First-Order Term-Indexing

Alexander Maringele

January 27th, 2016

References

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

Clausal form

$$\{\ \mathsf{P}(\mathsf{f}(x)) \lor \mathsf{f}(x) \not\approx \mathsf{a},\ \ \mathsf{g}(x,y) \approx \mathsf{a} \lor \neg\mathsf{Q}(x,y),\ \ \mathcal{C}_3\ \}$$

$$\equiv$$

$$\forall x\,(\mathsf{P}(\mathsf{f}(x)) \lor \mathsf{f}(x) \not\approx \mathsf{a})$$

$$\land$$

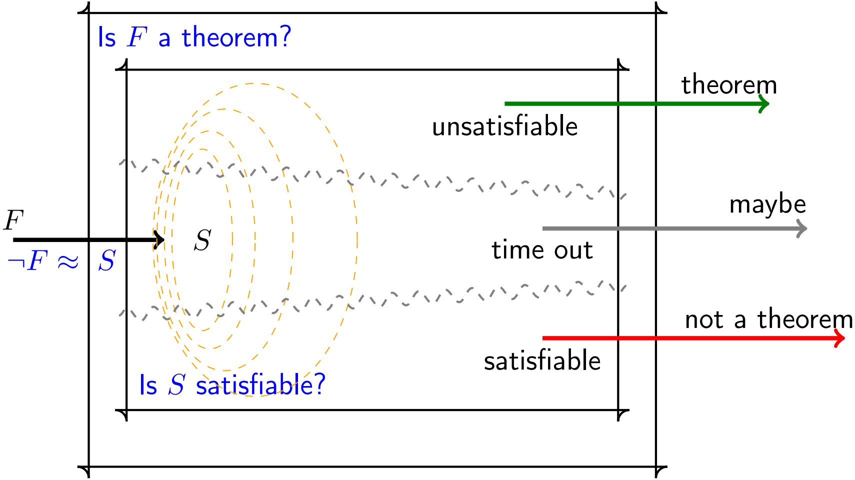
$$\forall xy\,(\mathsf{g}(x,y) \approx \mathsf{a} \lor \neg\mathsf{Q}(x,y))$$

$$\land$$

$$\forall\,\mathcal{V}\mathsf{ar}(\mathcal{C}_3)\,(\mathcal{C}_3)$$

Refutation

FOL Theorem Proving



Goal

A sound and refutation complete calculus.

Resolution (without equality)

Resolve and factor all clauses and literals in an unsatisfiable set.

$$\frac{A \vee \mathcal{C} \quad \neg B \vee \mathcal{D}}{(\mathcal{C} \vee \mathcal{D})\sigma}\ (\sigma)\ \text{resolution} \qquad \frac{A \vee B \vee \mathcal{C}}{(A \vee \mathcal{C})\sigma}\ (\sigma)\ \text{factoring}$$

$$\sigma = \mathrm{mgu}(A, B)$$

With a fair strategy the empty clause will be derived eventually.

Observation

Usually the set grows too fast to obtain a result.

Goal

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

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

Example (forward subsumption)

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

t_1 subsumes t_3

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

Resolution

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

InstGen/SMT

Goal

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

3. Quickly find
   - *variants* — variant removal
   - *instances* — backward subsumption
   - *generalizations* — forward subsumption
   - *unifiable terms* — resolution, etc.

   of a query term in a given set of terms.

Observation

Deduction rate drops quickly with linear search.

Improvement

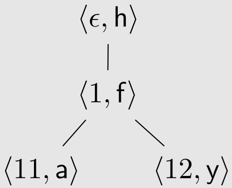
Term-Indexing

Position-Strings

Positions of a term

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

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



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

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

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

Variable names

Ignore names

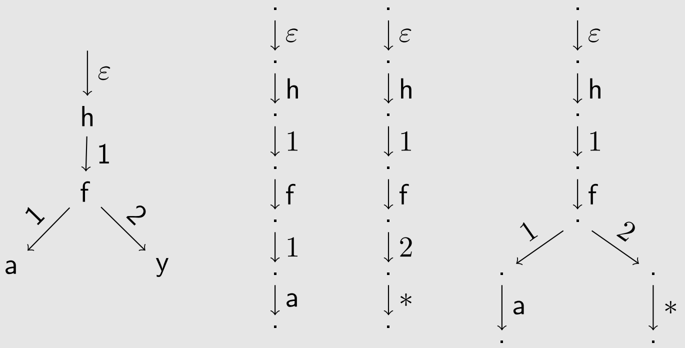
$$\text{path}: \mathsf{f}(x,y) \mapsto \{\mathsf{f}1*, \mathsf{f}2*\}$$
$$\mathsf{f}(z,z) \mapsto \{\mathsf{f}1*, \mathsf{f}2*\}$$
$$\text{pre-order}: \mathsf{f}(x,y) \mapsto \mathsf{f}**$$
$$\mathsf{f}(z,z) \mapsto \mathsf{f}**$$

