A Theoretical Limit for Safety Verification Techniques with Regular Fix-point Computations

Yohan Boichut, Pierre-Cyrille Heam

INRIA

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

A Theoretical Limit for Safety Verification Techniques with Regular Fix-point Computations

Yohan Boichut, PAREO — Pierre-Cyrille Héam, CASSIS

N° 6411

Janvier 2008

Thème SYM

ISSN 0249-6399 ISRN INRIA/RR--6411--FR+ENG

A Theoretical Limit for Safety Verification Techniques with Regular Fix-point Computations

Yohan Boichut, PAREO *, Pierre-Cyrille Héam, CASSIS †

Thème SYM — Systèmes symboliques
Équipes-Projets CASSIS et PAREO

Rapport de recherche n° 6411 — Janvier 2008 — 6 pages

Abstract: In computer aided verification, the reachability problem is particularly relevant for safety analyses. Given a regular tree language L, a term t and a relation R, the reachability problem consists in deciding whether a sequence of terms, beginning with a term of L and terminating on t and such that two successive terms of this sequence are in relation according to R, is constructable. In this case, the term t is said to be reachable, otherwise it is said unreachable. This problem is decidable for particular kinds of relations, but it is known to be undecidable in general, even if L is finite. Several approaches to tackle the unreachability problem are based on the computation of an R-closed regular language containing L. In this paper we show a theoretical limit to this kind of approaches for this problem.

Key-words: Reachability problem, regular tree languages, undecidable, theoretical limit.

* Laboratoire Lorrain de Recherche en Informatique et ses Applications
† Laboratoire Informatique de l'université de Franche-Comté

Centre de recherche INRIA Nancy – Grand Est
LORIA, Technopôle de Nancy-Brabois, Campus scientifique,
615, rue du Jardin Botanique, BP 101, 54602 Villers-Lès-Nancy
Téléphone : +33 3 83 59 30 00 — Télécopie : +33 3 83 27 83 19

Une limite théorique pour les techniques de vérification point-fixe

Résumé : Le problème d'atteignabilité est un point central pour les techniques dites de vérification. Etant donné un langage d'arbre L, un terme t et une relation R, le problème d'atteignabilité se résume à la question suivante : peut-on construire une séquence de termes partant d'un terme de L et terminant sur t telle que, deux éléments successifs de cette séquence soient en relation par R ? Ce problème est, dans le cas général, indécidable, même si L est un ensemble fini. Si en effet il existe une telle séquence, alors le terme est atteignable. Dans le cas contraire, il est dit inatteignable. Plusieurs approches proposent de calculer un langage régulier contenant L et clos par R dans le but de démontrer qu'un terme est inatteignable. Dans ce papier, nous montrons que ces méthodes sont limitées théoriquement pour cette problématique.

Mots-clés : Langages réguliers d'arbres, atteignabilité, indécidable, limites théoriques.

We assume that the reader is familiar with basic notions and notations on terms and on bottom-up tree automata. For a general reference see [4].

1 Introduction

In this paper we show a theoretical limit of regular fix-point techniques used for reachability analyses.

Automatic verification of software systems is one of the most challenging research problems in computer aided verification. In this context, regular model-checking has been proposed as a general framework for analysing and verifying infinite state systems. In this framework, systems are modelled using regular representations: configurations of the systems are modelled by finite words or trees (of unbounded size) and the dynamic of the systems is modelled by a relation $\mathcal{R}$ (in practice a transducer or a (term) rewriting system). Then, safety analysis of the system is reduced to the computation of regular languages closed under a relation $\mathcal{R}$: given a regular language L, a relation $\mathcal{R}$ and a regular set L_P of *bad configurations*, the question is to decide whether $R^\star(L) \cap L_p = \emptyset$ where $\mathcal{R}^\star$ is the reflexive transitive closure of $\mathcal{R}$. Since $\mathcal{R}^\star(L)$ is in general neither regular nor computable, several approaches handle restricted cases for this problem [6, 5, 10, 13].

However, modelling real systems leads in general out of decidable cases. In this context, several regular fix-point automatic [3] or human guided techniques [11, 8, 7] were developed in order to prove safety properties. The goal of these techniques is to compute a regular language K_{over} containing L and which is $\mathcal{R}$-closed. The language K_{over} is an over approximation of $\mathcal{R}^\star(L)$ (for language inclusion) and if $K_{\text{over}} \cap L_p = \emptyset$, then $\mathcal{R}^\star(L) \cap L_p = \emptyset$. This approach has been successfully used in order to prove safety of security protocols [9, 15, 14, 12, 2] or recently for static analysis of JAVA programs [1].

