SGGS Decision Procedures

Maria Paola Bonacina and Sarah Winkler
Università degli Studi di Verona

10th International Joint Conference on Automated Reasoning
3 July 2020

Overview

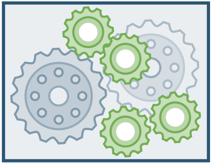
conflict-driven

model-constructing

Automatic Theorem Prover

$Q(0,0,0,0)$
$\neg Q(x,y,z,0) \vee Q(x,y,z,1)$
$\neg Q(x,y,0,1) \vee Q(x,y,1,0)$
$\neg Q(x,0,1,1) \vee Q(x,1,0,0)$
$\neg Q(0,1,1,1) \vee Q(1,0,0,0)$
$\neg Q(1,1,1,1)$

$\longrightarrow$



Semantically Guided Goal Sensitive Reasoning (SGGS)

refutationally complete
model complete

$\longrightarrow$ SAT

?

$\longrightarrow$ UNSAT

Decidable Fragment

- subset of first-order formulas for which there exists a decision procedure
- examples: Ackermann, monadic, guarded, EPR, PVD, FO^2, ...

This Talk

term rewriting to recognize decidable problems

- SGGS as decision procedure: stratified, restrained, and PVD fragments,
- restrained fragment: new decidable class
- SGGS implementation in prover Koala

Contents

SGGS: Ingredients

- set of input clauses S in many-sorted logic
- initial interpretation $\mathcal{I}$ $\qquad \mathcal{I}^-(\mathsf{P}) = \cdots = \mathcal{I}^-(\mathsf{R}) = \bot$
  $\mathcal{I}^+(\mathsf{P}) = \cdots = \mathcal{I}^+(\mathsf{R}) = \top$
- Herbrand constraints $\qquad top(x) \neq \mathsf{f} \wedge x \not\equiv y$
- constrained clause $A \rhd C$ is clause C with Herbrand constraint A, one literal selected per clause $\qquad top(x) \neq \mathsf{f} \rhd \neg\mathsf{P}(\mathsf{a}) \vee \mathsf{Q}(\mathsf{a}, x)$
- trail Γ is sequence of constrained clauses
- inference system $\vdash$ on trails Γ, parameterized by $\mathcal{I}$

Model representation: $\mathcal{I}[\Gamma]$

for trail $\Gamma = A_1 \rhd C_1[L_1], \ldots, A_n \rhd C_n[L_n]$ without conflict:
interpretation $\mathcal{I}[\Gamma]$ satisfies $\bigcup_i Gr(A_i \rhd L_i)$ and defaults to $\mathcal{I}$ otherwise

Theorem (Completeness) **(Bonacina & Plaisted 2014,2017)**

for fair derivation $\Gamma_0 \vdash \Gamma_1 \vdash \Gamma_2 \vdash \ldots$ from S with initial interpretation $\mathcal{I}$

- *if S is satisfiable then $\mathcal{I}[\Gamma_\infty] \models S$*
- *otherwise $\bot \in \Gamma_k$ for some k*

Example (SGGS as a Game)

$\mathsf{P}(\mathsf{a})$ $\quad \neg\mathsf{P}(x) \vee \mathsf{Q}(x,y) \vee \mathsf{R}(x)$ $\quad \neg\mathsf{Q}(x,x)$ $\quad \neg\mathsf{R}(\mathsf{b})$

Yay, it's satisfiable!

SGGS inference sequence using initial interpretation $\mathcal{I}^-$:

$$
\begin{array}{llr}
\epsilon \vdash & [\mathsf{P}(\mathsf{a})] & \text{extend} \\
\vdash & [\mathsf{P}(\mathsf{a})],\ \neg\mathsf{P}(\mathsf{a}) \vee [\mathsf{Q}(\mathsf{a},y)] \vee \mathsf{R}(\mathsf{a}) & \text{extend} \\
\vdash & [\mathsf{P}(\mathsf{a})],\ \neg\mathsf{P}(\mathsf{a}) \vee [\mathsf{Q}(\mathsf{a},y)] \vee \mathsf{R}(\mathsf{a}),\ [\neg\mathsf{Q}(\mathsf{a},\mathsf{a})] & \text{extend} \\
\vdash & [\mathsf{P}(\mathsf{a})],\ top(y) \neq \mathsf{a} \rhd \neg\mathsf{P}(\mathsf{a}) \vee [\mathsf{Q}(\mathsf{a},y)] \vee \mathsf{R}(\mathsf{a}), & \\
 & \neg\mathsf{P}(\mathsf{a}) \vee [\mathsf{Q}(\mathsf{a},\mathsf{a})] \vee \mathsf{R}(\mathsf{a}),\ [\neg\mathsf{Q}(\mathsf{a},\mathsf{a})] & \text{split} \\
\vdash & [\mathsf{P}(\mathsf{a})],\ top(y) \neq \mathsf{a} \rhd \neg\mathsf{P}(\mathsf{a}) \vee [\mathsf{Q}(\mathsf{a},y)] \vee \mathsf{R}(\mathsf{a}), & \\
 & [\neg\mathsf{Q}(\mathsf{a},\mathsf{a})],\ \neg\mathsf{P}(\mathsf{a}) \vee [\mathsf{Q}(\mathsf{a},\mathsf{a})] \vee \mathsf{R}(\mathsf{a}) & \text{move} \\
\vdash & [\mathsf{P}(\mathsf{a})],\ top(y) \neq \mathsf{a} \rhd \neg\mathsf{P}(\mathsf{a}) \vee [\mathsf{Q}(\mathsf{a},y)] \vee \mathsf{R}(\mathsf{a}),\ [\neg\mathsf{Q}(\mathsf{a},\mathsf{a})],\ \neg\mathsf{P}(\mathsf{a}) \vee [\mathsf{R}(\mathsf{a})] & \text{resolve}
\end{array}
$$

Semantically Guided Goal-Sensitive Reasoning

SGGS: Ingredients

- set of input clauses S in many-sorted logic
- initial interpretation $\mathcal{I}$ $\qquad \mathcal{I}^-(\mathsf{P}) = \cdots = \mathcal{I}^-(\mathsf{R}) = \bot$
  $\mathcal{I}^+(\mathsf{P}) = \cdots = \mathcal{I}^+(\mathsf{R}) = \top$
- Herbrand constraints $\qquad top(x) \neq \mathsf{f} \wedge x \not\equiv y$
- constrained clause $A \rhd C$ is clause C with Herbrand constraint A, one literal selected per clause $\qquad top(x) \neq \mathsf{f} \rhd \neg\mathsf{P}(\mathsf{a}) \vee [\mathsf{Q}(\mathsf{a}, x)]$
- trail Γ is sequence of constrained clauses
- inference system $\vdash$ on trails Γ, parameterized by $\mathcal{I}$

Model representation: $\mathcal{I}[\Gamma]$

for trail $\Gamma = A_1 \rhd C_1[L_1], \ldots, A_n \rhd C_n[L_n]$ without conflict:
interpretation $\mathcal{I}[\Gamma]$ satisfies $\bigcup_i Gr(A_i \rhd L_i)$ and defaults to $\mathcal{I}$ otherwise

Theorem (Completeness) **(Bonacina & Plaisted 2014,2017)**

for fair derivation $\Gamma_0 \vdash \Gamma_1 \vdash \Gamma_2 \vdash \ldots$ from S with initial interpretation $\mathcal{I}$

- *if S is satisfiable then $\mathcal{I}[\Gamma_\infty] \models S$*
- *otherwise $\bot \in \Gamma_k$ for some k*

Contents

Definition

basis is finite subset $\mathcal{B}$ of Herbrand base of input clause set S

Definition

- trail $A_1 \rhd C_1, \ldots, A_n \rhd C_n$ is in $\mathcal{B}$ if all atoms in $Gr(A_i \rhd C_i)$ are in $\mathcal{B}$
- SGGS derivation is in $\mathcal{B}$ if all its trails are

Lemma

If fair SGGS derivation $\Gamma_0 \vdash \Gamma_1 \vdash \ldots \vdash \Gamma_j \vdash \ldots$ is in $\mathcal{B}$, then $|\Gamma_j| \leqslant |\mathcal{B}|+1 \quad \forall j$

