Term-Indexing
in
First Order Theorem Proving

Alexander Maringele

January 27th, 2016

References

- Alexandre Riazanov and Andrei Voronkov, *Efficient instance retrieval with standard and relational path indexing*, Automated Deduction – CADE-19 (Franz Baader, ed.), Lecture Notes in Computer Science, vol. 2741, Springer Berlin Heidelberg, 2003, pp. 380–396 (English).
- R. Sekar, I. V. Ramakrishnan, and Andrei Voronkov, *Term indexing*, Handbook of Automated Reasoning (Alan Robinson and Andrei Voronkov, eds.), Elsevier Science Publishers B. V., Amsterdam, The Netherlands, The Netherlands, 2001, pp. 1853–1964.

Outline

1. $\Sigma = (\Sigma_{\mathsf{P}}, \Sigma_{\mathsf{f}}, V)$ signature
2. $T ::= V \mid F \mid F(T, \ldots, T)$ terms
3. $A ::= P \mid P(T, \ldots, T) \mid T \approx T$ atoms, i.e. predicate terms
4. $\ell ::= A \mid \neg A$ literals
5. $C ::= \Box \mid \ell \mid C \vee \ell$ clauses
6. $S ::= \emptyset \mid S \cup C$ set of clauses
   $\{C_1, \ldots, C_n\} \equiv \forall\, \mathcal{V}(C_1) C_1 \wedge \ldots \forall\, \mathcal{V}(C_1) C_1$

A set of clauses is equivalent to a conjunction of universally quantified disjunctions of literals and equivalent to a universally quantified conjunction of variable distinct disjunctions of literals.

Saturation-based theorem provers

Procedure

1. Transform the negation of a conjecture F into a equisatisfiable set of clauses S.
2. Expand S with derived clauses in a sufficient way.
3. Stop if
    1. a proof for unsatisfiability has been found — F is a theorem
    2. the set is saturated, hence satisfiable — not a theorem
    3. time's up or space's out — we don't know

   Otherwise, continue with 2.

Example (Superposition)

$$S = \{{}^{1:}\mathsf{f}(\mathsf{h}(x)) \approx \mathsf{c} \lor \mathsf{h}(\mathsf{h}(x)) \not\approx \mathsf{a},\ {}^{2:}\mathsf{h}(y) \approx y,\ {}^{3:}\mathsf{f}(\mathsf{a}) \not\approx \mathsf{c}\}$$

$$\dfrac{\dfrac{{}^{2:}\mathsf{h}(y) \approx y \qquad \dfrac{{}^{2:}\mathsf{h}(y) \approx y \qquad \dfrac{\dfrac{\dfrac{{}^{2:}\mathsf{h}(y) \approx y \quad {}^{1:}\mathsf{f}(\boxed{\mathsf{h}(x)}) \approx \mathsf{c} \lor \mathsf{h}(\mathsf{h}(x)) \not\approx \mathsf{a}}{\mathsf{f}(x) \approx \mathsf{c} \lor \mathsf{h}(\mathsf{h}(x)) \not\approx \mathsf{a}}\ \alpha \qquad {}^{3:}\boxed{\mathsf{f}(\mathsf{a})} \not\approx \mathsf{c}}{\mathsf{h}(\mathsf{h}(\mathsf{a})) \not\approx \mathsf{a} \lor \mathsf{c} \not\approx \mathsf{c}}\ \beta}{\mathsf{h}(\boxed{\mathsf{h}(\mathsf{a})}) \not\approx \mathsf{a}}\ \{\}}{\boxed{\mathsf{h}(\mathsf{a})} \not\approx \mathsf{a}}\ \gamma}{\mathsf{a} \not\approx \mathsf{a}}\ \gamma}{\square}\ \{\}$$

$$\alpha = \{y \mapsto x\},\ \beta = \{x \mapsto \mathsf{a}\},\ \gamma = \{y \mapsto \mathsf{a}\}$$