In this direction we cannot get away from the question to know whether this kind of fix-point approaches can always be used to prove safety of systems in the following sense: given the model of a system by a regular language L and a relation $\mathcal{R}$, for any language L_p such that $\mathcal{R}^\star(L) \cap L_p = \emptyset$, does there exist an $\mathcal{R}$-closed regular language K_{over} containing L and satisfying $K_{\text{over}} \cap L_p = \emptyset$? This issue can also be formalised as follows: does the following equality hold

$$L = \bigcap_{\mathcal{R}^\star(L) \subseteq K,\ \mathcal{R}(K) \subseteq K} K,$$

where the intersection is restricted to regular languages?

In this paper we give a negative answer to this question.

2 Main result

Proposition 1 *Let* $L = \{f(A, A)\}$, $\mathcal{R} = \{f(x, y) \to f(h(x), h(y)),\ f(h(x), h(y)) \to f(x, y),\ f(h(x), A) \to A, f(A, h(x)) \to A\}$ *where* x *and* y *are variables. One has* $A \notin R^\star(L)$ *but*

$$A \in \bigcap_{L \subseteq K,\ R(K) \subseteq K} K.$$

PROOF. Let $H = \{f(h^k(A), h^k(A)) \mid k \in \mathbb{N}\}$. First we claim that $\mathcal{R}^\star(L) = H$. Starting from $f(A, A)$ and using the rule $f(x, y) \to f(h(x), h(y))$, one has $L \subseteq H \subseteq \mathcal{R}^\star(L)$. Moreover, H is obviously closed by the rule $f(h(x), h(y)) \to f(x, y)$. Therefore, since the two rules $f(h(x), A) \to A$ and $f(A, h(x)) \to A$ cannot be applied to terms in H, it follows that $\mathcal{R}^\star(L) = H$, proving the claim. Furthermore, $A \notin R^*(L(\mathcal{A}))$ Moreover one can easily prove that $\mathcal{R}^\star(L)$ is not regular using classical pumping arguments.

Secondly, let K_{over} be a regular language such that $L(\mathcal{A}) \subseteq K_{\text{over}}$ and $R(K_{\text{over}}) \subseteq K_{\text{over}}$. Let also S be the regular language $\{f(h^k(A), h^\ell(A)) \mid k \geq 0, \ell \geq 0\}$. Since $\mathcal{R}^\star(L) \subseteq K_{\text{over}}$, $\mathcal{R}^\star(L) \cap S \subseteq K_{\text{over}} \cap S$. Using the claim, one has $\mathcal{R}^\star(L) \cap S = \mathcal{R}^\star(L)$. Consequently $\mathcal{R}^\star(L) \subseteq K_{\text{over}} \cap S$. Now, it is well known that the intersection of two regular tree languages is regular too. Thus $K_{\text{over}} \cap S$ is regular. However $\mathcal{R}^\star(L)$ is not regular. Consequently, the inclusion $\mathcal{R}^\star(L) \subset K_{\text{over}} \cap S$ is strict. So let t be an element of $K_{\text{over}} \cap S \setminus \mathcal{R}^\star(L)$. The term t is of the form $t = f(h^k(A), h^\ell(A))$ with $k \neq \ell$. Without lost of generality, we may assume that $k > \ell$. Since K_{over} is $\mathcal{R}$-closed and using the rule $f(h(x), h(y)) \to f(x, y)$, the term $f(h^{k-\ell}(A), A)$ is in K_{over}. Now the rule $f(h(x), A) \to A$ can be applied on $f(h^{k-\ell}(A), A)$. Consequently, K_{over} being $\mathcal{R}$-closed, it follows that $A \in K$, which concludes the proof. □

Thus, we have shown that A will be in every over-approximation computed by a regular fix-point technique. So, we won't be able to show it unreachable.

3 Conclusion

Undoubtedly, regular fix-point techniques mentioned previously have led to great results as seen in introduction. Nevertheless, they are disarmed against the problem illustrated in Proposition 1. This raises several open questions:

- Can we decide whether
$$\mathcal{R}^\star(L) = \bigcap_{L \subseteq K,\ \mathcal{R}(K) \subseteq K, K \text{ regular}} K?$$
- If the answer is no, does there exist decidable conditions on L and $\mathcal{R}$ such that the above equality holds?
- How regular fix-point approaches may be extended in order to handle more cases?

References

[1] Y. Boichut, T. Genet, T. Jensen, and L. Le Roux. Rewriting Approximations for Fast Prototyping of Static Analyzers. In *Proc. 18th RTA Conf., Paris (France)*, volume 4533 of *Lecture Notes in Computer Science*, pages 48–62, 2007.

[2] Y. Boichut, P.-C. Héam, and O. Kouchnarenko. Handling algebraic properties in automatic analysis of security protocols. In *Int. Col. on Theorical Aspects of Computing , ICTAC-06*, volume 4281 of *Lecture Notes in Computer Science*, pages 153–167. Springer Berlin/Heidelberg, 2006.

