Relational Constraint Solving in SMT

Paul Meng, Andrew Reynolds, Cesare Tinelli, Clark Barrett

Midwest Verification Day September 2018

Relational Reasoning

Many problems can be modeled **relationally**

- Ontologies
- Network systems
- High-level system design
- …

Relational logic is well suited for reasoning about structurally rich problems

A Motivating Example

Modeling a Toy File System

 $*$ – reflexive-transitive closure \wedge – transitive closure

Technical Preliminaries

Satisfiability Modulo Theories (SMT)

Satisfiability Modulo Theories (SMT)

Decide the satisfiability of many-sorted first-order **logic** formulas with respect to combinations of background **theories**

Satisfiability Modulo Theories (SMT)

- A **theory** $\mathcal{T} = (\Sigma, I)$ defines
	- A **signature** Σ: a set of non-logical symbols
	- A class of Σ-interpretations I
	- **Examples**: integer arithmetic, strings, finite sets, ...
		- \triangleright A simple **theory**: $\Sigma_s = \{0, 1, +, \equiv$
		- \triangleright A **formula** in the theory \mathcal{T}_s : $x + 0 = 1$

Related Work

Alloy

A declarative language based on first-order relational **logic** created at MIT

Model and analyze **structurally-rich** systems

SAT-based analysis by the Alloy Analyzer

- **Checks the consistency** of an Alloy Specification
- Can **disprove but only prove a given property** for an Alloy specification **within a given bounds**

Analysis of Alloy Specifications via SMT

El Ghazi et al. [8, 9, 10] **translates** the Alloy kernel language to SMT-LIB language and **solves using SMT solvers** (**AlloyPE**)

The resulting SMT formulas are difficult to solve due to **heavy usage of quantifiers** in the translation

Description Logics (DLs)

Fragments of relational logic for efficient **knowledge representation and reasoning**

Consider on purpose only unary and binary relations

OWL: a standardized semantic web ontology language based on **description logics**

• **Efficient solvers**: KONCLUDE, HermiT, FaCT++ and etc.

A Theory of Finite Set $\mathcal{T}_{\mathcal{S}}$ in SMT

A theory $\mathcal{T}_{\mathcal{S}}$ of **finite sets** was introduced by Kshitij Bansal et al. [3]

Signature \sum_{S} of \mathcal{T}_{S}

- **Singleton set constructor**: $\lceil \underline{\hspace{6pt}} \rceil$: $\alpha \rightarrow \text{Set}(\alpha)$
- **Subset**: \sqsubseteq : $Set(\alpha) \times Set(\alpha) \rightarrow Bool$
- **Membership**: \in : $\alpha \times$ Set(α) \rightarrow Bool
- **Union, intersection, set difference:** $\sqcap,\sqcup,\setminus\; : \mathsf{Set}(\alpha) \times \mathsf{Set}(\alpha) \to \mathsf{Set}(\alpha)$

A modular **set solver** was implemented in **CVC4**

My Research

A Theory of Finite Relations $\mathcal{T}_{\mathcal{R}}$

Type Notations

 $\text{Tup}_n(\alpha_1, ..., \alpha_n)$: a parametric tuple sort (n > 0)

 $Set(Tup_n(\alpha_1, ..., \alpha_n))$: a relational sort abbreviated as $\text{Rel}_n(\alpha_1, \dots, \alpha_n)$

Relational Signature $\Sigma_{\cal R}$ of $\mathcal{T}_{\cal R}$

Tuple constructor:

 $\langle _1, \ldots, _ \rangle : \alpha_1 \times \cdots \times \alpha_n \rightarrow \text{Tup}_n(\alpha_1, \ldots, \alpha_n)$

 \triangleright **Example**: $\langle 1, 2 \rangle$ a binary integer tuple constant

Singleton relation constructor:

 \lceil : Tup_n $(\alpha_1, ..., \alpha_n) \rightarrow$ Rel_n $(\alpha_1, ..., \alpha_n)$

Ø **Example**: ⟨1, "Hello"⟩ a singleton set of integer and string binary tuple

Relational Signature Σ_R of \mathcal{T}_R

Product: $*$: $\text{Rel}_m(\alpha) \times \text{Rel}_n(\beta) \rightarrow \text{Rel}_{m+n}(\alpha, \beta)$

