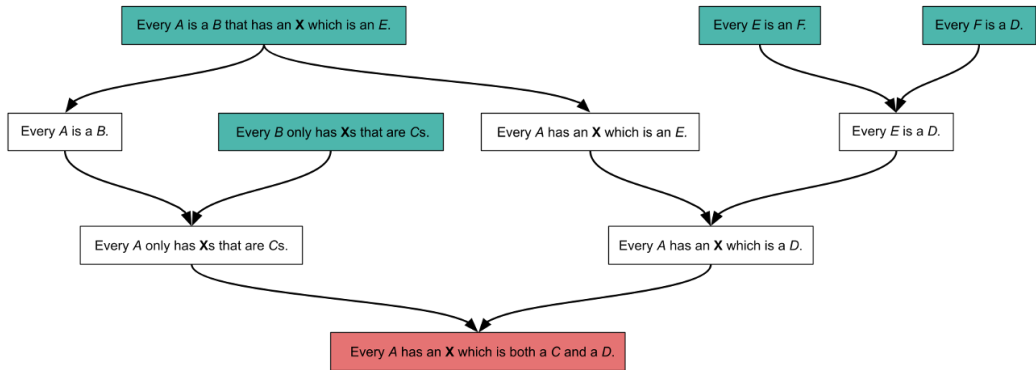
Designing user studies

domain-specific / abstract performance / preferences experts / students / laypeople

$$\begin{array}{c}
 \frac{\text{CClt} \sqsubseteq \exists \text{ct}. \text{C} \sqcap \forall \text{ct}. \text{C}}{\text{CClt} \sqsubseteq \exists \text{ct}. (\text{At} \sqcap \text{C})} \quad \frac{\frac{\text{CClt} \sqsubseteq \text{MaObj} \quad \text{MaObj} \sqsubseteq \exists \text{ct}. \text{At}}{\text{CClt} \sqsubseteq \exists \text{ct}. \text{At}} \quad \frac{\frac{\text{C} \sqsubseteq \text{Cmp}}{\text{At} \sqcap \text{C} \sqsubseteq \text{At} \sqcap \text{Cmp}} \quad \text{At} \sqcap \text{Cmp} \sqsubseteq \perp}{\text{At} \sqcap \text{C} \sqsubseteq \perp}}{\text{CClt} \sqsubseteq \perp}
 \end{array}$$

Designing user studies

domain-specific / abstract performance / preferences experts / students / laypeople



Designing user studies

domain-specific / abstract performance / preferences experts / students / laypeople

Since every woal is a luxi that is munted with a xylo,
every woal is a luxi.

Since every luxi is a kake,
every woal is a kake.

Moreover, every kake is only munted with kakes,
and thus every woal is only munted with kakes.

Every xylo is a pire
and every pire is an atis,
which means that every xylo is an atis.

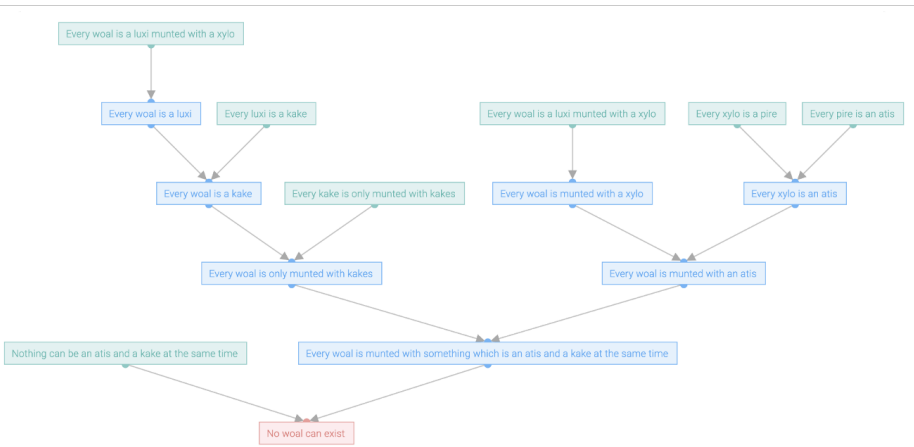
Since every woal is a luxi that is munted with a xylo,
every woal is munted with a xylo.

