Proceedings of
SAT COMPETITION 2018
Solver and Benchmark Descriptions

Marijn J. H. Heule, Matti Järvisalo, and Martin Suda (editors)
PREFACE

The area of Boolean satisfiability (SAT) solving has seen tremendous progress over the last years. Many problems (e.g., in hardware and software verification) that seemed to be completely out of reach a decade ago can now be handled routinely. Besides new algorithms and better heuristics, refined implementation techniques turned out to be vital for this success. To keep up the driving force in improving SAT solvers, SAT solver competitions provide opportunities for solver developers to present their work to a broader audience and to objectively compare the performance of their own solvers with that of other state-of-the-art solvers.


SC 2018 consisted of four tracks: In addition to the Main Track for sequential SAT solvers, a Random SAT track (for solvers focusing on efficiently solving satisfiable randomly generated SAT instances), Parallel track (for parallel SAT solvers designed for computers with multiple CPUs or CPU cores), and a special No-Limits track where solver source code and solution certificates are not required and portfolios of any nature are allowed) were organized. Additionally, “Glucose hacks”, i.e., solvers based on small modifications to the Glucose 3.0 CDCL SAT solver, were encouraged to be submitted to the main track with the intention of separately awarding the best Glucose hack.

There were two ways of contributing to SC 2018: by submitting one or more solvers for competing in one or more of the competition tracks, and by submitting interesting benchmark instances on which the submitted solvers could be evaluated on in the competition. Following the tradition put forth by SAT Challenge 2012, the rules of SC 2018 invited all contributors to submit a short, 1-2 page long description as part of their contribution. This book contains these non-peer-reviewed descriptions in a single volume, providing a way of consistently citing the individual descriptions.

Participants of the main track of the SAT Competitions are required to submit at least ten interesting benchmarks since 2017. This policy resulted in a significant increase in the number of submitted formulas. The organizers planned to modify all known benchmarks in the 2018 suite in a satisfiability-preserving way to avoid overfitting. However, the number of submitted benchmarks allowed selecting a benchmark suite of 400 formulas that consists only of unknown (submitted in 2018 or earlier, but never used) instances. The hardness of all submitted instances was determined by average the solving time of a mix of SAT Competition 2017 solvers. The selected benchmark suite is a balanced mixed between medium, hard, and challenging (too hard) instances as well as an intended 50%-50% mix of SAT and UNSAT instances. The exact balance cannot be determined as several formulas are too hard for existing solvers.

Successfully running SC 2018 would not have been possible without active support from the community at large. We would like to thank the StarExec initiative (http://www.starexec.org) for the computing resources needed to run the Main, No-Limits, and Random tracks. Many thanks go to Aaron Stump and Patrick J. Hawks for their invaluable help in setting up StarExec to accommodate for the competition’s needs. We also acknowledge the Texas Advanced Computing Center (TACC, http://www.tacc.utexas.edu) at The University of Texas at Austin for providing grid resources for running the Parallel Track. Finally, we would like to emphasize that a competition does not exist without participants: we thank all those who contributed to SC 2018 by submitting either solvers or benchmarks and the related description.

Marijn J. H. Heule, Matti Järvisalo, & Martin Suda
SAT Competition 2018 Organizers
# Contents

Preface ................................................................. 3

## Solver Descriptions

AbcdSAT and Glucose hack: Various Simplifications and Optimizations for CDCL SAT solvers  
*Jingchao Chen* .................................................. 10

CaDiCaL, Lingeling, Plingeling, Treengeling and YalSAT Entering the SAT Competition 2018  
*Armin Biere* ..................................................... 13

System Description of Candy Kingdom — A Sweet Family of SAT Solvers  
*Markus Iser and Felix Kutzner* ................................... 15

CBPeneLoPe2017 and CCSPeneLoPe2017 at the SAT Competition 2018  
*Tomohiro Sonobe* ................................................ 16

The CryptoMiniSat 5.5 set of solvers at the SAT Competition 2018  
*Mate Soos* ............................................................. 17

COMiniSatPS Pulsar and GHackCOMSPS in the 2018 Competition  
*Chanseok Oh* ...................................................... 19

dimetheus  
*Oliver Gableske* .................................................. 20

Description of expSAT Solvers  
*Md Solimul Chowdhury, Martin Müller, and Jia-Huai You* ............................................ 22

Glucose and Syrup: Nine years in the SAT competitions  
*Gilles Audemard and Laurent Simon* .................................. 24

GluHack  
*Aolong Zha* ......................................................... 26

ManyGlucose 4.1-2  
*Yuuya Gotou and Hidetomo Nabeshima* ............................................. 27

Maple_LCM_M1 and Maple_LCM+B Restart  
*Zhen Li and Kun He* ................................................. 28

Maple_LCM_Dist ChronoBT: Featuring Chronological Backtracking  
*Vadim Ryuchin and Alexander Nadel* ...................................... 29

Maple_LCM_Scavel and Maple_LCM_Scavel_200  
*Yang Xu, Guanfeng Wu, Qingshan Chen, and Shuwei Chen* 30
MapleCOMSPS_LRB_VSIDS and MapleCOMSPS_CHB_VSIDS in the 2018 Competition
  Jia Hui Liang, Chanseok Oh, Krzysztof Czarnecki, Pascal Poupart, and Vijay Ganesh .................................................. 31
painless-maplecomsps and painless-maplecomsps-sym
  Ludovic Le Frioux, Hakan Metin, Souheib Baarir, Maximilien Colange, Julien Sopena, and Fabrice Kordon ............................................. 33
probSAT
  Adrian Balint and Uwe Schöning .................................................. 35
Riss 7.1
  Norbert Manthey .................................................. 36
SparrowToRiss 2018
  Adrian Balint and Norbert Manthey .................................................. 38
BreakIDGlucose and BreakIDGlucoseSEL
  Jo Devriendt and Bart Bogaerts .................................................. 40
inIDGlucose
  Jo Devriendt .................................................. 41
ScaLoPe, PeneLoPe_MDLC and Glucose-3.0_PADC in SC18
  Rodrigue Konan Tchinda and Clémentin Tayou Djianeugi .................................................. 42
Maple_CM, Maple_CM_Dist, Maple_CM_ordUIP and Maple_CM_ordUIP+ in the SAT Competition 2018
  Mao Luo, Pan Xiao, Chu-Min Li, Felip Manyà, Zhipeng Lü, and Yu Li .................................................. 44
Glucose Hacks and TopoSAT2 Description
  Thorsten Ehlers and Dirk Nowotka .................................................. 47
Varisat, a SAT Solver Written in Rust
  Jannis Harder .................................................. 49
satUZK-ddc at SAT Competition 2018
  Alexander van der Grinten .................................................. 50
ReasonLS
  Shaowei Cai and Xindi Zhang .................................................. 52

Benchmark Descriptions

Generating the Uniform Random Benchmarks
  Marijn J. H. Heule .................................................. 55
Divider and Unique Inverse Benchmarks Submitted to the SAT Competition 2018
  Armin Biere .................................................. 56
Reversing Elementary Cellular Automata
  Jannis Harder .................................................. 57
Verifying Simple Floating-Point Programs
  Joseph Scott, Jia Hui Liang, and Vijay Ganesh .................................................. 58
GrandTourObs Puzzle as a SAT Benchmark
  Md Solimul Chowdhury, Martin Müller, and Jia-Huai You .................................................. 59
Polynomial Multiplication
  Chu-Min Li, Pan Xiao, Mao Luo, Felip Manyà, Zhipeng Lü, and Yu Li .................................................. 61
Encodings of Relativized Pigeonhole Principle Formulas with Auxiliary Constraints
  Jingchao Chen .................................................. 63
Random k-SAT q-planted solutions
   Adrian Balint .................................................. 64
k-colorability benchmarks
   Jo Devriendt .................................................... 65
Searching for a Unit-Distance Graph with Chromatic Number 6
   Marijn J. H. Heule ............................................. 66
SATcoin — Bitcoin mining via SAT
   Norbert Manthey and Jonathan Heusser ...................... 67
Logical Cryptanalysis Benchmarks for Classical and Modern Hash Functions
   Alexander Scheel ................................................. 68
Benchmark: Assigning Time-Slots to Students with Optimal Preference Satisfaction
   Markus Iser ..................................................... 70
SAT-Encodings of Tree Decompositions
   Max Bannach, Sebastian Berndt, Thorsten Ehlers, and Dirk Nowotka ...... 72
CBMC Formulae
   Norbert Manthey and Michael Tautschnig ...................... 73

Solver Index ............................................................. 75
Benchmark Index ....................................................... 76
Author Index ........................................................... 77
SOLVER DESCRIPTIONS
AbcdSAT and Glucose hack: Various Simplifications and Optimizations for CDCL SAT solvers

Jingchao Chen
School of Informatics, Donghua University
2999 North Renmin Road, Songjiang District, Shanghai 201620, P. R. China
chen-jc@dh.edu.cn

Abstract—This article focuses on the decision variable branching heuristic and learnt clause management for CDCL (conflict-driven clause-learning) SAT solvers. We improve the existing variable branching heuristics and learnt clause reduction via various simplification tricks. On the basis of that, we develop multiple improved versions of abcdSAT and a hack version of Glucose, which are submitted to main, no-limit, Glucose hack and parallel track at the SAT Competition 2018.

I. INTRODUCTION

In this article, we focus on how to simplify the existing techniques such as variable branching heuristic and learnt clause reduction etc. VSIDS (Variable State Independent Decay Sum) is a prevalent decision variable branching policy. The learning rate based branching heuristic (LRB) [9] proposed in the recent year can outperform in some cases. So we decide to simplify the LRB heuristic. The existing three-tiered learnt clause management scheme is sophisticated and difficult to use. Here we simplify it into a one-tiered learnt clause management scheme so that a general CDCL SAT solver can use it also. In addition, we introduce a new re-learning technique.

II. SIMPLIFYING HYBRID BRANCHING HEURISTIC

The learning rate based branching heuristic (LRB) [9] is a variant of the conflict history-based branching heuristic (CHB) [8]. The score $A_v$ of each variable $v$ is computed using the following reinforcement learning formula.

$$A_v = (1 - \alpha)A'_v + \alpha r_v$$

where $A'_v$ is the old score of variable $v$. The difference between LRB and CHB is the computation of $r_v$. In CHB, $r_v$ is a constant that is either 1 or 0.9. However, in LRB, $r_v$ is defined as

$$r_v = \frac{C_v + P_v}{T}$$

where $C_v$ and $P_v$ is the number of conflict clauses and reason clauses $v$ participated in since $v$ is assigned, and $T$ is the interval time that is defined as $T = \text{conflictCounter} - \text{assignedTime}[v]$. Our SAT solver modifies $r_v$ as

$$r_v = \frac{C'_v + S_v + P'_v}{T'}$$

where $C'_v$, $S_v$ and $P'_v$ is the number of conflict clauses, seen clauses and reason clauses $v$ participated in since $v$ is picked, $T' = \text{conflictCounter} - \text{pickedTime}[v]$. Notice, the time $v$ is assigned is not necessarily equal to the time it is picked.

In general, $S_v$ and $C'_v$ overlap and interweave. Computing $C'_v + S_v + P'_v$ is easier than computing $C_v + P_v$. In analyzing a conflict clause, we compute $C'_v$. In collecting reason clauses, we compute $S_v + P'_v$. In the detailed implementation, we use one counter to store the value of $C'_v + S_v + P'_v$.

III. SIMPLIFYING LBD BASED THREE-TIERED LEARNT CLAUSE MANAGEMENT SCHEME

LBD (literal block distance) is defined as the number of decision variables in a clause. According to LBD values, the CoMiniSatPS [6] solver classifies learnt clauses into three categories low-LBD, mid-LBD and high-LBD. The low-LBD clause is also called core clause, whose LBD value is less than 4. The high-LBD clause is also called local clause, whose LBD value is greater than 6. CoMiniSatPS uses three lists learnts_core, learnts_tier2 and learnts_local to store separately learnt clauses. Our SAT solver uses also the three-tiered learnt clause management scheme [7]. However, we merge three lists into one list called learnts, and set a mark for each learnt clause to distinguish which category a learnt clause belongs. In the other words, although we use the same learnt clause management scheme as CoMiniSatPS, our data structure is simpler than that of CoMiniSatPS. Except that each clause has an additional mark, our data structure storing learnt clauses is the same as that of Glucose.

CoMiniSatPS has two database maintenance subroutines. One tries to halve the number of local-tier learnt clauses at every 15,000 conflicts. The other checks mid-tier clauses for reduction at every 10,000 conflicts. The mid-tier clauses not used in the past 30,000 conflicts are moved to local-tier.
We merge such two subroutines into one. In details, at every 15,000 conflicts, we halve the number of local-tier learnt clauses, and move simultaneously mid-tier clauses not used in the past 26,000 conflicts to local-tier.

V. RE-LEARNING

Our re-learning notion is similar to learnt clause minimization (LCM) given in [10]. Let \( F \) be a current formula, \( C = x_1 \lor x_2 \lor \cdots \lor x_n \) be a learnt clause. The basic idea of LCM is that if \( (F - C) \cup \{\neg x_1, \neg x_2, \ldots, \neg x_n\} \) derives an empty clause, we can obtain a new learnt clause \( L \) by analyzing this conflict. If the LBD value of \( L \) is smaller than that of \( C \), we replace \( C \) with \( L \). In [10], this idea is applied to only the learnt clauses with small LBD. In our re-learning policy, we apply this idea to all the learnt clauses. The other difference between LCM in [10] and our re-learning policy is that we can remove some redundant learnt clauses by a subsumption operation on-the-fly, while LCM in [10] cannot. As an inprocessing, our re-learning policy cannot be applied to the whole solving procedure. We apply the re-learning policy only when the number of conflicts is less than \( 5 \times 10^5 \).

VI. ABCDSAT r18

This is submitted to the main track. Compared to abcdSAT r17 [5], it adds a simplified three-tiered learnt clause management scheme, a hybrid branching heuristic and a re-learning strategy given above. The solver runs in the Minisat-VSIDS [3] scoring scheme for the first \( 5 \times 10^4 \) conflicts. Afterwards the scoring scheme is switched to the LRB scoring scheme. When the number of conflicts reaches \( 5 \times 10^6 \) for a large formula or \( 1.5 \times 10^7 \) for a small formula, the scoring scheme is switched to the Glucose-VSIDS [4] scoring scheme. Like the r17 version, in the tree-based search, the solver produces also DRAT proofs. However, its tree-based search branching is different from that of the r17 version. This version uses two scoring policies ACE (Approximation of the Combined lookahead Evaluation) [2] and LRB. In details, it selects a tree node variable using ACE scores when the average LBD is greater than 11, and LRB scores otherwise. In general, the solver does not adopt a bit-encoding phase selection strategy [1] except that it runs in the Glucose-VSIDS scoring scheme.

VII. SMALLSAT

Smallsat is a simplified version of abcdSAT r18. It is also submitted to the main track. Compared to abcdSAT r18, it removes inprocessing techniques such as lifting, probing, distillation, elimination, complex hyper binary resolution, equivalent literal search, unhiding redundancy etc. The variable branching heuristic based on blocked clause decomposition is given up also. This solver has no Glucose-VSIDS scoring scheme. It runs in the Minisat-VSIDS scoring scheme when the number of conflicts is less than \( 5 \times 10^4 \) or greater than \( 1.8 \times 10^7 \), and in the LRB scoring scheme otherwise. It contains the tree-based search, but uses only the LRB scoring policy, excluding the ACE scoring policy.

VIII. ABCDSAT n18

This is similar to abcdSAT r18, but does not output a DRAT proof. So this solver is submitted to the no-limit track. Except for the symmetry breaking preprocessing and XOR Gaussian elimination etc, the simplification technique used in this solver is basically the same as one used in abcdSAT r18. Like the previous no-limit version [5], it divides the whole solving process into three phases. In the first phase, it uses a Minisat-VSIDS scoring scheme and a LRB scoring scheme to search a solution. The second phase simplifies the formula generated in the first phase, using various simplification technique including XOR and cardinality constraint simplification. The second phase uses a Glucose-VSIDS scoring scheme and a LRB scoring scheme to solve the simplified problem. The tree-based search is almost same as that in abcdSAT r18.

IX. ABCDSAT p18

This is a parallel version of abcdSAT n18. Compared with the last year’s the parallel version, this year’s version does not use the master-thread to solve the original problem. This solver uses at most 25 threads. Let the \( i \)-th pivot be \( p_i \), and input formula \( F \). 20 out of 25 threads solve the subproblem \( F \land p_i \). The other 4 threads solve either the original problem or the simplified problem. Once the thread of a subproblem ended, we re-use it to solve the simplified problem with learnt clauses generated so far. We use C++ pthread_cond_timedwait to detect which thread has terminated already. However, it failed sometimes to detect. This may be a C++ bug. For this reason, we use also the solving status of each thread to detect whether a thread has ended already. Only one thread applies bit-encoding phase selection strategies [1]. The symmetry breaking preprocessor is also applied to only another thread. Except the two policies, abcdSAT p18 uses almost the same inprocessing techniques as version p17 [5].

X. glu_mix

glu_mix is a hack version of Glucose. It made modifications on learnt clause management and branching heuristic. glu_mix uses a simplified three-tiered learnt clause management scheme and a hybrid branching heuristic given above. It uses a VSIDS scoring scheme when the number of conflicts is less than \( 2 \times 10^6 \) or greater than \( 2 \times 10^7 \), and a LRB scoring scheme otherwise. To be able to implement simultaneously such two policies in edit distance 1000 benefits from our simplification techniques here.

REFERENCES

Abstract—This note documents the versions of our SAT solvers submitted to the SAT Competition 2018, which are CADiCAL, Lingeling, its two parallel variants Treengeling and Plingeling, and our local search solver YALSAT. Lingeling, Plingeling, Treengeling, YALSAT

Compared to the version of Lingeling submitted last year [1] we added Satisfaction Driven Clause Learning (SDCL) [2], which however due to its experimental nature is disabled (option “--prune=0”). We further disabled blocked clause removal (option “--block=0”) [3], binary blocked clause addition (option “--bca=0”) [4], as well as on-the-fly subsumption (option “--otfs=0”) [5], since all of them can not be combined with SDCL style pruning.

As in our new version of CADiCAL we also experimented with bumping reason side literals too, as suggested in [6]. See below for the motivation to include this feature. There is also a slight change in the order how literals are bumped: previously they were bumped in trail order and are now bumped in variable score order.

Since already last year’s version of Lingeling [1] was almost identical to that from the SAT 2016 Competition [7], it is fair to say that Lingeling and also its parallel extensions Plingeling and Treengeling essentially did not change since 2016. This applies even more to the submitted version of YALSAT, which surprisingly won the random track in 2017, even though it did not change since 2016.

CADiCAL

As explained in our last year’s solver description [1] the goal of developing CADiCAL was to produce a radically simplified CDCL solver, which is easy to understand and change. This was only partially achieved, at least compared to Lingeling.

On the feature side not much changed, since CADiCAL still does not have a complete incremental API (assumptions are missing). However, the non-incremental version was used as back-end of Boolector in the SMT 2017 Competition [8] in the quantifier-free bit-vector track (QF_BV), where it contributed to the top performance of Boolector (particularly compared to the version with Lingeling as back-end).

Our analysis of the SAT 2017 Competition [9] results revealed that the technique of bumping reason side literals [6] of MapleSAT [6] and successors [10], [11] has an extremely positive effect on the selected benchmarks. It consists of going over the literals in learned (minimized 1st UIP) clauses and “bumping” [12] all other literals in their reason clauses too. Even though MapleSAT actually only uses this technique with the new variable scoring scheme proposed in [6], it is already effective in combination with the VMTF scheme [12] used in CADiCAL (and probably for VSIDS too).

Last year’s success of the MapleLCM solver [11], which is an extension of MapleSAT by a different set of authors, also showed that vivification [13] of learned clauses as described by the authors of MapleLCM in their IJCAI paper [14] can be quite useful. In last year’s version of CADiCAL we already had a fast implementation of vivification [1], but only applied it to irredundant clauses. During inprocessing [15] our new version of CADiCAL has two vivification phases, the first phase working on all including redundant clauses and the second phase works as before only on irredundant clauses.

Furthermore, all the top performing configurations of MapleSAT and MapleLCM made use of the observation of Chanseok Oh [16], that a CDCL solver should alternate between “quiet” no-restart phases and the usual fast restart scheduling [17]. This also turns out to be quite beneficial for last year’s selection of benchmarks and we added such “stabilizing” phases scheduled in geometrically increasing conflict intervals.

Then, we experimented with “rephasing”, which in arithmetically increasing conflict intervals overwrites all saved phases [18] and either (i) restores the initial phase (default true in CADiCAL), (ii) flips the current saved phase, (iii) switches to the inverted initial phase (thus false), or (iv) picks a completely random phase. This technique gives another (but smaller) boost to the performance of CADiCAL on last year’s benchmarks compared to the other new techniques above.
Finally, we observed, that for very long running instances (taking much longer than the 5000 seconds time limit used in the competition), the standard arithmetic increase [19] of the limit on kept learned clauses increases memory usage over time substantially and slows down propagation. Therefore we flush all redundant clauses (including low glucose level clauses) in geometrically increasing conflict intervals too. This should happen less than a dozen of times during each competition run though.

**LICENSE**

The default license of YALSAT, LINELING, PLINGELING and TREENELING did not change in the last three years. It allows the use of these solvers for research and evaluation but not in a commercial setting nor as part of a competition submission without explicit permission by the copyright holder. However, as part of our new open source release of BOOLECTOR 3.0 [20] we also plan to release LINELING under an open source MIT style license, which for CADDICAL continues to be the case.

**REFERENCES**


System Description of Candy Kingdom –
A Sweet Family of SAT Solvers

Markus Iser∗ and Felix Kutzner †
Institute for Theoretical Computer Science, Karlsruhe Institute of Technology
Karlsruhe, Germany
Email: *markus.iser@kit.edu, †felix.kutzner@qpr-technologies.de

Abstract—Candy is a branch of the Glucose 3 SAT solver and started as a refactoring effort towards modern C++. We replaced most of its custom lowest-level data structures and algorithms by their C++ standard library equivalents and improved or reimplemented several of its components. New functionality in Candy is based on gate structure analysis and random simulation.

I. INTRODUCTION

The development of our open-source SAT solver Candy† started as a branch of the well-known Glucose [1], [2] CDCL SAT solver (version 3.0). With Candy, we aim to facilitate the solver’s development by refactoring the Glucose source code towards modern C++ and by reducing dependencies within the source code. This involved replacing most custom lowest-level data structures and algorithms by their C++ standard library equivalents. The refactoring effort enabled high-level optimizations of the solver such as inprocessing and cache-efficient clause memory management. We also increased the extensibility of Candy via static polymorphism, e.g. allowing the solver’s decision heuristic to be customized without incurring the overhead of dynamic polymorphism. This enabled us to efficiently implement variants of the Candy solver. Furthermore, we modularized the source code of Candy to make its subsystems reusable. The quality of Candy is assured by automated testing, with the functionality of Candy tested on different compilers (Clang, GCC, Microsoft C/C++) and operating systems (Linux, Apple macOS, Microsoft Windows) using continuous integration systems.

II. CLAUSE MEMORY MANAGEMENT

Unlike Glucose, we use regular pointers to reference clauses in Candy. To reduce the memory access overhead, we introduced a dedicated cache-optimized clause storage system. To this end, we reduced the memory footprint of clauses by shrinking the clause header, in which only the clause’s size, activity and LBD values as well as a minimal amount of flags are stored. For clauses containing 500 literals or less, our new clause allocator prelocates clauses in buckets of same-sized clauses. Clauses larger than 500 literals are individually allocated on the heap. Buckets containing small clauses are regularly sorted by their activity in descending order to group frequently-accessed clauses, thereby concentrating memory accesses to smaller memory regions. Moreover, the watchers are regularly sorted by clause size and activity.

III. IMPROVED INCREMENTAL MODE

We enabled several clause simplifications in Candy’s incremental mode that had been deactivated in Glucose’s incremental mode. Also, certificates for unsatisfiability can be generated in incremental mode for sub-formulas not containing assumption literals. This is achieved by suppressing the emission of learnt clauses containing assumption literals as well as the output of the empty clause until no assumptions are used in the resolution steps by which unsatisfiability is deduced.

