OTTO VON GUERICKE UNIVERSITÄT MAGDEBURG
FACULTY OF COMPUTER SCIENCE

Logic-Independent Premise Selection for Automated Theorem Proving

Eugen Kuksa

AITP 2017
Obergurgl

Outline

1. Introduction
2. Theoretical Foundation: Entailment Relations
3. Case Study
4. There's More: Signatures and Logic Translations.
5. Conclusion

Introduction

Automated Theorem Proving

- A theorem proving problem consists of Axioms and a conjecture.
- An automated theorem prover (ATP) runs an algorithm to find a proof.
- A typical ATP is efficient on small problems.
- Large problems lead to combinatorial explosion.
  - ATP reach their time or memory limit.
  - Return with no result.

Example

- Suggested Upper Merged Ontology (SUMO)
- Formalised in FOF and THF in the TPTP library
- Problems with (tens of) thousands of axioms
- Pick *CSR119^3* (THF): ≈5000 Axioms
- Higher-Order Prover Leo-II runs into a timeout (60 seconds)

Solution: Premise Selection

- Reduce the set of axioms for the proving task
- Proving time decreases or proving even becomes possible at all
- The SUMO example *CSR119^3* passes in less than a second
- The axiom set was reduced to 390 out of over 5000 axioms by SInE

Logic Dependence

Problem:

- There are many premise selection algorithms
- Implemented only for FOF or some higher order logics
- . . . even though some are *described* logic-independently

Solution:

- Lift the algorithms to logic-independence
- Run them in an abstract notion of ‘logic’
- Transfer results to the concrete logic

Tool Support

Problem:

Many provers operate on one logic/syntax only

- FaCT, Pellet: Description logic with OWL
- Darwin, E-Prover, Geo-III, SPASS, Vampire: First-order logic with TPTP/FOF
- Leo-II, Satallax, Isabelle: Higher-order logic with TPTP/THF
- Isabelle/HOL's own logic
- ...

Solution:

- Lift the algorithms to logic-independence
- Run them in an abstract notion of 'logic'
- Transfer results to the concrete logic

Theoretical Foundation: Entailment Relations

Entailment relation

An entailment relation with symbols $(\mathrm{Sen}, \mathrm{Sym} \vdash, \mathrm{symbols})$ consists of

- A set of sentences Sen
- A set of symbols Sym
- A relation $\vdash \subseteq \mathcal{P}(\mathrm{Sen}) \times \mathrm{Sen}$ which is
  - reflexive (Axioms are theorems)
  - transitive (We may use lemmas)
  - monotonic (We may use premise selection)
- A function $\mathrm{symbols} : \mathrm{Sen} \to \mathcal{P}(\mathrm{Sym})$ giving the symbols that occur in a sentence

Case Study

Implementation: Ontohub

- Web application: `https://ontohub.org`
- Version controlled repository for ontologies/specifications/theories
- Version control (git)
- Integrated editor for small files
- Analyses theories
- Has interfaces with ATPs
- Back-end: Hets

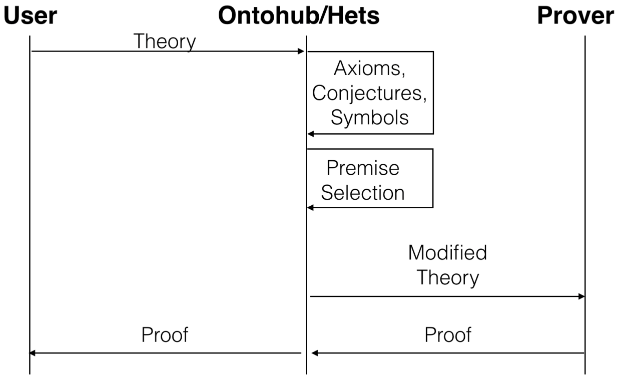
Implementation: Ontohub cont'd

- Supports different logics
  - Propositional Logic
  - OWL
  - FOL / TPTP-FOF
  - FOL + Induction
  - CASL
  - Modal Logic
  - Common Logic
  - HOL / TPTP-THF
  - Isabelle/HOL
  - ...
- Brings tool support
  - FaCT, Pellet
  - CVC4, Darwin, E-Prover, Geo-III, SPASS, Vampire
  - Leo-II, Satallax, Isabelle
  - ...