Theorem

A fair SGGS derivation in a finite basis is finite

Small model property

... is obtained if size of $\mathcal{B}$ can be computed

SGGS Decides the Stratified Fragment

Definition

signature $\mathcal{F}$ is stratified, if $\exists$ well-founded ordering $<_s$ on sorts such that all $f : s_1 \times \cdots \times s_n \to s$ in $\mathcal{F}$ satisfy $s <_s s_i$ for all $1 \leqslant i \leqslant n$

Example

- $\mathrm{P}(0,0,0,0) \wedge (\neg \mathrm{P}(x,y,z,0) \vee \mathrm{P}(x,y,z,1))$ EPR
  $0 : s_1 \quad 1 : s_1 \quad \mathrm{P} : s_1 \times s_1 \times s_1 \times s_1$ ✓
- $(\mathrm{Q}(\mathrm{f}(\mathrm{a}),y) \vee \mathrm{Q}(x,\mathrm{a})) \wedge \neg \mathrm{Q}(\mathrm{b},y)$
  $\mathrm{f} : s_2 \to s_1 \quad \mathrm{a} : s_2 \quad \mathrm{b} : s_1 \quad \mathrm{Q} : s_1 \times s_2 \quad s_1 <_s s_2$ ✓
- $\mathrm{R}(x) \vee \mathrm{R}(\mathrm{f}(x))$ ✗

Decidability (Abadi *et al* 2001)

for clause set S over stratified signature, Herbrand base is finite

Theorem

Any fair SGGS derivation from stratified clause set S halts,

- *is refutation if S unsatisfiable,*
- *constructs model if S satisfiable.*

Example (MSC015-n: Exponentially long EPR derivations)

given $k+1$ clauses encoding a binary counter:

$$Q(\bar{0}_k) \qquad \neg Q(\bar{x}_m, 0, \bar{1}_{k-m-1}) \vee Q(\bar{x}_m, 1, \bar{0}_{k-m-1}) \qquad \neg Q(\bar{1}_k)$$

SGGS derivation guided by $\mathcal{I}^-$ needs more than 2^k steps:

$$\begin{array}{llr}
\Gamma_0 : \varepsilon & \vdash \Gamma_1 : [Q(\bar{0}_k)] & \text{extend} \\
& \vdash \Gamma_2 : \ \dots, \neg Q(\bar{0}_k) \vee [Q(\bar{0}_{k-1}, 1)] & \text{extend} \\
& \vdash \Gamma_3 : \ \dots, \neg Q(\bar{0}_{k-1}, 1) \vee [Q(\bar{0}_{k-2}, 1, 0)] & \text{extend} \\
& \dots & \dots \\
& \vdash \Gamma_{2^k} : \ \dots, \neg Q(\bar{1}_{k-1}, 0) \vee [Q(\bar{1}_k)] & \text{extend} \\
& \vdash \Gamma_{2^k+1} : \ \dots, \neg Q(\bar{1}_{k-1}, 0) \vee [Q(\bar{1}_k)],\ [\neg Q(\bar{1}_k)] & \text{extend} \\
& \vdash \Gamma_{2^k+2} : \ \dots, [\neg Q(\bar{1}_k)],\ \neg Q(\bar{1}_{k-1}, 0) \vee [Q(\bar{1}_k)] & \text{move} \\
& \vdash \Gamma_{2^k+3} : \ \dots, [\neg Q(\bar{1}_k)],\ [\neg Q(\bar{1}_{k-1}, 0)] & \text{resolve} \\
& \dots & \dots \\
& \vdash \Gamma_{2^{k+2}+1} : \bot, \dots & \text{resolve}
\end{array}$$

- InstGen, SCL also behave exponentially, but resolution admits linear proof
- InstGen decides the stratified class, resolution does not decide EPR directly

Contents

Definition

clause C is ground-preserving if every variable in C occurs in negative literal

Example

- $\neg \mathsf{P}(\mathsf{s}(\mathsf{s}(x)), y) \vee \mathsf{P}(x, \mathsf{s}(y))$ ✓
- $\neg \mathsf{R}(x, y) \vee \neg \mathsf{R}(y, x) \vee \mathsf{R}(z, z)$ ✗

Lemma

SGGS with $\mathcal{I}^-$ generates only ground clauses from ground preserving clause set

Restrainedness: Basic Idea

for any LPO $\succ$

clause set S is ground-preserving:

$$\mathsf{P}(\mathsf{s}^{10}(0), \mathsf{s}^{9}(0)) \qquad \neg \mathsf{P}(\mathsf{s}(\mathsf{s}(x)), y) \succ \mathsf{P}(x, \mathsf{s}(y)) \qquad \neg \mathsf{P}(\mathsf{s}(0), 0)$$

selected literals have decreasing number of symbols!

