

# 1 Próximos passos

Para podermos melhorar nosso campo de atuação, é importante sabermos usar as ferramentas de que dispomos para estender outras de modo adequado, abrindo espaço para novas aplicações. Como conclusão deste texto, buscaremos mostrar isso com um breve exemplo, de extensão de *solvers* para SMT por CLP.

Antes, convém conhecer o problema *SAT*<sup>\*</sup>. SAT é o problema de determinar se existe uma interpretação (atribuição de valores *booleanos* a variáveis) que satisfaça uma fórmula booleana. Foi o primeiro problema a ser provado NP-completo[8] mas, apesar disso, tem se mostrado de grande utilidade para diversos problemas práticos. Como é de se esperar da formulação, SAT é um problema central em ciência da computação. Assim, tem recebido muita atenção e conta com implementações de *solvers* eficientes para grandes classes de sub-problemas mais específicos.

Uma instância de SAT é o de obter valores booleanos para  $A, B$  e  $C$  tal que a fórmula  $A \wedge B \vee (\neg C \vee (A \wedge \neg B))$  seja satisfeita.

*SMT*<sup>†</sup> é uma generalização do problema SAT para lidar com outras teorias e de forma mais geral. Exemplos de teorias abrangidas por SMT são sobre inteiros, listas e vetores<sup>‡</sup>. Uma instância de problema SMT é como uma de problema SAT, exceto que algumas (ou todas) as variáveis booleanas são substituídas por predicados sob variáveis não-binárias, como em  $(3 + X \leq 2) \wedge (4Y - \pi \geq X)$ .

Semelhante ao caso já visto com problemas em CLP, um *solver* para SMT terá, no caso geral, algoritmos especializados para lidar com teorias diferentes (lembre-se que uma das vantagens de CLP é facilitar essa unificação de diferentes métodos). Devido em grande parte à sua generalidade e à eficiência desenvolvida (depois de anos de pesquisa), o uso de *solvers* para SMT tem ganho grande popularidade para lidar com problemas de diversas áreas, como de análise de programas[12], síntese de programas[2], verificação de hardware[5], automação de design eletrônico[5], segurança de computadores[10], IA, pesquisa operacional[4] e biologia[11].

Existem atualmente vários *solvers* para SMT, dos quais os mais populares são Z3 da Microsoft e CVC4<sup>§</sup>, mantido por um grupo de pesquisadores (em grande parte, da Universidade de Stanford). Neste texto, focaremos no Z3<sup>¶</sup>, um *solver* considerado estado-da-arte, que se tornou um projeto de código aberto em 2012 (veja [13] para detalhes).

Para que se tenha alguma ideia de como se usa um *solver* para SMT, segue um exemplo (retirado de [15]) de código para Z3<sup>\*</sup>:

---

<sup>\*</sup>SAT vem de *satisfiability*.

<sup>†</sup>SMT vem de *satisfiability modulo theories*.

<sup>‡</sup>Existem várias outras, mas as teorias específicas utilizadas variam de *solver*.

<sup>§</sup>De *cooperating validity checking*.

<sup>¶</sup>Explicaremos apenas o básico para que seja compreensível. Para mais informações, cheque, por exemplo, [14].

<sup>\*</sup>Vale notar, ele está expresso em *Sezps*, mas não em Scheme.

```

; declare a mutually recursive parametric datatype
(declare-datatypes (T) ((Tree leaf (node (value T)
                                           (children
                                            TreeList))
                                           )
                        (TreeList nil (cons (car Tree)
                                             (cdr TreeList)
                                             )))))

(declare-const t1 (Tree Int))
(declare-const t2 (Tree Bool))
; we must use the 'as' construct to distinguish the
; leaf (Tree Int) from leaf (Tree Bool)
(assert (not (= t1 (as leaf (Tree Int)))))
(assert (> (value t1) 20))
(assert (not (is-leaf t2)))
(assert (not (value t2)))
(check-sat)
(get-model)

```

que avalia para:

```

sat
(model
  (define-fun t2 () (Tree Bool)
    (node false (as nil (TreeList Bool))))
  (define-fun t1 () (Tree Int)
    (node 21 (as nil (TreeList Int)))))

```

O *sat* é de *satisfiable*, a resposta ao (`check-sat`) (seria *unsat*, se o modelo posto não fosse satisfazível). SMT é um problema mais geral do que SAT (note que SAT é um caso particular de SMT), que é NP-completo. Assim, como é de se esperar, existem vários problemas para os quais ele não funciona muito bem. Então, parte do desafio ao (buscar) resolver um problema SMT está em expressá-lo de forma que fique na “parte fácil” para o *solver*.