IV. INPROCESSING

We improved the architecture of clause simplification such that Candy can now perform simplification based on clause subsumption and self-subsuming resolution during search. The original problem’s clauses are included as well as learnt clauses that are persistent in the learnt clause database, i.e. clauses of size 2 and clauses with an LBD value no larger than 2.

V. IMPROVEMENTS CANDY 2018

The approach to modularize the Glucose codebase was further pursued and realized. We improved a lot in separation of concerns and building of separate components for heuristics and algorithmic methods.

Candy now initializes variable ordering based on problem structure by default. This should make its default heuristic configuration more stable against problem scrambling.

REFERENCES


https://github.com/udopia/candy-kingdom
CBPeneLoPe2018 and CCSPeneLoPe2018 at the SAT Competition 2018

Tomohiro Sonobe
National Institute of Informatics, Japan
JST, ERATO, Kawarabayashi Large Graph Project, Japan
Email: tominlab@gmail.com

Abstract—In this description, we provide a brief introduction of our solvers: CBPeneLoPe2018 and CCSPeneLoPe2018, in the SAT Competition 2018. CBPeneLoPe2018 and CCSPeneLoPe2018 are based on the parallel SAT solver PeneLoPe. Both solvers use SatELite [4] as a preprocessor.

I. CBPeneLoPe2018

CBPeneLoPe2018 is a parallel portfolio SAT solver based on PeneLoPe [2] and a new version of ones submitted in the SAT Competition 2014, SAT Race 2015, SAT Competition 2016, and SAT Competition 2017. CBPeneLoPe2018 implements community branching [7], a diversification [5] technique using community structure of SAT instances [1]. Community branching assigns a different set of variables (community) to each worker and forces them to select these variables as decision variables in early decision levels, aiming to avoid overlaps of search spaces between the workers more vigorously than the existing diversification methods.

In order to create communities, we construct a graph where a vertex corresponds to a variable and an edge corresponds to a relation between two variables in the same clause, proposed as Variable Incidence Graph (VIG) in [1]. After that, we apply Louvain method [3], one of the modularity-based community detection algorithms, to identify communities of a VIG. Variables in a community have strong relationships, and a distributed search for different communities can benefit the whole search.

In addition, CBPeneLoPe2018 uses community-based learnt clause sharing (CLCS) for learnt clause sharing between workers. CLCS restricts the sharing of each learnt clause to workers that conduct the search for the variables related with communities in the target learnt clause. By combining community branching, CLCS distributes the target clauses to the workers with related communities. For example, if a learnt clause \((a \lor b \lor c)\) is to be shared among the workers, and the variable \(a\) and \(b\) belong to a community \(C_1\) and the variable \(c\) belongs to a community \(C_2\), this clause is distributed only to the workers that are assigned the community \(C_1\) or \(C_2\) by community branching.

II. CCSPeneLoPe2018

CCSPeneLoPe2018 is a parallel portfolio solver based on PeneLoPe. The features of CCSPeneLoPe2018 are as follows.

- Conflict history-based branching heuristic (CHB) [6] for some workers
- CLCS prioritizing high VSIDS or CHB scores

CHB is good at cryptographic instances in [6]. In CCSPeneLoPe2018, some workers use this heuristic with different sets of its parameters. For CLCS, each worker calculates an average activity score (VSIDS or CHB) of variables for each community and chooses the highest scored community as a “desired community”. CLCS distributes the target clause to the workers that desire to share that clause (i.e., including the variables that belong to the desired community).

III. ACKNOWLEDGMENT

This work was supported by JSPS KAKENHI Grant Number 17K12742.

REFERENCES

The CryptoMiniSat 5.5 set of solvers at the SAT Competition 2018
Mate Soos, National University of Singapore

I. INTRODUCTION

This paper presents the conflict-driven clause-learning SAT solver CryptoMiniSat v5.5 (CMS) as submitted to SAT Competition 2018. CMS aims to be a modern, open-source SAT solver that allows for multi-threaded in-processing techniques while still retaining a strong CDCL component. In general, CMS is an inprocessing SAT solver that uses optimised data structures and finely-tuned timeouts to have good control over both memory and time usage of simplification steps. Below are the changes to CMS compared to the SAT Competition 2016 version.

II. MAJOR IMPROVEMENTS

A. Careful code review

Over the years, much cruft has accumulated in CryptoMiniSat. This has left serious bugs in the implementation in important parts of the solver such as clause cleaning and restarting. This has lead to low performance. A code review of the most important parts of the solver such as bounded variable elimination, restarting, clause cleaning and variable activities has been conducted.

B. Integration of ideas from Maple, LCM, Dist

Some of the ideas from Maple, LCM, Dist[2][4] have been included into CMS. In particular, the clause cleaning system, the radical in-process distillation and the Maple-based variable activities are all used.

C. Cluster Tuning

The author has been generously given time on the ASPIRE-1 cluster of the National Supercomputing Centre Singapore[1]. This allowed experimentation and tuning that would have been impossible otherwise. CMS has not been tuned on a cluster for over 6 years and the difference shows. A slightly interesting side-effect is that the parameters suggested by the cluster are non-intuitive, such as not simplifying the CNF straight away, but rather CDCL-solving it first. Another interesting effect is that intree probing[3] seems to be very important.

D. Parallel Solving

As in previous competitions, CMS only shares unit and binary clauses, and shares them very rarely. The different threads, however, are run with very different, hoping to be orthogonal, parameters varying everything from clause cleaning strategies to default polarities.

E. Automatic Tuning

The "autotune" version of the solver measures internal solving parameters and re-configures itself after a preset number of conflicts to a configuration that has been suggested by the parameters and the machine learning algorithm C4.5[5].

III. GENERAL NOTES

A. On-the-fly Gaussian Elimination

On-the-fly Gaussian elimination is again part of CryptoMiniSat. This is explicitly disabled for the competition, but the code is available and well-tested. This allows for special uses of the solver that other solvers, without on-the-fly Gaussian elimination, are not capable of.

B. Robustness

CMS aims to be usable in both industry and academia. CMS has over 150 test cases and over 2000 lines of Python just for fuzzing orchestration, and runs without fault under both the ASAN and UBSAN sanitisers of clang. It also compiles and runs under Windows, Linux and MacOS X. This is in contrast many academic winning SAT solvers that produce results that are non-reproducible, cannot be compiled on anything but a few select systems, and/or produce segmentation faults if used as a library. CryptoMiniSat has extensive fuzzing setup for library usage and is very robust under strange/unexpected use cases.

IV. THANKS

This work was supported in part by NUS ODPRT Grant, R-252-000-685-133. The computational work for this article was performed on resources of the National Supercomputing Centre, Singapore[1]. The author would also like to thank all the users of CryptoMiniSat who have submitted over 400 issues and many pull requests to the GitHub CMS repository[6].

REFERENCES


COMiniSatPS Pulsar and GHackCOMSPS in the 2018 Competition

Chanseok Oh
Google
New York, NY, USA

Abstract—COMiniSatPS is a patched MiniSat generated by applying a series of small diff patches to the last available version (2.2.0) of MiniSat that was released several years ago. The essence of the patches is to include only minimal changes necessary to make MiniSat sufficiently competitive with modern SAT solvers. One important goal of COMiniSatPS is to provide these changes in a highly accessible and digestible form so that the necessary changes can be understood easily to benefit wide audiences, particularly starters and non-experts in practical SAT. As such, the changes are provided as a series of incrementally applicable diff patches, each of which implements one feature at a time. COMiniSatPS has many variations. The variations are official successors to an early prototype code-named SWDiA5BY that saw great successes in the past SAT-related competitive events.

I. INTRODUCTION

It has been shown in many of the past SAT-related competitive events that very simple solvers with tiny but critical changes (e.g., MiniSat [1] hack solvers) can be impressively competitive or even outperform complex state-of-the-art solvers [2]. However, the original MiniSat itself is vastly inferior to modern SAT solvers in terms of actual performance. This is no wonder, as it has been many years since the last 2.2.0 release of MiniSat. To match the performance of modern solvers, MiniSat needs to be modified to add some of the highly effective techniques of recent days. Fortunately, small modifications are enough to bring up the performance of any simple solver to the performance level of modern solvers. COMiniSatPS [3] adopts only simple but truly effective ideas that can make MiniSat sufficiently competitive with recent state-of-the-art solvers. In the same minimalistic spirit of MiniSat, COMiniSatPS prefers simplicity over complexity to reach out to wide audiences. As such, the solver is provided as a series of incremental patches to the original MiniSat. Each small patch adds or enhances one feature at a time and produces a fully functional solver. Each patch often changes solver characteristics fundamentally. This form of source distribution makes MiniSat sufficiently competitive with recent state-of-the-art solvers. COMiniSatPS is a base solver of the MapleCOMSPS solver series [7], [8], [9] that participated in SAT Competition 2016, 2017, and 2018.

II. COMINI SATPS PULSAR

This year’s solver is basically identical to the last year’s solver, fixing only two minor bugs:

• Correctly reports UNSAT when a problem is determined to be UNSAT during CNF parsing if Gaussian elimination is enabled. (The Gaussian elimination is enabled when not generating UNSAT proof.)
• Correctly generates UNSAT proof when a problem is solved by pre-processing alone.

III. GHBAC KCOMSPS

This year’s solver is identical to the last year’s solver [10] (which is in turn identical to the 2016 version). GHackCOMSPS qualifies as a Glucose hack.

IV. AVAILABILITY AND LICENSE

Source is available for download for all the versions described in this paper. Note that the license of the M4RI library (used to implement the Gaussian elimination) is GPLv2+.

ACKNOWLEDGMENT

We thank specifically the authors of Glucose, GlueMiniSat, Lingeling, CryptoMiniSat, and MiniSat.

REFERENCES

Abstract—This document describes the dimetheus SAT solver as submitted to the random SAT track of the SAT Competition 2018.

I. INTRODUCTION

Please note that this article must be understood as a rather brief overview of the dimetheus SAT solver. Additional information regarding its functioning, a comprehensive quick-start guide, as well as the source-code of the latest version of the solver can be found on the authors website. Additionally, the author elaborates on the theoretical background of the solver in his Ph.D. thesis [1] which can be found online. A preliminary overview of the applied techniques can be found in [2], [3].

This article will first cover the main techniques that the solver applies in Section II. Afterwards, a brief overview of the parameter settings are discussed in Section III. This is followed by a brief explanation of the programming language and the compiler relevant parameters in Section IV. Additionally, several SAT Competition relevant details are discussed in Section V. The article is concluded by a few remarks on the availability and the license of the solver in Section VI.

II. MAIN TECHNIQUES

The dimetheus solver runs in various phases as depicted in Figure 1.

In each phase, the solver must fulfill a pre-defined task. The first four of these phases (reset, params, guidance, startup) are not discussed here in detail. At the end of the startup phase the solver has loaded the formula and is able to work with it. The solver will then execute the classification phase in order to determine determine what type of CNF formula was provided. Since the solver is submitted to the random SAT track of the SAT Competition it will determine what type of random formula it has to solve (e.g., it will determine the size of the formula, the clause lengths, the clauses-to-variables ratio). The result of the classification phase is then forwarded to the adaptation phase. In this phase, the solver adapts a wide variety of internal parameters in order to initialize its internal heuristics and search algorithm.

Afterwards, preprocessing is performed. The preprocessing is kept simple and includes pure literal elimination and the removal of duplicate clauses.

Preprocessing is then followed by the search phase in which the solver tries to find a satisfying assignment for the formula (inprocessing is turned off when the solver solves random formulas). The approach that the solver applies is best understood as bias-based decimation followed by stochastic local search. The bias-based decimation applies a Message Passing algorithm to calculate biases for individual variables. These biases indicate how likely it is to observe a variable assigned to one or zero when taking into account the models of the formula. For more information see [1]. Afterwards, a fraction of the variables with the largest bias are assigned and unit propagation (UP) is performed which then leads to a simplified remaining formula. The bias calculation and the UP-based assignment of variables with the largest bias is repeated until one of two cases occurs. First, a model is found. In this case the solver merely outputs the model and terminates. Second, UP runs into a conflict. In this case the solver will undo all assignments and initializes an SLS solver. The starting assignment for the SLS is comprised of all the assignments made until the conflict arose as well as random assignments to the remaining variables. From this point onwards the SLS takes place until either a time-out is hit or a model is found. The dimetheus solver, as it runs in the SAT Competition 2018, is therefore an incomplete solver that cannot detect unsatisfiability.

III. MAIN PARAMETERS

The solver is started with the two following parameters.

-formula STRING: The STRING points to the file that contains the formula in DIMACS CNF input format.
This tells the classifier that it can assume the formula to be a random formula when determining what specific type of formula it is.

As mentioned in the previous section, the solver will use the information that was gathered in the classification phase in order to enforce an optimal parameter setting to a variety of internal parameters [1]. Unfortunately, it is not possible to correctly explain the abundance of parameters here which is why the reader is referred to the given reference for details.

IV. IMPLEMENTATION DETAILS

The dimetheus solver is implemented in C. The Message Passing algorithm that is applied to calculate the biases is an interpolation of Belief Propagation and Survey Propagation [1], [4]. The SLS search follows the probSAT approach [5].

V. SAT COMPETITION 2018 SPECIFICS

The dimetheus solver was submitted to the random SAT track. It was compiled on the StarExec Cluster using gcc with the compile flags -std=c99 -O3 -static -fexpensive-optimizations -flto -fwhole-program -march=native -Wall -pedantic. The result is a 64-bit binary.

VI. AVAILABILITY AND LICENSE INFORMATION

The dimetheus solver is publicly available and can be downloaded from https://www.gableske.net/dimetheus. The solver is provided under the Creative-Commons Non-Commercial Share-Alike license version 3.0 (CCBYNCSA3).

ACKNOWLEDGMENTS

The author hereby expresses his gratitude towards the organizers of the SAT Competition and the SAT Conference. Your hard work is much appreciated.

REFERENCES

Description of \textit{expSAT} Solvers

Md Solimul Chowdhury  
Department of Computing Science  
University of Alberta  
Edmonton, Alberta, Canada  
mdsolimu@ualberta.ca

Martin Müller  
Department of Computing Science  
University of Alberta  
Edmonton, Alberta, Canada  
mmlueller@ualberta.ca

Jia-Huai You  
Department of Computing Science  
University of Alberta  
Edmonton, Alberta, Canada  
jyou@ualberta.ca

\textbf{Abstract—} \textit{expSAT} is a novel CDCL SAT solving method, which performs random-walk based explorations of the search space w.r.t the current search state to guide the search. It uses new branching heuristics, called \textit{expVSIDS}, which combines the standard variable selection heuristic VSIDS, which is based on search performance so far, with heuristic scores derived from random samples of possible future search states. This document describes the \textit{expSAT} approach and four CDCL SAT solvers based on this approach, which we have submitted for the SAT competition-2018.

I. THE \textit{expSAT} APPROACH

This section presents the \textit{expSAT} approach, part of which is to appear in [1].

A. \textit{expSAT} algorithm

Given a CNF SAT formula $F$, let $\text{vars}(F)$, $\text{uVars}(F)$ and $\text{assign}(F)$ denote the set of variables in $F$, the set of currently unassigned variables in $F$ and the current partial assignment, respectively. In addition to $F$, \textit{expSAT} also accepts five \textit{exploration parameters} $nW, lW, \theta_{\text{stop}}, p_{\text{exp}}$ and $\omega$, where $1 \leq nW, lW \leq \text{uVars}(F)$, $0 < \theta_{\text{stop}}, p_{\text{exp}} \leq 1$. These parameters control the exploration aspects of \textit{expSAT}. The details of these parameters are given below.

Given a CDCL SAT solver, \textit{expSAT} modifies it as follows: (I) Before each branching decision, if the search-height, $\frac{\text{assign}(F)}{\text{vars}(F)} \leq \theta_{\text{stop}}$, with probability $p_{\text{exp}}$, \textit{expSAT} performs an \textit{exploration episode}, consisting of a fixed number $nW$ of random walks. Each walk consists of a limited number of \textit{random steps}. Each such step consists of (a) the uniform random selection of a currently unassigned \textit{step variable} and assigning a boolean value to it using the standard CDCL polarity heuristic, and (b) a followed by Unit Propagation (UP). A walk terminates either when a conflict occurs during UP, or after a fixed number $lW$ of random steps have been taken. Figure 1 illustrates an exploration episode. (II) In an exploration episode of $nW$ walks of maximum length $lW$, the \textit{exploration score} $\text{expScore}$ of a decision variable $v$ is the average of the walk scores $\text{ws}(v)$ of all those random walks within the same episode in which $v$ was one of the randomly chosen decision variables. $\text{ws}(v)$ is computed as follows: (a) $\text{ws}(v) = 0$ if the walk ended without a conflict. (b) Otherwise, $\text{ws}(v) = \frac{1}{\text{wldic}}$, with decay factor $0 < \omega \leq 1$, $\text{ldic}(v)$ the LBD score of the clause $c$ learned for the current conflict, and $d \geq 0$ the \textit{decision distance} between variable $v$ and the conflict which ended the current walk: If $v$ was assigned at some step $j$ during the current walk, and the conflict occurred after step $j' \geq j$, then $d = j' - j$. We assign credit to all the step variables in a walk that ends with a conflict and give higher credit to variables closer to the conflict. (III) The novel branching heuristic $\text{expVSIDS}$ adds VSIDS score and $\text{expScore}$ of the unassigned variables. At the current state of the search, the variable bumping factor of VSIDS is $g^z$, where $g > 1$ and $z \geq 1$ is the count of conflicts in the search so far. To achieve a comparable scale for $\text{expScore}$ and VSIDS score, we scale up the $\text{expScore}$ by $g^z$ before adding these scores. A variable $v^*$ with maximum combined score is selected for branching. (IV) All other components remain the same as in the underlying CDCL SAT solver.

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{expSAT_diagram.png}
\caption{An exploration episode with $nW = 3$ walks and a maximum of $lW = 3$ random steps per walk. $(v, i, j)$ represents that the variable $v$ is randomly decided at the $j^{th}$ step of $i^{th}$ walk.}
\end{figure}

B. Exploration Parameter Adaptation

In \textit{expSAT}, $(nW, lW, \theta_{\text{stop}}, p_{\text{exp}}, \omega)$, the set of exploration parameters, governs the exploration. The first two parameters dictate \textit{how} much exploration to perform per episode, the third and fourth parameter dictate \textit{when} to trigger an exploration episode. $\omega$ controls how exploration scores are computed.

How to adapt these parameters during the SAT search is an interesting question, which is not addressed in [1]. The \textit{expSAT} based solvers submitted for this competition uses a simple local search algorithm to adapt the first four exploration parameters $P = (nW, lW, \theta_{\text{stop}}, p_{\text{exp}})$ to dynamically control when to trigger exploration episodes and how much
exploration to perform in an exploration episode. This local search algorithm executes in parallel to the SAT search in expSAT.

a) The Adaptation Algorithm: The idea of this algorithm is to start with an initial value \( val(P, 1) = (nW^1, lV^1, \theta_{stop}^1, P_{exp}) \) of the exploration parameters and iteratively update the values between two consecutive restarts, based on the performance of exploration in the previous restarts.

Assume expSAT has just performed the \( j \)th \( (j \geq 2) \) restart. Let \( \sigma_j \) the performance of exploration between \((j-1)\)th and \(j\)th restarts. We define \( \sigma_j \) as follows:

\[
\sigma_j = w_1 * \frac{c^j}{rSteps^j} + w_2 * \frac{g^j}{rSteps^j} + w_3 * \frac{1}{rsLBD^j}
\]

Here, \( rSteps^j \) is the number of random steps taken during the exploration episodes, which has occurred between \((j-1)\)th and \(j\)th restarts, \( c^j, g^j, rsLBD^j \) are the number of conflicts, number of glue-clauses and the mean LBD value of the (learned) clauses identified in these \( rSteps^j \) steps, respectively. \( w_1, w_2 \) and \( w_3 \) are three fixed weights.

After restart \( j \), just before starting SAT search, the algorithm updates the exploration parameter values by comparing \( \sigma_j \) and \( \sigma_{j-1} \). Let \( val(P, j) = (nW^j, lV^j, \theta_{stop}^j, P_{exp}) \) is the updated value of exploration parameters before the search begins, just after the \((j-1)\)th restart.

• If \( \sigma_j < \sigma_{j-1} \), then the performance of exploration deteriorates after the \((j-1)\)th restart. In this case, we perform two operations on the exploration parameters after \( j \)th restart:
  - **Decrement**: Let \( dp \in P \) the parameter whose value was increased from \( x \) to \( x' \) after the \((j-1)\)th restart. We attribute this deterioration of performance to this update. We revert the value of \( dp \) back to \( x \) from \( x' \).
  - **Increment**: Randomly select a parameter \( rp \in P \) and increase its value to \( y' \) from \( y \), where \( rp \neq dp \).

• If \( \sigma_j > \sigma_{j-1} \) = 0, then we only perform the **Increment** operation, as we do not know whom to blame for the stall. The updated value of these parameters remain effective until the \((j+1)\)th restart.

• If \( \sigma_j > \sigma_{j-1} \), then the performance of exploration is increasing after the \((j-1)\)th restart and we do not change any parameter value as the current value of the exploration parameters leads to better performance.

For changing the value of a parameter \( x \in P \), we associate a step size \( s_x \) with \( x \). Also, in order to prevent the unbounded growth/shrink of the parameters we associate a lower and upper bound with each of the parameters. That is, for \( x \in P \), we have a \([l_x, u_x] \). Whenever the value of \( x \) exceeds \( u_x \) or the value of \( x \) is less than \( l_x \), then the value of \( x \) is reset to its initial value \( x^j \).

II. expSAT Solvers

We have submitted four CDCL SAT solvers based on the expSAT approach, which are implemented on top of Glucose, MapleCOMPSPS_LRB and MapleCOMPSPS. In the following, we describe our solvers:

a) **expGlucose**: expGlucose is an extension of Glucose, where we replace VSIDS by expVSIDS and have kept everything else the same as in Glucose.

b) **expMC_LRB_VSIDS_Switch**: The corresponding baseline system MapleCOMPSPS_LRB switches between branching heuristics LRB and VSIDS in between restarts.

In expMC_LRB_VSIDS_Switch, we replace VSIDS with expVSIDS and have kept everything else the same as in MapleCOMPSPS_LRB.

c) **expMC_LRB_VSIDS_Switch_2500**: The corresponding baseline system MapleCOMPSPS_LRB has three switches between VSIDS and LRB (i) VSIDS for initialization (first 50,000 conflicts), (ii) then run LRB for 2,500 seconds, and (iii) then switches to VSIDS for rest of the execution of the solver. In expMC_LRB_VSIDS_Switch_2500, we replace VSIDS with expVSIDS for (iii) and have kept everything else the same as in MapleCOMPSPS_LRB.

d) **expV_MLSVSIDS_Switch_2500**: This system is a variant of expMC_LRB_VSIDS_Switch_2500. It uses expVSIDS for the first 2,500 seconds and then switches to LRB for the rest of its execution.

REFERENCES

Glucose and Syrup: Nine years in the SAT competitions

Gilles Audemard
Univ. Lille-Nord de France
CRIL/CNRS UMR8188
audemard@cril.fr

Laurent Simon
Bordeaux-INP / Univ. Bordeaux
LaBRI/CNRS UMR 5800
lsimon@labri.fr

Abstract—Glucose is a CDCL solver developped on top of Minisat nine years ago, with a special focus on removing useless clauses as soon as possible, and an original restart scheme based on the quality of recent learnt clauses. Syrup is the parallel version of Glucose, with an original lazy clauses exchanges policy, thanks to a one-watched scheme. We describe in this short solver description the small novelties introduced this year for the SAT 2018 competition.

I. INTRODUCTION

Glucose is a CDCL (Conflict Driven Clause Learning) solver introduced in 2009 that tries to center all the components of the SAT solver around a measure of learnt clause quality, called LBD, for Literal Block Distance. This measure allows to delete a lot of learnt clauses from the beginning of the computation. From a practical point of view, it seems that this feature allows Glucose to produce more shorter proofs, which probably explains why Glucose and Syrup won a number of competitions in the last 9 years. A recent survey paper summarizes most of the improvements we added to the original Glucose [1]. Of course, the current short description does not mean to be exhaustive and the interested reader should refer to the previous paper.