Premise Selection: The Algorithm 'SInE'

- Developed by Kryštof Hoder
- Fully automatic with a few user-defined parameters
- Operates on syntax
- Selects recursively the axioms that share a symbol with the conjecture or an already selected axiom
- The shared symbol that allows to select an axiom must hold more conditions
- Selection stops after n recursion steps
- We implemented SInE in Ontohub

Data Flow


User
Ontohub/Hets
Prover
Theory
Axioms,
Conjectures,
Symbols
Premise
Selection
Modified
Theory
Proof
Proof

BETA Ontohub | Repositories | Ontologies | Logics | Mappings | More | Help | Eugen Kuksa

Do you want to support the development of Ontohub by testing? Visit test.ontohub.org for more information

Repositories

107 repositories currently available

Create Repository

There is also a complete list of repository clone-urls

Featured Repositories

- BFO This repository contains an axiomatization of Basic Formal Ontology (BFO) 2.0. This axiomatization is the result of a collaboration between Barry Smith and Fabian Neuhaus. NOTE: At this stage the axiomatization is incomplete and under change.
- Bioportal Mirror of http://bioportal.bioontology.org/ with all ontologies below 5 megabytes.
- FOIS Ontology Competition The aim of the FOIS 2014 ontology competition is (1) to encourage ontology authors to make their ontologies publicly available and (2) to subject them to evaluation according to a set of pre-determined criteria. See http://fois2014.inf.ufes.br/p/call-for-ontology-competition.html

Common Repositories

- AdHoc Domain-specific ontology to detect messages on key personnel turnover in the ad hoc publicity of public companies.
- AlgebraBlending
- Amalgams Music Initial theories for testing chord blending in CASL
- Amalgams OWL examples owl ontologies for amalgams
- AppliedOntologyOntohubPaper These are examples that are used in the paper "Ontohub -- A semantic repository for heterogeneous ontologies"
- bfo-test
- Bioportal_mappings Mappings from Bioportal
- BOC2016 Ontology Building Competition 2016 at Technical University of Cluj-Napoca, Romania Intelligent Systems Group Assoc. Prof. Adrian Groza http://cs-gw.utcluj.ro/~adrian

Eugen's Repositories

- bfo-test
- Eugens Test 1
- zzz-push-test
- zzz-issue-982-test
- eugens-test-5
- Hets-Lib
- eugens-test-3
- eugens-test-4
- eugens-test-6
- AppliedOntologyOntohubPaper
- OntohubAOpaperExamples
- eugens-test2-tptp
- eugens-test3-tptp
- eugens-test-tptp
- eugens-test-2

© 2012-2015 Universität Magdeburg and Universität Bremen About Sources Open Ontology Repository (OOR) Initiative

BETA Ontohub | Repositories | Ontologies | Logics | Mappings | More | Help | Eugen Kuksa

Do you want to support the development of Ontohub by testing? Visit test.ontohub.org for more information

grundlagen

Overview | Ontologies | File browser | History | Settings

File browser

In this page, you have direct access to the file system where the repository is stored. Below is the list of files in the current directory, independent of whether they contain ontologies or not. In order to find a particular file, you can navigate through the file system.

grundlagen / bushy-by — Create Subdirectory | Upload file | History of this directory

l_all_th1	l_all_th1.p
l_and_th1	l_and_th1.p
l_comiff	l_comiff.p
l_e_changef_p	l_e_changef_p.p
l_e_inj_t1	l_e_inj_t1.p
l_e_isf	l_e_isf.p
l_e_isst1	l_e_isst1.p
l_e_ite_t14	l_e_ite_t14.p
l_e_ite_t4	l_e_ite_t4.p
l_e_notis_th5	l_e_notis_th5.p
l_e_refis	l_e_refis.p
l_e_st_emptyi	l_e_st_emptyi.p
l_e_st_eq_10_t7	l_e_st_eq_10_t7.p
l_e_st_eq_11_t4	l_e_st_eq_11_t4.p
l_e_st_eq_1_th3	l_e_st_eq_1_th3.p

76fcb55 © 2012-2016 Universität Magdeburg and Universität Bremen | About | Sources | Open Ontology Repository (OOR) Initiative

BETA Ontohub | Repositories | Ontologies | Logics | Mappings | More | Help | Eugen Kuksa