► Example: R1 = $($ 1, 2 $)$, $($ 3, 4 $)$; R2 = $($ $($ 5, 6 $)$] $R1 * R2 = [(1, 2, 5, 6), (3, 4, 5, 6)]$

Join:
$$
\bowtie
$$
 : $\text{Rel}_{p+1}(\alpha, \gamma) \times \text{Rel}_{q+1}(\gamma, \beta) \to \text{Rel}_{p+q}(\alpha, \beta)$
with $p + q > 0$

► **Example**: R1 = [(1, "Hello"), (2, "Hi")];
R2 = [("Hello", 3), ("World", 4)];
R1
$$
\bowtie
$$
 R2 = [(1, 3)]

Relational Signature $\Sigma_{\mathcal{R}}$ of $\mathcal{T}_{\mathcal{R}}$

Transpose: $\mathcal{L}^{-1}: \text{Rel}_m(\alpha_1, \cdots, \alpha_m) \rightarrow \text{Rel}_m(\alpha_m, \cdots, \alpha_1)$

Ø **Example**: R = ⟨1, "Hello"⟩,⟨2, "Hi"⟩ ; R-1 = ⟨"Hello", 1⟩,⟨"Hi" , 2⟩ ;

Transitive Closure: L^+ : Rel₂ (α , α) \rightarrow Rel₂ (α , α)

$$
\triangleright \text{ Example: } R = [(1, 2), (2, 3)]
$$

$$
R^+ = [(1, 2), (2, 3), (1, 3)]
$$

A Calculus $\mathcal{C}_{\mathcal{R}}$ for $\mathcal{T}_{\mathcal{R}}$

A Compact Calculus for T_s

INTER UP	INTER Down		
$x \in s \in S^*$	$x \in t \in S^*$	$s \sqcap t \in \mathcal{T}(S)$	$x \in s \sqcap t \in S^*$
$S := S, x \in s \sqcap t$	$S := S, x \in s, x \in t$		
UNION UP	UNION Down		
$x \in u \in S^*$	$u \in \{s, t\}$	$s \sqcup t \in \mathcal{T}(S)$	$x \in s \sqcup t \in S^*$
$S := S, x \in s \sqcup t$	$s \sqcup t \in \mathcal{T}(S)$	$x \in s \sqcup t \in S^*$	
$S := S, x \in s \parallel S := S, x \in t$			

Derivation rules for intersection and union

A Compact Calculus for \mathcal{T}_{s}

Derivation rules for set difference, singleton, disequality and contradiction

TRANSPOSE Derivation Rule $\binom{-1}{-}$

$$
\frac{\langle x_1, \dots, x_n \rangle \in R \in \mathcal{S}^* \quad R^{-1} \in \mathcal{T}(\mathcal{S})}{\mathcal{S} := \mathcal{S}, \langle x_n, \dots, x_1 \rangle \in R^{-1}}
$$

$$
TransP DOWN
$$

\n
$$
\langle x_1, \dots, x_n \rangle \in R^{-1} \in S^*
$$

\n
$$
S := S, \langle x_n, \dots, x_1 \rangle \in R
$$

JOIN Derivation Rule (M)

Join UP

\n
$$
\frac{\langle x_1, \ldots, x_m, z \rangle \in R_1, \langle z, y_1, \ldots, y_n \rangle \in R_2 \in \mathcal{S}^* \quad m + n > 0 \quad R_1 \bowtie R_2 \in \mathcal{T}(\mathcal{S})}{\mathcal{S} := \mathcal{S}, \langle x_1, \ldots, x_m, y_1, \ldots, y_n \rangle \in R_1 \bowtie R_2}
$$

Join Down
\n
$$
\frac{\langle x_1, \ldots, x_m, y_1, \ldots, y_n \rangle \in R_1 \bowtie R_2 \in S^* \text{ ar}(R_1) = m + 1}{S := S, \langle x_1, \ldots, x_m, z \rangle \in R_1, \langle z, y_1, \ldots, y_n \rangle \in R_2}
$$

z is a fresh variable

PRODUCT Derivation Rule (*)

PROD UP $\langle x_1,\ldots,x_m\rangle \in R_1 \in \mathcal{S}^* \quad \langle y_1,\ldots,y_n\rangle \in R_2 \in \mathcal{S}^* \quad R_1 * R_2 \in \mathcal{T}(\mathcal{S})$ $S := S, \langle x_1, \ldots, x_m, y_1, \ldots, y_n \rangle \in R_1 * R_2$

