Optimization of an abductive reasoner for description logics

Master thesis

Katarína Fabianová
Adviser: Mgr. Júlia Pukancová, PhD.
Consultant: RNDr. Martin Homola, PhD.

FMFI, 2018

Aims

- Descripton logics
  - $\mathcal{ALC}$, $\mathcal{EL}$++
  - DL Tableau algorithm
- Abduction
  - ABox abduction
  - Minimal Hitting Set algorithm
- Reasoning
  - ELK, JFact, Hermit, Pellet
- Implementation
  - Reiter's algorithm with optimizations
  - Adjusted MergeXPlain algorithm
- Evaluation of results

Description logics

- Family of knowledge representation languages
- Every description logic has different expressivity
- Each type of expressivity supports different constructors
- We will use $\mathcal{ALC}$ and $\mathcal{EL}{++}$ DL

Description logics (DL)

Syntax

- $\mathcal{ALC}$ DL is shaped by 3 mutually disjoint sets:

$$N_I = \{a, b, c, ...\}$$

$$N_C = \{A, B, C, ...\}$$

$$N_R = \{R_1, R_2, R_3, ...\}$$

- Individual, concept (atomic or complex)
- $\mathcal{ALC}$ DL consists of the following constructors:

$$\neg, \sqcup, \sqcap, \forall, \exists$$

DL conceptualization: Everybody who is sick, is not happy.

$$\mathit{Sick} \sqsubseteq \neg \mathit{Happy} \text{ (axiom)}$$

- Ontology describes relationships between entities in a specific area
- Knowledge base $\mathcal{KB} = (\mathcal{T}, \mathcal{A})$
- $\mathcal{T}$ stands for TBox, $\mathcal{A}$ stands for ABox
- TBox contains axioms that model ontology
- ABox contains assertion axioms

$$\mathcal{KB} = \left\{ \begin{array}{l} \mathit{Profesor} \sqcup \mathit{Scientist} \sqsubseteq \mathit{Academician} \\ \mathit{AssocProfesor} \sqsubseteq \mathit{Professor} \\ \mathit{jack} : \mathit{Academician} \end{array} \right\}$$

Reasoning

- Reasoning problems: consistency, satisfiability, inference
- Algorithm: DL Tableau algorithm
- Aims: finding model, classification
- Reasoners we work with: ELK, JFact, Hermit, Pellet

Abduction

- Knowledge base and observation is known
- Search for explanations

$$\mathcal{KB} = \left\{ \begin{array}{l} \mathit{Profesor} \sqcup \mathit{Scientist} \sqsubseteq \mathit{Academician} \\ \mathit{AssocProfesor} \sqsubseteq \mathit{Professor} \end{array} \right\}$$

$$\mathcal{O} = \{\mathit{jack} : \mathit{Academician}\}$$

- We use adjusted minimal HS algorithm to find minimal explanations
- Algorithm finds these explanations:

$$\mathcal{E}_1 = \{\mathit{jack} : \mathit{Professor}\}$$

$$\mathcal{E}_2 = \{\mathit{jack} : \mathit{Scientist}\}$$

$$\mathcal{E}_3 = \{\mathit{jack} : \mathit{AssocProfessor}\}$$

Adjusted Reiter's algorithm: Minimal Hitting Set

- Algorithm:
  - Generating a tree using breadth-first search (queue)
  - Root node: negation of model from $\mathcal{KB} \cup \neg\mathcal{O}$
  - Checking conditions to determine node value (explanation, possible explanation (continue with negation of model))
  - Finish if queue is empty
- Implementation with more reasoners (ELK/JFact/Hermit/Pellet)
- Elk should be quicker thanks to lower expressivity
- Optimizations:
  - Observation can not be an explanation
  - Explanation that is not minimal does not even have to be considered

Progress

- Implementation of Reiter's algorithm with optimizations for simple concepts:
  - Works for observation given as simple concept yet
  - Tested on 2 ontologies with correct results
- Started implementation of MergeXPlain algorithm for simple concepts:
  - Designed for observation given as simple concept yet

What is next?

- Modify implementation of Reiter's algorithm with optimizations:
  - To accept complex concept as observation
  - To accept more complex concepts as observation
  - Add more optimizations
- Modify implementation of MergeXPlain algorithm:
  - To process correctly observation given as simple concept
  - To accept complex concepts and possibly more complex concepts as observation
- Evaluation of results (ELK/JFact/Hermit/Pellet)

References

Articles

- Yevgeny Kazakov, Markus Krötzsch, František Simančík. ELK Reasoner: Architecture and Evaluation
- Júlia Pukancová, Martin Homola. Tableau-Based ABox Abduction for Description Logics: Preliminary Report
- Júlia Pukancová, Martin Homola. Tableau-Based ABox Abduction for the $\mathcal{ALCHO}$ Description Logic
- Júlia Pukancová, Martin Homola. ABox Abduction for Description Logics: The Case of Multiple Observations
- Raymond Reiter. A Theory of Diagnosis from First Principles
- Russell Greiner, Barbara A. Smith, Ralph W. Wilkerson. A Correction to the Algorithm in Reiter's Theory of Diagnosis
- Franz Wotawa. A variant of Reiter's hitting-set algorithm
- Kostyantyn Shchekotykhin, Dietmar Jannach and Thomas Schmitz. MERGEXPLAIN: Fast Computation of Multiple Conflicts for Diagnosis

Thank you for your attention