Overview | Ontologies | File browser | History | Settings

File browser

In this page, you have direct access to the file system where the repository is stored. Below are the contents of the file presented as text as well as its path relative to the repository home directory. If you want to see the contents of this file as ontologies, click on the ontology mappings above the text box.

eugens-test-tptp / AGT001+1.p

Edit file | Update file | Download file | History of this file | Delete File

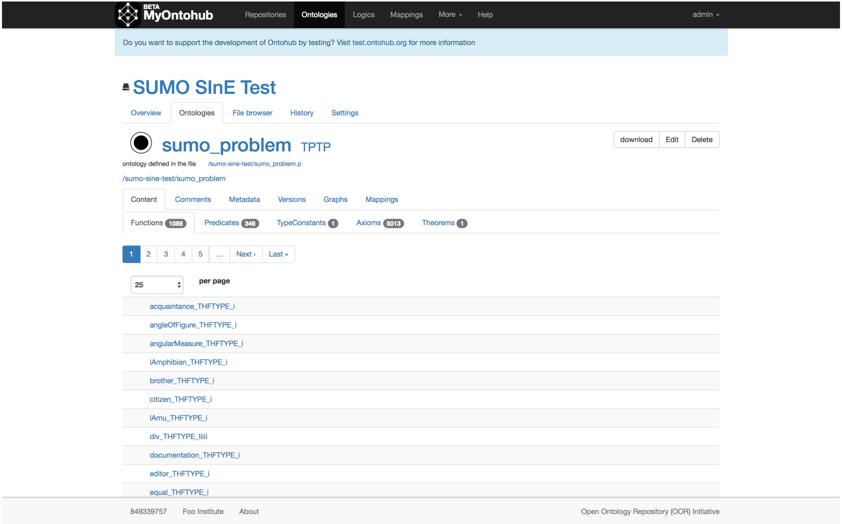
AGT001+1.p

```
%------------------------------------------------------------------------------
% File     : AGT001+1 : TPTP v6.4.0. Bugfixed v3.1.0.
% Domain   : Agents
% Problem  : Problem for the CPlanT system
% Version  : [Bar03] axioms : Especial.
% English  :

% Refs     : [Bar03] Barta, J. (2003), Email to G. Sutcliffe
%          : [BT+03] Barta et al. (2003), Meta-Reasoning in CPlanT Multi-Ag
%          : [TBP03] Tozicka et al. (2003), Meta-reasoning for Agents' Priv
% Source   : [Bar03]
% Names    :

% Status   : Theorem
% Rating   : 0.03 v6.4.0, 0.08 v6.3.0, 0.12 v6.1.0, 0.10 v6.0.0, 0.09 v5.5.0, 0.07 v5.4.0, 0.0
% Syntax   : Number of formulae    :  556 ( 524 unit)
%            Number of atoms       :  656 (   2 equality)
%            Maximal formula depth :    8 (   1 average)
%            Number of connectives :  143 (  43 ~  ;   1  |;  67  &)
%                                         (  16 <=>;  16 =>;   0 <=)
%                                         (   0 <~>;   0 ~|;   0 ~&)
%            Number of predicates  :   17 (   0 propositional; 1-4 arity)
%            Number of functors    :  290 ( 286 constant; 0-2 arity)
%            Number of variables   :   70 (   0 singleton;  70 !;   0 ?)
%            Maximal term depth    :    5 (   1 average)
% SPC      : FOF_THM_RFO_SEQ

% Comments :
% Bugfixes : v3.0.0 - Bugfixes in NUM005+1.ax
%          : v3.1.0 - Changes to NUM005 axioms
%------------------------------------------------------------------------------
%----Include axioms of CPlanT
include('Axioms/AGT001+0.ax').
%----Include events of CPlanT
include('Axioms/AGT001+1.ax').
%----Include axioms for RDN and RDN less
include('Axioms/NUM005+0.ax').
include('Axioms/NUM005+1.ax').
%------------------------------------------------------------------------------
fof(query_1,conjecture,
    ( accept_team(countryamedicalorganization,countryacivilorganization,towna,n6) )).

%------------------------------------------------------------------------------
```

Defining 1 ontology

- AGT001+1

© 2012-2015 Universität Magdeburg and Universität Bremen About | Sources | Open Ontology Repository (OOR) Initiative


