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 logic

- 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 logic (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 following constructors:

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

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

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

- ontology describes relationships between entities in a specific area
- contains 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

$$\mathcal{KB} = \left\{ \begin{array}{l} Sick \sqsubseteq \neg Happy \\ mary : Sick \end{array} \right\}$$

Reasoning

- Reasoning problems: consistency, satisfiability, inference
- Algorithm: DL Tableau algorithm
- Aims: finding model, classification
- Current reasoners: Elk, JFact, Hermit, Pellet

Abduction

- Knowledge base and observation is known
- Search for explanations

$$\mathcal{KB} = \{ \ \mathit{Sick} \sqsubseteq \neg \mathit{Happy} \ \}$$

$$\mathcal{O} = \{\mathit{mary} : \neg \mathit{Happy}\}$$

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

$$\mathcal{E}_1 = \{\mathit{mary} : \mathit{Sick}\}$$

Reiter's algorithm: Minimal hitting set

- Reiter's algorithm computes minimal hitting sets
- Definitions: Hitting set, HS-tree
- Algorithm: Generate pruned HS-tree
- Optimizations:
  - Observation can not be an explanation
  - Explanation that is not minimal does not even have to be considered

Progress

Ontology data set

- Academy

$$\mathcal{KB} = \left\{ \begin{array}{c} Professor \sqcup Scientist \sqsubseteq Academician \\ AssocProfessor \sqsubseteq Professor \end{array} \right\}$$

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

$$\mathcal{E} = \{\ jack : Professor,\ jack : Scientist,\ jack : AssocProfessor\ \}$$

- Test

$$\mathcal{KB} = \left\{ \begin{array}{c} E \sqsubseteq C \\ F \sqsubseteq D \end{array} \right\}$$

$$\mathcal{O} = \{\ a : C\ \}$$

$$\mathcal{E} = \{\ a : E\ \}$$

- Implementation of Reiter's algorithm with optimizations for simple concepts:
  - works for observation given as simple concept yet
  - tested on both ontologies with correct results
- Started implementation of MergeXPlain algorithm for simple concepts:
  - designed for observation given as simple concept yet
  - not working quite good 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