Term-Indexing

for Instantiation-Based First Order Theorem Proving

Alexander Maringele

January 27th, 2016

References

- Robert Nieuwenhuis, Thomas Hillenbrand, Alexandre Riazanov, and Andrei Voronkov, *On the evaluation of indexing techniques for theorem proving*, Automated Reasoning (Rajeev Goré, Alexander Leitsch, and Tobias Nipkow, eds.), Lecture Notes in Computer Science, vol. 2083, Springer Berlin Heidelberg, 2001, pp. 257–271.

- 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, 2001, pp. 1853–1964.

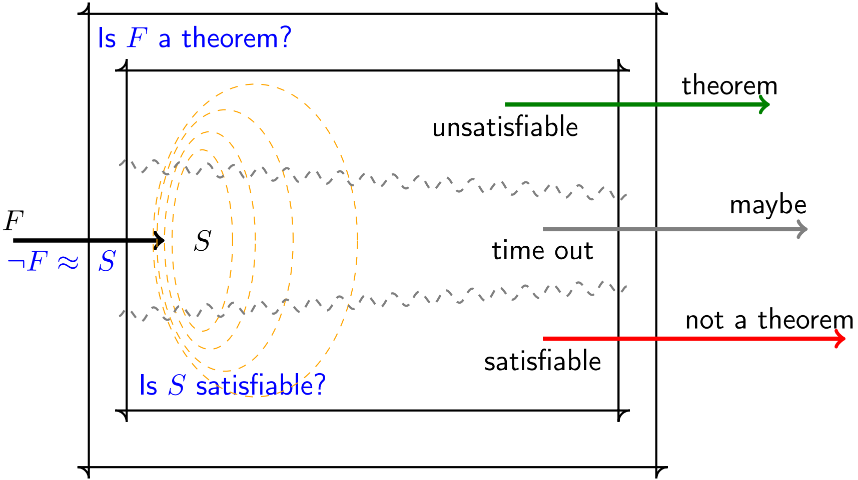
Outline

Notation (First Order Logic)

- $\Sigma = (\mathcal{V}, \mathcal{F}, \mathcal{P})$ signature
- $\mathcal{T} = \mathcal{V} \cup \{\ f(t_1, \ldots, t_n) \mid f \in \mathcal{F}_{(n)}, t_i \in \mathcal{T}\}$ terms
- $\mathcal{A} = \{P(t_1, \ldots, t_n) \mid P \in \mathcal{P}_{(n)}, t_i \in \mathcal{T}\} \cup \{s \approx t \mid s, t \in \mathcal{T}\}$ atoms
- $\mathcal{L} = \mathcal{A} \cup \{\neg A \mid A \in \mathcal{A}\}$ literals
- $\mathcal{C} = 2^{\mathcal{L}}$, e.g. $\mathsf{f}(x) \approx \mathsf{a} \vee \mathsf{a} \not\approx \mathsf{b}$ clauses
- $\mathcal{S} = 2^{\mathcal{C}}$, e.g. $\{\mathsf{f}(x) \approx \mathsf{a} \vee \mathsf{a} \not\approx \mathsf{b}, \mathsf{f}(x) \approx \mathsf{b}\}$ clause sets

$$\{C_1, \ldots, C_n\} \equiv \forall \mathcal{V}\mathsf{ar}(C_1) C_1 \wedge \ldots \wedge \forall \mathcal{V}\mathsf{ar}(C_1) C_n$$

Resolution


Is F a theorem?
theorem
unsatisfiable
maybe
F
S
$\neg F \approx S$
time out
not a theorem
satisfiable
Is S satisfiable?

Goal

A sound, refutational complete, and *effective* calculus.

- *Reduce* search space
  - e.g. Ordered Resolution
  - ...selection functions for clauses and literals
- *Reduce* redundancy
  - Ignore clause D, if C subsumes D, i.e. $C\tau \subseteq D$?
  - ...depends on the calculus

Example

$$S = \{{}^{1:}\mathsf{P}(x,y), {}^{2:}\neg\mathsf{P}(\mathsf{a},z)\} \cup \{{}^{3:}\mathsf{P}(\mathsf{a},z')\}$$

t_1 subsumes t_3

$$S\bot = \{\mathsf{P}(\bot,\bot), \neg\mathsf{P}(\mathsf{a},\bot), \mathsf{P}(\mathsf{a},\bot)\}$$

InstGen/SMT

$$\frac{\mathsf{P}(x,y) \qquad \neg\mathsf{P}(\mathsf{a},z)}{\square}\ \{x \mapsto \mathsf{a}, y \mapsto z\}$$