> **PROD DOWN** $\langle x_1,\ldots,x_m,y_1,\ldots,y_n\rangle \in R_1 * R_2 \in \mathcal{S}^*$ $\text{ar}(R_1) = m$ $S := S, \langle x_1, \ldots, x_m \rangle \in R_1, \langle y_1, \ldots, y_n \rangle \in R_2$

TRANSITIVE CLOSURE Derivation Rule (_⁺)

$$
\begin{array}{ll}\n\text{TCLos UP I} & \text{TCLos UP II} \\
\langle x_1, x_2 \rangle \equiv R \in \mathcal{S}^* & R^+ \in \mathcal{T}(\mathcal{S}) & \langle x_1, x_2 \rangle \equiv R^+, \langle x_2, x_3 \rangle \equiv R^+ \in \mathcal{S}^* \\
\hline\n\mathcal{S} := \mathcal{S}, \langle x_1, x_2 \rangle \equiv R^+ & \mathcal{S} := \mathcal{S}, \langle x_1, x_3 \rangle \equiv R^+ \\
\text{TCLOS Down} & \langle x_1, x_2 \rangle \equiv R^+ \in \mathcal{S}^* \\
\hline\n\mathcal{S} := \mathcal{S}, \langle x_1, x_2 \rangle \equiv R & || & \mathcal{S} := \mathcal{S}, \langle x_1, z \rangle \equiv R, \langle z, x_2 \rangle \equiv R\n\end{array}
$$

 $\parallel \mathcal{S} := \mathcal{S}, \langle x_1, z_1 \rangle \in R, \langle z_1, z_2 \rangle \in R^+, \langle z_2, x_2 \rangle \in R, z_1 \not\approx z_2$

 z, z_1, z_2 are fresh variables

An Example

 $S = \{ \langle a, b \rangle \in R^+, \langle a, b \rangle \notin R, \langle a, b \rangle \notin R \bowtie R \}$ $S := S \cup \{ \langle a, b \rangle \in \mathbb{R} \} \cup \{ \langle a, k \rangle \in \mathbb{R}, \langle k, b \rangle \in \mathbb{R} \}$ **JOINEUR UNSAT** $(a, b) \notin R \Join R$ **EQ UNSAT** $\mathsf{EQ} \mathsf{UNSAT} \otimes \mathsf{GL} \otimes \mathsf{B}$ **TCLOS DOWN** $\mathcal{S} := \mathcal{S} \cup \{ \langle a, b \rangle \in \mathbb{R} \text{ and } R$ ules apply **UNSAT** $\mathcal{S} \rightarrow \mathcal{S} \rightarrow \mathcal{S} \rightarrow \mathcal{N} \rightarrow \mathcal{$ (After exhaustively applying JOIN-UP) **SAT**

Calculus C_R Correctness

Calculus C_R Correctness

Refutation Sound – a closed derivation tree proves that input constraints are **UNSAT**

Model Sound – from a saturated branch of a derivation tree one can extract a model for input constraints

Detailed proof can be found in Meng et al. [21]

Termination for a Fragment of $\mathcal{T}_{\mathcal{R}}$

(element) (unary relation) (binary relation) (constraint)

$$
e := x
$$

\n
$$
u := x | [] | u_1 \sqcup u_2 | u_1 \sqcap u_2 | [\langle e \rangle] | b \bowtie u
$$

\n
$$
b := x | [] | b_1 \sqcup b_2 | b_1 \sqcap b_2 | [\langle e_1, e_2 \rangle] | b^{-1}
$$

\n
$$
\varphi := e_1 \approx e_2 | \langle e \rangle \sqsubseteq u | \langle e_1, e_2 \rangle \sqsubseteq b | \neg \varphi_1
$$

Termination: If S is a finite set of constraints generated by the grammar above, then all derivation trees are finite.

