FLEA

first order proving with equality
master project

Alexander Maringele
Supervisor: Georg Moser

Sep 2016 – Obergurgl

Project overview

Goals and requirements

- Goals
  - ATP for FOL with equality
  - Master thesis
- Requirements
  - Input: problems in clausal normal form
  - Data: clauses, literals, terms, indices, etc.
  - Algorithms: substitution, unification, etc.
  - **Proof search**
- Non-Goals and Non-Requirements
  - CASC
  - FOF $\approx$ CNF

Clausal normal form

PUZ001-1

```
...
cnf(same_hates, hypothesis, ( ~hates(agatha,X) | hates(butler,X) )).
...
```

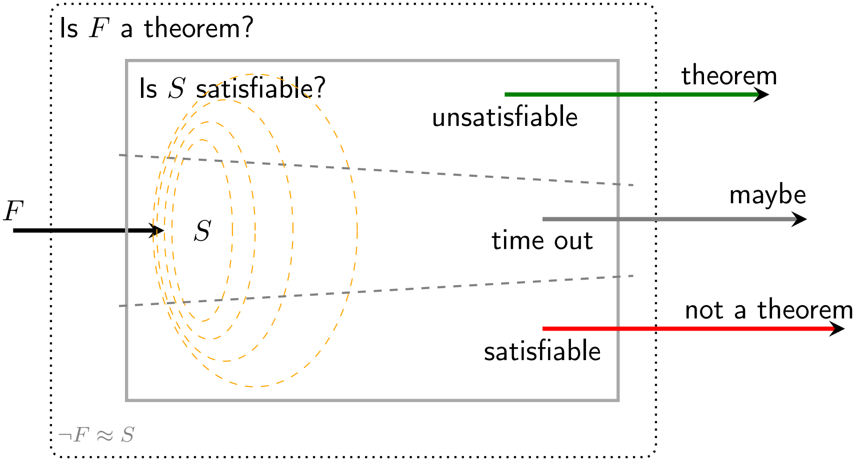
$$\{\ \ldots,\ \neg\mathsf{hates}(\mathsf{agatha}, x) \vee \mathsf{hates}(\mathsf{butler}, x),\ \ \ldots\}$$

$$\equiv$$

$$\ldots\ \wedge\ \forall x\ (\neg\mathsf{hates}(\mathsf{agatha}, x) \vee \mathsf{hates}(\mathsf{butler}, x))\ \wedge\ \ldots$$

G. Sutcliffe, *The TPTP Problem Library and Associated Infrastructure: The FOF and CNF Parts, v3.5.0*, Journal of Automated Reasoning **43** (2009), no. 4, 337–362.

Refutation


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

Ordered Resolution

$$\frac{A \vee C \quad \neg B \vee D}{(C \vee D)\sigma}$$

where

$$A\sigma \text{ strictly maximal in } \mathcal{C}\sigma,\ \neg B\sigma \text{ maximal in } \mathcal{D}\sigma,\ \sigma = \mathrm{mgu}(A, B).$$

$$S = \{\mathsf{P}(x) \vee \neg\mathsf{P}(y), \neg\mathsf{P}(\mathsf{a}), \mathsf{P}(\mathsf{b})\}$$

$$\dfrac{\dfrac{\mathsf{P}(x) \vee \neg\mathsf{P}(y) \quad \neg\mathsf{P}(\mathsf{a})}{\neg\mathsf{P}(y)}\ x \mapsto \mathsf{a} \qquad \mathsf{P}(\mathsf{b})}{\square}\ y \mapsto \mathsf{b}$$

Inst-Gen

$$\frac{A \vee \mathcal{C} \quad \neg\mathsf{c} \vee D}{(A \vee \mathcal{C})\sigma \quad (\neg B \vee \mathcal{D})\sigma}$$

where

$$\mathrm{sel}(A \vee \mathcal{C}) = A \qquad \mathrm{sel}(\neg B \vee \mathcal{D}) = \neg B \qquad \sigma = \mathrm{mgu}(A, B)$$