Resolution

Unit Superposition Inference Rules

$$\frac{s \approx t \quad L[s']}{(L[t]) \cdot \sigma} \text{ unit paramodulation}$$

where $\sigma = \mathrm{mgu}(s, s')$, $s' \notin \mathcal{V}$, $t\sigma \not\succeq s\sigma$

$$\frac{s \approx t \quad u[s'] \not\approx v}{(u[t] \not\approx v) \cdot \sigma} \text{ unit superposition } \frac{s \approx t \quad u[s'] \approx v}{(u[t] \approx v) \cdot \sigma}$$

where $\sigma = \mathrm{mgu}(s, s')$, $s' \notin \mathcal{V}$, $t\sigma \not\succeq s\sigma$, $v\sigma \not\succeq u[s']\sigma$

$$\frac{s \not\approx t}{\Box} \text{ unit equality resolution} \qquad \frac{A \quad \neg B}{\Box} \text{ unit resolution}$$

where s and t (A and B respectively) are unifiable

Term retrieval problems

Definition

In a given set of terms

- find terms that are variants of a given term t
  $\mathrm{variant}(s,t) \Leftrightarrow \exists\sigma\ s\sigma = t$ and σ is renaming
- find terms that are instances of a given term
  $\mathrm{instance}(s,t) \Leftrightarrow \exists\sigma\ s = t\sigma$ backward subsumption
- find terms that are generalizations of a given term
  $\mathrm{generalization}(s,t) \Leftrightarrow \exists\sigma\ s\sigma = t$ forward subsumption
- find terms that are unifiable with a given term
  $\mathrm{unifiable}(s,t) \Leftrightarrow \exists\sigma\ s\sigma = t\sigma$ resolution, etc.

Position-Strings

Positions in a term

$$\mathcal{P}\mathsf{os}(t) = \begin{cases} \{\epsilon\} & \text{if } t = x \in \mathcal{V} \\ \{\epsilon\} \cup \{ip \mid 1 \leq i \leq n \wedge p \in \mathcal{P}\mathsf{os}(t_i)\} & \text{if } t = f(t_1, \ldots, t_n) \end{cases}$$

$\mathsf{h}(\mathsf{f}(\mathsf{a}, y))$

$$\begin{tikzcd}[column sep=small] & \langle \epsilon, \mathsf{h} \rangle \arrow[d] & \\ & \langle 1, \mathsf{f} \rangle \arrow[dl] \arrow[dr] & \\ \langle 11, \mathsf{a} \rangle & & \langle 12, \mathsf{y} \rangle \end{tikzcd}$$

$$\mathcal{P}\mathsf{os}(\mathsf{h}(\mathsf{f}(\mathsf{a}, \mathsf{y}))) = \{\epsilon, 1, 11, 12\}$$

$$\mathsf{h}(\mathsf{f}(\mathsf{a}, \mathsf{y}))|_{12} = y \qquad \langle 12, y \rangle$$

$\langle \epsilon, \mathsf{h} \rangle \langle 1, \mathsf{f} \rangle \langle 12, y \rangle$	path to leaf	$(\mathsf{h}1\mathsf{f}2y)$
$\langle \epsilon, \mathsf{h} \rangle \langle 1, \mathsf{f} \rangle \langle 11, \mathsf{a} \rangle \langle 12, y \rangle$	pre-order	$(\mathsf{hfa}y)$

Normalization of variables

Amnesia

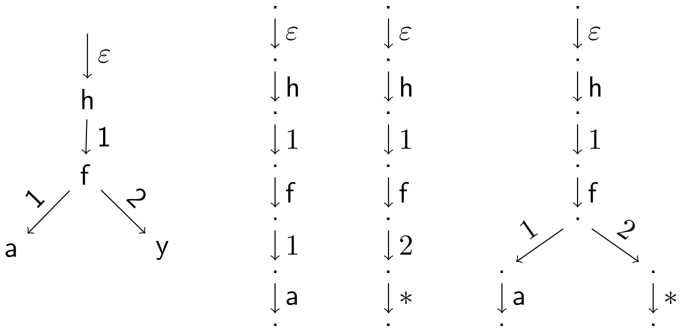
$$\text{path} : \mathsf{f}(x, y) \mapsto \{\mathsf{f}1*, \mathsf{f}2*\}$$
$$\mathsf{f}(x, x) \mapsto \{\mathsf{f}1*, \mathsf{f}2*\}$$
$$\text{pre-order} : \mathsf{f}(x, y) \mapsto \mathsf{f}**$$
$$\mathsf{f}(x, x) \mapsto \mathsf{f}**$$