## 1.1 Breve introdução a SMT e sobre o problema da estratégia

*Solvers* para SMT se baseiam fortemente em *solvers* para SAT, baseados em DPLL<sup>†</sup>. Para SMT, temos o DPLL(T), que é um formalismo para descrever como *solvers* da teoria T devem ser integrados com os *solvers* SAT.

---

<sup>†</sup>O algoritmo de Davis–Putnam–Logemann–Loveland é um algoritmo de busca completo, baseado em *backtracking*, para decidir a se dadas fórmulas de lógica proposicional em forma normal conjuntiva podem ser satisfeitas, introduzido em 1962, que ainda forma a base de *solvers* eficientes para SAT.

Uma dificuldade é que, para um *solver* para SMT de alta performance, parte importante da implementação não está no esquema formal do DPLL(T), mas sim em heurísticas. Essas heurísticas são frequentemente projetadas para funcionar muito bem para algumas classe de problemas, tendendo a funcionar mal para outras. À medida que *solvers* para SMT ganham a atenção de cientistas e engenheiros, passou a se tornar claro que isso é um problema, porque há uma necessidade de grande controle sobre o *solver* para que ele se comporte de forma eficiente, o que significa expor até centenas de parâmetros para que usuários decidam quais heurísticas usar e como. Esse é um problema posto por Leonardo de Moura *et. al*, um criador do Z3, no artigo [3]. Nesse artigo, ele define o problema da estratégia como o de prover ao usuário meios adequados de dirigir a busca (já que, de uma forma ou de outra, a resolução de uma instância do problema SMT envolve busca). Mais em geral, ele define estratégia (neste contexto) como “*adaptations of general search mechanisms which reduce the search space by tailoring its exploration to a particular class of problems*” [3] e, assim, o problema da estratégia se traduziria como o de prover uma linguagem adequada que o usuário possa usar para realizar sua busca.

Existe mais de uma forma de buscar resolver esse problema. A abordagem que tomaremos (de forma resumida) aqui é a adotada por Nada Amin e William Byrd, de juntar SMT com CLP, em um CLP(SMT), usando *miniKanren* como base\*. Para checar esse trabalho, veja [6] ou [7] (esse último, na linguagem Clojure). Vale notar que, no momento da escrita deste texto (em outubro de 2018), isto é trabalho em progresso.

A ideia é pensar no *solver* para SMT como um implementador de restrições de baixo nível, com o qual o usuário interage com o *miniKanren*, de forma mais abstrata (assim como fizemos com outras técnicas para otimização e para busca de satisfação de restrições). Isso pode ser útil de várias formas. Primeiro, nota-se que *solvers* para SMT não lidam bem com recursões em geral, o que não é problema para linguagens de programação lógica ou de programação por restrições. Assim, por exemplo, se o seguinte for submetido ao Z3, depois de algum tempo, ele retorna (no momento da escrita deste texto) *unknown*:

```
(declare-fun fact (Int) Int)
(assert (= (fact 0) 1))
(assert (forall ((n Int))
  (=> (> n 0) (= (fact n) (* n (fact (- n 1)))))))
(declare-const r6 Int)
(check-sat)
```

Mas, usando Z3 a partir do *miniKanren*<sup>†</sup> (com o software em [6], por exemplo), podemos fazer o seguinte:

---

\*Vale notar, o que implementamos no capítulo passado não lida com restrições no sentido usual, mas ele pode ser facilmente estendido com relações de CLP. Veja [1].