Since every xylo is an atis,
we know that every woal is munted with an atis.

We have inferred that every woal is only munted with kakes
and that every woal is munted with an atis.
Therefore, every woal is munted with something which is both an atis and a kake.
However, nothing can be an atis and a kake at the same time
which lets us conclude that woals cannot exist.

Designing user studies

domain-specific / abstract performance / preferences experts / students / laypeople



Designing user studies

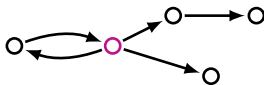
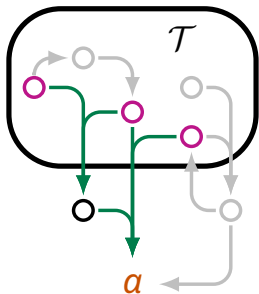
domain-specific / abstract performance / preferences experts / students / laypeople

Hear more from Alisa in 45 minutes:

In the Heart of the Beholder:
User-Tailored Explanations for Description Logics

Outline

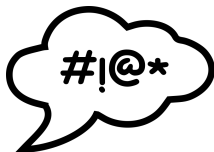
Proofs



Missing entailments



User studies



Verbalization

Drone handover scenario



Handover from autonomous drone to human operator
requires **timely, short warning** of a critical situation

(Wiehr, Hirsch, Schmitz, Knieriemen, Krüger, Kovtunova, B, Chang, Demberg, Steinmetz, and Hoffmann 2021)

Drone handover scenario



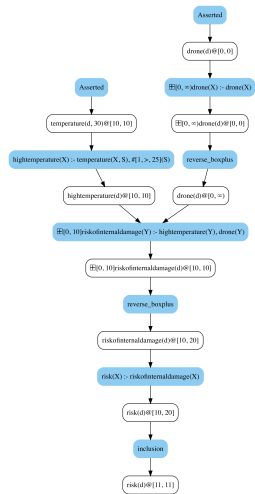
Handover from autonomous drone to human operator
requires **timely, short warning** of a critical situation

(Wiehr, Hirsch, Schmitz, Knieriemen, Krüger, Kovtunova, B, Chang, Demberg, Steinmetz, and Hoffmann 2021)

Verbalization of entailments and **explanations**:

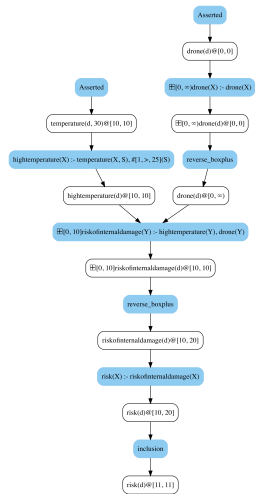
- controlled natural language (Power 2012)
- template-based verbalization (Schiller, Schiller, and Glimm 2017)
- transformer models (Chang, Kovtunova, B, Demberg, Chapman, and Yeh 2022)
- large language models (Jobanputra, Kovtunova, Balthes, Pogulskiy, Wang, B, and Demberg 2025)

Proof verbalization with LLMs



(Jobanputra, Kovtunova, Balthes, Pogulskiy, Wang, B, and Demberg 2025)
Proofs provide more information than **justifications**

Proof verbalization with LLMs

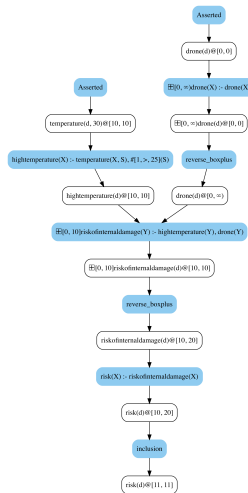


(Jobanputra, Kovtunova, Balthes, Pogulskiy, Wang, B, and Demberg 2025)