Detailed proof can be found in Meng et al. [21]

A Relational Solver in CVC4

A Relational Solver in CVC4

- Allows us to solve constraints from a combination of relations and other domains
- Extend SMT-LIB/CVC4 native language with support for relations
- Enables natural mappings from several relational modeling languages to SMT
- Brings to those languages the power of SMT solvers and their ability to reason efficiently about built-in types

Applications of $\mathcal{T}_{\mathcal{R}}$

Application 1: Alloy to CVC4

Support Alloy kernel language in **SMT** natively

Finite model finding of CVC4 can efficiently reason about problems with presence of quantifiers

Built a **translator** from Alloy kernel language to SMT

Can **disprove and prove properties** with respect to Alloy specifications

Evaluation on Alloy Benchmarks

Evaluated CVC4 with two configurations

- **CVC4**: enables full native support for relational operators
- **CVC4+AX:** encodes all relational operators as uninterpreted functions with axioms

Compared with **Alloy Analyzer** and **AlloyPE** on two sets of benchmarks:

- 1. From AlloyPE and
- 2. From an academic course

Evaluation on Alloy Benchmarks

Compared to the **Alloy Analyzer**

- **CVC4** is **overall slower for SAT benchmarks**
- **CVC4 solves UNSAT benchmarks**, whereas the Alloy Analyzer can only answer bounded UNSAT

Compared to **AlloyPE**

- **CVC4 solves SAT benchmarks**, whereas AlloyPE solves none
- **CVC4 solves most of AlloyPE's benchmarks**

Compared to **CVC4+AX**

- **CVC4 solves SAT benchmarks**, whereas CVC4+AX solves none
- **CVC4 solves significantly more UNSAT benchmarks**

Evaluation on SAT Benchmarks

Evaluation on UNSAT Benchmarks

Application 2: OWL DL to SMT

OWL DL based on an **expressive** description logic fragment

Built a translator from OWL DL to SMT in $T_{\mathcal{R}}$

Check **logical consistency** of OWL models using CVC4

Evaluation on OWL Benchmarks

Evaluated on **OWL models** from 4th OWL Reasoner Evaluation competition

Compared with a state of the art **DL reasoner HermiT**

For the ones (4269) we both solved:

- **CVC4** takes **2.62s** per benchmark and solves faster on **1617** benchmarks
- **HermiT** takes **1.76s** per benchmark and solves faster on **2652** benchmarks

Comparison with HermiT

X-Values: CVC4 Y-Values: HermiT (in Log Scale)

Summary

- Introduced **a theory of finite relations in SMT**
- Developed a refutation-sound and model-sound **calculus** for the theory of relations
- Demonstrated useful **applications** in Alloy and OWL
- Shown **promising experimental results** on Alloy and OWL benchmarks

Thank you!

1. F. Baader. The description logic handbook: Theory, implementation and applications. Cambridge university press, 2003.

2. F. Baader, I. Horrocks, and U. Sattler. Description logics. In V. L. Frank van Harmelen and B. Porter, editors, Handbook of Knowledge Representation, volume 3 of Foundations of Artificial Intelligence, pages $135 - 179$. Elsevier, 2008.

3. K. Bansal, A. Reynolds, C. W. Barrett, and C. Tinelli. A new decision procedure for finite sets and cardinality constraints in SMT. In Proceedings of IJCAR'16, volume 9706 of LNCS, pages 82–98. Springer, 2016.

4. C. Barrett, C. L. Conway, M. Deters, L. Hadarean, D. Jovanovi'c, T. King, A. Reynolds, and C. Tinelli. CVC4. In Proceedings of CAV'11, volume 6806 of LNCS, pages 171–177. Springer, 2011.

5. C. Barrett, P. Fontaine, and C. Tinelli. The SMT-LIB standard—Version 2.6. In A. Gupta and D. Kroening, editors, SMT 2010, 2010.

6. C. Barrett, R. Sebastiani, S. Seshia, and C. Tinelli. Satisfiability modulo theories. In A. Biere, M. J. H. Heule, H. van Maaren, and T. Walsh, editors, Handbook of Satisfiability, volume 185, chapter 26, pages 825–885. IOS Press, February 2009.

7. B. Dutertre and L. D. Moura. The YICES SMT solver. Technical report, SRI International, 2006.

8. A. A. E. Ghazi and M. Taghdiri. Analyzing alloy constraints using an SMT solver: a case study. In 5th International Workshop on Automated Formal Methods (AFM), 2010.

9. A. A. E. Ghazi and M. Taghdiri. Relational reasoning via SMT solving. In Proceedings of FM'11, volume 6664 of LNCS, pages 133–148. Springer, 2011.

10. A. A. E. Ghazi, M. Taghdiri, and M. Herda. First-order transitive closure axiomatization via iterative invariant injections. In Proceedings of NFM'15, volume 9058 of LNCS. Springer, 2015.