Renaming

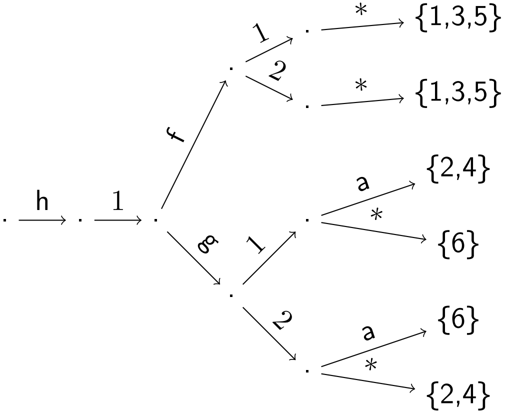
$$\text{path} : \mathsf{f}(x, y) \mapsto \{\mathsf{f}1*_1, \mathsf{f}2*_2\}$$
$$\mathsf{f}(x, x) \mapsto \{\mathsf{f}1*_1, \mathsf{f}2*_1\}$$
$$\text{pre-order} : \mathsf{f}(x, y) \mapsto \mathsf{f}*_1*_2$$
$$\mathsf{f}(x, x) \mapsto \mathsf{f}*_1*_1$$

Path-Strings

The path-strings of $t = \mathsf{h}(\mathsf{f}(\mathsf{a}, y))$ are $\epsilon.\mathsf{h}.1.\mathsf{f}.1.\mathsf{a}$ and $\epsilon.\mathsf{h}.1.\mathsf{f}.1.2.*$:

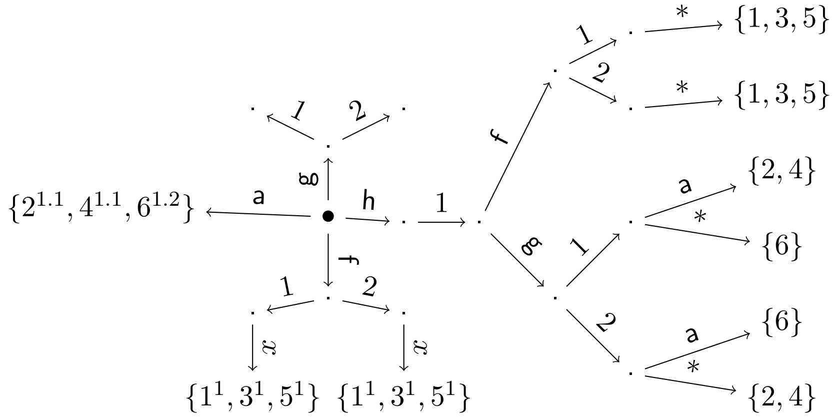


$$\{{}^{1:}\mathsf{h}(\mathsf{f}(x,x)),{}^{2:}\mathsf{h}(\mathsf{g}(\mathsf{a},x)),{}^{3:}\mathsf{h}(\mathsf{f}(y,z)){}^{4:}\mathsf{h}(\mathsf{g}(\mathsf{a},y)),{}^{5:}\mathsf{h}(\mathsf{f}(y,x)),{}^{6:}\mathsf{h}(\mathsf{g}(y,a))\}$$

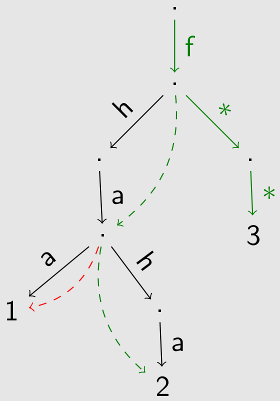


$\mathsf{h}(\mathsf{g}(y,x)) \mapsto \{\mathsf{h}.1.\mathsf{g}.1.*, \mathsf{h}.1.\mathsf{g}.2.*\}$

$$\{ {}^{1:}\mathsf{h}(\mathsf{f}(x,x)), {}^{2:}\mathsf{h}(\mathsf{g}(\mathsf{a},x)), {}^{3:}\mathsf{h}(\mathsf{f}(y,z)) {}^{4:}\mathsf{h}(\mathsf{g}(\mathsf{a},y)), {}^{5:}\mathsf{h}(\mathsf{f}(y,x)), {}^{6:}\mathsf{h}(\mathsf{g}(y,a)) \}$$



Non-linear terms

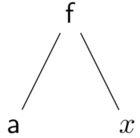


$\{{}^{1:}\mathsf{f}(\mathsf{h}(\mathsf{a}),\mathsf{a}), {}^{2:}\mathsf{f}(\mathsf{h}(\mathsf{a}),\mathsf{h}(\mathsf{a}))\}, {}^{3:}\mathsf{f}(x,y)\}$

The terms $\mathsf{f}(x,y)$ and $\mathsf{f}(z,z)$
share the preorder-term $\mathsf{f}{*}{*}$.

But $\mathsf{f}(\mathsf{h}(\mathsf{a}),\mathsf{a})$ is not an instance of $\mathsf{f}(z,z)$.

$\mathsf{f}(\mathsf{a}, x)$



$\mathsf{f}(\mathsf{a}, x) \approx \mathsf{f}(x, \mathsf{a})$