Proofs provide more information than **justifications**

Challenge: Find the important parts of the proof

Proof verbalization with LLMs



(Jobanputra, Kovtunova, Balthes, Pogulskiy, Wang, B, and Demberg 2025)

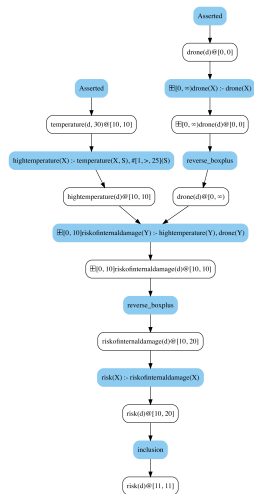
Proofs provide more information than **justifications**

Challenge: Find the important parts of the proof

“At time 10, there is a risk of internal damage to your drone (d) because it has a high temperature of 30 degrees Celsius. **This is likely because the drone is flying.**”

Inconsistent outputs, **hallucinations**

Proof verbalization with LLMs



Hear more from Alisa in 60 minutes:

ProofTeller: Exposing Recency Bias in LLM Reasoning and Its Side Effects on Communication (Extended Abstract)

Summary

- **Justifications** are often enough, but sometimes **proofs** are needed
- **Missing entailments** also require explanations
- Computing explanations can take a long time
- Formal explanations are mainly useful for experts

Summary

- **Justifications** are often enough, but sometimes **proofs** are needed
- **Missing entailments** also require explanations
- Computing explanations can take a long time
- Formal explanations are mainly useful for experts

Future work:

- Protégé proofs with concrete domains
- Explain description logics to laypeople



An aerial photograph of a historic city, likely Lisbon, Portugal. The image shows a dense cluster of buildings with red-tiled roofs. In the background, a large, ornate cathedral with two prominent towers is visible. The sky is overcast and grey. The text "Thank you!" is overlaid in white on the left side of the image.

Thank you!

References I

- Alrabbaa, Christian, Stefan B, Tom Frieze, Anke Hirsch, Nina Knieriemen, Patrick Koopmann, Alisa Kovtunova, Antonio Krüger, Alexej Popovic, and Ida S. R. Siahaan (2024). "Explaining Reasoning Results for OWL Ontologies with Eevee". In: *Proceedings of the 21st International Conference on Principles of Knowledge Representation and Reasoning (KR)*. DOI: 10.24963/KR.2024/67.
- Alrabbaa, Christian, Stefan B, Tom Frieze, Patrick Koopmann, and Mikhail Kotlov (2023). "Why Not? Explaining Missing Entailments with Eevee". In: *DL*. URL: <https://ceur-ws.org/Vol-3515/paper-1.pdf>.
- Alrabbaa, Christian, Stefan B, Philipp Herrmann, and Markus Krötzsch (2025). "The Shape of EL Proofs: A Tale of Three Calculi". In: *Proceedings of the 38th International Workshop on Description Logics (DL)*. URL: <https://ceur-ws.org/Vol-4091/paper45.pdf>.
- Alrabbaa, Christian, Stefan B, Anke Hirsch, Nina Knieriemen, Alisa Kovtunova, Anna Milena Rothermel, and Frederik Wiehr (2022). "In the Head of the Beholder: Comparing Different Proof Representations". In: *RuleML+RR 2022*. DOI: 10.1007/978-3-031-21541-4_14.
- Alrabbaa, Christian, Stefan B, Patrick Koopmann, and Alisa Kovtunova (2022). "Finding Good Proofs for Answers to Conjunctive Queries Mediated by Lightweight Ontologies". In: *DL*. URL: <https://ceur-ws.org/Vol-3263/paper-3.pdf>.
- Alrabbaa, Christian, Franz Baader, Stefan B, Patrick Koopmann, and Alisa Kovtunova (2020). "Finding Small Proofs for Description Logic Entailments: Theory and Practice". In: *LPAR*. DOI: 10.29007/nhpp.
- (2021). "Finding Good Proofs for Description Logic Entailments using Recursive Quality Measures". In: *CADE*. DOI: 10.1007/978-3-030-79876-5_17.
 - (2023). "Combining Proofs for Description Logic and Concrete Domain Reasoning". In: *RuleML+RR*. DOI: 10.1007/978-3-031-45072-3_4.