In a few words, however, Glucose enters SAT competitions/races [2], [3] every years since its creation. Glucose is based on the internal architecture of minisat [4] (especially for the VSIDS implementation, the 2-Watched scheme and the memory management of clauses (garbage collection, ...)). It is based on the notion of Literal Block Distance, as aforementioned, a measure that is able to estimate the quality of learnt clauses [5]. This measure simply counts the number of distinct decision levels of literals occurring in learnt clauses, at the time of their creation. Thanks to that, new strategies for deleting clauses were proposed. Moreover, the solver constantly watch the quality of the last learnt clauses and triggers a restart when the quality is worst than the average. Recent developments includes a way of postponing restarts when the number of assigned literals suddenly increases without conflicts (a SAT solution may be then expected). In the last version of Glucose, a special strategy allowed the solver to decide which strategy to use with respect to a set of identified extreme case [6].

Indeed, learnt clauses removal, restarts, small modifications of the VSIDS heuristic are based on the concept of LBD. The core engine of Glucose (and Syrup) is 8 years old. Syrup is a major improvement of Glucose on which we focused most of our efforts in the last years.

II. NEW COMPONENTS

The 2018 version of Glucose and Syrup are very similar to the 2016 ones, with two improvements. The main modifications are based on the extension of the recent LCM strategies proposed last year [7] (which “revived” the vivification technique [8]). We observed that the LCM strategy was not always performed on clauses of small LBD only, because LCM was not triggered right after clause database reduction, and thus the order of clauses traversed by the LCM was not based on a sorted order of learnt clauses. However, we observed that LCM was more efficient when not always run on good clauses only (LCM can replace clauses, and thus may delete a good clause). We observed that LCM was more efficient when active clauses were kept, in addition to clauses of small LBD. Glucose is now keeping 10% of the most active clauses in addition to the usual LBD based ranking. In addition, we integrated the LCM technique into Syrup (good clauses found during the LCM reduction were not shared initially). The integration into Syrup was not obvious because LCM seemed to be much more efficient on sequential solvers only. LCM was thus not activated on all the cores in the parallel version.

III. PARALLEL VERSION OF GLUCOSE

We used a version with 24 and 48 cores this year for the parallel versions of Glucose (called Glucose-Syrup).

IV. ALGORITHM AND IMPLEMENTATION DETAILS

Glucose uses a special data structure for binary clauses, and a very limited self-subsumption reduction with binary clauses, when the learnt clause is of interesting LBD. The parallel version uses a special data structure for sharing clauses that may prevent some clause to be shared when it is full.

V. ACKNOWLEDGMENTS

Recent developments of Glucose-Syrup is supported by the French Project SATAS ANR-15-CE40-0017. We would also like to thanks the organizers of the competition for their efforts.

REFERENCES


Abstract—GluHack is a SAT solver submitted to the Glucose Hack Track of the SAT Competition 2018. It updates Glucose 3.0 in the following aspects: searching watch list, conflict learning and learnt clause database reduction.

I. INTRODUCTION

Gluco [1] is an open-source CDCL-based SAT solver [2] that has achieved numerous excellent performance in past SAT Competition. In Unit Propagation (UP), the head of trail queue of assigned literal will be specified as a watched literal for searching its corresponding watch list, which is in order to propagate the next unit unassigned literal in a clause. This procedure will not stop until all variables are assigned, unless a conflict occurs. Let’s consider the situation which is to decide literal \(a\) is assigned to true in UP. Then by searching watch list we know that all literals in clause \(C_1\) have been assigned to false including literal \(\neg a\). Obviously, a conflict occurs. Glucose stops searching watch list, returns the reason of this conflict, the clause \(C_1\), and goes to the learning phase. Under current assignment gluHack keeps searching watch list until all conflicts are detected, then stores all corresponding reasons \(C_1, C_2, \ldots, C_n\) into a vector of clauses and returns it.

II. IMPLEMENTATION

GluHack stops unit propagating (assigning), but keeps searching watch list when first conflict occurs. All detected conflicts are stored into a vector of clauses.

Modification 1 1  [CRef Solver::propagate()]

Initialize: bool firstConflOccur ← true;
vec (CRef) conflCRefList;

1: · · ·
2: // Did not find watch – clause is unit under assignment:
3: *j++ ← w;
4: if value(first) = 1 then
5: if firstConflOccur = true then
6: firstConflOccur ← false;
7: confl ← cr;
8: end if
9: conflCRefList.push(cr);
10: qhead ← trail.size();
11: else if firstConflOccur = true then
12: uncheckedEnqueue(first, cr);
13: end if
14: NextClause::;
15: · · ·
16: return confl;

The analyze() function will generate corresponding learnt clauses for each conflict in this vector. Finally, we evaluate and filter some effective learnt clauses, whose the number of literal contained \(\leq 2\), and add them into the learnt clause database.

Modification 2 1 [bool Solver::search(int)]

Initialize: vec (Lit)* learnt_clauseList ←
new vec (Lit)[conflCRefList.size()];

1: · · ·
2: for i = 1 to conflCRefList.size() do
3: analyze(conflCRefList[i], learnt_clause, selectors, backtrack_level, nblevels, szWoutSelectors); · · ·
4: for j = 0 to learnt_clause.size() do
5: learnt_clauseList[i].push(learnt_clause[j]);
6: end for
7: end for
8: analyze(confl, learnt_clause, selectors, backtrack_level, nblevels, szWoutSelectors);
9: · · ·
10: if learnt_clause.size() = 1 then
11: uncheckedEnqueue(learnt_clause[0]); nbUn++;
12: else
13: for i = 1 to conflCRefList.size() do
14: if learnt_clauseList[i].size() \(\leq 2\) then
15: CRef cr = ca.alloc(learnt_clauseList[i], true); · · ·
16: end if
17: end for
18: CRef cr = ca.alloc(learnt_clause, true); · · ·
19: end if
20: · · ·

In order to keep high quality of learnt clauses, we change the rate of learnt clause database reduction from 50% to 70%.

Modification 3 1 void Solver::reduceDB()

1: · · ·
2: int limit = learnts.size() / 10 * 7; · · ·

The different between this hack version and original sources of Glucose 3.0 is total 779 non-space characters.

REFERENCES

ManyGlucose 4.1-2

Yuuya Gotou
University of Yamanashi, JAPAN

Hidetomo Nabeshima
University of Yamanashi, JAPAN

Abstract—ManyGlucose 4.1-2 is a deterministic parallel SAT solver based on Glucose syrup 4.1 with an efficient clause exchange mechanism that is a refinement of ManySAT algorithm.

I. INTRODUCTION

ManyGlucose 4.1-2 is a deterministic parallel SAT solver. Given an instance, a deterministic solver has reproducible results in terms of solution (satisfying assignment or proof of unsatisfiability) and running time. ManyGlucose supports such reproducible behavior. ManySAT 2.0 [2] is a representative deterministic parallel solver which is built on MiniSat 2.2 [3] with a deterministic clause exchange algorithm [1]. ManyGlucose 4.1-2 has a refined algorithm for clause exchange and is built on Glucose syrup 4.1 [4].

II. MAIN TECHNIQUES

ManySAT 2.0 has reproducible behavior by synchronizing among threads and exchanging learning clauses among them after synchronization. Each thread synchronizes every execution interval called period, which is defined as the number of conflicts. Since the generation speed of conflicts depends on the search space of each thread, the execution time of a period is different in each thread. Thus, each thread often has an idle time for synchronization.

In order to reduce the idle time, we introduce two improvements to ManySAT algorithm.

1) Refinement of period: a period is defined as the number of scanned literals in unit propagations instead of the number of conflicts.
2) Lazy clause exchange: each thread receives learned clauses obtained in \( m \) periods ago of the other threads. This eliminates the need to wait if the gap of the period of each thread is less than or equal to \( m \).

III. MAIN PARAMETERS

We define a period as 2 million scanned literals and use 20 as the margin for lazy clause exchange. Each thread uses different random seeds to hold the diversity of solvers. We submit ManyGlucose 4.1-2 with 24 threads and with 48 threads to Parallel track.

ACKNOWLEDGMENT

This research is supported in part by Grant-in-Aid for Scientific Research (No. 17K00300 and 16H02803) from Japan Society for the Promotion of Science.

REFERENCES


Maple_LCM_M1 and Maple_LCM+BCrestart

Zhen Li  
School of Computer Science & Technology  
Huazhong University of Science & Technology  
Wuhan, China  
violetcrestfall@hotmail.com

Kun He  
School of Computer Science & Technology  
Huazhong University of Science & Technology  
Wuhan, China  
brooklet60@hust.edu.cn

Abstract—Maple_LCM_M1 is a patched Maple_LCM solver by slightly modifying the conflict analyzing function. Maple_LCM+BCrestart adds a new restarting strategy to the Maple_LCM solver. Maple_LCM+BCrestart_M1 combines both patches in one solver.

I. INTRODUCTION

The Maple_LCM solver submitted to the SAT Competition 2017, developed by Chu-Min Li's team, won the first prize of the 2017 competition. Our team is trying to improve its performance by applying small patches to the original solver. In Maple_LCM_M1, one of the new versions, we adjusted the coefficients of activity bumping process for each variable involved in conflicts. The farther the clause containing the variable from the conflict clause, the less activity it will gain. In Maple_LCM_BCrestart, we present a new restarting strategy that the solver will restart once the average conflicts per decision level reach threshold. The third new solver Maple_LCM+BCrestart_M1 combines the modifications in the two new solvers.

II. MAPLE_LCM_M1

When a conflict is reached, the original behavior of Maple_LCM is to bump the activity of all variables involved in the conflict 0.5 times var_inc. In the Maple_LCM_M1 version, the multiplier 0.5 varies from clause to clause. The initial value of the multiplier is 0.5 when processing variables of the initial conflict clause. It will be assigned to 0.9 times itself at each round of iteration during the learnt clause generation process.

We think the modification makes sense, as the farther an assignment of a variable in the conflict, the less impact it has to contribute the conflict. And, the multiplier decaying rate might be adjusted for a better performance.

III. MAPLE_LCM+BCRESTART

Maple_LCM+BCrestart is also based on Maple_LCM. It calculates the total conflicts reached after the last restart and checks the average conflicts per decision level. If the average conflicts reaches some threshold, we will run a restart. Maple_LCM+BCrestart divide the search into three phases. The first phase let the search to learn from the conflicts, then in the second phase, Maple_LCM+BCrestart will restart when the average of conflicts per decision level reach a suit threshold. At last, when the search comes to the leaves of search tree, reduce the restarts. The value of average conflicts illustrates the impact of a decision variable on constructing new conflicts. The search tree is better when it is shorter and more balanced. Given the limit amount of learnt information, this number can be used to determine whether the current search space is worth for keeping the current searching process.

IV. MAPLE_LCM+BCRESTART_M1

This solver just combines the modifications in Maple_LCM_M1 and Maple_LCM+BCrestart.

REFERENCES

Maple_LCM_Dist_ChronoBT: Featuring Chronological Backtracking

Vadim Ryvchin and Alexander Nadel
Intel Corporation, P.O. Box 1659, Haifa 31015 Israel
Email: {vadim.ryvchin,alexander.nadel}@intel.com

Abstract—This is the system description of the solver Maple_LCM_Dist_ChronoBT, submitted to the SAT Competition 2018. We have integrated chronological backtracking [3] into the SAT Competition 2017 [1] winner, Maple_LCM_Dist [2].

I. DESCRIPTION

The goal of our submission is to test the performance of Chronological Backtracking (CB) [3] in the settings of the latest SAT competition. In our solver Maple_LCM_Dist_ChronoBT, we updated Maple_LCM_Dist [2] with CB (configuration {T = 100, C = 4000}) based on the results in [3].

REFERENCES

Maple_LCM_Scavel and Maple_LCM_Scavel_200

1st Yang Xu, 4th Shuwei Chen
School of Mathematics
National-Local Joint Engineering Laboratory of System Credibility
Automatic Verification, Southwest Jiaotong University
Chengdu, China
xuyang@home.swjtu.edu.cn, swchen@home.swjtu.edu.cn

2nd Guanfeng Wu, 3rd Qingshan Chen
School of Information Science and Technology
National-Local Joint Engineering Laboratory of System Credibility
Automatic Verification, Southwest Jiaotong University
Chengdu, China
wl520gx@gmail.com, qschen@home.swjtu.edu.cn

Abstract—This document describes Maple_LCM_Scavel SAT solver.

I. INTRODUCTION

Maple_LCM_Scavel is an improved SAT solver based on Maple_LCM_Dist. The improvements are made mainly from the following two aspects. First, we changed the evaluation method of VSIDS, so that the better branch decision variable can be selected. The second is that we improved the evaluation method of learning clauses so that better clauses can be preserved.

II. TWO IMPROVEMENT

A. Dynamic comprehensive variable activity evaluation

LIANG et al. [1] analyzed the role of the number of conflicts in the global conflict, where the variable was used in the latest conflict analysis, and proposed the Conflict History-based Branching strategy (CHB). The increment of variable activity is calculated based on the formula (1).

\[ s' = (1 - \alpha) \cdot s + \alpha \cdot r \]  

(1)

Where \( \alpha \) is the interval of increment whose initial value is 0.4 and the attenuation rate is 10\(^{-6}\) after each conflict, \( r \) is the reward value. The variable is used for conflict analysis and derivation of learning clauses, which should be endowed dynamic activity. Therefore, we put forward a heuristic branch decision algorithm that provides a comprehensive evaluation of the decision level of the effective decision variable and its role in the conflict, called as Dynamic Activity algorithm (DA)[2]. Its framework is similar to that of EVSIDS, and is expressed as follows.

DA set an activity counter ‘s’ for every variable, the increment of ‘s’ can be calculated as (3).

\[ s' = s + W(v) \cdot f \quad (0 < f < 1) \]  

(2)

Where \( f \) is the increment factor, \( W \) is a function of decision level and conflict level, and is defined as (4).

\[ W(v) = \alpha \cdot \frac{v_{Level}}{n_{Levels}} + (1 - \alpha) \cdot \frac{v_{Conflict}}{n_{Conflicts}} \quad (0 \leq \alpha \leq 1) \]  

(3)

Where \( v_{Level} \) is the latest decision level when the variable is treated as decision variable, \( n_{Levels} \) is the current decision level, \( v_{Conflict} \) is the total number of conflicts when taking this decision variable, \( n_{Conflicts} \) is the global number of conflicts. \( \alpha \) is the regulation factor, and normally takes the value of 0.7. The \( v_{Level} \) of a variable is set to 0 when its value is unassigned. If a variable is in the learn clauses or is used in the conflict analysis, its activity will be updated automatically.

B. Improved learning clause management strategy

Both the activity evaluation and LBD evaluation adopted a more aggressive deletion strategy. According to the experiments, LBD[3] or Activity[4] based solvers continually produce a large number of learning clauses, at the same time delete clauses in frequently. The clauses which have high priority to be deleted have less opportunity to be used in conflict analysis. In other words, the probability of the good clauses to be preserved will usually become low. It can be inferred that the probability of a clause being deleted is closely related to the number of times it is used.

Maple_LCM_Scavel sets a threshold for the number of the learning clauses being used in conflict analysis, \( N_{Based} \), named “\( NB\_threshold\_value \)”. The initial value of \( N_{Based} \) of each learned clauses is 0. If the learned clause is used in conflict analysis, increase the value of \( N_{Based} \) by 1. In the compression procedure (in “reducedDB” method), when the clause’s \( N_{Based} \) value reaches the specified threshold \( NB\_threshold\_value \), it will be deleted. The experimental results show that, the threshold value being 150 or 200 will get best results.

REFERENCES


MapleCOMSPS_LRB_VSIDS and MapleCOMSPS_CHB_VSIDS in the 2018 Competition

Jia Hui Liang∗‡, Chanseok Oh†‡, Krzysztof Czarnecki*, Pascal Poupart*, Vijay Ganesh*
* University of Waterloo, Waterloo, Canada
† Google, New York, United States
‡ Joint first authors

Abstract—This document describes the SAT solvers MapleCOMSPS_LRB_VSIDS and MapleCOMSPS_CHB_VSIDS that implement our machine learning branching heuristics called the learning rate branching heuristic (LRB) and the conflict history-based branching heuristic (CHB).

I. INTRODUCTION

A good branching heuristic is vital to the performance of a SAT solver. Glancing at the results of the previous competitions, it is clear that the VSIDS branching heuristic is the de facto branching heuristic among the top performing solvers. We are submitting two unique solvers with a new branching heuristic called the learning rate branching heuristic (LRB) [1] and another solver with the conflict history-based branching heuristic (CHB) [2].

Our intuition is that SAT solvers need to prune the search space as quickly as possible, or more specifically, learn a high quantity of high quality learnt clauses. In this perspective, branching heuristics can be viewed as a bi-objective problem to select the branching variables that will simultaneously maximize both the quantity and quality of the learnt clauses generated. To simplify the optimization, we assumed that the first-UIP clause learning scheme will generate good quality learnt clauses. Thus we reduced the two objectives down to just one, that is, we attempt to maximize the quantity of learnt clauses.

II. LEARNING RATE BRANCING

We define a concept called learning rate to measure the quantity of learnt clauses generated by each variable. The learning rate is defined as the following conditional probability, see our SAT 2016 paper for a detailed description [1].

\[ \text{learningRate}(x) = \mathbb{P}(\text{Participates}(x) | \text{Assigned}(x) \land \text{SolverInConflict}) \]

If the learning rate of every variable was known, then the branching heuristic should branch on the variable with the highest learning rate. The learning rate is too difficult and too expensive to compute at each branching, so we cheaply estimate the learning rate using multi-armed bandits, a special class of reinforcement learning. Essentially, we observe the number of learnt clauses each variable participates in generating, under the condition that the variable is assigned and the solver is in conflict. These observations are averaged using an exponential moving average to estimate the current learning rate of each variable. This is implemented using the well-known exponential recency weighted average algorithm for multi-armed bandits [3] with learning rate as the reward.

Lastly, we extended the algorithm with two new ideas. The first extension is to encourage branching on variables that occur frequently on the reason side of the conflict analysis and adjacent to the learnt clause during conflict analysis. The second extension is to encourage locality of the branching heuristic [4] by decaying unplayed arms, similar to the decay reinforcement model [5], [6]. We call the final branching heuristic with these two extensions the learning rate branching heuristic.

III. CONFLICT HISTORY-BASED BRANCING

The conflict history-based branching heuristic (CHB) precedes our LRB work. CHB also applies the exponential recency weighted average algorithm where the reward is the reciprocal of the number of conflicts since the assigned variable last participated in generating a learnt clause. See our paper for more details [2].

IV. SOLVERS

All the solvers are modifications of COMiniSatPS [7]. We used the same COMiniSatPS version that also participates in the competition [8]. This year’s solvers are basically identical to the last year’s solvers, fixing only a few minor bugs:

- Fixed the bug in the last year’s MapleCOMSPS_LRB_VSIDS [9] that crippled the on-the-fly probing techniques.
- Fixes two minor bugs in the base solver COMiniSatPS [8].

V. AVAILABILITY AND LICENSE

Source is available for download for all the versions described in this paper. All the solvers use the same license as COMiniSatPS. Note that the license of the M4RI library (which COMiniSatPS uses to implement Gaussian elimination) is GPLv2+. 
ACKNOWLEDGMENT

We thank specifically the authors of Glucose, GlueMiniSat, Lingeling, CryptoMiniSat, and MiniSat.

REFERENCES


painless-mcomsps and painless-mcomsps-sym

Ludovic Le Frioux*†, Hakan Metin†, Souheib Baarir*† ‡, Maximilien Colange*, Julien Sopena†§, Fabrice Kordon†

*LRDE, EPITA, Kremlin-Bicêtre, France
†Sorbonne Université, LIP6, CNRS, UMR 7606, Paris, France
‡Université Paris Nanterre, France
§INRIA, Delys Team, Paris, France

Abstract—This paper describes the solvers painless-mcomsps, and painless-mcomsps-sym submitted to the parallel track of the SAT Competition in 2018. They are parallel solvers instantiated with PArallel INStantiable LE Sat Solver (PaInleSS) framework and using MapleCOMPS as core sequential solver.

I. INTRODUCTION

painless-mcomsps and painless-mcomsps-sym are parallel SAT solvers built by instantiating components of the PaInleSS parallel framework [1]. They are Portfolio based solvers implementing a diversification strategy, fine control of learnt clause exchanges, and using MapleCOMPS [2] as a core sequential solver. Moreover, painless-mcomsps-sym included dynamic symmetry breaking [3] by using the Cosy library.

Section II gives an overview on PaInleSS framework. Section III details the implementation of painless-mcomsps using PaInleSS and MapleCOMPS. Section IV explains how dynamic symmetry breaking has been incorporated in painless-mcomsps to give the solver painless-mcomsps-sym.

II. DESCRIPTION OF PAINLESS

PaInleSS is a framework that aims at simplifying the implementation and evaluation of parallel SAT solvers for many-core environments. Thanks to its genericity and modularity, the components of PaInleSS can be instantiated independently to produce new complete solvers.

The main idea of the framework is to separate the technical components (e.g., those dedicated to the management of concurrent programming aspects) from those implementing heuristics and optimizations embedded in a parallel SAT solver. Hence, the developer of a (new) parallel solver concentrates his efforts on the functional aspects, namely parallelization and sharing strategies, thus delegating implementation issues (e.g., data concurrent access protection mechanisms) to the framework.

Three main components arise when treating parallel SAT solvers: sequential engines, parallelization, and sharing. These form the global architecture of PaInleSS.

A. Sequential Engines

The core element that we consider in our framework is a sequential SAT solver. This can be any CDCL state-of-the-art solver. Technically, these engines are operated through a generic interface providing basics of sequential solvers: solve, interrupt, add clauses, etc.

Thus, to instantiate PaInleSS with a particular solver, one needs to implement the interface according this engine.

B. Parallelization

To built a parallel solver using the aforementioned engines, one needs to define and implement a parallelization strategy. Portfolio and Divide-and-Conquer are the basic known ones. Also, they can be arbitrary composed to form new strategies.

In PaInleSS, a strategy is represented by a tree-structure of arbitrary depth. The internal nodes of the tree represent parallelization strategies, and leaves are core engines. Technically, the internal nodes are implemented using WorkingStrategy component and the leaves are instances of SequentialWorker component.

Hence, to develop its own parallelization strategy, the user should create one or more strategies, and build the required tree-structure.

C. Sharing

In parallel SAT solving, the exchange of learnt clauses warrants a particular focus. Indeed, beside the theoretical aspects, a bad implementation of a good sharing strategy may dramatically impact the solver’s efficiency.

In PaInleSS, solvers can export (import) clauses to (from) the others during the resolution process. Technically, this is done by using lockfree queues [4]. The sharing of these learnt clauses is dedicated to particular components called Sharers. Each Sharer in charge of sets of producers and consumers and its behaviour reduces to a loop of sleeping and exchange phases.

Hence, the only part requiring a particular implementation is the exchange phase, that is user defined.

III. PAINLESS-MCOMSPS

This section describes the overall behaviour of our competing instantiation named painless-mcomsps. Its architecture is highlighted in Fig. 1.
IV. PAINLESS-MCOMSPS-SYM

This section describes the overall behaviour of our competing instantiation named painless-mcomspssym. Its architecture is highlighted in Fig. 1.

A. Dynamic Symmetry Breaking

The idea we bring is to break symmetries on the fly: when the current partial assignment cannot be a prefix of a leader (of an orbit), a constraint called esbp is generated. This constraint prunes this forbidden assignment and all its extensions.

B. Integration to painless-mcomspssym

Cosy, a C++ library, provides dynamic symmetry breaking primitives. We integrated the library into MapleCOMSPS, and we added a parameter to activate or not dynamic symmetry breaking mode.

In painless-mcomspssym, there is only one solver that used dynamic symmetry breaking, we call it S in the Fig. 1. This solver uses the VSIDS heuristics.

The solver S receives clauses from all the others, but it does not export clauses.

ACKNOWLEDGMENT

We would like to thank Jia Hui Liang, Chanseok Oh, Vijay Ganesh, Krzysztof Czarnecki, and Pascal Poupart, the authors of MapleCOMSPS.

REFERENCES


Abstract—We describe some details about the SLS solver probSAT, a simple and elegant SLS solver based on probability distributions, a heuristic first presented in the SLS solver Sparrow [3].

I. INTRODUCTION

The probSAT solver is an efficient implementation of the probSAT algorithm presented in [2] with slightly different parameterization and implementations.