[3] A. Bouajjani, P. Habermehl, A. Rogalewicz, and T. Vojnar. Abstract regular tree model checking. In *Proceedings of 7th International Workshop on Verification of Infinite-State Systems – INFINITY 2005*, number 4 in BRICS Notes Series, pages 15–24, 2005.

[4] H. Comon, M. Dauchet, R. Gilleron, F. Jacquemard, D. Lugiez, S. Tison, and M. Tommasi. Tree automata techniques and applications. `http://www.grappa.univ-lille3.fr/tata/`, 2002.

[5] J.-L. Coquidé, M. Dauchet, R. Gilleron, and V. S. Bottom-up tree pushdown automata and rewrite systems. In R. V. Book, editor, *Rewriting Techniques and Applications, 4th International Conference, RTA-91*, LNCS 488, pages 287–298, Como, Italy, Apr. 10–12, 1991. Springer-Verlag.

[6] Dauchet and Tison. The theory of ground rewrite systems is decidable. In *LICS: IEEE Symposium on Logic in Computer Science*, 1990.

[7] G. Feuillade, T. Genet, and V. Viet Triem Tong. Reachability Analysis over Term Rewriting Systems. *JAR*, 33 (3-4):341–383, 2004.

[8] T. Genet. Decidable approximations of sets of descendants and sets of normal forms. In *Proc. 9th RTA Conf., Tsukuba (Japan)*, volume 1379 of *LNCS*, pages 151–165. Springer-Verlag, 1998.

[9] T. Genet and F. Klay. Rewriting for Cryptographic Protocol Verification. In *In Proc. CADE'2000*, volume 1831 of *LNAI*. Springer-Verlag, 2000.

[10] R. Gilleron and S. Tison. Regular tree languages and rewrite systems. *Fundam. Inform*, 24(1/2):157–174, 1995.

[11] F. Jacquemard. Decidable approximations of term rewriting systems. In H. Ganzinger, editor, *Proc. 7th RTA Conf., New Brunswick (New Jersey, USA)*, pages 362–376. Springer-Verlag, 1996.

[12] M. Nesi and G. Rucci. Formalizing and Analyzing the Needham-Schroeder Symmetric-Key Protocol by Rewriting. In *In Proceedings of the 2nd Workshop on Automated Reasoning for Security Protocol Analysis*, 2005.

[13] P. Réty and J. Vuotto. Regular sets of descendants by leftmost strategy. *Electr. Notes Theor. Comput. Sci*, 70(6), 2002.

[14] T. Takai. A verification technique using term rewriting systems and abstract interpretation. In V. van Oostrom, editor, *Rewriting Techniques and Applications, 15th International Conference, RTA-04*, Lecture Notes in Computer Science 3091, pages 119–133, Valencia, Spain, June 3-5, 2004. Springer.

[15] T. Takai, Y. Kaji, and H. Seki. Right-linear finite-path overlapping term rewriting systems effectively preserve recognizability. In *Proc. 11th RTA Conf., Norwich (UK)*, volume 1833 of *LNCS*. Springer-Verlag, 2000.

Contents

Centre de recherche INRIA Nancy – Grand Est
LORIA, Technopôle de Nancy-Brabois - Campus scientifique
615, rue du Jardin Botanique - BP 101 - 54602 Villers-lès-Nancy Cedex (France)

Centre de recherche INRIA Bordeaux – Sud Ouest : Domaine Universitaire - 351, cours de la Libération - 33405 Talence Cedex
Centre de recherche INRIA Grenoble – Rhône-Alpes : 655, avenue de l'Europe - 38334 Montbonnot Saint-Ismier
Centre de recherche INRIA Lille – Nord Europe : Parc Scientifi que de la Haute Borne - 40, avenue Halley - 59650 Villeneuve d'Ascq
Centre de recherche INRIA Paris – Rocquencourt : Domaine de Voluceau - Rocquencourt - BP 105 - 78153 Le Chesnay Cedex
Centre de recherche INRIA Rennes – Bretagne Atlantique : IRISA, Campus universitaire de Beaulieu - 35042 Rennes Cedex
Centre de recherche INRIA Saclay – Île-de-France : Parc Orsay Université - ZAC des Vignes : 4, rue Jacques Monod - 91893 Orsay Cedex
Centre de recherche INRIA Sophia Antipolis – Méditerranée : 2004, route des Lucioles - BP 93 - 06902 Sophia Antipolis Cedex

Éditeur
INRIA - Domaine de Voluceau - Rocquencourt, BP 105 - 78153 Le Chesnay Cedex (France)
http://www.inria.fr
ISSN 0249-6399