Normalization

$$\text{path}: \mathsf{f}(x,y) \mapsto \{\mathsf{f}1x_1, \mathsf{f}2x_2\}$$
$$\mathsf{f}(z,z) \mapsto \{\mathsf{f}1x_1, \mathsf{f}2x_1\}$$
$$\text{pre-order}: \mathsf{f}(x,y) \mapsto \mathsf{f}x_1x_2$$
$$\mathsf{f}(z,z) \mapsto \mathsf{f}x_1x_1$$

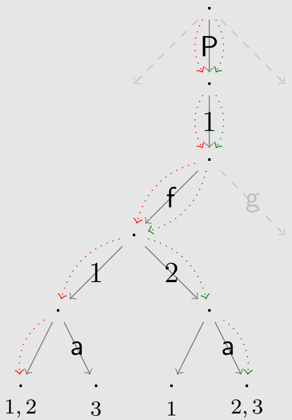
Path-Strings

The path-strings of $\mathsf{h}(\mathsf{f}(\mathsf{a}, y))$ are h1f1a and h1f2∗.



Path-Index

Insert

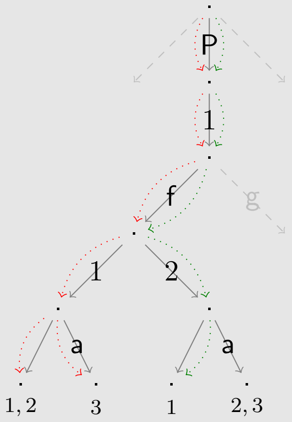


$^{1:}$P(f(x, y))

$^{2:}$P(f(x, a))

$^{3:}$P(f(a, a))

Retrieve



[1:] $\mathsf{P}(\mathsf{f}(x, y))$

[2:] $\mathsf{P}(\mathsf{f}(x, \mathsf{a}))$

[3:] $\mathsf{P}(\mathsf{f}(\mathsf{a}, \mathsf{a}))$

$$u : \mathsf{P}(\mathsf{f}(x', \mathsf{b})) \mapsto \{1, 2, 3\} \cap \{1, 3\}$$

$$i : \mathsf{P}(\mathsf{f}(x', \mathsf{b})) \mapsto \{1, 2, 3\} \cap \{\}$$

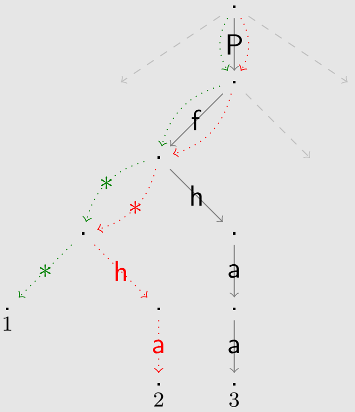
$$g : \mathsf{P}(\mathsf{f}(x', \mathsf{b})) \mapsto \{1, 2\} \cap \{1\}$$

$$v : \mathsf{P}(\mathsf{f}(x', \mathsf{b})) \mapsto \{1, 2\} \cap \{\}$$

$$v : \mathsf{P}(\mathsf{f}(x', x')) \mapsto \{1, 2\} \cap \{1\}$$

Discrimination Trees

Insert

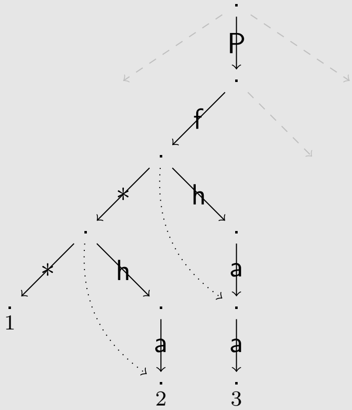


[1:] $\mathsf{P}(\mathsf{f}(x, y))$

[2:] $\mathsf{P}(\mathsf{f}(x), \mathsf{h}(\mathsf{a})))$

[3:] $\mathsf{P}(\mathsf{f}(\mathsf{h}(\mathsf{a})), \mathsf{a})$

Retrieve



$$^{1:}\mathsf{P}(\mathsf{f}(x, y))$$

$$^{2:}\mathsf{P}(\mathsf{f}(x), \mathsf{h}(\mathsf{a})))$$

$$^{3:}\mathsf{P}(\mathsf{f}(\mathsf{h}(\mathsf{a})), \mathsf{a})$$

$$u : \mathsf{P}(\mathsf{f}(x', y')) \mapsto$$