BETA MyOntohub
Repositories
Ontologies
Logics
Mappings
More
Help
admin
Do you want to support the development of Ontohub by testing? Visit test.ontohub.org for more information
SUMO SInE Test
Overview
Ontologies
File browser
History
Settings
sumo_problem TPTP
download
Edit
Delete
ontology defined in the file /sumo-sine-test/sumo_problem.p
/sumo-sine-test/sumo_problem
Content
Comments
Metadata
Versions
Graphs
Mappings
Functions 1088
Predicates 346
TypeConstants 1
Axioms 5013
Theorems 1
1 2 3 4 5 … Next › Last »
25
per page
acquaintance_THFTYPE_i
angleOfFigure_THFTYPE_i
angularMeasure_THFTYPE_i
lAmphibian_THFTYPE_i
brother_THFTYPE_i
citizen_THFTYPE_i
lAmu_THFTYPE_i
div_THFTYPE_liiil
documentation_THFTYPE_i
editor_THFTYPE_i
equal_THFTYPE_i
849339757
Foo Institute
About
Open Ontology Repository (OOR) Initiative

BETA Ontohub · Repositories · Ontologies · Logics · Mappings · More · Help · Eugen Kuksa

eugens-test-tptp

Overview · Ontologies · File browser · History · Settings

AGT001+1 TPTP

download · Edit · Delete

ontology defined in the file /eugens-test-tptp/AGT001+1.p

/eugens-test-tptp/AGT001+1

Content · Comments · Metadata · Versions · Graphs · Mappings

Constants 286 · Functors 4 · Predicates 16 · Axioms · Theorems 1

query_1 THM

Definition of query_1

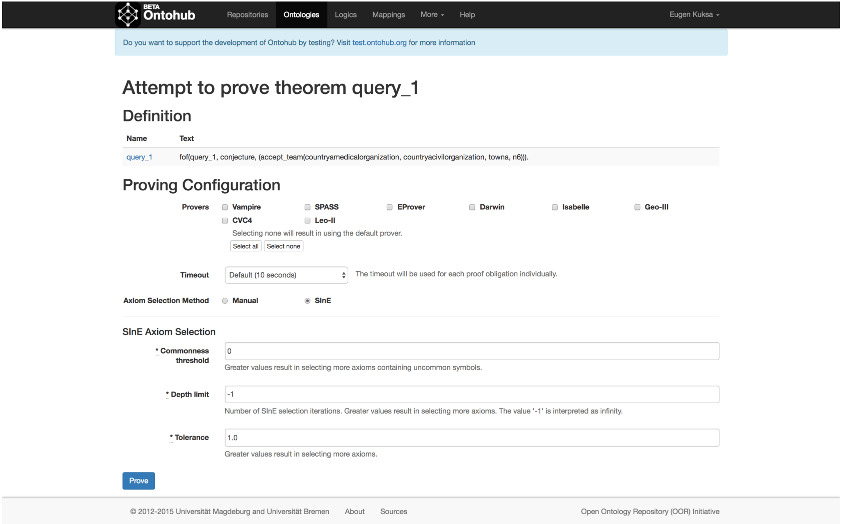
fof(query_1, conjecture, (accept_team(countryamedicalorganization, countryacivilorganization, towna, n6))).

Proof Attempts

Prove this theorem

Number	Date	Time taken	Prover	Proof Status	Evaluation State
9	a month ago	66 seconds	Isabelle	THM	done (a month ago)
8	a month ago	0 seconds	Leo-II	THM	done (a month ago)
7	a month ago	0 seconds	CVC4	THM	done (a month ago)
6	a month ago	1 seconds	Geo-III	UNK	done (a month ago)
5	a month ago	0 seconds	Darwin	THM	done (a month ago)
4	a month ago	0 seconds	EProver	THM	done (a month ago)
3	a month ago	0 seconds	Satallax	ERR	done (a month ago)
2	a month ago	0 seconds	SPASS	THM	done (a month ago)
1	a month ago	0 seconds	Vampire	THM	done (a month ago)

© 2012-2015 Universität Magdeburg and Universität Bremen · About · Sources · Open Ontology Repository (OOR) Initiative