Selection

$S = \{\mathsf{P}(x) \vee \neg\mathsf{P}(y), \neg\mathsf{P}(\mathsf{a}), \mathsf{P}(\mathsf{b})\}$ unsatisfiable

$P = (p_* \vee \neg p_*) \wedge \neg p_a \wedge p_b$ satisfiable

$$S_0 = \{\mathsf{P}(x) \vee \neg\mathsf{P}(y), \neg\mathsf{P}(\mathsf{a}), \mathsf{P}(\mathsf{b})\}$$ unsatisfiable

$$P_0 = (p_* \vee \neg p_*) \wedge \neg p_a \wedge p_b$$ satisfiable

$$\frac{\mathsf{P}(x) \vee \neg\mathsf{P}(y)) \quad \neg\mathsf{P}(\mathsf{a})}{\mathsf{P}(\mathsf{a}) \vee \neg\mathsf{P}(y)}\ x \mapsto \mathsf{a}$$

$$P_1 = (p_* \vee \neg p_*) \wedge \neg p_a \wedge p_b \wedge (p_a \vee \neg p_*)$$ satisfiable

$$\frac{\mathsf{P}(\mathsf{b}) \quad \mathsf{P}(\mathsf{a}) \vee \neg\mathsf{P}(y)}{\mathsf{P}(\mathsf{a}) \vee \neg\mathsf{P}(\mathsf{b})}\ y \mapsto \mathsf{b}$$

$$P_2 = (p_* \vee \neg p_*) \wedge \neg p_a \wedge p_b \wedge (p_a \vee \neg p_*) \wedge (p_a \vee \neg p_b)$$ unsatisfiable

Equality

$$S = \{^{1:}\mathsf{P}(\mathsf{a}),^{2:} \neg\mathsf{P}(\mathsf{f}(x,\mathsf{b})),^{3:} \mathsf{f}(x,\mathsf{b}) = x\}$$ saturated

$$P = p_1 \wedge \neg p_2 \wedge e_3$$ satisfiable

Equality Axioms

$$\frac{x = x \quad s \neq s \vee \mathcal{C}}{s = s \qquad \mathcal{C}} \; x \mapsto s$$

reflexivity

$$\frac{x \neq y \vee y = x \quad s \neq t \vee \mathcal{C}}{t \neq s \vee s = t} \; x \mapsto t, y \mapsto s$$

symmetry

$$\frac{x \neq y \vee y \neq z \vee x = z \quad s \neq t \vee \mathcal{C}}{s \neq y \vee y \neq t \vee s = t} \; x \mapsto s, y \mapsto t$$

transitivity

$$\frac{x \neq y \vee \mathsf{f}(x) = \mathsf{f}(y) \quad \mathsf{f}(s) \neq \mathsf{f}(t)}{s \neq t \vee \mathsf{f}(s) = \mathsf{f}(t)} \; x \mapsto s, y \mapsto t$$

f-congruence

$$\frac{x \neq y \vee \mathsf{L}^c(x) \vee \mathsf{L}(y) \quad \mathsf{L}(s) \vee \mathcal{C}}{s \neq y \vee \mathsf{L}^c(s) \vee \mathsf{L}(y)} \; x \mapsto s$$

P-congruence

$$S_0 = \{^{1:}\mathsf{P}(\mathsf{a}), ^{2:}\neg\mathsf{P}(\mathsf{f}(x,\mathsf{b})), ^{3:}\mathsf{f}(x,\mathsf{b}) = x\}$$ unsatisfiable

$$P_0 = p_1 \wedge \neg p_2 \wedge e_3$$ satisfiable

$$^{4:}x \neq \mathsf{f}(x,\mathsf{b}) \vee \neg\mathsf{P}(x) \vee \mathsf{P}(\mathsf{f}(x,\mathsf{b}))$$ $^{2:1}\mathsf{P}_{\neg}$-congruence