Example (Inst-Gen-Eq)

$$S_0 = \{^{1:}\mathsf{f}(\mathsf{h}(x)) \approx \mathsf{c} \vee \mathsf{h}(\mathsf{h}(x)) \not\approx \mathsf{a}, ^{2:}\,\mathsf{h}(y) \approx y, ^{3:}\,\mathsf{f}(\mathsf{a}) \not\approx \mathsf{c}\}$$

$$S_0\bot = \{\mathsf{f}(\mathsf{h}(\bot)) \approx \mathsf{c} \vee \mathsf{h}(\mathsf{h}(\bot)) \not\approx \mathsf{a}, \mathsf{h}(\bot) \approx \bot, \mathsf{f}(\mathsf{a}) \not\approx \mathsf{c}\}$$

$$\cfrac{\cfrac{^{2^\ell:}\mathsf{h}(y) \approx y \qquad ^{1^\ell:}\mathsf{f}(\boxed{\mathsf{h}(x)}) \approx \mathsf{c}}{\mathsf{f}(x) \approx \mathsf{c}}\;\sigma \qquad ^{3^\ell:}\mathsf{f}(\mathsf{a}) \not\approx \mathsf{c}}{\square}\;\tau$$

$$S_1 = \{\ldots, \mathsf{f}(\mathsf{a}) \not\approx \mathsf{c}, ^{2\sigma\tau:}\,\mathsf{f}(\mathsf{h}(\mathsf{a})) \approx \mathsf{c} \vee \mathsf{h}(\mathsf{h}(\mathsf{a})) \not\approx \mathsf{a}, ^{3\tau:}\,\mathsf{h}(\mathsf{a}) \approx \mathsf{a}\}$$

$$S_1\bot = \{\ldots, \mathsf{f}(\mathsf{a}) \not\approx \mathsf{c}, \mathsf{f}(\mathsf{h}(\mathsf{a})) \approx \mathsf{c} \vee \mathsf{h}(\mathsf{h}(\mathsf{a})) \not\approx \mathsf{a}, \mathsf{h}(\mathsf{a}) \approx \mathsf{a}\}$$

$$\sigma = \{y \mapsto x\},\ \tau = \{x \mapsto \mathsf{a}\}$$

Remark

Superposition [Bachmair, Ganzinger], InstGenEq (Ganzinger, Korovin) are sound and refutational complete with a *fair* strategy.

$\mathsf{f}(x, y) \approx c, \ldots$

Term retrieval problems

- Find terms that are variants of a given term.
  $\mathrm{variant}(\ell, t) \Leftrightarrow \exists\sigma\ \ell\sigma = t$ and σ is renaming.
- Find terms that are unifiable with a given term.
  $\mathrm{unifiable}(\ell, t) \Leftrightarrow \exists\sigma\ \ell\sigma = t\sigma$
- Find terms that are instances of a given term.
  $\mathrm{instance}(\ell, t) \Leftrightarrow \exists\sigma\ \ell = t\sigma$
- Find terms that are generalizations of a given term.
  $\mathrm{generalization}(\ell, t) \Leftrightarrow \exists\sigma\ \ell\sigma = t$

Definition

Definition

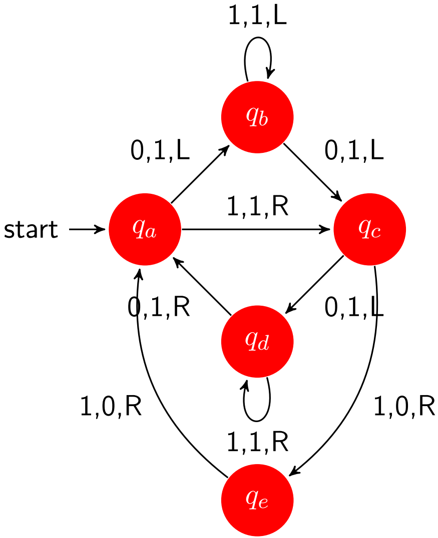
A position is a sequence of positive integers. The empty sequence ε denotes the root position, pq denotes the concatenation of positions. $\mathrm{pos}(t)$ denotes the set of positions in term t, and $t|_p$ denotes the subterm of t at positon $p \in \mathrm{pos}(t)$.

Definition

A postion string is a nonempty string of the form $\langle p_1, s_1\rangle \dots \langle p_n, s_n\rangle$ where p_i are positions and s_i are function or variable symbols and

1. if p_i is a proper prefix of p_j then $i < j$
2.

Definition



www.texample.net/tikz/examples/state-machine/

to do