SGGS with $\mathcal{I}^-$ generates finite derivation:

$$\begin{aligned}
\varepsilon &\vdash [\mathsf{P}(10, 9)] \\
&\vdash [\mathsf{P}(10, 9)],\ \neg \mathsf{P}(10, 9) \vee [\mathsf{P}(8, 10)] \\
&\vdash [\mathsf{P}(10, 9)],\ \neg \mathsf{P}(10, 9) \vee [\mathsf{P}(8, 10)],\ \neg \mathsf{P}(8, 10) \vee [\mathsf{P}(6, 11)] \vdash \dots
\end{aligned}$$

Definition (Restraining ordering)

quasi-ordering $\succeq$ on terms and atoms is restraining if

- it is stable under substitutions
- strict ordering $\succ = \succeq \setminus \preceq$ is well-founded
- equivalence $\succeq \cap \preceq$ has finite classes

Definition (Restrained clause)

ground-preserving clause C is (strictly) restrained wrt restraining ordering $\succeq$ if

$$\forall \text{ non-ground } L \in C^+ \quad \exists\, \neg M \in C^- \quad \text{such that} \quad M \succeq L \quad (M \succ L)$$

and clause set S is restrained with respect to $\succeq$ if all its clauses are

Example

- previous slide: strictly restrained wrt LPO
  quasi-order needed
  $\mathsf{P}(\mathsf{s}^{10}(0), \mathsf{s}^9(0)) \qquad \neg\mathsf{P}(\mathsf{s}(\mathsf{s}(x)), y) \succ \mathsf{P}(x, \mathsf{s}(y)) \qquad \neg\mathsf{P}(\mathsf{s}(0), 0)$
- binary counter problem: strictly restrained wrt LPO with $0 \succ 1$
  $\mathsf{Q}(\bar{0}_k) \qquad \neg\mathsf{Q}(\bar{x}_m, 0, \bar{1}_{k-m-1}) \succ \mathsf{Q}(\bar{x}_m, 1, \bar{0}_{k-m-1}) \qquad \neg\mathsf{Q}(\bar{1}_k)$
- PLA030-1 contains $\neg\mathrm{diff}(x, y) \succeq \mathrm{diff}(y, x)$: restrained wrt AC-RPO

SGGS Decides the Restrained Fragment

Notation

- $\mathcal{A}_S$ is set of ground atoms occurring in S
- $\mathcal{A}_S^{\preceq}$ is subset of the Herbrand base upper-bounded by $\mathcal{A}_S$: finite

$$\mathcal{A}_S^{\preceq} = \{L \mid L \in \mathcal{A} \text{ such that } \exists M \in \mathcal{A}_S \text{ with } M \succeq L\}$$

Key Lemma

Any fair SGGS-derivation from restrained clause set S using $\mathcal{I}^-$ is in $\mathcal{A}_S^{\preceq}$.

Theorem

any fair SGGS-derivation with $\mathcal{I}^-$ from restrained clause set S halts,

- *is refutation if S unsatisfiable,*
- *constructs model if S satisfiable*

Remarks

- SGGS also decides PVD
- . . . but does not decide (Ackermann, monadic, FO^2): does not halt on

$$\mathsf{P}(0) \qquad \mathsf{P}(x) \lor \mathsf{P}(\mathsf{f}(x)) \qquad \neg\mathsf{P}(x) \lor \neg\mathsf{P}(\mathsf{f}(x))$$

Positive Resolution Decides the Restrained Fragment

Positive Resolution

- ordered resolution using $>$
- such that positive literals are $>$-maximal only in positive clauses

Key Lemma

if S is restrained, then for all $C \in R^*_>(S)$ and all $L \in C^+$ either

(i) $L \in \mathcal{A}^{\preceq}_S$, *or* *(ii)* $M \succeq L$ *for some* $\neg M \in C^-$

Theorem

Any fair ordered resolution run using $>$ from restrained set S terminates, and is a refutation if S is unsatisfiable.