II. MAIN TECHNIQUES

The probSAT solver is a pure stochastic local search solver based on the following algorithm:

Algorithm 1: ProbSAT
Input : Formula F, maxTries, maxFlips
Output: satisfying assignment a or UNKNOWN

1 for i = 1 to maxTries do
2   a ← randomly generated assignment
3 for j = 1 to maxFlips do
4     if (a is model for F) then
5         return a
6     Cu ← randomly selected unsat clause
7     for x in Cu do
8         compute f(x, a)
9     var ← random variable x according to probability \( \sum_{z \in Cu} f(z, a) \)
10    flip(var)
11 return UNKNOWN;

ProbSAT uses only the break values of a variable in the probability functions \( f(x, a) \), which can have an exponential or a polynomial shape as listed below.

\[
\begin{align*}
\text{k} & \quad \text{fct} & \quad \text{cb} & \quad \text{epsilon} \\
3 & 0 & 2.06 & 0.9 \\
4 & 1 & 3 & - \\
5 & 1 & 3.88 & - \\
6 & 1 & 4.6 & - \\
\geq 7 & 1 & 4.6 & - \\
\end{align*}
\]

where \( k \) is the size of the longest clause found in the problem during parsing. The parameters of probSAT have been found using automated tuning procedures included in the EDACC framework [1].

IV. FURTHER DETAILS

ProbSAT is implemented in C and uses a new XOR implementation scheme for the flip procedure described in detail in [4].

ACKNOWLEDGMENT

This project was funded by the Deutsche Forschungsgemeinschaft (DFG) under the number SCHO 3029-1.

REFERENCES

Riss 7.1

Norbert Manthey
nmanthey@conp-solutions.com
Dresden, Germany

Abstract—The sequential SAT solver Riss combines a heavily modified Minisat-style solving engine of GLUCOSE 2.2 with a state-of-the-art preprocessor COPROCESSOR and adds many modifications to the search process. Riss allows to use preprocessing based on COPROCESSOR. Based on this Riss, we create a parallel portfolio solver PRISS, which allows clause sharing among the incarnations, as well as sharing information about equivalent literals.

I. INTRODUCTION

The CDCL solver Riss is a highly configurable SAT solver based on MINISAT [1] and GLUCOSE 2.2 [2], [3], implemented in C++. Many search algorithm extensions have been added, and Riss is equipped with the preprocessor COPROCESSOR [4]. Furthermore, Riss supports automated configuration selection based on CNF formulas features, emitting DRAT proofs for many techniques and comments why proof extensions are made, and incremental solving. The solver is continuously tested for being able to build, correctly solve CNFs with several configurations, and compile against the IPASIR interface. For automated configuration, Riss is also able to emit its parameter specification on a detail level specified by the user. The repository of the solver provides a basic tutorial on how it can be used, and the solver provides parameters that allow to emit detailed information about the executed algorithm in case it is compiled in debug mode (look for “debug in the help output”). While Riss also implements model enumeration, parallel solving, and parallel model enumeration, this document focusses only on the differences to Riss 7, which has been submitted to SAT Competition 2017.

II. SAT COMPETITION SPECIFICS

While last years submissions did not make full use of the implemented formula simplification techniques, the configuration submitted to the NoLimit track of the competition now uses XOR reasoning [5] and cardinality reasoning [6] again. These techniques have been disabled last year, as they cannot print DRAT proofs efficiently.

III. MODIFICATIONS OF THE SEARCH - LCM

Last years winning solver family was the “Maple LCM” solvers, which are based on learned clause minimization (LCM) [7]. The implementation of LCM in Riss has a few modifications to those in the original publication, namely:

1) apply LCM after every second reduction
2) when simplifying a clause, try to simplify it in reverse order as well
3) when a clause could be reduced, use a resolution based simplification to reduce the size further

The first modification helps to reduce the overhead LCM might introduce on clauses that would be removed in the next clause removal phase. The second modifications uses a Bloom filter like effect following the assumption: if a clause can be reduced by performing vivification in one particular order, then using the reverse order might allow to drop even more literals. On the other hand, for clauses that cannot be reduced, no additional cost is introduced. Hence, the second modification focusses on clauses that can be reduced. Finally, the last modification makes the reduction more effective: while vivification stops when a conflict is found and proceeds with the current set of literals, applying conflict analysis with resolution allows to remove further redundant literals. Cycles for reduction are only spend on clauses that could be reduced in the first place.

IV. MODIFICATIONS OF THE SIMPLIFIER

To be able to emit DRAT proofs for the main track, many simplification techniques of Coprocessor had to be disabled, among them reasoning with XORs and cardinality constraints [6]. In the NoLimit track, these techniques are enabled again.

V. AVAILABILITY

The source of the solver is publicly available under the LGPL v2 license at https://github.com/conp-solutions/riss. The version with the git tag “v7.1.0” is used for the submission. The submitted starexec package can be reproduced by running “/scripts/make-starexec.sh” on this commit.

ACKNOWLEDGMENT

The author would like to thank the developers of GLUCOSE 2.2 and MINISAT 2.2. The development of this project was supported by the DFG grant HO 1294/11-1.

REFERENCES


Abstract—SparrowToRiss is a combination of the solver Sparrow and Riss. SparrowToRiss is first trying to solve the problem with Sparrow, limiting its execution to $5 \cdot 10^8$ flips and then passes the assignment found to the CDCL solver Riss, which uses this information for initialization and then tries to solve the problem.

The SLS solver Sparrow is the same version as used in 2014. The solver Riss is used in version 7.1, which is also submitted to this competition.

I. INTRODUCTION

While in 2014, where this solver combination was submitted for the first time, the benchmark was split in industrial and combinatorial families, recent competitions did not insist on this split any more. As we do not know the origin of submitted and used benchmarks for this years competition, and did not see solver from previous competitions that performed well on crafted formulas, we submit SparrowToRiss again, to validate whether the combination of a SLS and a CDCL solver is still a reasonable solving approach. The submission is furthermore supported by the results found in [1], where formula simplification is used to boost the efficiency of SLS solvers on crafted families. The best found technique together with Sparrow represents the basis of our solver Sparrow+CP3. As SLS solvers cannot show unsatisfiability, we run a CDCL solver after a fixed amount of $5 \cdot 10^8$ flips, so that the overall solver behavior stays deterministic.

II. MAIN TECHNIQUES

Sparrow is a clause weighting SLS solvers that uses promising variables and probability distribution based selection heuristics. It is described in detail in [2]. Compared to the original version, the one submitted here is updating weights of unsatisfied clauses in every step where no promising variable can be found.

The built-in preprocessor CP3 is an extension of Coprocessor 2 [20], and received updates. Compared to the submitted version of Riss to the SAT competition 2017, no new techniques have been added.

The CDCL solver Riss uses the Minisat search engine [17], more specifically the extensions added in Glucose 2.2 [18], [19]. Furthermore, Riss is equipped with the preprocessor Coprocessor.

The combination of the Sparrow and Riss, called SparrowToRiss, does not simply execute the two solvers after each other, but also forwards information from the SLS solver to the CDCL solver: when Sparrow terminates, it outputs its last full assignment in chronological order (i.e. the oldest variable first), which is used to initialize the phase saving of Riss, such that the first decisions of Riss follow this assignment. In a brief empirical evaluation this communication turned out to be useful. The solvers are also able to forward the information about the age of the variables in the SLS search. This data could be used to initialize the activities of the variables inside Riss. However, this feature is not enabled in the used configuration.

III. MAIN PARAMETERS

Sparrow is using the same parameters as Sparrow 2011. The configuration of CP3 has been tuned for Sparrow in [1] on the SAT Challenge 2012 satisfiable hard combinatorial benchmarks. The configuration used in 2018 is the same configuration used in the version of 2014.

The main parameters of Riss control how the formula simplification of CP3 is executed. A major modification of Riss is the addition learned clause minimization [7], which has been slightly modified. The configuration of CP3 has been tuned for Glucose 2.2 in [1] on the SAT Challenge 2012 application benchmark. The final setup of the preprocessor inside Riss uses the following techniques: UP, SUB+STR (producing all resolvents for ternary clauses), Unhide without hidden literal elimination [10] and 5 iterations, BVE without on the fly BCE. Furthermore, if no proof should be emitted, Gaussian Elimination and Cardinality Constraint [21] reasoning is applied, as well as Covered Literal Elimination [22].

For SparrowToRiss it can be chosen whether to forward the last assignment, or the activity information.

IV. IMPLEMENTATION DETAILS

Sparrow is implemented in C. The solver Riss is build on top of Minisat 2.2 and Glucose 2.2, and is implemented in C++.

V. AVAILABILITY

The source code of Riss (including CP3) is available at https://github.com/conp-solutions/riss under LGPL v2.1. Sparrow is available at https://github.com/adrianopolus/Sparrow.

REFERENCES


BreakIDGlucose and BreakIDGlucoseSEL

Jo Devriendt
University of Leuven
Leuven, Belgium

Bart Bogaerts
University of Leuven
Leuven, Belgium

Abstract—BreakIDGlucose and BreakIDGlucoseSEL combine the Glucose SAT solver with the symmetry detection tool BreakID. The former breaks symmetry statically by employing classic symmetry breaking formulas, while the latter handles it dynamically by employing symmetric explanation learning.

I. INTRODUCTION

Many real-world problems exhibit symmetry, but the SAT competition and SAT race seldomly feature solvers who are able to exploit symmetry properties. Similarly to 2013 and 2016, we submit a static symmetry breaking approach BreakIDGlucose. This year, we also submit the dynamic symmetry handling approach BreakIDGlucoseSEL.

II. MAIN TECHNIQUES

As symmetry breaking preprocessor we use BreakID 2.3 [1]. Compared to BreakID 2.2 used in 2016, version 2.3 fixes a small bug in the automorphism graph construction routine. As SAT-solver we employ Glucose 4.0 [2] which is modified to support symmetric explanation learning (SEL) [3].¹ The difference between the approaches is that BreakIDGlucose employs BreakID’s symmetry breaking formulas and disables SEL, while BreakIDGlucoseSEL has SEL activated and forgoes BreakID’s symmetry breaking formulas.

The reason we use the modified version of Glucose as backend for BreakIDGlucose is that we want to compare both static and dynamic versions of symmetry handling as equally as possible, which requires exactly the same backend solver.

III. MAIN PARAMETERS

The main user-provided parameters control:

- How much time is allocated to symmetry detection. The builtin graph automorphism tool Saucy [4] gets 100 seconds to detect symmetry generators.
- How large the symmetry breaking formulas are allowed to grow, measured in the number of auxiliary variables introduced by a symmetry breaking formula. We limit this to 50 auxiliary variables.
- How many generators BreakIDGlucoseSEL uses to handle row interchangeability symmetry groups. We employ a quadratic number of row-swaps (e.g., swapping every two pigeons of a pigeonhole problem). The alternative would have been a linear amount of swaps (e.g., swapping every two consecutive pigeons of a pigeonhole problem).

IV. SPECIAL ALGORITHMS, DATA STRUCTURES, AND OTHER FEATURES

BreakIDGlucoseSEL employs a second symmetrical clause store for clauses symmetrical to the ones that are asserting in the current search state. These symmetrical clauses σ(c) are added to the main learned clause store only when they become unit or conflicting, and otherwise are quickly forgotten after a backjump causes the original clause c to revert to non-unit status. As usual, a two-watched literal scheme keeps track of the truth value of any clause in the symmetrical clause store.

V. SAT COMPETITION 2018 SPECIFICS

BreakIDGlucose and BreakIDGlucoseSEL participate in the No-Limit track since BreakIDGlucoseSEL constructs proofs using a symmetry rule not present in the DRAT format.

VI. AVAILABILITY

Source code and documentation for BreakID is available under a non-commercial license [5]. Source code and documentation for the extension of Glucose with symmetric explanation learning is freely available [6].

ACKNOWLEDGMENT

We would like to thank the authors of Saucy, MiniSat [7] and Glucose.

REFERENCES

Abstract—inIDGlucose (pronounced “init-glucose”) is a submission to the Glucose Hack Track that initializes variable activity and phase based on a weighted literal occurrence count on the original CNF.

I. Main Technique

The activity of a variable is the priority given to the variable when the solver selects a variable to decide. The phase of a variable is the value assigned to it when the solver decides the variable.

MiniSat-based solvers have an initial phase of false and an initial activity of 0 for all variables. As this does not take any information from the instance at hand into account, we hope to improve upon this by counting the occurrences of each literal in a clause the CNF, taking into account clause length as well.

The general idea behind our activity initialization is that variables whose literals occur both positively and negatively in short clauses are probably hard to decide a good value for, and might be part of a lot of failing search branches. As activity tracks the amount of conflicts a variable is part of, such variables seem good candidates to start out with a high activity.

The general idea behind our phase initialization is that we want to maximize the number of satisfied clauses by assigning true or false to a variable, since this might speed up the derivation of satisfying assignments. Variables that occur mostly positively (resp. negatively) then should be assigned true (resp. false). Generally speaking, long clauses are easier to satisfy than short ones, so again the occurrence count of literals in short clauses should be more important than in long clauses.

II. Special Algorithms, Data Structures, and Other Features

We weigh the occurrence of a literal $l$ by the inverse square of the length of the clause $c$, as this strongly reduces the importance of occurrences in long clauses. The total weighted occurrence ($two$) for a literal then is:

$$two(l) = \sum_{c \text{ s.t. } l \in c} \frac{1}{\text{length}(c)^2}$$

The initial activity for a variable $v$ simply is the product

$$two(v) \cdot two(\bar{v})$$

which is a measure for both the total weighted occurrence of a variable as well as the difference between the positive and negative occurrences of the variable. E.g., a variable $v$ occurring as a pure literal will have an initial activity of 0, as either $two(v)$ or $two(\bar{v})$ will be 0.

The initial phase for a variable $v$ is the truth value of

$$two(v) > two(\bar{v})$$

which measures whether the variable occurs mostly positively or mostly negatively. Note that in case of a tie, the initial phase is negative.

Note that the activity and phase are only initialized with the above values. During search, Glucose runs its customary phase caching and activity updating schemes.
Abstract—We provide in this paper a short description of our solvers SCaLoPe, PeneLoPe_MDLC and GLUCOSE-3.0_PADC submitted to the SC18. The first solver SCaLoPe is a simplified implementation of the one we submitted to the SC17 [1]. The second solver PeneLoPe_MDLC is a new one also based on PeneLoPe 2014 [2] which aims to minimize clause duplication by using two learned clause databases: one for the clauses derived from conflicts analysis and the other for the imported clauses. This latter is subject to a special cleaning strategy since it can contain duplicate or already subsumed clauses. The third solver GLUCOSE-3.0_PADC is a GLUCOSE3.0 hack with periodical steps of deep database cleaning.

I. INTRODUCTION

SAT solvers have become very efficient today. This efficiency is the result of a subtle combination of several features such as unit propagation through watched literals, restarts strategies, dynamic branching and polarity heuristics, conflict clause analysis and clause learning. This latter is one of the most important features of modern SAT solvers. In fact after each found conflict, a procedure is invoked to analyze it in order to produce an asserting clause. This asserting clause is then added to a learned clause database (i.e. this clause is learned) and is used to redirect the back-jumping level as well as to prevent the same conflict in the future as long as this clause is kept in the learned clause database. However during the search there can be a huge amount of learned and kept all these clauses will have a negative impact on the performances of the solver by slowing down unit propagations. In parallel solvers, this issue is even more important since beside learned clause there are also imported clauses coming from other threads of the portfolio that are also added to the learned clause database. To cope with this issue, modern SAT solvers use learned clause database cleaning strategies where the common clause quality measures are the size or the LBD of the clause. Our solvers SCaLoPe, PeneLoPe_MDLC and GLUCOSE-3.0_PADC presented in this paper mainly use several forms of learned clause database management to achieve different objectives. These Solvers are described in details in the subsequent sections.

II. SCaLoPe

The solver SCaLoPe [1] that we submitted to SC17 was designed to improve the scalability of the portfolio by organizing threads in teams. It allowed an intensive communication between threads of the same team and limited communication between threads of different teams. This version used an explicit division of thread in team by providing to each team a different cooperation object. Hence there were in each team a particular thread that was in charge of exporting and importing information from and to other teams. Once the information were imported by this thread, the others could simply retrieve them via classical communications within the team. This year, we changed this implementation. Instead, we use one Cooperation object and dedicated channels between each pair of threads in order to share learned clauses as in PeneLoPe. However the amount of clauses shared via a channel depends on the proximity of the threads i.e. whether they are in the same team or not. This simplification really reduce the complexity of the code. Furthermore inter-team communication is no longer restricted to unit clauses sharing as it was in the previous version. We now use the LBD [3] as a discriminant when sharing clauses. Hence only clauses that have good LBD scores are exported outside the team. In the current version, every learned clause with the LBD score lower than or equal to 2 is exported outside the teams.

III. PeneLoPe_MDLC

PeneLoPe_MDLC is a built on top of PeneLoPe [2]. Unlike classical SAT solvers, PeneLoPe_MDLC manages two separate learned clause databases per thread: one for the clauses that are derived from conflicts by the thread itself and the other for the clauses that were imported from other threads. PeneLoPe_MDLC allows the move of a clause from the imported clause database to the learned clause database where the lifetime is greater for good quality clauses. Let \( \Delta \) be the learned clause database and \( \Gamma \) the imported clause database. A clause \( c \) is moved from \( \Gamma \) to \( \Delta \) whenever it propagates or is conflicting while none clause in \( \Delta \) could do so. Hence no clause in \( \Delta \cup \Gamma \) will be duplicated nor subsumed by another one. We could then apply different cleaning strategy on \( \Delta \) and \( \Gamma \). As such, it might be possible to have a more aggressive strategy in \( \Gamma \) since it can possibly contains duplicated or subsumed clauses. In the current version of PeneLoPe_MDLC, \( \Gamma \) is implemented as a Queue which means that the FIFO strategy is used to manage the imported clause database. At each clause database cleaning step, a fraction (which we called \( \text{reduceFactor} \) and which is such that \( 0 \leq \text{reduceFactor} \leq 1 \)) of the total number of clauses to delete is removed from \( \Delta \) on the less useful clauses and the rest is removed from \( \Gamma \) on the oldest clauses. The rationale is...
that if a clause has not propagated any literal since its entrance in the queue, there can be more chances that the literals it could propagate are already propagated by other clauses present in \( F \cup \Delta \). This can be the case when the clause is duplicated or is subsumed by another one. An Exception is however made on clauses with \( \text{LBD} \leq 2 \). These latter are given a longer lifetime in \( \Gamma \). Concretely, just before insertion in \( \Gamma \) during imports, each clause with \( \text{LBD} \leq 2 \) is given a maximum number of deletion attempts — which we call its lifetime — before definitely delete it. After each unsuccessful deletion attempt the clause is reinserted in the queue while decreasing its lifetime. This procedure is sketched in algorithm 1. Algorithm 2 as far as it is concerned, describes the propagation phase of each solver in PENELOPE_MDLC. PENELOPE_MDLC also includes some optimization techniques such as learned clause minimization with binary clause resolution. It also uses a special data structure to handle binary clauses — as in GLUCOSE — in order to check them first during unit propagations.

### Algorithm 1: Reduce DB

**Input**: The Learned clause database \( \Delta \) and the imported clause database \( \Gamma \)

```
begin
  totalClauseToDelete := (|\( \Delta \)| + |\( \Gamma \)|)/2;
  nLt := totalClauseToDelete + reduceFactor;
  nImp := totalClauseToDelete - nLt;
  if |\( \Gamma \)| < nImp then
    nImp := |\( \Gamma \)|;
    nLt := totalClauseToDelete - nImp;
  for (i := 0; i < nImp ; i++) do
    if lbd(c) > 2 or c.lifetime \leq 0 then
      delete(c);
    else
      \( \Gamma \).insert(c);
      c.lifetime --;
  sort \( \Delta \) according to clauses’ LBD;
  remove from \( \Delta \) the nLt clauses with bad LBD;
end
```

### Algorithm 2: Propagation phase

**Input**: A formula \( F \), a learned clause database \( \Delta \) and the imported clause database \( \Gamma \)

```
begin
  repeat
    while \( \exists \) a unit or a falsified clause \( c \) in \( F \cup \Delta \) do
      if \( c \) is unit then
        add the unassigned literal of \( c \) to the current interpretation;
      else
        return \( c \);
      if \( \exists \) a unit or conflicting clause \( c \in \Gamma \) then
        move \( c \) to \( \Delta \);
    until no new clause has been moved to \( \Delta \);
  return noConflictClauseFound;
end
```

PENELOPE_MDLC is concerned, we tune it to use the reduceFactor of 0.6 and each imported clause with \( \text{LBD} \leq 2 \) was given a lifetime of 1.

## VI. ACKNOWLEDGEMENT

We would like to thanks authors of MiniSat[5] , GLUCOSE[3], ManySat[6] and PENELOPE[2] for the great work they have done.

## REFERENCES

Maple_CM, Maple_CM_Dist, Maple_CM_ordUIP and Maple_CM_ordUIP+ in the SAT Competition 2018

Mao Luo¹, Fan Xiao¹, Chu-Min Li², Felip Manyá³, Zhipeng Liu¹, Yu Li²
¹School of Computer Science, Huazhong University of Science and Technology, Wuhan, China
{maoluo, fanxiao, zhipeng.lv}@hust.edu.cn
²MIS, University of Picardie Jules Verne, Amiens, France
{chu-min.li, yu.li}@u-picardie.fr, corresponding authors
³Artificial Intelligence Research Institute (IIIA-CSIC), Barcelona, Spain. felip@iiia.csic.es

1. Introduction

The CDCL SAT solver Maple_LCM won the gold medal of the main track of the SAT Competition 2017. It was implemented on top of the solver MapleCOMPS_DRU [1], [2] by integrating the effective learnt clause minimization approach described in [3]. In the current competition, we propose Maple_CM, which is a new solver on top of Maple_LCM that extends clause minimization to original DRUP. Maple_CM_Dist is Maple_CM but reorders the first UIPs in conflicts selected under specific conditions; and Maple_CM_ordUIP+ is Maple_CM_ordUIP but selects the conflicts to reorder the first UIPs in function of the branching heuristics VSIDS and LRB. All these solvers are described in the remaining sections.

2. Clause Minimization in Maple_CM

Clause minimization based on unit propagation (UP) can be described as follows: Given a clause \( \neg l_1 \lor \neg l_2 \lor \cdots \lor \neg l_r \) that is derived by UP, \( \neg l_1 \lor \neg l_2 \lor \cdots \lor \neg l_r \) is the subset of literals in \( \{ \neg l_1, \neg l_2, \ldots, \neg l_r \} \) that are responsible of the conflict, we replace \( C \) by \( \{ l'_1 \lor l'_2 \lor \cdots l'_{r'} \} \). This clause minimization is not applied to every clause at every restart because it is costly. It works as follows in Maple_CM:

- During preprocessing, each original clause is minimized. The minimization process stops when the total number of unit propagations is greater than \( 10^8 \).
- During the search, Maple_CM organizes the learnt clauses in three sets as MapleCOMPS_DRU: CORE, TIER2 and LOCAL. The sets CORE and TIER2 roughly store the learnt clauses with LBD \( \leq 6 \), where LBD refers to the number of decision levels in a clause [4]. It also identifies a subset of original clauses called useful clauses that are used to derive at least one learnt clause of LBD \( \leq 20 \) since the last clause minimization. Then, before a restart, Maple_CM minimizes each clauses \( C \) such that function liveClause(\( C \)) returns true, provided that the number of clauses learnt since the last clause minimization is greater than or equal to \( \alpha + 2 \times \beta \times \sigma \), where \( \alpha = 1000 \) and \( \sigma \) is the number of minimizations executed so far. Algorithm 1 defines liveClause(\( C \)).

