Sujet de thèse - Campagne EDITE 2023 :

OCRA : One Clausebase to Rule them All

Directeur : Souheib Baarir (LRE / EPITA)

Encadrants : Julien Sopena (LIP6 / SU) et Thibault Lejemble (LRE / EPITA)

Résumé

During this thesis, we will focus on the memory sharing between cores in parallel SAT-solving. We will design new concurrent data structures and algorithm to solve this dificult problems, allowing to scale to a hundred cores on a single machine.

Keywords — parallelism, Boolean satisfiability, memory sharing, data structure.

This thesis will deal with the Boolean satisfiability problem (SAT). Given a set of constraints expressed as a propositional formula, the SAT-solving consists in finding an assignment of the variables that satisfies all the constraints. A formula is expressed as a conjunction of clauses. Each clause is a disjunction of literals, which are either variables or negations of variables. This problem is extremely important in that it is one of the most famous NP-complete problems, but which also has many applications : software and hardware verification [2], automated theorem proving [3] or bioinformatics [7]. Widely discussed in [1], this problem has already been studied upon on from the parallelism perspective, using methods such as Divide and Conquer and Portfolio [4].

However, the current state of the art on the parallel SAT-solving problem deals with multicore, but does not scale on manycore machines, which we are focusing on. Indeed, all the knowledge learnt during the resolution of the problem is owned by each instance of the solver, and shared using a master, which makes it unusable in practice on manycore machines (it takes up to 500GB of memory for 36 cores, while we are focusing on machines with 80 to 120 cores).

In order to tackle this issue, we propose to study the problem of memory sharing between the cores used on the machine. Is it possible to share memory among cores in SAT-solving without losing performances ? This requires to design new algorithms, new concurrent data structures, and to have a strong understanding of the way the current algorithms such as CDCL (Conflict-Driven Clause Learning) [8] behaves. This PhD will especially focus on low-level programming, with a highlight on parallel memory management and cache management.

Preliminary studies show that the most important structure to work on is the learnt-clause database, which takes a lot of space, for which a copy is owned by each core, and that should be shared. To that extent, we intend to relate our work with the data structure studied upon in concurrent BDDs [9]. The data structures need to be readable in parallel, but also writable in parallel, with the smallest contention possible. To that extent we need to find some placing heuristics that allow parallel cache line access on the clauses.

This PhD aims at focusing on fundamental aspects of SAT-solving, with proved algorithms, and an implementation of our solution which will be tested during the SAT-solving competitions [5, 6].

References

[1] Armin Biere, Marijn Heule, and Hans van Maaren. Handbook of Satisfiability. Vol. 185. IOS press, 2009.

[2] Armin Biere et al. *"Symbolic model checking without BDDs"*. In : Proceedings of the 5th International Conference on Tools and Algorithms for the Construction and Analysis of Systems (TACAS). Springer, 1999, pp. 193-307.
[3] Marijn JH Heule, Oliver Kullmann, and Victor W Marek. *"Solving and verifying the boolean pythagorean triples problem via cube-and-conquer"*. In : Proceedings of the 19th International Conference on Theory and Applications of Satisfiability Testing (SAT). Springer, 2016, pp. 228-245.
[4] Ludovic Le Frioux et al. *"Painless : a framework for parallel SAT solving"*. In : Proceedings of the 20th International Conference on Theory and Applications of Satisfiability Testing (SAT). Springer, 2017, pp. 233-250.
[5] Ludovic Le Frioux et al. *"painless-maplecomsps"*. In : Proceedings of SAT Competition 2017 : Solver and Benchmark Descriptions. Department of Computer Science, University of Helsinki, Finland, 2017, pp. 26-27.
[6] Ludovic Le Frioux et al. *"painless-mcomsps and painless-mcomsps-sym"*. In : Proceedings of SAT Competition 2018 : Solver and Benchmark Descriptions. Department of Computer Science, University of Helsinki, Finland, 2018, pp. 33-34.
[7] Inês Lynce and Joao Marques-Silva. *"SAT in bioinformatics : Making the case with haplotype inference"*. In : Proceedings of the 9th International Conference on Theory and Applications of Satisfiability Testing (SAT). Springer, 2006, pp. 136-141.
[8] J. P. Marques Silva and K. A. Sakallah. *"GRASP-A new search algorithm for satisfiability"*. In : Proceedings of International Conference on Computer Aided Design. Nov. 1996, pp. 220-227. doi : 10.1109/ICCAD.1996.569607.
[9] Tom van Dijk. *"Sylvan : multi-core decision diagrams"*. English. CTIT Ph.D. thesis series no. 16-398 IPA dissertation series no. 2016-09. PhD thesis. University of Twente, July 2016. isbn : 978-90-365-4160-2. doi : 10.3990/1.9789036541602.