References II

- Alrabbaa, Christian, Franz Baader, Raimund Dachzelt, Alisa Kovtunova, and Julián Méndez (2025). "The Concrete Evonne: Visualization Meets Concrete Domain Reasoning". In: *Frontiers of Combining Systems - 15th International Symposium (FroCoS)*. DOI: 10.1007/978-3-032-04167-8_1.
- Alrabbaa, Christian and Willi Hieke (2022). "Explaining Non-Entailment by Model Transformation for the Description Logic EL". In: *IJCKG*. DOI: 10.1145/3579051.3579060.
- B, Stefan (2021). *Concise Justifications Versus Detailed Proofs for Description Logic Entailments*. LTCS-Report 21-05. doi: 10.25368/2023.225.
- B, Stefan, Vera Demberg, Mayank Jobanputra, Alisa Kovtunova, and Duy Nhu (2024). "Explaining Critical Situations Over Sensor Data Streams Using Proofs and Natural Language". In: *DL*. URL: <https://ceur-ws.org/Vol-3739/paper-2.pdf>.
- Bauer, Johannes, Ulrike Sattler, and Bijan Parsia (2009). "Explaining by Example: Model Exploration for Ontology Comprehension". In: *DL*. URL: https://ceur-ws.org/Vol-477/paper_37.pdf.
- Borgida, Alexander, Diego Calvanese, and Mariano Rodríguez-Muro (2008). "Explanation in the DL-Lite Family of Description Logics". In: *OTM*. DOI: 10.1007/978-3-540-88873-4_35.
- Calvanese, Diego, Magdalena Ortiz, Mantas Simkus, and Giorgio Stefanoni (2013). "Reasoning about Explanations for Negative Query Answers in DL-Lite". In: *J. Artif. Intell. Res.* 48. DOI: 10.1613/jair.3870.
- Chang, Ernie, Alisa Kovtunova, Stefan B, Vera Demberg, Kathryn Chapman, and Hui-Syuan Yeh (2022). "Logic-Guided Message Generation from Raw Real-Time Sensor Data". In: *LREC*. URL: <https://aclanthology.org/2022.lrec-1.745>.

References III

- Dachsel, Raimund, Lukas Gerlach, Philipp Hanisch, Alex Ivliev, Markus Krötzsch, Maximilian Marx, and Julián Méndez (2026). "Declarative Debugging for Datalog with Aggregation". In: *Proceedings of the Workshops of the EDBT/ICDT 2026 Joint Conference*. URL: <https://ceur-ws.org/Vol-4192/TGD-paper4.pdf>.
- Eiter, Thomas, Magdalena Ortiz, Mantas Simkus, Trung-Kien Tran, and Guohui Xiao (2012). "Query Rewriting for Horn-SHIQ Plus Rules". In: *AAAI*. DOI: 10.1609/aaai.v26i1.8219.
- Haifani, Fajar, Patrick Koopmann, Sophie Touret, and Christoph Weidenbach (2022). "Connection-Minimal Abduction in *EL* via Translation to FOL". In: *IJCAR*. DOI: 10.1007/978-3-031-10769-6_12.
- Horridge, Matthew, Samantha Bail, Bijan Parsia, and Uli Sattler (2013). "Toward cognitive support for OWL justifications". In: *Knowl.-Based Syst.* 53. DOI: 10.1016/j.knosys.2013.08.021.
- Horridge, Matthew, Bijan Parsia, and Ulrike Sattler (2010). "Justification Oriented Proofs in OWL". In: *ISWC*. DOI: 10.1007/978-3-642-17746-0_23.
- Ivliev, Alex, Lukas Gerlach, Simon Meusel, Jakob Steinberg, and Markus Krötzsch (2024). "Nemo: Your Friendly and Versatile Rule Reasoning Toolkit". In: *Proceedings of the 21st International Conference on Principles of Knowledge Representation and Reasoning (KR)*. DOI: 10.24963/KR.2024/70.
- Jobanputra, Mayank, Alisa Kovtunova, Brisca Balthes, Fedor Grigoryevich Pogulskiy, Yifan Wang, Stefan B, and Vera Demberg (2025). "ProofTeller: Exposing recency bias in LLM reasoning and its side effects on communication". In: *Proceedings of the 14th International Joint Conference on Natural Language Processing and the 4th Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics, (IJCNLP-AAACL)*. DOI: 10.18653/V1/2025.IJCNLP-LONG.80.