Algorithm 1: liveClause(\( C \))

<table>
<thead>
<tr>
<th>Line</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>begin</td>
</tr>
<tr>
<td>2</td>
<td>if ( C ) is a learnt clause then</td>
</tr>
<tr>
<td>3</td>
<td>if ( C ) is in CORE or TIER2 then</td>
</tr>
<tr>
<td>4</td>
<td>if ( C ) was never minimized, or the LBD of ( C ) is decreased 2 times since its last minimization, or the LBD of ( C ) is decreased to 1 since its last minimization then</td>
</tr>
<tr>
<td>5</td>
<td>[ Return true; ]</td>
</tr>
<tr>
<td>6</td>
<td>Return false;</td>
</tr>
<tr>
<td>7</td>
<td>else</td>
</tr>
<tr>
<td>8</td>
<td>if ( C ) was used to derive at least one learnt clause of LBD ( \leq 20 ) since the last clause minimization then</td>
</tr>
<tr>
<td>9</td>
<td>if ( C ) was never minimized during search, or the LBD of ( C ) is decreased 3 times since its last minimization, or the LBD of ( C ) is decreased to 1 since its last minimization then</td>
</tr>
<tr>
<td>10</td>
<td>[ Return true; ]</td>
</tr>
<tr>
<td>11</td>
<td>Return false;</td>
</tr>
<tr>
<td>12</td>
<td>end</td>
</tr>
</tbody>
</table>

The condition to select a restart for triggering a clause minimization process in Maple_CM is the same as in
Maple_LCM. However, Maple_CM selects the clauses to be minimized differently from Maple_LCM. First, Maple_LCM only minimizes the learnt clauses in CORE and TIER2, whereas Maple_CM also minimizes the useful original clauses, because original clauses can also contain redundant literals. Second, a learnt clause is minimized at most once in Maple_LCM, whereas a clause, either learnt or original, can be minimized more than once in Maple_CM under some conditions specified in terms of the decrease of its LBD.

The rationale behind the re-minimization of a clause is that further redundant literals can be detected, using unit propagation, after adding additional learnt clauses since its last minimization. Maple_CM re-minimizes a learnt (original) clause if its LBD was decreased two (three) times since its last minimization, because UP probably becomes more powerful in this case. The condition to re-minimize an original clause is stronger because an original clause presumably contains fewer redundant literals.

A particular case is a clause with LBD 1. This clause is probably very powerful in unit propagation and the LBD value cannot be decreased anymore. So, a clause will be re-minimized if its LBD becomes 1 since its last minimization, no matter how many times the LBD value was decreased.

3. Distance: A New Branching Heuristic Based on Implication Graphs in Maple_CM_Dist

Standard branching heuristics in CDCL solvers, such as VSIDS [5] and LRB [1], select a variable based on the behaviour of this variable in the past. However, at the beginning of search, very few things have happened. So, Maple_CM_Dist uses a new heuristic, different from VSIDS and LRB, at the beginning of the search.

Maple_CM_Dist constructs a complete implication graph for each conflict of the first 50,000 conflicts detected during the search. Each vertex in the implication graph corresponds to a variable and also to a clause. For each variable in the graph, we can collect the set \( D \) of clauses in all paths from the variable to the conflict. We then say that the variable depends on the clauses in \( D \) to contribute to the conflict. A variable depending on fewer clauses to contribute to a conflict is probably more powerful to contribute to a future conflict. Thus, the cardinality \( |D| \) can be viewed as a measure of the strength with which a variable contributes to a conflict. Unfortunately, we are not aware of any linear-time algorithm that computes \( |D| \) for all the variables.

Therefore, Maple_CM_Dist constructs the complete implication graph and computes the number of vertices in the longest path, denoted by \( \text{longDist}(x) \), from each variable \( x \) to the conflict, based on the observation that a variable depending on many clauses probably needs a longer path to reach the conflict. Then, it computes the Distance score for each variable \( x \), denoted by \( \text{distAct}(x) \) and initialized to 0, as follows: When \( x \) contributes to a conflict, Maple_CM_Dist calls Algorithm 2 to increment \( \text{distAct}(x) \) by \( \text{inc} \times 1/\text{longDist}(x) \), where \( \text{inc} \) is a global variable initialized to 1, and \( \text{dist_Decay} \) is intended to give more importance to recent conflicts, similarly to the \( \text{var_decay} \) parameter in VSIDS. The default value of \( \text{dist_Decay} \) is 0.95 as the default value of \( \text{var_decay} \) in MiniSAT.

Algorithm 2: updateDistanceScore\((C)\)

<table>
<thead>
<tr>
<th>Input: ( C ), a clause in which all literals are falsified by the current partial assignment</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 begin</td>
</tr>
<tr>
<td>2 Construct the complete implication graph ( G ) that falsifies ( C ) and compute ( \text{longDist}(x) ) for each variable ( x ) occurring in ( G );</td>
</tr>
<tr>
<td>3 for each variable ( x ) occurring in ( G ) do</td>
</tr>
<tr>
<td>4 ( \text{distAct}(x) \leftarrow \text{distAct}(x) + \text{inc} \times 1/\text{longDist}(x); )</td>
</tr>
<tr>
<td>5 ( \text{inc} \leftarrow \text{inc} \times \text{dist_Decay}; )</td>
</tr>
<tr>
<td>6 end</td>
</tr>
</tbody>
</table>

Maple_CM_Dist computes \( \text{distAct}(x) \) and branches on the variable \( x \) with maximum \( \text{distAct}(x) \) only for the first 50,000 conflicts. We limit this heuristic to the first 50,000 conflicts because constructing the complete implication graph is time-consuming, and VSIDS and LRB perform well after 50,000 conflicts. Maple_CM_Dist behaves like Maple_CM after the first 50,000 conflicts.

4. Blocking Restarts and Re-ordering UIPs in Maple_CM_ordUIP

The block-restart mechanism was introduced in Glucose 2.1 [6] for postponing a restart, when the solver is estimated to be approaching a global solution using a heuristic, in order to find quickly the global solution.

In Maple_CM_ordUIP, we push further the block-restart mechanism of Glucose. When the solver derives a conflict but is estimated to be approaching a global solution, it is forbidden to restart the search at least for the next 50 conflicts. In addition, it constructs a complete implication graph of the conflict to collect the first UIP (Unique Implication Point [7]) of every decision level involved in the conflict, in order to re-order these first UIPs and re-produce the conflict as described below.

Recall that a UIP in a decision level is a literal \( l \) such that every path from the decision literal of the level to the conflict goes through \( l \). The first UIP (denoted by 1UIP) in the level is the closest UIP to the conflict in the level. Literal \( l_1 \) is said to imply literal \( l_2 \) if \( l_2 \) occurs in a path from \( l_1 \) to the conflict. When the solver identifies the 1UIP in a level, it estimates also the total number of vertices it implies.

Then, the solver backtracks so that all decisions in the complete implication graph are canceled, and continues the search from there by picking the decision literals from the collected 1UIPs in decreasing order of the total number of literals they imply, breaking ties in favor of the 1UIP with the smallest decision level in the implication graph.

Concretely, each time Maple_CM_ordUIP derives a conflict after the first 10\(^5\) conflicts, it calls Algorithm 3 before
analyzing the conflict. In the algorithm, nbC_For_restart and nbC_For_reOrder are two global variables that are incremented by 1 upon each conflict. Restart is possible only if nbC_For_Restart is greater than 50. In other words, when the 1UIPs are re-ordered, the next restart or the next 1UIP re-ordering is possible only after the next 50 conflicts. If unit propagation derives an empty clause \( C' \), function unitPropagate() returns \( C' \); otherwise, it returns “no conflict”.

Algorithm 3: reOrderUIPs(C)

Input: \( C \), a clause in which all literals are falsified by the current partial assignment
Output: \( C' \), a clause in which all literals are falsified by the current partial assignment or the value “no conflict”

1 begin
2     if nbC_For_ReOrder > 50 and the number of assigned variables is two times greater than the average number of assigned variables in the previous 5000 conflicts then
3         nbC_For_Restart ← 0;
4         nbC_For_ReOrder ← 0;
5         Empty the vector UIPs;
6         Construct the complete implication graph \( G \) from \( C \) and push the 1UIP of each decision level of \( G \) into vector UIPs;
7         Sort vector UIPs in decreasing order of the number of literals the 1UIPs imply, breaking ties in favor of the 1UIP with the smallest decision level;
8         Backtrack to cancel all decisions in \( G \);
9         for \( i \) ← 0 to UIPs.size() do
10            if UIPs[i] is free then
11                \( C' \) ← unitPropagate(UIPs[i]);
12                if \( C' \) is a falsified clause then
13                    Return \( C' \);
14            end
15     end

If Algorithm 3 returns a falsified clause \( C' \), Maple_CM_ordUIP analyzes \( C' \) instead of \( C \). The conflict analysis stops at the last picked 1UIP, which implies fewer literals among the picked literals and may result in a shorter learnt clause. Otherwise, it continues the search by picking the decision literals among the free literals as Maple_CM.

The intuition behind re-ordering the 1UIPs can be described as follows: When a solver is approaching a global solution but derives a conflict, the conflict may simply be due to the ordering of the 1UIPs in the complete implication graph. In fact, the 1UIPs constitute the direct reason of the conflict. The 1UIP implying the greatest number of literals is probably the most constrained and should be satisfied first, leading more easily to the global solution.

5. Blocking Restarts and Re-ordering UIPs in Maple_CM_ordUIP

Maple_CM, like MapleCOMSPS_DRUP and Maple_LCM, first uses the LRB heuristic for 2500 seconds and switches to the VSIDS heuristic for the remaining run time. It appears that the LRB heuristic is more powerful to solve satisfiable instances, whereas the VSIDS heuristic is more powerful to solve unsatisfiable instances. So, we implement Maple_CM_ordUIP+, that is Maple_CM_ordUIP but selects the conflict to re-order the 1UIPs differently in function of the branching heuristic used to derive the conflict.

- In a restart using VSIDS, Maple_CM_ordUIP+ reorders the 1UIPs implying the first conflict as in Algorithm 3.
- In a restart using LRB, Maple_CM_ordUIP+ behaves as Maple_CM_ordUIP.

Intuitively, the first conflict is the most important in a restart, because it determines the search direction in the restart. Maple_CM_ordUIP+ aims at reinforcing the capability of the VSIDS heuristic to find a global solution of the instance, by branching first on the 1UIPs implying the greatest number of literals in a restart.

Acknowledgments

Research supported by the National Natural Science Foundation of China (Grants no. 61370183, 61370184, 61472147), MINECO-FEDER project RASO TIN2015-71799-C2-1-P, EU Horizon 2020 project LOGISTAR and the Matrics platform of University of Picardie Jules Verne.

References

Glucose Hacks and TOPOSAT2 Description

Thorsten Ehlers
Kiel University
Kiel, Germany
Dirk Nowotka
Kiel University
Kiel, Germany

Abstract—This document describes the parallel SAT solver TOPOSAT2 and the Glucose Hack which we submit to the SAT Competition 2018.

I. INTRODUCTION

This short paper describes the GLUCOSE hack and the version of our parallel solver TOPOSAT2 submitted to the SAT Competition 2018.

II. GLUCOSE HACK

Our Glucose Hack is considered with the preprocessor, which mainly is code from MINISAT [1] and SATelite [2]. This code still contains old comments like

// FIX: this is too inefficient but would be nice to have (properly implemented)
// if (!find(subsumption_queue, &c))

We suggest to improve the preprocessor code at three places.

1) We add a boolean flag to the clause header which prevents clauses from being put on the subsumption queue several times. Though not a big issue on many formulas, this actually is a big problem on some formulas which may cause the preprocessor to be stuck for hours.

2) The subsumption check for clauses \(a, b\) has a worst-case running time of \(|a| \cdot |b|\). Using a simple lookup-array, this can be reduced to \(|a| + |b|\).

3) The same applies for the merge-procedure which computes the resolvent of two clauses.

We think that this is particularly interesting as many current solvers are built on top of MINISAT and GLUCOSE. The required changes are rather small (the edit distance is 892), and can easily be applied to every solver which is built upon MINISAT and GLUCOSE.

III. TOPOSAT2

TOPOSAT2 was mainly designed to run in a massively-parallel environment (> 1000 solver threads). Thus, we are curious to see how it performs on a shared-memory system. It is built on top of Glucose 3.0, but uses a bug-fixed version of the lockless clause sharing mechanism from ManySAT [3] for communication on one compute node rather than the lock-based implementation from Glucose Syrup. The communication between nodes uses MPI, but this is not used for the competition.

It comes with two features which we hope will be especially useful in the competition.

A. Diversification

The first portfolio solvers used different sequential solvers, or different settings of one sequential solver. We somewhat go back to the roots and diversify the search of the solver threads by the following parameters.

- Branching: Some solver threads use VSIDS, whereas other use LRB [4], as this branching heuristic was quite successful in the past SAT competitions. As VSIDS still works better on a significant amount of benchmarks, we use both.

- Restarts: We use the inner/outer restart scheme [5], Luby restarts, and the adaptive restart strategy from GLUCOSE [6].

- Learnt Clause DB management: Some solver threads use a scheme similar to the one suggested in [7]: Clauses with very low LBD (≤ 3) are stored permanently. Clauses of intermediate LBD are stored at least for some time, and there is a small activity-based clause storage. The LBD of clauses imported from other solver threads are initialised with the size of the clause. Thus, the clause must be used in order to update its LBD, and allowing it to be stored for a longer time. Some other solver threads use the default clause management strategy from GLUCOSE [8].

B. Lifting exported clauses

Wieringa et. al suggested to use some threads of a parallel SAT solver to strengthen learnt clauses [9]. Similarly, in [10] some of the learnt clauses are strengthened during search. We use this technique when exporting clauses. Whenever one solver thread learns a clause of sufficiently low LBD, it is stored in an extra buffer. After the next restart, the clauses from this buffer are strengthened, and the results are exported to the other solver threads.

C. Submitted versions

We submit one version of TOPOSAT2 to the SAT Competition. However, we submit two different scripts to start it with different parameters. These parameters are concerned with the way in which clauses are import by the solver. The first version uses a variation of the clause import strategy of MANYSAT. During search, the trail size is monitored. Clauses are imported when some time has passed and the solver is somewhat close to the root of the search tree. In this way, we try to prevent the solver from backtracking too often when imported clauses are unit under the current assignment.
The second version imports more often. However, clauses which are unit under the current assignment and would thus require a backtrack are imported as one-watched clauses as in Glucose Syrup [11].

ACKNOWLEDGMENT

This work is funded by the German Federal Ministry of Education and Research, combined project 01IH15006A.

REFERENCES

Varisat, a SAT Solver Written in Rust

Jannis Harder
Email: me@jix.one

Abstract—Varisat is a CDCL-based SAT solver written from scratch in the programming language Rust. It is still in an early stage of development, offering little beyond the essentials. Nevertheless, it hopefully shows that Rust is a suitable programming language for implementing a SAT solver.

I. INTRODUCTION

Most state-of-the-art SAT solvers are implemented in C or C++. This is not surprising, as optimizing data layout, memory access patterns and other low-level optimizations are important for the performance of SAT solvers.

Recently the programming language Rust [1] has seen increased use in areas where before mostly C and C++ were popular. Rust is a systems programming language with a focus on memory safety and performance, which is sponsored by Mozilla Research.

Varisat is my attempt to implement a SAT solver in Rust, which I began on one hand to become familiar with the details of implementing a modern SAT solver and on the other hand to show that Rust is a suitable language for implementing a SAT solver.

II. IMPLEMENTED TECHNIQUES

Varisat is a CDCL-based SAT solver [2] written from scratch. It is still in an early stage of development and implements only basic techniques. Among the already implemented techniques are: VSIDS branching heuristic, Luby series restarts, minimization of learned clauses [3], LBDs [4] and a 3-Tiered learned clause database [5].

No pre- or inprocessing techniques are implemented yet. Also Varisat has seen very little benchmarking and thus very little parameter tuning.

III. SOURCE CODE

The source code is licensed under the MIT license and available at https://jix.one/sc18/varisat. It uses some Rust features that are not yet part of the stable release. The submitted version was developed using the nightly release 2018-02-08.

REFERENCES

satUZK-ddc at SAT Competition 2018
Alexander van der Grinten
University of Cologne

Abstract—We describe the current version of our distributed satUZK-ddc solver. Since SAT Competition 2017, we have augmented the solver with a parallel preprocessing technique that we call parallel distillation. This algorithm performs the same CNF simplifications as the conventional CNF distillation algorithm, namely removing or shortening clauses of the input formula according to certain redundancy properties. While sequential CNF distillation is known to be expensive, due to parallelism, it becomes feasible to preprocess even large real-world CNF formulas with our algorithm.

I. OVERVIEW
We submitted two solver configurations to the SAT Competition 2018. Both configurations are based on the satUZK-ddc engine that was also submitted to last year’s SAT Competition [1]. The first configuration is the current default configuration of satUZK-ddc. This configuration corresponds to the satUZK-ddc configuration from last year, with minor adjustments and bug fixes. The second configuration (satUZK-ddc --ddist) is augmented with our parallel distillation preprocessing technique. In contrast to last year, we did not submit a sequential version as our sequential algorithm is mostly unchanged from the SAT Competition 2017 version.

II. DDC ALGORITHM
The distributed divide and conquer (DDC) algorithm was shortly described in [1] and is described in more detail in [2]. We summarize the main ideas of the algorithm here.

The DDC algorithm uses a parallel lookahead technique to partition the search space until there are at least as many subproblems as processors. Those subproblems are solved by CDCL. In contract to Cube and Conquer [3], the DDC algorithm does not rely on work stealing for load balancing; instead, the algorithm maintains a distributed divide and conquer tree and routes individual processors through this tree. As the routing procedure uses only local information and does not require extensive synchronization, this approach is able to fully utilize all processors at every step of the algorithm.

The basic algorithm is extended with preprocessing, inprocessing, an LBD [4] score that takes the incremental nature of the algorithm into account, clause sharing and diversification. Again, for details on those extensions, we refer to [2]. satUZK-ddc is an MPI-based implementation of the DDC algorithm that builds on our earlier CDCL implementation in satUZK.

III. PARALLEL DISTILLATION
Our most significant contribution to the SAT Competition 2018 is the parallel distillation algorithm described below.

CNF distillation: In the following, we shortly review the traditional distillation [5] procedure1 and state it in the language of [7], [8]. Let \( \phi \) be the input formula (in CNF) and let \( C \in \phi \) be a clause. If \( C' \subseteq C \) is a subset of \( C \) and unit propagation on \( \phi \setminus \{C\} \cup \{\overline{C} \colon C \in C'\} \) assigns all literals of \( C \) to false, then \( C \) can be shortened to \( C' \) while preserving equivalence. This technique is called asymmetric literal elimination (ALE). Note that \( C' \) is not necessarily unique for each \( C \in \phi \) (not even if we require \( C' \) to be minimal). Thus, implementations of ALE usually pick an arbitrary suitable \( C' \).

On the other hand, if unit propagation on \( \phi \setminus \{C\} \cup \{\overline{C} \colon C \in C'\} \) leads to a conflict, then \( C \) can be removed from \( \phi \) while preserving equivalence. This process is called asymmetric tautology elimination (ATE).

Distillation is a systematic procedure to perform ALE and ATE. To accomplish this, unit propagation is used to detect opportunities for ALE and ATE. In particular, unit propagation is applied at least once for each clause in \( \phi \). In the sequential case, this is the main bottleneck of the algorithm; therefore, our parallel version tries to parallelize the detection of ALE and ATE opportunities.

Commutation properties: In order to parallelize the distillation algorithm, we observe the following commutation properties (see [2] for details):

ALE vs. ALE If ALE transforms \( \phi \cup \{C_1\} \to \phi \cup \{C_2\} \) and also \( \phi \cup \{C_2\} \to \phi \cup \{C_3\} \) (i.e. ALE shortens the clauses \( C_1 \) to \( C'_1 \) and \( C_2 \) to \( C'_2 \), respectively), then the two formulas \( \phi \cup \{C_1, C_2\} \) and \( \phi \cup \{C'_1, C'_2\} \) are equivalent2.

ALE vs. ATE If ALE transforms \( \phi \cup \{C\} \to \phi \cup \{C'\} \) and ATE transforms \( \phi \cup \{D\} \to \phi \) (i.e. ATE removes the clause \( D \)), then the two formulas \( \phi \cup \{C, D\} \) and \( \phi \cup \{C'\} \) are equivalent. Specifically, ALE can only expand the set of literals that are fixed by unit propagation on \( \phi \cup \{\overline{\ell} \colon \ell \in D\} \).

ATE vs. ATE On the other hand, if ATE transforms \( \phi \cup \{D_1\} \to \phi \) and also \( \phi \cup \{D_2\} \to \phi \), the two formulas \( \phi \cup \{D_1, D_2\} \) and \( \phi \) are not necessarily equivalent. However, they are not equivalent only if \( D_1 \) participates in the conflict3 which is derived by applying unit propagation to \( \phi \cup \{\overline{\ell} \colon \ell \in D_2\} \), or vice versa. Thus, this situation must be detected to determine if replacing \( \phi \cup \{D_1, D_2\} \) by \( \phi \) is sound.

Parallel algorithm: Let \( p \) denote the number of processors that are available to the solver. Initially, each processor has its own copy of the input formula \( \phi \). The parallel distillation algorithm now works as follows: The set of clauses of \( \phi \) is

---

1 We use the term distillation synonymously with vivification [6].
2 This is not particularly surprising; in fact, this property holds true for every equivalence-preserving simplification technique that only shortens clauses.
3 We say that a clause participates in a conflict if it is part of the conflict graph [9].
partitioned into $p$ subsets $\phi_1, \ldots, \phi_p$. $\phi_i$ will be the set of clauses that are processed by processor $i$. For each clause $C \in \phi_i$, processor $i$ tries to shorten $C$ to a subset $S(C) \subseteq C$ by applying ALE to $\phi$. If $C$ cannot be shortened by ALE, we set $S(C) = C$. Furthermore, processor $i$ determines a flag $r(C)$ that is true if and only if $C$ can be removed from $\phi$ using ATE. If that is the case, a set $D(C)$ is determined. $D(C)$ consists of exactly those clauses that participate in the conflict which is derived by unit propagation on $\phi \cup \{\{\ell_i\} : \ell_i \in C\}$. We call $D(C)$ the set of dependencies of $C$. Indeed, because of the last commutation property, the removal of $C$ depends on the clauses from $D(C)$ in the following sense: $C$ can be safely removed from $\phi$ if all clauses from $D(C)$ are retained in $\phi$. We remark that our algorithm parallelizes at the clause level; like in the case of traditional distillation, ALE and ATE checks are performed by standard unit propagation.

After the computation is finished on all processors, the information $S(C)$, $r(C)$ and $D(C)$ are gathered on all processors, for every clause $C \in \phi$. The processors agree on an order of those clauses and all processors inspect the clauses in the same order. Now, for each clause $C \in \phi$, the flag $r(C)$ is checked to determine if $C$ is a candidate for removal. If $r(C)$ is true, all processors remove $C$ if and only if (i) no dependency in $D(C)$ has already been removed and (ii) $C$ was not a dependency of any clause that has already been removed. If $C$ is not removed, we replace $C$ by $S(C)$. The commutation properties discussed above ensure that this algorithm results in a formula that is equivalent to the input formula $\phi$. As all processors apply modifications in the same order, the resulting formulas are identical. Note that the effectiveness of the algorithm potentially depends on the order in which clauses are inspected; however, optimizing this order did not turn out to be necessary in practice.

Optimizations: We discuss some modifications of the parallel distillation algorithm that aim to improve its empirical performance. First, we modify the algorithm to not remove binary clauses using ATE. This allows us to omit binary clauses from the $D(C)$ sets. In the case of real-world CNF instances, this often yields a significant size reduction, as more than half of the clauses are binary in many of those instances. We note that this optimization only marginally decreases the strength of distillation when it is combined with the (considerably faster) unhiding [10] algorithm: In fact, we can expect that many binary clauses that are removed by ATE can also be removed by the weaker hidden tautology elimination (HTE) technique. The established failed literal elimination (FLE) and hyper-binary-resolution (HBR) [11] techniques allow to turn additional ATEs into HTEs.

Secondly, we want to avoid removing useful clauses using ATE. For example, note that ATE would be strong enough to remove all learned clauses from a formula. Certainly, removing all learned clauses impairs the performance of CDCL solvers. Thus, instead of removing clauses after ATE, we only mark them as non-redundant and depend on the usual clause database reduction heuristics to remove those clauses eventually.

Furthermore, to increase the opportunities for ALE and ATE, we apply a parallel FLE algorithm before running parallel distillation. This FLE algorithm is a probing-based algorithm that relies on the tree-based lookahead approach [11] to check, for each literal $\ell$, if $\ell$ is a failed literal. If that is the case, the clause $\{\ell\}$ is added to the input formula $\phi$. As adding the clauses to $\phi$ can only enlarge the set of failed literals, it is possible to traverse multiple lookahead trees in parallel while preserving correctness.