11. I. Horrocks and U. Sattler. Decidability of shiq with complex role inclusion axioms. Artificial Intelligence, 160(1-2):79-104, 2004.

12. D. Jackson. Alloy: a lightweight object modelling notation. ACM Trans. Softw. Eng. Methodol., 11(2):256–290, 2002.

13. D. Jackson. Software Abstractions - Logic, Language, and Analysis. MIT Press, 2006.

14. R. Nieuwenhuis, A. Oliveras, and C. Tinelli. Solving SAT and SAT Modulo Theories: from an abstract Davis-Putnam-Logemann-Loveland Procedure to DPLL(T). Journal of the ACM, 53(6):937-977, Nov. 2006.

15. A. Reynolds, C. Tinelli, A. Goel, and S. Krstic. Finite model finding in SMT. In Proceedings of CAV'13, volume 8044 of LNCS, pages 640–655. Springer, 2013.

16. A. Steigmiller, T. Liebig, and B. Glimm. Konclude: System description. Web Semantics: Science, Services and Agents on the World Wide Web, 27(1), 2014.

17. E. Torlak and D. Jackson. Kodkod: a relational model finder. In Proceedings of TACAS'07, volume 4424 of LNCS, pages 632–647. Springer, 2007.

18. D. Tsarkov and I. Horrocks. Fact++ description logic reasoner: system description. In Proceedings of IJCAR'06, volume 4130 of LNCS. Springer, 2006.

19. D. Tsarkov and I. Palmisano. Chainsaw: a metareasoner for large ontologies. In I. Horrocks, M. Yatskevich, and E. Jim'enez-Ruiz, editors, ORE, 2012.

20. W3C. OWL 2 web ontology language, https://www.w3.org/2007/OWL/wiki/Syntax.

21. Baoluo Meng, Andrew Reynolds, Cesare Tinelli, and Clark Barrett. Relational Constraint Solving in SMT. In Proceedings of the 26th International Conference on Automated Deduction, Gothenburg, Sweden

A Toy File System Specification in Alloy

```
abstract sig FSO {}
sig File extends FSO {}
sig Dir	extends FSO {
   contents:	Set FSO
}
-- contents relation is acyclic
fact {all d:	Dir |	not (d	in d.^contents)}
-- Every file system object only has one location
assert oneLocation {
 all o	:	FSO |	lone d	:	FSO	|	o	in d.contents
}
```
check oneLocation for 7

An Example

$$
S = \{(a, b) \notin \mathbb{R}^{-1}, \mathbb{R} \approx \mathbb{Q}, \boxed{(a) \in \mathbb{P}, \langle b \rangle \in \mathbb{P}, \mathbb{P} * \mathbb{P}} \approx \mathbb{Q} \sqcap \mathbb{T} \}
$$
\n
$$
S := S \cup \{ \boxed{(a, b) \in \mathbb{P} * \mathbb{P}, \langle b, a \rangle \in \mathbb{P} * \mathbb{P}, \langle a, a \rangle \in \mathbb{P} * \mathbb{P}, \dots \}}
$$
\n
$$
S := S \cup \{ \langle a, b \rangle \in \mathbb{Q}, \boxed{(b, a) \in \mathbb{Q}, \langle a, a \rangle \in \mathbb{Q}, \langle b, b \rangle \in \mathbb{Q}, \dots \}
$$
\n
$$
\langle a, b \rangle \notin \mathbb{R}^{-1}, \mathbb{R} \approx \mathbb{Q} \text{ TRANS UP}
$$
\n
$$
S := S \cup \{ \boxed{(a, b) \in \mathbb{R}^{-1}, \dots \}} \xrightarrow{\text{EQ UNSAT}}
$$
\n
$$
\{ \text{INIST} \text{ INSAT} \}
$$

Satisfiability Modulo Theories (SMT)

- A **theory** $\mathcal{T} = (\Sigma, I)$ defines
	- A **signature** Σ: a set of non-logical symbols
	- A class of Σ-interpretations I
	- **Examples**: integer arithmetic, strings, finite sets, ...
		- \triangleright A simple **theory**: $\Sigma_s = \{0, 1, +, \equiv$
		- \triangleright A **formula** in the theory \mathcal{T}_s : $x + 0 = 1$