$$P_1 = P_0 \wedge (\neg e_4 \vee \neg p_* \vee p_2)$$

$$^{5:}\mathsf{f}(x,\mathsf{b}) \neq x \vee x = \mathsf{f}(x,\mathsf{b})$$ $^{4:1}$symmetry

$$P_2 = P_0 \wedge (\neg e_4 \vee \neg p_* \vee p_2) \wedge (\neg e_3 \vee e_4)$$

$$^{6:}\mathsf{a} \neq \mathsf{f}(\mathsf{a},\mathsf{b}) \vee \neg\mathsf{P}(\mathsf{a}) \vee \mathsf{P}(\mathsf{f}(\mathsf{a},\mathsf{b}))$$ $^{4:1,1:1}\ x \mapsto \mathsf{a}$

$$^{7:}\neg\mathsf{P}(\mathsf{f}(\mathsf{a},\mathsf{b}))$$ $^{2:1,6:1}\ x \mapsto \mathsf{a}$

$$P_3 = P_2 \wedge (\neg e_6 \vee \neg p_1 \vee p_6) \wedge \neg p_6$$

$$^{8:}\mathsf{f}(\mathsf{a},\mathsf{b}) \neq \mathsf{a} \vee \mathsf{a} = \mathsf{f}(\mathsf{a},\mathsf{b})$$ $^{1:1}$-symmetry

$$^{9:}\mathsf{f}(\mathsf{a},\mathsf{b}) = \mathsf{a})$$ $^{3:1,8:1}\ x \mapsto a$

$$P_4 = P_3 \wedge (\neg e_8 \vee e_6) \wedge e_8$$ unsatisfiable

Inst-Gen-Eq

Find a tree proof for the empty clause from selected literals.
Instantiate contributing clauses with substitutions from the tree.

Unit-Superposition

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

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

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

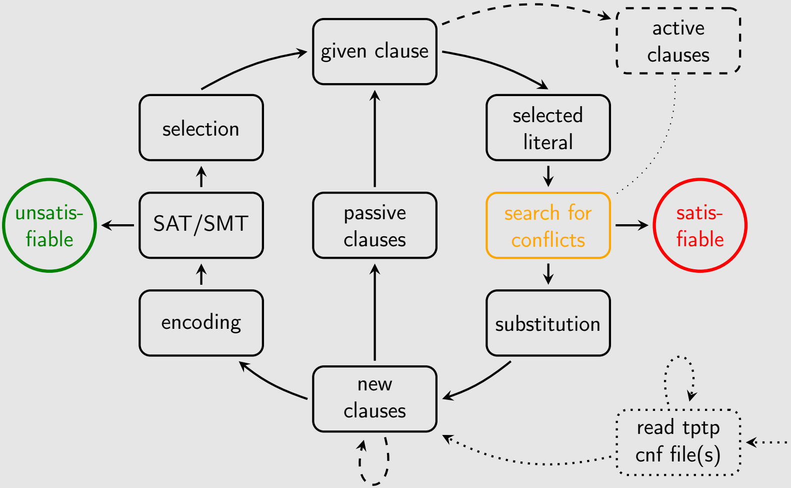
where $\sigma = \mathrm{mgu}(s, s')$, $s' \notin \mathcal{V}$, $t\sigma \not\succeq s\sigma$, $v\sigma \not\succeq u[s']\sigma$

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

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

- requirements
  - Linux, macOS
  - Flex, Bison, Clang, Swift
  - Yices, Z3
- structure
  - wrapper for C-APIs
  - data and algorithms
  - sequential processing of growing list of clauses
- optimizations
  - indexing of literals
  - indexing of clauses
  - sharing of terms
- in progress = missing
  - maximal completion
  - experiments
  - written thesis

Run-loop



given clause
active clauses
selection
selected literal
unsatis-fiable
SAT/SMT
passive clauses
search for conflicts
satis-fiable
encoding
substitution
new clauses
read tptp cnf file(s)