Finally, instead of partitioning $\phi$ into $p$ fixed subsets, we employ a work-stealing load balancer to assign clauses to processors. As we expect the running time of unit propagations to differ substantially depending on affected variables and clauses, this improves processor utilization compared to static partitioning.

REFERENCES


ReasonLS

Shaowei Cai and Xindi Zhang
State Key Laboratory of Computer Science, Institute of Software, Chinese Academy of Sciences
School of Computer and Control Engineering, University of Chinese Academy of Sciences

Abstract—This note describes the SAT solver “ReasonLS”, which is a hybrid solver that combines a semi-exact CDCL based solver and a local search solver.

I. INTRODUCTION

An interesting direction for solving SAT is to combine the CDCL method and the local search method, as these two methods have different advantages. Previous works in this direction usually use a CDCL solver as a black box. On the other hand, our recent works show that initial assignments generated by unit propagation based procedures can significantly improve the performance of local search on solving industrial MaxSAT instances, even leading the algorithm to find the optimal solution [1].

ReasonLS is a hybrid solver that combines a CDCL based semi-exact solver and a local search solver. We develop a backtracking procedure by modifying a CDCL solver to find a complete assignment with high consistency, which is then handed to a local search solver trying to find a solution. ReasonLS works by interleaving between these two solvers. In each round (except for the first one), some restarting strategies are used to restart the backtracking procedure. ReasonLS uses a parameter (a real number from 0 to 1) to control to what level of consistency that we require the backtracking style procedure to find, and once the procedure finds such a complete assignment, it returns the assignment. Then ReasonLS calls a local search to find a solution, using the assignment found by the backtracking procedure.

ReasonLS can also be seen as a generalization of both CDCL solvers and local search solvers. When the parameter is set to 0, ReasonLS becomes a local search solver; when the parameter is set to 1, it becomes a CDCL solver.

II. A SEMI-EXACT PROCEDURE

We modify a CDCL solver to make it allow conflicts in some cases. We use a parameter $p$ (a real number from 0 to 1), and if the solver finds a path that assigns at least a portion $p$ of the variables without generating a conflict, then it will finish the path by assigning the remaining variables, even if it may lead to conflicts.

In ReasonLS, we develop our semi-exact solver by modifying the CDCL solver named Maple_LCM_Dist [2], which is a state-of-the-art CDCL solver and won the main track of SAT Competition 2017.

III. THE LOCAL SEARCH ALGORITHM

As for the local search solver used in ReasonLS, we develop a variant of CCAnr [3] by forbidding the aspiration mechanism. The algorithm is described as follows.

Starting from an initial assignment, the algorithm flips a variable in each step. Firstly, if there are configuration changed variables, it picks a configuration changed variable with the greatest score, breaking ties by favoring the oldest variable. Otherwise, the clause weights are updated according to a Threshold-based Smoothed Weighting (TSW) scheme; then, it picks a random unsatisfied clause and selects a variable in the clause with the greatest score, breaking ties by favoring the oldest variable.

IV. MAIN PARAMETERS

There is one parameter $p$ for controlling the cooperation of the backtracking style procedure and the local search solver. There are three parameters in the local search solver: the average weight threshold parameter $\gamma$, and the two factor parameters $\rho$ and $q$. All of the three parameters are for the TSW weighting scheme. The parameters are set as follows: $\gamma = 50$; $\rho = 0.3$; $q = 0.7$. Meanwhile, there are a button $ASP$ to control whether turn on or turn off the process of aspiration of local search, and a parameter $VSIDS$ to decide when change search strategy of CDCL to VSIDS.

In our solver, we call the ReasonLS solver to solve an instance with different parameter settings as follow. Each thread runs the solver with one setting.

1) $p = 1$; $VSIDS = 2500s$ (ReasonLS becomes the CDCL solver named Maple_LCM_Dist, thus no need to set the parameters for local search.)
2) $p = 1$; $VSIDS = 0s$.
3) $p = 0$; $\gamma = 50$; $VSIDS = 2500s$; $ASP = on$; $\rho = 0.3$; $q = 0.7$.
4) $p = 0$; $\gamma = 50$; $VSIDS = 2500s$; $ASP = on$; $\rho = 0.3$; $q = 0$.
5) $p = 0$; $\gamma = 50$; $VSIDS = 2500s$; $ASP = off$; $\rho = 0.3$; $q = 0.7$.
6) $p = 0$; $\gamma = 50$; $VSIDS = 2500s$; $ASP = off$; $\rho = 0.3$; $q = 0$.
7) $p = 0$; $\gamma = 300$; $VSIDS = 2500s$; $ASP = on$; $\rho = 0.3$; $q = 0.7$.
8) $p = 0$; $\gamma = 300$; $VSIDS = 2500s$; $ASP = on$; $\rho = 0.3$; $q = 0$.
9) $p = 0$; $\gamma = 300$; $VSIDS = 2500s$; $ASP = off$; $\rho = 0.3$; $q = 0.7$. 
10) $p = 0$; $\gamma = 300$; $VSIDS = 2500s$; $ASP = off$; $\rho = 0.3$; $q = 0$.
11) $p = 0.9$; $\gamma = 50$; $VSIDS = 2500s$; $ASP = on$; $\rho = 0.3$; $q = 0.7$.
12) $p = 0.9$; $\gamma = 50$; $VSIDS = 2500s$; $ASP = on$; $\rho = 0.3$; $q = 0$.
13) $p = 0.9$; $\gamma = 50$; $VSIDS = 2500s$; $ASP = off$; $\rho = 0.3$; $q = 0.7$.
14) $p = 0.9$; $\gamma = 50$; $VSIDS = 2500s$; $ASP = off$; $\rho = 0.3$; $q = 0$.
15) $p = 0.9$; $\gamma = 300$; $VSIDS = 2500s$; $ASP = on$; $\rho = 0.3$; $q = 0.7$.
16) $p = 0.9$; $\gamma = 300$; $VSIDS = 2500s$; $ASP = on$; $\rho = 0.3$; $q = 0$.
17) $p = 0.9$; $\gamma = 300$; $VSIDS = 2500s$; $ASP = off$; $\rho = 0.3$; $q = 0.7$.
18) $p = 0.9$; $\gamma = 300$; $VSIDS = 2500s$; $ASP = off$; $\rho = 0.3$; $q = 0$.
19) $p = 0.9$; $\gamma = 50$; $VSIDS = 2500s$; $ASP = on$; $\rho = 0.3$; $q = 0.7$. (increase the number of flips of local search, and so dose the 20th.)
20) $p = 0.9$; $\gamma = 50$; $VSIDS = 2500s$; $ASP = on$; $\rho = 0.3$; $q = 0.7$. (set more strict limit on the time of local search.)
21) $p = 0.7$; $\gamma = 50$; $VSIDS = 2500s$; $ASP = on$; $\rho = 0.3$; $q = 0.7$.
22) $p = 0.7$; $\gamma = 50$; $VSIDS = 0s$; $ASP = on$; $\rho = 0.3$; $q = 0.7$.
23) $p = 0.5$; $\gamma = 50$; $VSIDS = 2500s$; $ASP = on$; $\rho = 0.3$; $q = 0.7$.
24) $p = 0.5$; $\gamma = 50$; $VSIDS = 0s$; $ASP = on$; $\rho = 0.3$; $q = 0.7$.

V. IMPLEMENTATION DETAILS

ReasonLS is implemented in C++. It is developed based on the codes of Maple_LCM_Dist [2] and CCAnr solver [3].

VI. SAT COMPETITION 2018 SPECIFIES

ReasonLS is submitted to “Parallel Track”. It is compiled by g++ with the ‘O3’ optimization option.

Its running command is: "/ReasonLS-run.sh $1". $1 is the absolute path of input file. For a given input file "~/sc/a.cnf", the call command is "/ReasonLS-run.sh ~/sc/a.cnf".

REFERENCES

BENCHMARK DESCRIPTIONS
Generating the Uniform Random Benchmarks

Marijn J. H. Heule
Department of Computer Science,
The University of Texas at Austin, United States

Abstract—The benchmark suite of the Random Track of SAT Competition 2018 can be partitioned into three parts. The first part consists of uniform random $k$-SAT instances described below. The second part consists of benchmarks generated by a tool by Tomáš Balyo [1]. These benchmarks are similar as the ones used in the 2016 SAT Competition. The third part consists of random benchmarks contributed by Adrian Balint described on page 61 of these proceedings.

INTRO

This description explains how the benchmarks were created of the uniform random categories of the SAT Competition 2018. These categories consists of uniform random $k$-SAT instances with $k \in 3, 5, 7$ – Boolean formulas for which all clauses have length $k$. For each $k$ the same number of benchmarks have been generated.

GENERATING THE SATISFIABLE BENCHMARKS

The satisfiable uniform random $k$-SAT benchmarks are generated for two different sizes: medium and huge. The medium-sized benchmarks have a clause-to-variable ratio equal to the phase-transition ratio\(^1\). The number of variables differs for all the benchmarks. The huge random benchmarks have a few million clauses and are therefore as large as some of the application benchmarks. For the huge benchmarks, the ratio ranges from far from the phase-transition ratio to relatively close, while for each $k$ the number of variables is the same. Table I shows the details.

No filtering was applied to construct the competition suite. As a consequence, a significant fraction (about 50%) of the medium-sized generated benchmarks is unsatisfiable.

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|}
\hline
$k$ & medium (40) & huge (20) \\
\hline
3 & $r = 4.267$ & $r \in \{3.86, 3.88, \ldots, 4.24\}$ \\
& $n \in \{5000, 5200, \ldots, 12800\}$ & $n = 1,000,000$ \\
\hline
5 & $r \approx 21.117$ & $r \in \{16, 16.2, \ldots, 19.8\}$ \\
& $n \in \{200, 210, \ldots, 590\}$ & $n = 250,000$ \\
\hline
7 & $r \approx 87.79$ & $r \in \{55, 56, \ldots, 74\}$ \\
& $n \in \{90, 92, \ldots, 168\}$ & $n = 50,000$ \\
\hline
\end{tabular}
\caption{Parameters of generating the satisfiable benchmarks}
\end{table}

\(^1\)The observed clause-to-variable ratio for which 50% of the uniform random formulas are satisfiable. For most algorithms, formula generated closer to the phase-transition ratio are harder to solve.

REFERENCES

Divide and Unique Inverse Benchmarks
Submitted to the SAT Competition 2018

Armin Biere
Institute for Formal Models and Verification
Johannes Kepler University Linz

Our benchmark submission for the SAT 2018 Competition consist of two sets of word-level properties originally formulated as SMT problems in the quantifier-free theory of bit-vectors in BTOR [1] or SMTLIB [2] format. We then use our SMT solver Boolector [3] to synthesize AIGs [4], which in turn were translated to DIMACS format.

DIVISION

The first set specifies word-level (modulo $2^n$) division using multiplication for various bit-widths $n$ in BTOR format [1].

We consider both unsigned and signed dividers. For unsigned division we check validity over unsigned $n$-bit bit-vectors ($\div_u$ denotes unsigned division):

$$y \neq 0 \Rightarrow (x - (x \div_u y) \cdot y) <_u y$$

As common in bit-vector logics arithmetic operators take two $n$-bit bit-vectors as input and produce one $n$-bit bit-vector as output, with the effect, that there is no difference between signed and unsigned versions of multiplication nor subtraction.

For signed division it is more complicated and we have to take signs into account (now $\div_s$ denotes signed division):

$$y \neq 0 \Rightarrow |x - (x \div_s y) \cdot y| <_u |y|$$

where “$|$” is actually implemented with an if-then-else operator testing the argument to be smaller than zero (using signed “$<$” comparison) and if so negating it (two-complement). These signed benchmarks are as a consequence much harder.

INVERSION

The second set of benchmarks checks that bit-vector multiplication modulo $2^n$ has unique inverses for odd numbers, which translates to the following SMT benchmark for $n = 32$ in SMTLIB format [2]:

```
(set-logic QF_BV)
(declare-fun x () (_ BitVec 32))
(declare-fun y () (_ BitVec 32))
(declare-fun z () (_ BitVec 32))
(assert (= (bvmul x y) (bvmul x z)))
(assert (distinct y z))
(check-sat)
(exit)
```

REFERENCES


Supported by FWF, NFN Grant S11408-N23 (RiSE).
Reversing Elementary Cellular Automata

Jannis Harder
Email: me@jix.one

I. INTRODUCTION

Elementary (i.e. one-dimensional two-state) cellular automata are systems with simple rules that can nevertheless show complex behavior. These benchmarks encode the problem of finding a sequence of predecessor states for a given target state and elementary automaton rule \( R \). To allow for a finite encoding of states, we assume that all states, target and predecessors, have a fixed period \( n \).

We first generate a pseudo-random state \( X \) from a seed \( s \). Starting from \( X \) and advancing it \( f \) times, we generate a target state \( Y \). We then ask whether there is an initial state \( I \), that evolves to \( Y \) in exactly \( r \) steps. The sequence of states from \( I \) to \( Y \) does not have to contain the state \( X \), but when \( r \leq f \) such a sequence does exist.

The generated instance for the parameters \( (R, n, f, r, s) \) is named `ecarev-R-n-f-r-s.cnf`

II. SAT-ENCODING

We need to consider the sequence of states \( S_1, \ldots, S_r \), where \( S_1 = I \) and \( S_r = Y \). Each state \( S_i \) is represented by the cell values for \( n \) consecutive cells \( S_{i,1}, \ldots, S_{i,n} \). These cell values correspond to variables in the SAT-encoding. Since the states are periodic, \( S_{i,j} = S_{i,j+n} \) holds and we will use both to represent the same variable in the encoding.

To encode the constraint that the states follow the automata rule, we need to encode the state transition \( S_{i+1} = T_R(S_i) \).

This relation can be decomposed into

\[
\bigwedge_{1 \leq j \leq n} S_{i+1,j} = t_R(S_{i,j-1}, S_{i,j}, S_{i,j+1})
\]

where \( t_R \) is specified by the automaton’s rule table. It is encoded as a conjunction of all implied minimal clauses.

We then constrain the states to follow the automata’s evolution by

\[
\bigwedge_{1 \leq i < r} S_{i,j} = S_{i,j+n}.
\]

Finally we add unit clauses to constrain \( S_r = Y \).

The implementation of this encoding is written in Python 3 and available at https://jix.one/sc18/ecarev.py.

III. SELECTED BENCHMARKS

The submitted benchmarks are for the rule-110 automaton, which is capable of universal computation. The state period \( n \) was chosen large enough to make an exhaustive search infeasible. Apart from that the parameters were manually chosen to make the instances not too easy or too hard and to make it difficult to guess the outcome from the parameters alone.

Verifying Simple Floating-Point Programs

Joseph Scott, Jia Hui Liang, Vijay Ganesh
University of Waterloo, Waterloo, ON, Canada

Abstract—A brief description of the instances we submitted to the SAT Competition 2018 encoding bounded model checking of floating-point C programs.

I. BACKGROUND

Floating point data types use a finite number of bits to represent the set of real numbers within a fixed interval. Unlike its integer counterpart, not every real number within the interval has an exact representation in its floating point type. For example, the real number 0.1 does not have an exact representation in common binary floating point types.

Computer programs with floating point arithmetic frequently suffer from a (quiet) inexact exception where the floating point operation results in a real value that does not have an exact representation in its floating point type. A rounding decision must be made to set the result of the operation to a representable floating point value.

Floating-point operations and rounding behavior are dictated by the IEEE 754 standard and behave unintuitively under certain circumstances. Programmers who treat floating-point numbers as real numbers are surprised to learn that basic properties of real numbers such as associativity and distributivity are not respected in the floating-point space due to the required rounding. This often leads to hard-to-debug errors that expose themselves in floating-point programs under rare conditions. Additionally, floating-point adds special representations like not-a-number (NaN), positive and negative zero, and positive and negative infinity, which can cause several subtle errors when assuming similar behavior to familiar integer data types.

In program verification over integer data types, SMT solvers with rich theories can be used to guide a SAT solver. Due to the unintuitive nature of floating point and lack of established properties, solvers must use a more direct translation to SAT which makes the solvability of floating point instances desirable.

II. BENCHMARKS

We use CBMC (C Bounded Model Checker) by Kroening and Tautschnig [1] to translate 46 simple C-programs written by us involving floating-points into to SAT. These programs test for:

- Associativity of + and × with various acceptable relative error bounds.
- Floating-point commutative property
- Floating-point distributive property
  - Exact distributivity.
  - Distributivity with various acceptable relative error bounds.
- Finding floating-point roots of a quadratic polynomial
  - Testing that the roots are exact.
  - Testing that the roots when evaluated are within a threshold from zero.
- Inexactness of the square-root of floating-point numbers
- Triangle Inequality
  1) Exact triangle inequality.
  2) Triangle inequality with various added fixed constants for acceptable error.
- Perceptron classification with floating-points over a fixed data set.

The original C-programs are available here: https://sites.google.com/a/gsd.uwaterloo.ca/maplesat/floatingpointsource.zip?attredirects=0&d=1. We used the command cbmc source.c --dimacs to generate the CNF files. The floating-point operations in the C programs are essentially bit-blasted down to propositional logic by CBMC. If the CNF is satisfiable, then the assertion in the C code is satisfiable at the point of execution. Otherwise the CNF is unsatisfiable.

REFERENCES

**GrandTour**\textsuperscript{obs} Puzzle as a SAT Benchmark

Md Solimul Chowdhury  
Department of Computing Science  
University of Alberta  
Edmonton, Alberta, Canada

Martin Müller  
Department of Computing Science  
University of Alberta  
Edmonton, Alberta, Canada

Jia-Huai You  
Department of Computing Science  
University of Alberta  
Edmonton, Alberta, Canada

---

**Abstract**—GrandTour\textsuperscript{obs} is a puzzle game, which is usually played in a rectangular grid of points, where a player is posed with the challenge to find a closed loop that goes through each point exactly once. In GrandTour, the player is allowed to connect any pair of points to solve the game. We consider a slight variation of the GrandTour puzzle game, where a player is allowed to connect a restricted subset of pairs of points, as there are obstacles between some of the points. We call this puzzle GrandTour with obstacles or GrandTour\textsuperscript{obs}. For SAT competition-2018, we provide 20 SAT instances of the GrandTour\textsuperscript{obs} puzzle.

I. Encoding and Instance Generation

a) GrandTour\textsuperscript{obs} as Hamiltonian Cycle Problem: The problem of determining whether a Hamiltonian cycle exists in a given graph is a well-known NP-complete problem. This problem deals with the following question: is there a cycle in a given graph, in which each vertex is visited exactly once? While solving a GrandTour puzzle is equivalent to finding a Hamiltonian cycle in a complete graph\textsuperscript{1} induced from its completely connected puzzle grid, solving GrandTour\textsuperscript{obs} problem is equivalent to finding a Hamiltonian cycle in an incomplete graph, which is induced from its incompletely connected puzzle grid. If a Hamiltonian cycle exists in the graph induced from the puzzle grid of a GrandTour\textsuperscript{obs} puzzle, then the puzzle is solvable, otherwise it is not.

b) Problem Encoding: A ground normal logic program \( P \) can be converted into a SAT formula \( S \), from the Clark’s completion \( C \) of \( P \). For example, the Clark’s completion of \( P = \{a \leftarrow b, a \leftarrow c\} \) is \( C = \{a \leftarrow b \lor c\} \) and \( C \) can be expressed into the SAT formula \( S \) in CNF form as: \((-\neg b \land \neg c) \lor a \equiv (-\neg b \lor a) \land (-\neg c \lor a)\). This technique is utilized in ASSAT\textsuperscript{1} [1], a solver for answer set programs using a SAT solver as the underlying inference engine. Given a ground normal logic program \( P \), ASSAT computes its stable models using a SAT solver. As part of its solving process, ASSAT produces a SAT formula \( S \) from \( P \). We exploit this feature of ASSAT to generate SAT instances of the GrandTour\textsuperscript{obs} problem from the normal logic program encoding of Hamiltonian cycle as proposed in [2].

c) Instance Generation: For the GrandTour\textsuperscript{obs}, a puzzle grid of size \( x \times y \) can be induced from a graph with \( n \) nodes, where \( n \) is an even number and \( n = x \times y \). The website for ASSAT\textsuperscript{3} contains 32 normal logic program instances for the Hamiltonian cycle problem from two graphs with 60 and 50 vertexes, respectively. We treat these instances as the GrandTour\textsuperscript{obs} instances of grid size \( 10 \times 6 \) and \( 10 \times 5 \), respectively.

TABLE I: Details of the 20 SAT instances for the GrandTour\textsuperscript{obs} benchmark; An instance \( gto_p\_pxy \) is based on a grid which has \( x \) points and \( y \) pairs of points are connected.

<table>
<thead>
<tr>
<th>Problem/Grid Size</th>
<th>Variables/Clauses</th>
<th>MiniSAT Time (s)</th>
<th>MiniSAT Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>( gto_{60c295} \times 10 \times 6 )</td>
<td>1011739</td>
<td>101,35</td>
<td>UNSAT</td>
</tr>
<tr>
<td>( gto_{60c345} / p50c345 \times 10 \times 6 )</td>
<td>1021359</td>
<td>101,76</td>
<td>UNSAT</td>
</tr>
<tr>
<td>( gto_{60c234} / p60c234 \times 10 \times 6 )</td>
<td>1037362</td>
<td>102,76</td>
<td>UNSAT</td>
</tr>
<tr>
<td>( gto_{60c231} / p60c231 \times 10 \times 6 )</td>
<td>1045721</td>
<td>103,94</td>
<td>UNSAT</td>
</tr>
<tr>
<td>( gto_{60c333} / p60c333 \times 10 \times 6 )</td>
<td>1053533</td>
<td>1077,71</td>
<td>UNSAT</td>
</tr>
<tr>
<td>( gto_{60c234} / p60c234 \times 10 \times 6 )</td>
<td>10627571</td>
<td>1077,71</td>
<td>UNSAT</td>
</tr>
<tr>
<td>( gto_{60c333} / p60c333 \times 10 \times 6 )</td>
<td>10705333</td>
<td>1060,01</td>
<td>UNSAT</td>
</tr>
<tr>
<td>( gto_{60c345} / p60c345 \times 10 \times 6 )</td>
<td>10853649</td>
<td>10272,67</td>
<td>UNSAT</td>
</tr>
<tr>
<td>( gto_{60c235} / p60c235 \times 10 \times 6 )</td>
<td>10932350</td>
<td>5000</td>
<td>UNSAT</td>
</tr>
<tr>
<td>( gto_{60c234} / p60c234 \times 10 \times 6 )</td>
<td>11015353</td>
<td>5000</td>
<td>UNSAT</td>
</tr>
<tr>
<td>( gto_{60c333} / p60c333 \times 10 \times 6 )</td>
<td>11102117</td>
<td>5000</td>
<td>UNSAT</td>
</tr>
<tr>
<td>( gto_{60c345} / p60c345 \times 10 \times 6 )</td>
<td>11203530</td>
<td>5000</td>
<td>UNSAT</td>
</tr>
<tr>
<td>( gto_{60c231} / p60c231 \times 10 \times 6 )</td>
<td>11303120</td>
<td>5000</td>
<td>UNSAT</td>
</tr>
<tr>
<td>( gto_{60c333} / p60c333 \times 10 \times 6 )</td>
<td>11403711</td>
<td>5000</td>
<td>UNSAT</td>
</tr>
<tr>
<td>( gto_{60c345} / p60c345 \times 10 \times 6 )</td>
<td>11503513</td>
<td>5000</td>
<td>UNSAT</td>
</tr>
<tr>
<td>( gto_{60c234} / p60c234 \times 10 \times 6 )</td>
<td>11610406</td>
<td>5000</td>
<td>UNSAT</td>
</tr>
</tbody>
</table>

These 32 instances are known to be hard for SAT solvers. In our experiment, MiniSat could not solve any of those within 5000 seconds. So, these instances are not interesting for the SAT competition-2018. As per the requirement of the benchmark submission, to generate 10 interesting SAT instances, we took a graph \( G \), namely hard.1.graph in the ASSAT website, and repeated the following steps until we get 10 interesting instances: (i) randomly remove a few arcs to produce a new graph \( G' \), (ii) generate the SAT instance \( S \) from the normal logic program encoding of the Hamiltonian cycle program along with \( G' \) by using ASSAT, and (iii) test \( S \) with MiniSat to see if it is interesting. The first 10 rows in Table 1 show these interesting instances. The next 10 rows give the details of the 10 hard instances taken from the ASSAT website.