Remark: Flip the Sign

SGGS using $\mathcal{I}^+$ and negative resolution decide negatively restrained class

Contents

Recognizing Restrained Sets

Observation

restrainedness is an undecidable property

Automation: Reduction to Termination of Rewriting

- given clause set S, generate term rewrite system $\mathcal{R}_S$:
  $\forall\ C \in S$ with non-ground $L \in C^+$ have rule in $\mathcal{R}_S$ such that $\neg M \in C^-$:
  $$M \to L$$
- if $\mathcal{R}_S$ terminates then S is strictly restrained with respect to $\to_{\mathcal{R}_S}^+$

Example

binary counter for four bits: $\mathcal{R}_S$ is terminating, e.g. by LPO with $0 > 1$

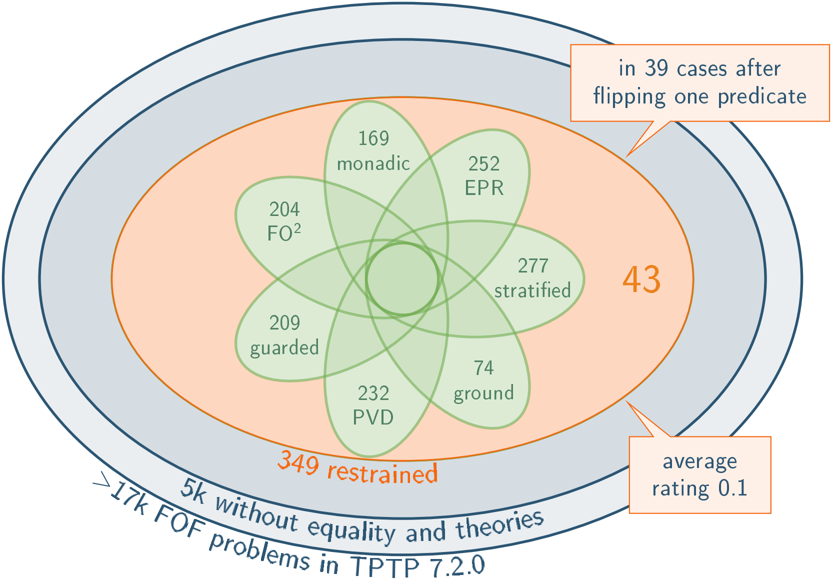
$\mathsf{P}(x, y, z, 0) \to \mathsf{P}(x, y, z, 1) \quad \mathsf{P}(x, y, 0, 1) \to \mathsf{P}(x, y, 1, 0) \quad \mathsf{P}(x, 0, 1, 1) \to \mathsf{P}(x, 1, 0, 0)$

- use termination tools for rewrite systems: $\mathsf{T_TT_2}$ or AProVE

Remark

use termination modulo (relative termination) for non-strict restrainedness

Recognizing Restrained Sets: Experiments



Tool

- implemented in OCaml, re-using some code of iProver:
  re-using data structures, discrimination trees, type inference
- prototype: (very) little optimization
- $\mathcal{I}^+$ or $\mathcal{I}^-$ as initial interpretation, depending on ground-preservingness
- performs type inference to compute sorts, take into account for constraints

Experiments

	Koala		E 2.4		iProver 3.1		Vampire 4.4	
	sat	unsat	sat	unsat	sat	unsat	sat	unsat
1246 stratified	277	643	145	709	333	891	271	872
349 restrained	50	283	47	289	51	294	51	298
351 PVD\restrained	76	232	44	226	85	252	63	252

TPTP 7.2.0, 300s timeout

- `http://profs.scienze.univr.it/winkler/sggsdp`

Conclusion

Discussion

- SGGS attractive as decision procedure: conflict-driven, model-constructing
- SGGS decides fragments with finite basis: stratified, restrained, PVD, ...
- restrained fragment: new decidable class ($\sim$ 10% of tested TPTP problems)
  —use termination tools to recognize restrainedness
- implementation of SGGS in prototype Koala:
  reasonable performance on satisfiable problems

Future Work

- SGGS with equality, extend restrainedness to equality
- use complexity tools for rewriting to automatically estimate model sizes
- improve Koala, find problem classes where conflict-drivenness is beneficial
- combine SGGS with CDSAT

Thanks!