References IV

- Kazakov, Yevgeny, Markus Krötzsch, and Frantisek Simancik (2014). "The Incredible ELK - From Polynomial Procedures to Efficient Reasoning with $\mathcal{EL}$ Ontologies". In: *J. Autom. Reason.* 53.1. DOI: 10.1007/s10817-013-9296-3.
- Koopmann, Patrick, Warren Del-Pinto, Sophie Turret, and Renate A. Schmidt (2020). "Signature-Based Abduction for Expressive Description Logics". In: *KR*. DOI: 10.24963/kr.2020/59.
- Peñaloza, Rafael and Francesco Ricca (2022). "Pinpointing Axioms in Ontologies via ASP". In: *Logic Programming and Nonmonotonic Reasoning - 16th International Conference, LPNMR*. DOI: 10.1007/978-3-031-15707-3_24.
- Peuter, Dennis (2024). "Applications for Symbol Elimination in Combination with Hierarchical Reasoning". PhD thesis. URL: <https://uko.opus.hbz-nrw.de/frontdoor/index/index/docId/2507>.
- Power, Richard (2012). "OWL Simplified English: A Finite-State Language for Ontology Editing". In: *CNL*. DOI: 10.1007/978-3-642-32612-7_4.
- Schiller, Marvin R. G., Florian Schiller, and Birte Glimm (2017). "Testing the Adequacy of Automated Explanations of EL Subsumptions". In: *DL*. URL: <https://ceur-ws.org/Vol-1879/paper43.pdf>.
- Schlobach, Stefan (2004). "Explaining Subsumption by Optimal Interpolation". In: *JELIA*. DOI: 10.1007/978-3-540-30227-8_35.
- Tena Cucala, David, Bernardo Cuenca Grau, and Ian Horrocks (2021). "Pay-as-you-go consequence-based reasoning for the description logic SROIQ". In: *Artif. Intell.* 298. DOI: 10.1016/j.artint.2021.103518.
- Wei-Kleiner, Fang, Zlatan Dragisic, and Patrick Lambrix (2014). "Abduction Framework for Repairing Incomplete EL Ontologies: Complexity Results and Algorithms". In: *AAAI*. DOI: 10.1609/aaai.v28i1.8858.

References V

Wiehr, Frederik, Anke Hirsch, Lukas Schmitz, Nina Knieriemen, Antonio Krüger, Alisa Kovtunova, Stefan B, Ernie Chang, Vera Demberg, Marcel Steinmetz, and Jörg Hoffmann (2021). "Why Do I Have to Take Over Control? Evaluating Safe Handovers with Advance Notice and Explanations in HAD". In: *ICMI*. DOI: [10.1145/3462244.3479884](https://doi.org/10.1145/3462244.3479884).

Images:

Lisbon tram by [Aayush Gupta on Unsplash](#), [Unsplash license](#)

Drones by [JanBaby on Pixabay](#), [CC0](#)

Users by [danilo on Openclipart](#), [CC0](#)

Speech bubble by [gramzon on Openclipart](#), [CC0](#)

Control tower by [Annabel_P on Pixabay](#), [CC0](#)

Lisbon roofs by [Liam McKay on Unsplash](#), [Unsplash license](#)

All other images/illustrations © CPEC