\textsuperscript{1}http://curiouscheetah.com/Museum/Puzzle/Grandtour

\textsuperscript{2}In a complete graph, each pair of vertexes are connected by an edge.

\textsuperscript{3}http://assat.cs.ust.hk/hardsat.html
REFERENCES


Abstract—Multiplying two polynomials of degree \( n - 1 \) can need \( n^2 \) coefficient products, because each polynomial of degree \( n - 1 \) has \( n \) coefficients. If the coefficients are real numbers, the Fourier transformation allows to reduce the number of necessary coefficient products to \( O(n \times \log(n)) \). However, when the coefficients are not real numbers (e.g., the coefficients can be a matrix), the Fourier transformation cannot be used. In this case, reducing the number of necessary coefficient products can significantly speed up the multiplication of two polynomials.

In this short paper, we reduce the problem of multiplying two polynomials of degree \( n - 1 \) with \( t \) (\( t \leq n^2 \)) coefficient products to SAT and provide 20 new crafted SAT instances.

1. Introduction

A simple example of polynomial multiplication can be expressed using Equation 1:

\[
(ax + b)(cx + d) = acx^2 + (ad + bc)x + bd
\] (1)

The trivial multiplication of the two polynomials of degree 1 needs 4 coefficient products: \{ac, ad, bc, bd\}. A smart multiplication of the two polynomials needs only 3 coefficient products \{ac, (a + b)(c + d), bd\}, as expressed in Equation 2:

\[
(ax + b)(cx + d) = acx^2 + (a + b)(c + d) - ac - bd)x + bd
\] (2)

In Equation 2, we need more addition and subtraction operations than in Equation 1. However, multiplication is much more costly than addition and subtraction. So, we can multiply two polynomials of degree 1 more quickly using Equation 2 than using Equation 1.

In the general case, we want to multiply two polynomials of degree \( n - 1 \) using fewer than \( n^2 \) coefficient products. If the coefficients are real numbers, the Fourier transformation allows to reduce the number of necessary coefficient products to \( O(n \times \log(n)) \). However, when the coefficients are not real numbers (e.g., the coefficients can be a matrix), the Fourier transformation cannot be used.

In the sequel, we describe how to reduce the problem of multiplying two polynomials of degree \( n - 1 \) using \( t \) (\( t \leq n^2 \)) coefficient products to SAT. When the obtained SAT instance is satisfiable, the SAT solution gives a way to multiply two polynomials of degree \( n - 1 \) using \( t \) coefficient products. When the obtained SAT instance is unsatisfiable, we know that more than \( t \) coefficient products are needed. We refer to [1], [2] for other efficient algorithms for polynomials.

2. SAT Encoding of polynomial Multiplication Using \( t \) Products

Consider two polynomials of degree \( n - 1 \):

\[
A(x) = a_{n-1}x^{n-1} + a_{n-2}x^{n-2} + \cdots a_1x + a_0
\]

\[
B(x) = b_{n-1}x^{n-1} + b_{n-2}x^{n-2} + \cdots b_1x + b_0
\]

Their product is

\[
A(x) \times B(x) = c_{2n-2}x^{2n-2} + c_{2n-3}x^{2n-3} + \cdots + c_1x + c_0
\]

We want to compute \( A(x) \times B(x) \) using \( t \) (\( t \leq n^2 \)) coefficient products: \( P_1, P_2, \ldots, P_t \), where each \( P_l \) (\( 1 \leq l \leq t \)) is of the form \((a_1^l + \cdots + a_k^l)(b_1^l + \cdots + b_k^l)\) with \( a_1^l, a_2^l, \ldots \in \{a_{n-1}, a_{n-2}, \ldots, a_0\} \) and \( b_1^l, b_2^l, \ldots \in \{b_{n-1}, b_{n-2}, \ldots, b_0\} \). Addition and subtraction of these products give the coefficients \( c_k \) (\( 0 \leq k \leq 2n - 2 \)) of \( A(x) \times B(x) \). The problem becomes to determinize \( a_1^l \) and \( b_1^l \) for each product. In order to solve the problem, we first define the following Boolean variables.

- \( a_{ij} = 1 \) iff \( a_i \) is involved in product \( P_j \);
- \( b_{ij} = 1 \) iff \( b_j \) is involved in product \( P_i \);
- \( c_k = 1 \) iff product \( P_l \) is used to compute \( c_k \);
- \( x_{ijkl} = 1 \) iff \( a_i \) and \( b_j \) are involved in product \( P_l \), and product \( P_l \) is used to compute \( c_k \);

We then define the clauses of the CNF, which encode the following properties:

- \( x_{ijkl} \equiv a_{ij} \wedge b_{lj} \wedge c_k \)
For each $i$ and $j$ ($0 \leq i, j \leq n - 1$) and for each $k$ ($0 \leq k \leq 2n - 2$) such that $i + j \neq k$, if $a_i$ and $b_j$ are involved in product $P_l$ (i.e., $a_i \land b_j$ is implied) and $P_l$ is used to produce $c_k$, then the product of $a_i$ and $b_j$ should be eliminated by subtraction using another product $P'_l$ involving $a_i$ and $b_j$. If $i + j = k$, one product of $a_i$ and $b_j$ should remain in $c_k$. So,

$$\sum_{l=1}^{t} x_{ijkl} \mod 2 = \begin{cases} 0 & \text{if } i + j = k \\ 1 & \text{otherwise} \end{cases}$$

3. Set of Submitted Instances

We generated 20 SAT instances, using the encoding of the previous section, by varying $n$ and $t$ as follows:

- $n = 8$, $t \in \{60, 61, 62, 63\}$
- $n = 11$, $t \in \{118\}$
- $n = 13$, $t \in \{165, 166\}$
- $n = 14$, $t \in \{194\}$
- $n \in \{23, 27, 29, 37, 39, 42, 44, 45, 49, 51, 52, 54\}$, $t = 6$

Each combination of $n$ and $t$ gives an instance $poln-t$. Table 1 shows, for each one of the 20 generated instances, its number of variables and clauses, the status of the formula (satisfiable, unsatisfiable or unknown), and the time needed by MiniSat [3] to solve the instance on a computer with Intel Westmere Xeon E5-2680 of 2.40GHz and 20GB of memory under Linux. The cutoff time is 3600 seconds.

<table>
<thead>
<tr>
<th>Instance</th>
<th>#Variables</th>
<th>#Clauses</th>
<th>Satisfiability</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nb8T60</td>
<td>114180</td>
<td>453120</td>
<td>unknown</td>
<td>timeout</td>
</tr>
<tr>
<td>Nb8T61</td>
<td>116131</td>
<td>460800</td>
<td>unknown</td>
<td>timeout</td>
</tr>
<tr>
<td>Nb8T62</td>
<td>118082</td>
<td>468480</td>
<td>SAT</td>
<td>2737</td>
</tr>
<tr>
<td>Nb8T63</td>
<td>120033</td>
<td>476160</td>
<td>unknown</td>
<td>timeout</td>
</tr>
<tr>
<td>Nb11T168</td>
<td>597127</td>
<td>3378576</td>
<td>unknown</td>
<td>timeout</td>
</tr>
<tr>
<td>Nb13T165</td>
<td>1389990</td>
<td>5543200</td>
<td>unknown</td>
<td>timeout</td>
</tr>
<tr>
<td>Nb13T166</td>
<td>1398491</td>
<td>5577000</td>
<td>unknown</td>
<td>timeout</td>
</tr>
<tr>
<td>Nb14T194</td>
<td>2048090</td>
<td>8170848</td>
<td>unknown</td>
<td>timeout</td>
</tr>
<tr>
<td>Nb23T6</td>
<td>214791</td>
<td>952200</td>
<td>UNSAT</td>
<td>243.4</td>
</tr>
<tr>
<td>Nb27T6</td>
<td>348375</td>
<td>1545480</td>
<td>UNSAT</td>
<td>467.6</td>
</tr>
<tr>
<td>Nb29T6</td>
<td>432123</td>
<td>1977480</td>
<td>UNSAT</td>
<td>694.1</td>
</tr>
<tr>
<td>Nb37T6</td>
<td>900315</td>
<td>3997480</td>
<td>UNSAT</td>
<td>1441</td>
</tr>
<tr>
<td>Nb39T6</td>
<td>1054983</td>
<td>4684680</td>
<td>UNSAT</td>
<td>1343</td>
</tr>
<tr>
<td>Nb42T6</td>
<td>1318710</td>
<td>5856480</td>
<td>UNSAT</td>
<td>2354</td>
</tr>
<tr>
<td>Nb44T6</td>
<td>1516938</td>
<td>6737280</td>
<td>UNSAT</td>
<td>2295</td>
</tr>
<tr>
<td>Nb45T6</td>
<td>1623099</td>
<td>1623099</td>
<td>UNSAT</td>
<td>2184</td>
</tr>
<tr>
<td>Nb49T6</td>
<td>2097243</td>
<td>9315880</td>
<td>UNSAT</td>
<td>4301</td>
</tr>
<tr>
<td>Nb51T6</td>
<td>2365527</td>
<td>10508040</td>
<td>UNSAT</td>
<td>3185</td>
</tr>
<tr>
<td>Nb52T6</td>
<td>2507850</td>
<td>11140480</td>
<td>UNSAT</td>
<td>3310</td>
</tr>
<tr>
<td>Nb54T6</td>
<td>2809398</td>
<td>12480480</td>
<td>UNSAT</td>
<td>3249</td>
</tr>
</tbody>
</table>

Acknowledgments

Research partially supported by the National Natural Science Foundation of China (Grants no. 61370183, 61370184, 61472147), the MINECO-FEDER project RASO TIN2015-71799-C2-1-P, and the MeCS platform of university of Picardie Jules Verne.

References

Encodings of Relativized Pigeonhole Principle Formulas with Auxiliary Constraints

Jingchao Chen
School of Informatics, Donghua University
2999 North Renmin Road, Songjiang District, Shanghai 201620, P. R. China
chen-jc@dhu.edu.cn

Abstract—Although a relativized pigeonhole principle (RPHP) formula has been proven to require resolution proofs of size roughly $n^k$, if using its symmetry, a CDCL SAT solver solve easily it. To destroy its symmetry, here we add extra constraints so that the formula generated is a non-symmetry CNF.

I. INTRODUCTION

The SAT-encoding of relativized pigeonhole principle (RPHP) formulas yields a symmetry CNF [1]. In theory, such a CNF requires resolution proofs of size roughly $n^k$ [2]. However, if using a preprocessor such as BreakID [3] to detect symmetry and construct symmetry breaking formulas, it has polynomial-size bounded-depth proofs, and is solved easily by a general CDCL SAT solver [4]. In this documentation, we encode RPHP with additional constraints. The resulting CNF is non-symmetry.

II. SAT ENCODING OF RPHP WITH EXTRA CONSTRAINTS

A relativized pigeonhole principle (RPHP) formula is defined as whether $k$ pigeons can fly into $k-1$ holes via $n$ “resting places”. We denote such a claim by $RPHP_{k,n}$. Let $[n] = \{1, 2, \ldots, n\}$. Let $p$ and $q$ be functions: $[k] \rightarrow [n]$ and $[n] \rightarrow [k]$, respectively. Furthermore, assume that $p$ is one-to-one and defined on $[k]$, and $q$ is one-to-one and defined on the range of $p$. We encode $p$ and $q$ and a superset of the range of $p$ by Boolean variables $p_{u,v}$, $q_{v,w}$ and $r_v$. $RPHP_{k,n}$ is encoded as follows.

$$
p_{u,1} \lor p_{u,1} \lor \cdots \lor p_{u,n} \quad 1 \leq u \leq k
$$

$$
\overline{p}_{u,v} \lor \overline{p}_{u',v} \quad \text{for all } u \neq u', v
$$

$$
\overline{q}_{v,w} \lor r_v \quad \text{for all } u, v
$$

$$
\overline{q}_{v,1} \lor \overline{q}_{v,1} \lor \cdots \lor \overline{q}_{v,k-1} \quad 1 \leq v \leq n
$$

$$
\overline{r}_v \lor \overline{r}_v \lor \overline{q}_{v,w} \lor \overline{q}_{v',w} \quad \text{for all } v \neq v', w
$$

Let $m = n/2$. We add the following constraint clauses to $RPHP_{k,n}$.

$$
\overline{p}_{u,v} \lor x_v \quad 1 \leq u \leq k, 1 \leq v \leq m
$$

$$
\overline{q}_v \lor \overline{q}_v \lor \overline{q}_w \lor \overline{q}_w \quad 1 \leq v \leq m < v' \leq n, 1 \leq w < k
$$

$$
\overline{r}_v \lor \overline{r}_v \lor \overline{r}_w \lor \overline{r}_w \quad 1 \leq v \leq m < v' \leq n, 1 \leq w < k
$$

where $x_v$ is a additional variable different from $p_{u,v}$, $q_{v,w}$ and $r_v$. Clearly, $RPHP_{k,n}$ with the above constraint clauses is not symmetry.

REFERENCES


Random k-SAT

q-planted solutions

Adrian Balint
independent researcher
Ulm, Germany

Hiding solutions in a randomly generated formula can be accomplished in different ways. An overview of the different possibilities is given in [1] Section 2.2.3.

One of the most promising approaches is the one proposed in [4]. The authors define a single parameter $q$. A randomly generated clause that has $t$ satisfied literals in the planted solution is accepted with a probability proportional to $q^t$. The authors call this model the $q$-hidden model. To generate balanced instances (positive and negative literal probability should be the same) the authors show that $q$ should be set to $q = (\sqrt{5} - 1)/2 = 0.618$, which is the golden ratio. To derive this value we need some equations. In case of 3-SAT we have seven possible clauses that have one, two or three satisfied literals in the planted solution. For each type of clause we assign the probabilities $p_1$ for 1-satisfied and accordingly $p_2$ and $p_3$. Due to probability normalization we have the equation:

$$3p_1 + 3p_2 + p_3 = 1 \quad (1)$$

For an arbitrary position in the clause we have to balance the probabilities of negative and positive occurrence of a variable. Setting the sum of these probabilities to be equal, we can add the additional equation:

$$p_1 + 2p_2 + p_3 = 2p_1 + p_2 \quad (2)$$

Adding to the Equations 1, 2 the additional equations:

$$p_2/p_1 = p_3/p_2 = q \quad (3)$$

and then solving this equation system, we will get the same value for $q$ as mentioned above.

The value of $q$, which depends on $k$, can be computed as the positive root of the equation:

$$(1 - q)(1 + q)^{k-1} - 1 = 0 \quad (4)$$

To compute the values of $p_t$ we can use the general form of the normalization Equation 1, which then results in:

$$p_t = \frac{q^t}{(1 + q)^k - 1} \quad (5)$$

Table I shows the values of $q$ for typical $k$-SAT. Smaller values for $q$ will result in problems where local search algorithms are more likely to move away from the hidden solution then being attracted by it. This type of problems is also called deceptive formulas. A systematic construction of deceptive formulas that were shown to be very hard for local search solvers are also presented in [3].

A simple uniform $k$-SAT generator can be altered to generate planted solution instances according to the $q$-hidden model by adding an additional condition that checks if the clause is to be accepted or should be dropped depending on the planting model. After computing the number of true literals in the clause, the clause will be accepted according to an acceptance probability, which can be computed from the probability distribution $(p_1, \ldots, p_k)$ by normalizing the $p$-values with respect to $p_1$. This will result in an acceptance probability distribution of $(1, q, q^2, \ldots, q^{k-1})$ where the value of $q$ can be found in Table I.

For the SAT Competition 2018 we submit ten 3-SAT instances ($r=4.267, n=7000$), ten 5-SAT instances ($r=21.117, n=250$) and ten 7-SAT instances ($r=87.79, n=120$) generated according to the $q$-hidden model.

**REFERENCES**


<table>
<thead>
<tr>
<th>$k$-SAT</th>
<th>$q$</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>0.6180</td>
</tr>
<tr>
<td>4</td>
<td>0.8392</td>
</tr>
<tr>
<td>5</td>
<td>0.9275</td>
</tr>
<tr>
<td>6</td>
<td>0.9659</td>
</tr>
<tr>
<td>7</td>
<td>0.9835</td>
</tr>
</tbody>
</table>

**TABLE I**

The positive roots of Equation 5 for different values of $k$. 

Abstract—Deciding the $k$-colorability of a graph is a well-known NP-complete problem with symmetry properties.

I. DESCRIPTION

The $k$-colorability problem consists of deciding whether there exists an assignment (a coloring) of a set of $k$ colors to a graph’s nodes, such that a different color is assigned to two neighboring nodes.

An instance of this problem is encoded to a CNF-formula by introducing Boolean variables $x_{ij}$ that denote whether node $i$ is assigned color $j$, by constructing for each node a clause that requires at least one color to be assigned to the node, and by introducing for each edge in the graph a binary clause excluding that its incident nodes are assigned the same color (for each color).

The $k$-colorability problem exhibits strong symmetry properties, as permuting the colors in a coloring has no influence on whether neighboring nodes are colored differently. In other words, the colors are interchangeable, and the larger the set of colors (the higher $k$), the more this symmetry property will hinder search.

Aside from symmetry induced by color interchangeability, certain node permutations might also induce symmetry, depending on the structure of the graph at hand.

To generate 20 $k$-coloring encodings, we took graphs from Michael Trick’s Operations Research Page. For each graph, we chose $k$ such that the instance became unsatisfiable: too few colors are available for a valid coloring. In Table I, the specific $k$-values for each instance are presented.

Finally, we composed an instance set where Glucose 4.0 was able to solve 9 instances on an Intel(R) Xeon(R) E3-1225 cpu using 16 GB of memory and a timeout of 5000 seconds.

<table>
<thead>
<tr>
<th>instance name</th>
<th>$k$-value</th>
<th>Glucose 4.0 solve time (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>anna.col.11.cnf</td>
<td>10</td>
<td>101</td>
</tr>
<tr>
<td>david.col.11.cnf</td>
<td>10</td>
<td>103</td>
</tr>
<tr>
<td>hack.col.11.cnf</td>
<td>10</td>
<td>145</td>
</tr>
<tr>
<td>jean.col.10.cnf</td>
<td>9</td>
<td>8</td>
</tr>
<tr>
<td>le450_15a.col.15.cnf</td>
<td>14</td>
<td>5000+</td>
</tr>
<tr>
<td>le450_15b.col.15.cnf</td>
<td>14</td>
<td>5000+</td>
</tr>
<tr>
<td>le450_15c.col.15.cnf</td>
<td>14</td>
<td>5000+</td>
</tr>
<tr>
<td>le450_25a.col.25.cnf</td>
<td>24</td>
<td>5000+</td>
</tr>
<tr>
<td>le450_25b.col.25.cnf</td>
<td>24</td>
<td>5000+</td>
</tr>
<tr>
<td>le450_25d.col.25.cnf</td>
<td>24</td>
<td>5000+</td>
</tr>
<tr>
<td>mycie5.col.6.cnf</td>
<td>5</td>
<td>17</td>
</tr>
<tr>
<td>queen10_10.col.10.cnf</td>
<td>9</td>
<td>8</td>
</tr>
<tr>
<td>queen11_11.col.11.cnf</td>
<td>10</td>
<td>116</td>
</tr>
<tr>
<td>queen12_12.col.12.cnf</td>
<td>11</td>
<td>2429</td>
</tr>
<tr>
<td>queen13_13.col.13.cnf</td>
<td>12</td>
<td>5000+</td>
</tr>
<tr>
<td>queen14_14.col.14.cnf</td>
<td>13</td>
<td>5000+</td>
</tr>
<tr>
<td>queen15_15.col.15.cnf</td>
<td>14</td>
<td>5000+</td>
</tr>
<tr>
<td>queen8_12.col.12.cnf</td>
<td>11</td>
<td>2081</td>
</tr>
<tr>
<td>school1_nsh.col.14.cnf</td>
<td>13</td>
<td>5000+</td>
</tr>
<tr>
<td>school1.col.14.cnf</td>
<td>13</td>
<td>5000+</td>
</tr>
</tbody>
</table>

1 mat.gsia.cmu.edu/COLOR/instances.html
Searching for a Unit-Distance Graph with Chromatic Number 6

Marijn J. H. Heule

Department of Computer Science,
The University of Texas at Austin, United States

Abstract—This benchmark suite consists of determining whether some unit-distance graphs can be colored with 5 colors. Starting with a hard to color unit-distance graph, smaller graphs are produced by removing dozens of vertices at a time. All instances are satisfiable.

INTRO

The chromatic number of the plane (CNP), a problem first proposed by Edward Nelson in 1950 [1], asks how many colors are needed to color all points of the plane such that no two points at distance 1 from each other have the same color. Early results showed that at least four and at most seven colors are required. By the de Bruijn–Erdős theorem, the chromatic number of the plane is the largest possible chromatic number of a finite unit-distance graph [2]. The Moser Spindle, a unit-distance graph with 7 vertices and 11 edges, shows the lower bound [3], while the upper bound is due to a 7-coloring of the entire plane by John Isbell [1].

In a recent breakthrough for this problem, Aubrey de Grey improved the lower bound by providing a unit-distance graph with 1581 vertices with chromatic number 5 [4]. This graph was obtained by shrinking the initial graph with chromatic number 5 consisting of 20 425 vertices. The 1581-vertex graph is almost minimal: at most 4 vertices can be removed without introducing a 4-coloring of the remaining graph. The discovery by de Grey started a Polymath project to find smaller unit-distance graphs with chromatic number 5 and a graph with chromatic number 6 or 7.

PRELIMINARIES

A graph for which all edges have the same length is called a unit-distance graph. A lower bound for CNP of $k$ colors can be obtained by showing that a unit-distance graph has chromatic number $k$.

Given two sets of points $A$ and $B$, the Minkowski sum of $A$ and $B$, denoted by $A \oplus B$, equals $\{a + b \mid a \in A, b \in B\}$. Consider the sets of points $A = \{(0, 0), (1, 0)\}$ and $B = \{(0, 0), (1/2, \sqrt{3}/2)\}$, then $A \oplus B = \{(0, 0), (1, 0), (1/2, \sqrt{3}/2), (3/2, \sqrt{3}/2)\}$.

Given a positive integer $i$, we denote by $\theta_i$ the rotation around point $(0, 0)$ with angle $\arccos(\frac{2i}{2k})$ and by $\theta_i^k$ the application of $\theta_i$ $k$ times. Let $p$ be a point with distance $\sqrt{t}$ from $(0, 0)$, then the points $p$ and $\theta_i(p)$ are exactly distance 1 (unit distance) apart and thus would be connected with an edge in a unit-distance graph.

REFERENCES

SATcoin – Bitcoin mining via SAT

Norbert Manthey
nmanthey@conp-solutions.com
Dresden, Germany

Jonathan Heusser
jonathan.heusser@gmail.com
London, England

Abstract—The main security properties of Bitcoin, censorship resistance and an immutable historical ledger of transactions is enforced by its mining algorithm. By design this "proof of work" requires a large amount of computational power to verify transactions. While current mining algorithms are based on brute force, this document briefly describes how the same problem can be solved via SAT. Most details of this document are taken from [1].

I. BITCOIN MINING

A Bitcoin mining program essentially performs the following (in pseudo-code):

```python
nonce = MIN
while(nonce < MAX):
    if sha(sha(block+nonce)) < target:
        return nonce
    nonce += 1
```

The task is to find a nonce which, as part of the bitcoin block header, hashes below a certain value. This is a brute force approach to something like a preimage attack on SHA-256. The process of mining consists of finding an input to a cryptographic hash function which hashes below or equal to a fixed target value. At every iteration the content to be hashed is slightly changed to find a valid hash; there’s no smart choice in the nonce. The choice is essentially random as this is the best one can do on such hash functions.

In [1], Heusser proposed an alternative mining algorithm which does not perform a brute force search. Instead, we utilize tools from the program verification domain to find bugs or prove properties of programs, see as example [2]. To find the correct nonce or prove the absence of a valid nonce, a model checker backed by a SAT solver is used on a C implementation of the hashing. In contrast to brute force, which actually executes and computes many hashes, the new approach is symbolically executing the hash function with added constraints that are inherent in the bitcoin mining process. The submitted benchmark is based on CNFs created by CBMC.