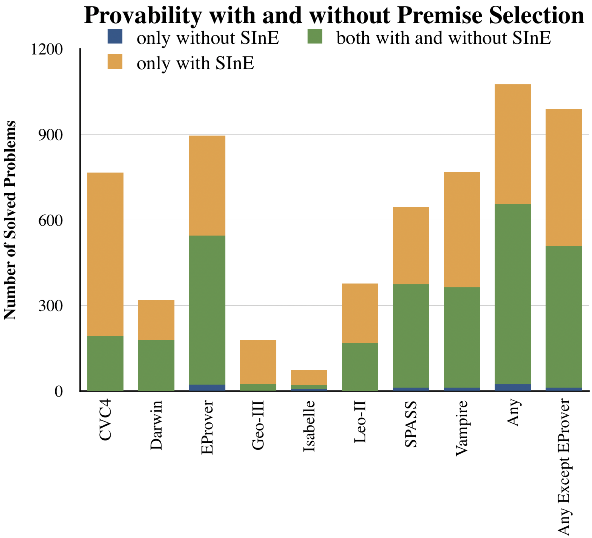

BETA Ontohub
Repositories
Ontologies
Logics
Mappings
More
Help
Eugen Kuksa
Do you want to support the development of Ontohub by testing? Visit test.ontohub.org for more information
Attempt to prove theorem query_1
Definition
Name
Text
query_1
fof(query_1, conjecture, (accept_team(countryamedicalorganization, countryacivilorganization, towna, n6))).
Proving Configuration
Provers
Vampire
SPASS
EProver
Darwin
Isabelle
Geo-III
CVC4
Leo-II
Selecting none will result in using the default prover.
Select all
Select none
Timeout
Default (10 seconds)
The timeout will be used for each proof obligation individually.
Axiom Selection Method
Manual
SInE
SInE Axiom Selection
* Commonness threshold
0
Greater values result in selecting more axioms containing uncommon symbols.
* Depth limit
-1
Number of SInE selection iterations. Greater values result in selecting more axioms. The value '-1' is interpreted as infinity.
* Tolerance
1.0
Greater values result in selecting more axioms.
Prove
© 2012-2015 Universität Magdeburg and Universität Bremen
About
Sources
Open Ontology Repository (OOR) Initiative

Experiments: Setup

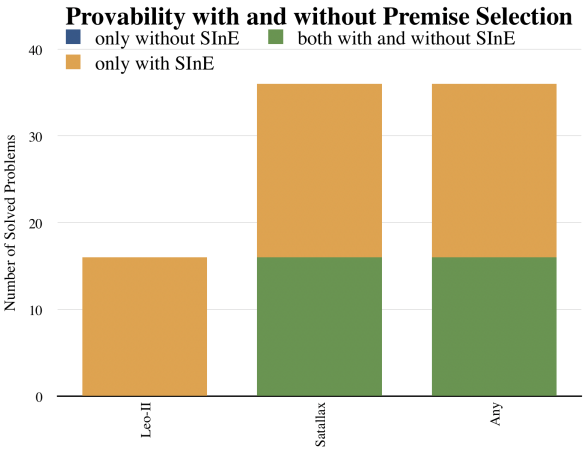
We applied our implementation of SInE to

- All 2078 problems of the MPTP2078 (FOF)
- A subset (501 problems) of a formalisation into THF0 of the Automath formalization of Landau's 'Grundlagen der Analysis'

Results: FOF


Provability with and without Premise Selection
only without SInE
both with and without SInE
only with SInE
Number of Solved Problems
1200
900
600
300
0
CVC4
Darwin
EProver
Geo-III
Isabelle
Leo-II
SPASS
Vampire
Any
Any Except EProver

Results: THF0



There's More: Signatures and Logic Translations.

Special handling of THF

Problem:

- THF (among other logics) is typed
- Symbols must be declared with a formula before their first use
- Such 'signature-defining delarations' must not be removed

Solution:

- Preserve the needed 'signature-defining declarations' after the premise selection

Theoretical Foundation: Entailment Systems

An entailment system with symbols $(\mathrm{Sign}, \mathrm{Sen}, \mathrm{Sym}, \vdash, \mathrm{symbols})$ consists of