<sup>†</sup>O que, já que as entradas e saídas são em *Sexprs*, não é difícil.

```

(define faco
  (lambda (n out)
    (conde ((z/assert '(= ,n 0))
            (z/assert '(= ,out 1)))
          ((z/assert '(> ,n 0))
            (fresh (n-1 r)
              (z/assert '(= (- ,n 1) ,n-1))
              (z/assert '(= (* ,n ,r) ,out))
              (faco n-1 r))))))

(test "faco-7"
  (run 7 (q)
    (fresh (n out)
      (faco n out)
      (== q '(,n ,out))))
  '((0 1) (1 1) (2 2) (3 6) (4 24) (5 120) (6 720)))

```

em que  $(\text{cond}^e (a_1 \dots a_i) \dots (f_1 \dots f_j))$  pode ser lido logicamente como  $(a_1 \wedge \dots \wedge a_i) \vee \dots \vee (f_1 \dots f_j)$  e **z/assert** faz a interface com o *solver*, tomando uma expressão booleana com variáveis (por padrão, inteiras). Além de **z/assert**, existem **z/**, **z/check** e **z/purge**, para interagir com o Z3 de forma diferente.

Ainda não é certo, para qual ou quais tipos de problemas esse tipo de abordagem é apropriado, ou mesmo se para algum, mas é um exemplo de tentativa de uso da ideia de CLP para algo diferente, de forma não óbvia.

## 1.2 Conclusão

Se você prestou atenção no caminho até aqui, deve ter adquirido algum conhecimento básico sobre programação por restrições e, assim esperamos, quando e se precisar de fazer uso de algumas das ideias desenvolvidas, será capaz de avaliar como fazê-lo. Mais importante ainda, esperamos que seja capaz de avaliar quando e como pode “ser preciso” fazer uso das ideias aqui desenvolvidas.

## Referências

- [1] Alvis, Claire E., Jeremiah J. Willcock, Kyle M. Carter, William E. Byrd, Daniel P. Friedman. “cKanren miniKanren with constraints.” (2011).
- [2] Beyene, Tewodros A., Swarat Chaudhuri, Corneliu Popeea, e Andrey Rybalchenko “Recursive games for compositional program synthesis.” Working Conference on Verified Software: Theories, Tools, and Experiments, pp. 19-39. Springer, Cham, 2015.
- [3] De Moura, Leonardo, and Grant Olney Passmore. “The strategy challenge in SMT solving.” Automated Reasoning and Mathematics. Springer, Berlin, Heidelberg, 2013. 15-44.
- [4] Li, Yi, Aws Albarghouthi, Zachary Kincaid, Arie Gurfinkel, Marsha Chechik. “Symbolic optimization with SMT solvers.” In ACM SIGPLAN Notices, vol. 49, no. 1, pp. 607-618. ACM, 2014.
- [5] Mukherjee, Rajdeep, Daniel Kroening, and Tom Melham “Hardware verification using software analyzers.” In VLSI (ISVLSI), 2015 IEEE Computer Society Annual Symposium on, pp. 7-12. IEEE, 2015.
- [6] CLP(SMT) miniKanren: <https://github.com/namin/clpsmt-miniKanren>
- [7] Explorations in logic programming: <https://github.com/namin/logically>
- [8] Schaefer, Thomas J. “The complexity of satisfiability problems” In: Proceedings of the tenth annual ACM symposium on Theory of computing. ACM, 1978. p. 216-226.
- [9] Trindade, Alessandro B., Lucas C. Cordeiro. “Applying SMT-based verification to hardware/software partitioning in embedded systems.” Design Automation for Embedded Systems 20, no. 1 (2016): 1-19.
- [10] Vanegue, Julien, Sean Heelan, Rolf Rolles. “SMT Solvers in Software Security.” WOOT 12 (2012): 9-22.
- [11] Yordanov, Boyan, Christoph M. Wintersteiger, Youssef Hamadi, Hillel Kugler. “Z34Bio: An SMT-based framework for analyzing biological computation.”. SMT’13 (2013).
- [12] Zheng, Yunhui, Xiangyu Zhang, e Vijay Ganesh. “Z3-str: a z3-based string solver for web application analysis.”, Proceedings of the 2013 9th Joint Meeting on Foundations of Software Engineering, pp. 114-124. ACM, 2013.
- [13] Z3 se torna *open source*: <http://leodemoura.github.io/blog/2012/10/02/open-z3.html>

- [14] Z3 wiki: <https://github.com/Z3Prover/z3/wiki>
- [15] Z3 Tutorial: <https://rise4fun.com/z3/tutorial>