II. BITCOIN MINING USING SAT SOLVING AND MODEL CHECKING

We take an existing C implementation of SHA-256 from a mining program and strip away everything but the actual hash function and the basic mining procedure of sha(sha(block)). This C file is the input to CBMC [3].

By adding the right assumptions and assertions to the implementation, i.e. MIN and MAX in the pseudo code above, we direct the SAT solver to find a nonce. Instead of a loop which executes the hash many times and a procedure which checks if we computed a correct hash, we add constraints that when satisfied implicitly have the correct nonce in its solution. The assumptions and assertions can be broken down to the following ideas:

- The nonce is modelled as a non-deterministic value
- The known structure of a valid hash, i.e. leading zeros, is encoded as assumptions in the model checker
- An assertion is added stating that a valid nonce does not exist

More details about the translation, and a basic solver performance comparison can be found in the full article [1].

III. SAT COMPETITION SPECIFICS

The used input file for CBMC implements the described analysis, and uses the so called genesis block of the bitcoin blockchain, the very first block of the chain. We provide a script that allows to produce a CNF based on the fact whether the formula should be satisfiable, or unsatisfiable. The formula is satisfiable, if the valid nonce is part of the given range. Hence, the range is constructed such that the valid nonce is right in the middle of the range to be analyzed. For unsatisfiable formulas, the range starts right after the valid nonce.

The code, as well as all scripts to create a benchmark are available at https://github.com/jheusser/satcoin.

REFERENCES


Logical Cryptanalysis Benchmarks for Classical and Modern Hash Functions

Alexander Scheel, Iowa State University, alexander.m.scheel@gmail.com

I. Introduction

This collection of benchmarks focuses on new techniques in logical cryptanalysis for analyzing the structure of various hash functions. We offer as a benchmark a new technique analyzing the security of a hash function of the Merkle-Damgard construction which are broadly applicable to all such constructions and several benchmarks analyzing the various security properties of the Keccak hash function. While the latter techniques have not yielded useful cryptanalysis results, they pose as a source of variable scalable benchmarks due to several choices of parameters (the bitwidth, w and the number of rounds).

II. Classical Hash Functions

The MD4 hash function is composed of 48 iterated rounds, each updating one of four 32-bit state variables [4]. While authors such as I. Mironov have applied SAT solvers to finding collisions given an existing differential path [3] and D. Jovanović has applied applied SAT solvers to brute force finding collisions and preimages in hash functions [2], these techniques either require pre-existing collisions or are too hard for SAT solvers on a large number of rounds.

One technique that is efficient for even a modest number of rounds (≤28) is the notion of a differential family search. In most cases, the differential path between two blocks b_1 and b_2 is the simple XOR difference between the intermediate rounds. For a given path p, there are often many such blocks which create a collision; I. Mironov showed that SAT solvers can find such blocks in MD4 and MD5 [3] relatively quickly. We can extend this concept to differential paths as well: given a differential path p, its family is the ordered tuple of indices of the rounds with a non-zero difference. For example, for the family of the differential path introduced by M. Schlaffer (in [6]) is:

(1, 2, 3, 5, 6, 7, 8, 9, 10, 11, 12, 15, 16, 19, 20, 35, 36)

Thus for a given family, f, there are multiple possible differential paths which have a structure described by f.

By using this concept of a differential family, we can divide the search space into multiple different SAT problems: for a given number of rounds, r, there are 2^{r-4} possible differential families (versus 2^{32w−32} possible differential paths). For r ≤ 28 it is possible to exhaustively search large portions of the family space; for 32 ≤ r ≤ 36 it becomes possible to find in select cases, and for r > 36, it remains impossible in nearly all cases (due to the large time limit). Our benchmarks are a sample of possible differential families for r = 24, mixing both SAT and UNSAT results. These benchmarks can be found in the families folder [5].

III. Modern Hash Functions

The notion of a differential family is not as useful for Keccak due to the interaction between the sponge function and the internal permutation functions causing the collision space of families to be entirely different between successive rounds. From an algebraic perspective, and to study the security margin of the XOF construct [1], it is important to show that the Keccak round functions are bijective and have high permutation order. Two of our benchmarks (bijection and orders) model these problems in SAT. In general, these instances are easy, but with a few notable outliers: for w ≥ 4, proving the order of θ is 3w is difficult. Note that this should simplify to showing that ∀x, x ≠ x is UNSAT, and thus should be relatively trivial; however w = 4 produces runtimes in excess of one hour. Further, anything later involving θ (such as ρ = θ, etc.) also becomes difficult for w ≥ 4.

In differences, we study the differential properties of the permutation functions: for a given parameters x, y ≤ 25w and round function f, we seek to find witnesses a and b such that:

#(1, a ⊕ b) = x
#(1, f(a) ⊕ f(b)) = y

That is, find an input with difference x which produces an output with difference y. There are a few interesting outliers in this model: for w ≥ 4, most round functions cannot be individually analyzed this way. Further, while π is a permutation of the order of the bits (and thus does not change the values of any bits), certain instances in w = 2 where x ≠ y produce runtimes in excess of an hour. This is surprising as the model is trivially SAT if and only if x = y (and the model merely involves changing the orders of variables).

The benchmarks in output-margins consider the effects of the sponge function. Since all of the round functions are bijective, if two inputs differ, the internal state of Keccak must also differ. However, to produce a collision, only the first y bits of the output must be the same. Thus we can create a model for a given number of differences x, z ≤ 25w, security margin y ≤ 25w; and round functions f:

#(1, a XOR b) = x
#(1, f(a) XOR f(b))[0 : y] = z

If such a witness a and b exist for z = 0, then the input difference x is possible of producing collisions at a security margin m. Our provided benchmarks sample the space for small values of w and relatively few round functions; for w ≥ 16 and for any set of functions including θ, the runtimes become exceedingly long.
The benchmarks in xof-state attempt to recreate the internal state of Keccak given a series of outputs from the XOF (extensible output function at a given margin). In general, this is possible for either small values of $w$ or small numbers of rounds (for larger values of $w$). However, care must be selected in choosing the base seed, otherwise, there can possibly be multiple satisfying seeds. However, for a set of inputs with unique solution, these benchmarks can be extended to contain the output from several rounds of Keccak. After a threshold dependent on the margin, these are redundant information and thus test the SAT solvers to work with larger models which overfit to the solution.

IV. Thanks

A special thanks to Mate Soos and his CryptoMiniSat solver [7] for suggesting these problems be submitted as benchmarks to the conference.

This work was completed as part of an Undergraduate Honors Research project at Iowa State University under the guidance of Eric W. Davis and as part of MATH 490: Independent Study of SHA-3 under the guidance of Clifford Bergman.

References

Benchmark: Assigning Time-Slots to Students with Optimal Preference Satisfaction

Markus Iser
Institute for Theoretical Computer Science,
Karlsruhe Institute of Technology
Karlsruhe, Germany
Email: markus.iser@kit.edu

Abstract—This family contains benchmarks from a tool that assigns time-slots to students by given individual preferences with local and global cost reduction. The problem is solved by encoding a sequence of SAT problems with decreasing local and global bounds.

I. INTRODUCTION
During each semester of our lecture, a number $m$ of tutors $S = \{s_0, \ldots, s_m\}$ is helping their fellow students to comprehend the subject matter through lessons in small tutorial groups. The trainer allocates $n$ free time-slots $T = \{t_0, \ldots, t_n\}$ in seminar rooms, in order to assign exactly one time-slot to each tutor. Before an assignment $z: S \rightarrow T$ can be determined, each tutor may order the available time-slots by their preference, such that preferred time-slot come first in their list. A cost function $c: S \times T \rightarrow \mathbb{N}$ can be deduced such that the cost of the assignment corresponds to its position in the ordered list.

The goal is to find an assignment $z$ of time-slots to tutors, such that their preferences are met as good as possible. In order to achieve this, we use two subsequent optimization rounds: one local and one global optimization round. In the local optimization round, we search the smallest $k_1$ and an assignment $z$ that satisfies $\max_{s \in S} c(s, z(s)) \leq k_1$. When the optimal $k_1$ is found, a second optimization round begins. In the global optimization round we search for the smallest $k_2$ and an assignment $z$ that satisfies $\sum_{s \in S} c(s, z(s)) \leq k_2$, and that still satisfies the local bound constraint with the previously determined $k_1$.

We use a tool [1] that generates a sequence of SAT problems to find an optimal solution for this problem. It follows a quick overview on the encoding.

II. ENCODING
So the input of our tool is a set of dates $T = \{t_0, \ldots, t_n\}$, a set of students $S = \{s_0, \ldots, s_m\}$ and the set of preferences $c: S \times T \rightarrow \mathbb{N}$.

The goal in the local optimization round is to find an assignment $z: S \rightarrow T$ and the minimal local bound $k_1$ such that $\max_{s \in S} c(s, z(s)) \leq k_1$ holds.

The goal in the subsequent global optimization round is to find an assignment $z: S \rightarrow T$ and the minimal global bound $k_2$ such that $\sum_{s \in S} c(s, z(s)) \leq k_2$ holds and the optimal local bound $k_1$ is still satisfied.

A. Boolean atoms
In our encoding we use $n \times m$ Boolean atoms $z_{st}$ for the assignment which shall be true iff date $t$ gets assigned to student $s$. For the costs we generate $n \times n \times m$ Boolean atoms $c_{sti}$ such that for each student $s$ and date $t$ we have $n$ cost atoms that form a unary representation of the cost of the assignment.

B. Constraints
Two basic constraints ensure that (A) $z$ assigns there is exactly one date to each student and (B) each assignment produces the costs given by the preferences $c$.

(A) $\forall s \in S, \sum_{t \in T} z_{st} = 1$.
(B) $\forall s \in S, \forall t \in T, z_{st} \implies \bigwedge_{i \leq c(s,t)} c_{sti}$.

C. Minimization
For optimization we use a cardinality encoding based on parallel counters [2] on the bound variables. Optimization is then realized incrementally calling a SAT solver with decreasing bounds.

For the local optimization round we encode $m$ cardinality constraints, one for each individual students cost.

$\forall s \in S, \sum_{t \in T} c_{sti} \leq k_1$, find smallest $k_1 \in [1, n]$

In the global optimization round we use 1 cardinality constraint over the sum of all costs.

$\sum_{s,t,i} c_{sti} \leq k_2$, find smallest $k_2 \in [m, n \times m]$

III. BENCHMARKS
We submitted a set of problems generated from real data with 34 students with individual preferences on 35 available dates. We sampled a sequence of 10 problems from the first optimization round with different local bounds and a sequence of 10 problems from the second optimization round with fixed local bound and different global bounds.
REFERENCES


SAT-Encodings of Tree Decompositions

Max Bannach
Institute for Theoretical Computer Science
Universität zu Lübeck
Email: bannach@tcs.uni-luebeck.de

Sebastian Berndt, Thorsten Ehlers and Dirk Nowotka
Department of Computer Science
University of Kiel
Email: {seb, the, dn}@informatik.uni-kiel.de

Abstract—We suggest some benchmarks based on a propositional encoding of tree decompositions of graphs.

I. INTRODUCTION

The treewidth of a graph is a fundamental property, often used e.g. in the field of fixed-parameter tractability (FPT). Intuitively speaking, it describes how tree-like a graph is. Many NP-hard problems on graphs can be solved efficiently if the input instance has bounded treewidth.

Recently, there have been two competitions seeking for efficient methods for computing the tree decomposition of graphs [1], seeking for both exact and heuristic algorithms.

Although the most successful exact approaches were based on a combination of dynamic programming and decomposition, we submit some SAT encodings of tree decompositions. Each of these formulas encodes the existence of a tree decomposition of a certain width. They were generated using Jdrasil [2]. The encoding basically describes an elimination ordering on a preprocessed version of the graph [3], from which a tree decomposition can be derived efficiently. This encoding is combined with some symmetry breaking constraints [2].

II. GRAPH SELECTION

For the generation of these formulas, we used graph instances which are publicly available at https://people.mmci.uni-saarland.de/~hdell/pace17/ex-instances-PACE2017-public-2016-12-02.tar.bz2. Out of these graphs, we selected those for which Jdrasil can compute an optimum tree decomposition within 1800 seconds. As the size of the SAT encoding is cubic with respect to the number of vertices, we only selected graphs with at most 200 nodes (after preprocessing). For each of them, we then created formulas which describe the existence of a tree decomposition of width $\text{opt} \pm 3$, where $\text{opt}$ is the treewidth of the respective graph.

III. BENCHMARK SELECTION

We generated 448 SAT formulas from 64 graphs (192 unsatisfiable, 256 satisfiable). For each of these formulas, we ran Glucose 3.0 for 5 hours, and partitioned the formulas into three categories:

- Easy: Solvable within 20 minutes.
- Medium: Solvable with 60 minutes.
- Hard: Solvable within 300 minutes.
- Very Hard: Not solved within 5 hours.

In order to provide formulas of different hardness, we then chose 5 satisfiable and unsatisfiable formulas from each of the first 3 categories. Furthermore, we chose the 5 smallest and largest formulas from the last category, as those might be especially interesting for the parallel track. This yields overall 40 formulas.

These formulas are available at http://www.informatik.uni-kiel.de/~the/sat_benchmarks_2018_ehlers_nowotka.tar.gz.

We are curious to see the difference of solver strengths on this kind of formulas.

ACKNOWLEDGMENT

This work is funded by the German Federal Ministry of Education and Research, combined project 01IH15006A.

REFERENCES


CBMC Formulae

Norbert Manthey
nmanthey@conp-solutions.com
Dresden, Germany

Michael Tautschnig
michael.tautschnig@qmul.ac.uk
Queen Mary University of London, UK

Abstract—SAT-based bounded model checking translates reachability questions over input programs into propositional satisfiability. This submission of benchmarks presents formulae generated by the C Bounded Model Checker (CBMC) on input programs from the TACAS Software Verification Competition (SV-COMP). CBMC uses Minisat 2.2 as its default back-end, and can optionally print the problems in DIMACS format. Profiling on SV-COMP benchmarks reports that from 10 to more than 50% of CBMC’s CPU time is spent in the SAT solver. Improvements in SAT solvers directly translate to reduced program analysis time.

I. THE CBMC TOOL

CBMC [1] is a bounded model checker for C, C++, and Java programs. It supports C89, C99, most of C11 and most compiler extensions provided by GCC and Visual Studio. It also supports SystemC using Scoot. More recently, support for Java bytecode has been added.

CBMC verifies array bounds (buffer overflows), pointer safety, arithmetic exceptions and user-specified assertions. Furthermore, it can check C and C++ for consistency with other languages, such as Verilog. The verification is performed by unwinding loops in the program and passing the resulting equation to a decision procedure.

CBMC has built-in support to translate bit-vector equation systems to gate level, and by default links in Minisat 2.2 [2] as a decision procedure. Interfaces to other SAT solvers are implemented as well, as is an IPASIR interface. As an alternative, CBMC supports external SMT solvers. Finally, CBMC can also dump the formula in DIMACS format instead of running a solver. This feature has been used to generate the benchmark set.

II. SAT COMPETITION BENCHMARK

The collection of formulae in this benchmark submission were generated from input problems of the TACAS Software Verification Competition (SV-COMP) [3]. Each problem is the output of CBMC for a given software benchmark, a bit-width (32 or 64 bits), and a loop unwinding bound. To determine these configuration parameters for the purpose of generating SAT benchmarks, we provide a script run.sh that first downloads the SV-COMP logs, extracts unwind values and bit width for CBMC, and finally runs CBMC and gzips the CNF. The script that is used to assemble the benchmark is maintained at https://github.com/tautschnig/sv-comp-sat.

To make the benchmarks more challenging, the unwind bound has been increased by 2. The generation time for CBMC has been limited to 30 seconds. Furthermore, any generated CNF formula is dropped as soon as Minisat 2.2 can solve it within 10 seconds.

III. AVAILABILITY

The source of CBMC can be found at https://github.com/diffblue/cbmc. For more details, please have a look at http://www.cprover.org/cbmc/.

REFERENCES


## Solver Index

<table>
<thead>
<tr>
<th>Solver</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>abcdSAT</td>
<td>10</td>
</tr>
<tr>
<td>abcdSAT n18</td>
<td>10</td>
</tr>
<tr>
<td>abcdSAT r18</td>
<td>10</td>
</tr>
<tr>
<td>BreakIDGlucose</td>
<td>40</td>
</tr>
<tr>
<td>BreakIDGlucoseSEL</td>
<td>40</td>
</tr>
<tr>
<td>CaDiCal</td>
<td>13</td>
</tr>
<tr>
<td>Candy</td>
<td>15</td>
</tr>
<tr>
<td>CBPeneLoPe2018</td>
<td>16</td>
</tr>
<tr>
<td>CCSPeneLoPe2018</td>
<td>16</td>
</tr>
<tr>
<td>COMiniSatPS Pulsar</td>
<td>17, 19</td>
</tr>
<tr>
<td>CryptoMiniSat 5.5 main-all4fixed</td>
<td>17</td>
</tr>
<tr>
<td>CryptoMiniSat 5.5 main-autotune-v19</td>
<td>17</td>
</tr>
<tr>
<td>CryptoMiniSat 5.5 main-glbreak-v19</td>
<td>17</td>
</tr>
<tr>
<td>CryptoMiniSat 5.5 main-otherconf-v20</td>
<td>17</td>
</tr>
<tr>
<td>CryptoMiniSat 5.5 main-v19</td>
<td>17</td>
</tr>
<tr>
<td>CryptoMiniSat 5.5 nolimits-autotune-v19</td>
<td>17</td>
</tr>
<tr>
<td>CryptoMiniSat 5.5 nolimits-glbreak-v19</td>
<td>17</td>
</tr>
<tr>
<td>CryptoMiniSat 5.5 nolimits-otherconf-v20</td>
<td>17</td>
</tr>
<tr>
<td>CryptoMiniSat 5.5 nolimits-v19</td>
<td>17</td>
</tr>
<tr>
<td>dimetheus</td>
<td>20</td>
</tr>
<tr>
<td>expGlucose</td>
<td>22</td>
</tr>
<tr>
<td>expGlucoseSilent</td>
<td>22</td>
</tr>
<tr>
<td>expMC_LRB_VSIDS_Switch</td>
<td>22</td>
</tr>
<tr>
<td>expMC_VSIDS_LRB_Switch_2500</td>
<td>22</td>
</tr>
<tr>
<td>GHackCOMSPS</td>
<td>19</td>
</tr>
<tr>
<td>glu_mix</td>
<td>10</td>
</tr>
<tr>
<td>Glucose 4.2.1</td>
<td>24</td>
</tr>
<tr>
<td>Glucose-3.0+width</td>
<td>44</td>
</tr>
<tr>
<td>Glucose-3.0_PADC_10</td>
<td>42</td>
</tr>
<tr>
<td>Glucose-3.0_PADC_3</td>
<td>42</td>
</tr>
<tr>
<td>Glucose_Hack_Kiel_fastBVE</td>
<td>47</td>
</tr>
<tr>
<td>GluHack</td>
<td>26</td>
</tr>
<tr>
<td>inIDGlucose</td>
<td>41</td>
</tr>
<tr>
<td>Lingeling</td>
<td>13</td>
</tr>
<tr>
<td>ManyGlucose 4.1-2</td>
<td>27</td>
</tr>
<tr>
<td>Maple_CM</td>
<td>44</td>
</tr>
<tr>
<td>Maple_CM_Dist</td>
<td>44</td>
</tr>
<tr>
<td>Maple_CM_ordUIP</td>
<td>44</td>
</tr>
<tr>
<td>Maple_CM_ordUIP+</td>
<td>44</td>
</tr>
<tr>
<td>Maple_LCM+BCrestart</td>
<td>28</td>
</tr>
<tr>
<td>Maple_LCM+BCrestart_M1</td>
<td>28</td>
</tr>
<tr>
<td>Maple_LCM_Dist_ChronoBT</td>
<td>29</td>
</tr>
<tr>
<td>Maple_LCM_M1</td>
<td>28</td>
</tr>
<tr>
<td>Maple_LCM_Scavel</td>
<td>30</td>
</tr>
<tr>
<td>Maple_LCM_Scavel_200</td>
<td>30</td>
</tr>
<tr>
<td>MapleCOMSPS_CHB_VSIDS</td>
<td>31</td>
</tr>
<tr>
<td>MapleCOMSPS_LRB_VSIDS</td>
<td>31</td>
</tr>
<tr>
<td>painless-maplecomsps</td>
<td>33</td>
</tr>
<tr>
<td>painless-maplecomsps-sym</td>
<td>33</td>
</tr>
<tr>
<td>PeneLoPe_MDLC</td>
<td>42</td>
</tr>
<tr>
<td>Plingeling</td>
<td>13</td>
</tr>
<tr>
<td>probSAT</td>
<td>35</td>
</tr>
<tr>
<td>ReasonLS</td>
<td>52</td>
</tr>
<tr>
<td>ReasonLS_h</td>
<td>52</td>
</tr>
<tr>
<td>Riss 7.1</td>
<td>36</td>
</tr>
<tr>
<td>satUZK-ddc</td>
<td>50</td>
</tr>
<tr>
<td>ScaLoPe</td>
<td>42</td>
</tr>
<tr>
<td>smallsat</td>
<td>10</td>
</tr>
<tr>
<td>SparrowToRiss</td>
<td>38</td>
</tr>
<tr>
<td>Syrup</td>
<td>24</td>
</tr>
<tr>
<td>TopoSAT2</td>
<td>47</td>
</tr>
<tr>
<td>Treengeling</td>
<td>13</td>
</tr>
<tr>
<td>Varisat</td>
<td>49</td>
</tr>
<tr>
<td>YalSAT</td>
<td>13</td>
</tr>
</tbody>
</table>
Benchmark Index

Bitcoin mining, 67
C bounded model checking, 73
Chromatic number of the plane, 66
Divider, 56
Floating-point program verification, 58
GrandTour\textsuperscript{Obs}, 59
Graph coloring, 65
Hash function cryptanalysis, 68
Planted Random SAT, 64
Polynomial multiplication, 61
Random SAT, 55
Reversing cellular automata, 57
Student timetabling, 70
Treewidth, 72
Unique inverse, 56
Author Index

Audemard, Gilles, 24
Baarir, Souheib, 33
Balint, Adrian, 35, 38, 64
Bannach, Max, 72
Berndt, Sebastian, 72
Biere, Armin, 13, 56
Bogaerts, Bart, 40

Cai, Shaowei, 52
Chen, Jingchao, 10, 63
Chen, Qingshan, 30
Chen, Shuwei, 30
Chowdhury, Md Solimul, 22, 59
Colange, Maximilien, 33
Czarnecki, Krzysztof, 31

Devriendt, Jo, 40, 41, 65
Djamegni, Clémentin Tayou, 42

Ehlers, Thorsten, 47, 72

Gableske, Oliver, 20
Ganesh, Vijay, 31, 58
Gotou, Yuuya, 27

Harder, Jannis, 49, 57
He, Kun, 28
Heule, Marijn J. H., 55, 66
Heusser, Jonathan, 67

Iser, Markus, 15, 70

Kordon, Fabrice, 33
Kutzner, Felix, 15

Lü, Zhipeng, 44, 61
Le Frioux, Ludovic, 33
Li, Chu-Min, 44, 61
Li, Yu, 44, 61
Li, Zhen, 28
Liang, Jia Hui, 31, 58
Luo, Mao, 44, 61

Müller, Martin, 22, 59
Manteney, Norbert, 36, 38, 67, 73
Manyà, Felip, 44, 61
Metin, Hakan, 33

Nabeshima, Hidemoto, 27
Nadel, Alexander, 29
Nowotka, Dirk, 47, 72

Oh, Chanseok, 19, 31
Poupart, Pascal, 31
Ryvchin, Vadim, 29
Schöning, Uwe, 35
Scheel, Alexander, 68
Scott, Joseph, 58
Simon, Laurent, 24
Sonobe, Tomohiro, 16
Soo, Mate, 17
Sopena, Julien, 33

Tautschnig, Michael, 73
Tchinda, Rodrigue Konan, 42
van der Grinten, Alexander, 50

Wu, Guanfeng, 30
Xiao, Fan, 44, 61
Xu, Yang, 30

You, Jia-Huai, 22, 59
Zha, Aolong, 26
Zhang, Xindi, 52