- a category Sign of signatures and signature morphisms
- a functor $\mathrm{Sen} : \mathrm{Sign} \to \mathrm{Set}$ giving the set of sentences over a signature
- a faithful functor $\mathrm{Sym} : \mathrm{Sign} \to \mathrm{Set}$ giving the set of symbols of a signature
- for each $\Sigma \in |\mathrm{Sign}|$ a relation $\vdash_\Sigma \subseteq \mathcal{P}(\mathrm{Sen}(\Sigma)) \times \mathrm{Sen}(\Sigma)$ which
  - is reflexive, transitive, monotonic
  - and satisfies $\vdash$-translation: Given a signature morphism $\sigma : \Sigma_1 \to \Sigma_2$, we have $\Gamma \vdash_{\Sigma_1} \varphi \Rightarrow \mathrm{Sen}(\sigma)(\Gamma) \vdash_{\Sigma_2} \mathrm{Sen}(\sigma)(\varphi)$
- a natural transformation $\mathrm{symbols} : \mathrm{Sen} \to \mathcal{P} \circ \mathrm{Sym}$ giving the symbols of a sentence

Tool Support for Logics

Problem:

- Some logics don't have direct tool support, e.g. CASL, Common Logic, modal logic
- People need to formalise the theory in tool-supported logics
- ...or cannot use ATP (with premise selection)

Solution:

- Run premise selection in the desired logic
- Translate the modified theory to logic with tool support
- Run the prover on the translation

Theoretical Foundation: Entailment relation morphism

An entailment relation morphism

$$\alpha : (\mathrm{Sen}_S, \mathrm{Sym}_S, \vdash_S, \mathrm{symbols}_S) \to (\mathrm{Sen}_T, \mathrm{Sym}_T, \vdash_T, \mathrm{symbols}_T)$$

is a function $\alpha : \mathrm{Sen}_S \to \mathrm{Sen}_T$ such that

for all $\Gamma \subseteq \mathrm{Sen}_S, \varphi \in \mathrm{Sen}_S : \Gamma \vdash_S \varphi$ implies $\alpha(\Gamma) \vdash_T \alpha(\varphi)$

α is called conservative if

for all $\Gamma \subseteq \mathrm{Sen}_S, \varphi \in \mathrm{Sen}_S : \Gamma \vdash_S \varphi$ if and only if $\alpha(\Gamma) \vdash_T \alpha(\varphi)$

Theoretical Foundation: Theoroidal entailment relation morphism

A conservative theoroidal entailment relation morphism (α, Δ) contains

- a function $\alpha : (\mathrm{Sen}_S, \mathrm{Sym}_S, \vdash_S, \mathrm{symbols}_S) \to (\mathrm{Sen}_T, \mathrm{Sym}_T, \vdash_T, \mathrm{symbols}_T)$
- a base set of sentences $\Delta \subseteq \mathrm{Sen}_T$

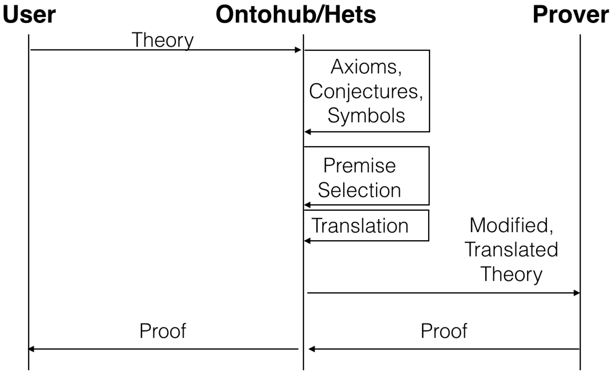
that hold

for all $\Gamma \subseteq \mathrm{Sen}_S, \varphi \in \mathrm{Sen}_S : \Gamma \vdash_S \varphi$ if and only if $\Delta \cup \alpha(\Gamma) \vdash_T \alpha(\varphi)$

Hets

- Evaluation component of Ontohub
- Actually analyses theories
- *Translates theories*
- Interfaces with provers

Data Flow


User
Ontohub/Hets
Prover
Theory
Axioms,
Conjectures,
Symbols
Premise
Selection
Translation
Modified,
Translated
Theory
Proof
Proof

Conclusion

Conclusion and Future Work

Conclusion

- Premise selection improves proving performance significantly
- Entailment relation morphisms allow its use with different logics
- And different reasoning tools
- SInE in Ontohub is only a proof of concept

Future Work

- Develop more premise selection algorithms and deploy them to Ontohub
- Learn from found proofs and disproofs (after premise selection)
- Use modular structure (signature morphisms)

Thank you for listening!

Questions?