Reasoner Verification

Michael Lee, Uli Sattler, Bijan Parsia
University of Manchester

Reasoners

- Vital they are correct
  - Reliant on them for classification, etc.
  - Not possible to do by hand
- Disagreements have occurred before (ORE)
  - An error must have occurred
- Different inferred class hierarchies
  - Complex input, complex output
  - Failing silently

Resolving Disagreements

- GOAL: Find disagreements, resolve them
- Have reasoners
  - Perform classification
  - Disagreement? Make justification for entailment
  - Evaluate each justification (reasoner/human)
- Can examine these small sets
  - Work out correct Class Hierarchy
  - Provide a minimal test case
- Exact details: See ISWC paper.

Method

GOAL: Check for disagreements,
then decide disagreements

FaCT++

HermiT

Method

GOAL: Check for disagreements,
then decide disagreements

$Classify(O)$

$Classify(O)$

FaCT++

HermiT

Method

GOAL: Check for disagreements,
then decide disagreements

$$O \vDash A \sqsubseteq B$$

FaCT++

HermiT

Method

GOAL: Check for disagreements,
then decide disagreements

$$O \models A \sqsubseteq B$$

...

FaCT++

HermiT

Method

GOAL: Check for disagreements,
then decide disagreements

FaCT++

HermiT

Method

GOAL: Check for disagreements,
then decide disagreements

$$O \models A \sqsubseteq B$$

What?

Why?

FaCT++

HermiT

Have reasoners
produce
justifications:

$$J = \{\alpha_1, \alpha_4, \alpha_5\}$$

$$J \models A \sqsubseteq B$$

Method

GOAL: Check for disagreements,
then decide disagreements

FaCT++

HermiT

Have
reasoners
examine
justifications:

$$J = \{\alpha_1, \alpha_4, \alpha_5\}$$

$$J \vDash A \sqsubseteq B$$

No

Method

GOAL: Check for disagreements, then decide disagreements

If unsure,
get a human!

$$J = \{\alpha_1, \alpha_4, \alpha_5\}$$

$$J \models A \sqsubseteq B$$

No

FaCT++

HermiT

Method

GOAL: Check for disagreements, then decide disagreements

If unsure,
get a human!

This is completely wrong!

$$J = \{\alpha_1, \alpha_4, \alpha_5\}$$

$$J \models A \sqsubseteq B$$

No

FaCT++

HermiT

Method

GOAL: Check for disagreements, then decide disagreements

If unsure,
get a human!

This is completely wrong!

FaCT++

HermiT

Experiment & Results

- Tested 4 reasoners on bio-portal corpus
  - FaCT++, JFact, HermiT, Pellet
  - 190 Ontologies (after filtering & processing)
- Class Hierarchy: 181 agree, 9 disagree
  - 1905 entailment disagreements
  - 1622 distinct justifications
- Errors with
  - Datatypes
  - Missing asserted axioms

Understanding Justifications

- Justification verification task
- Using techniques to aid this:
  - Orderings
  - Laconic form
  - Lemmas
- Concerned only with logic not modelling
  - Results have general application
- Will need help from you!

Take Away Messages

- Use more than one reasoner
  - Compare Class Hierarchies
- Method gives
  - Confidence Class Hierarchy is correct (agreement)
  - How to find correct Hierarchy (disagreement)
- Reasoners generally stable
  - 95% level of agreement
- Possible future service
  - May need your help!