

The Challenge of Unifying Semantic and Syntactic Inference Restrictions

Christoph Weidenbach

Max Planck Institute for Informatics, Saarland Informatics Campus E1 4
66123 Saarbrücken, Germany

weidenb@mpi-inf.mpg.de

While syntactic inference restrictions don't play an important role for SAT, they are an essential reasoning technique for more expressive logics, such as first-order logic, or fragments thereof. In particular, they can result in short proofs or model representations. On the other hand, semantically guided inference systems enjoy important properties, such as the generation of solely non-redundant clauses. I discuss to what extent the two paradigms may be unifiable.

1 Introduction

In [29] I discussed the differences between simple-syntactic portfolio solvers and sophisticated-semantic portfolio solvers. In this paper I present the challenges in combining a solver based on syntactic inference restrictions with a solver based on the semantic guidance of inferences. More concretely, I investigate the differences between CDCL-style solvers building an explicit partial model assumption and solvers based on ordering and selection restrictions on inferences.

2 NP-Complete Problems

The prime calculus for SAT is CDCL (Conflict Driven Clause Learning) [26, 19, 6, 7]. The calculus can be viewed as a resolution variant where resolution inferences are selected via explicit partial model assumptions. The CDCL calculus operates on a five tuple (M, N, U, k, C) where M is a sequence of literals representing the current model assumption, called the *trail*, N the clause set under consideration, U a set of learned clauses, i.e., clauses derived via resolution, k the number guessed/decided literals in M , called the *decision level*, and C a clause that is \top if no conflict has occurred yet, some non-empty clause representing a found conflict, or \perp in case of an overall contradiction. Consider the clause set

$$N = \{P \vee Q \vee R, \neg R \vee S, \neg S \vee P \vee Q, \dots\}$$

where a CDCL run starting from the empty trail ε may result in

$$(\varepsilon; N; \emptyset; 0; \top) \\ \Rightarrow_{\text{CDCL}}^* ([\neg P^1 \neg Q^2 R^{P \vee Q \vee R} S^{\neg R \vee S}]; N; \emptyset; 2; \neg S \vee P \vee Q)$$

where the literals $\neg P$, $\neg Q$ were decided (guessed) at decision levels 1, 2, respectively; R is a propagated literal via the clause $P \vee Q \vee R$; S is propagated via $\neg R \vee S$ and finally in the partial assignment $[\neg P^1 \neg Q^2 R^{P \vee Q \vee R} S^{\neg R \vee S}]$ the clause $\neg S \vee P \vee Q$ is false. The conflict is solved by resolving the false clause with clauses propagating literals from the trail. First with the clause $\neg R \vee S$ resulting in the clause

$\neg R \vee P \vee Q$ and then with the clause $P \vee Q \vee R$ resulting in the clause $P \vee Q$. Now this clause is learned yielding the new CDCL state

$$\Rightarrow_{\text{CDCL}}^* ([\neg P^1 Q^{P \vee Q}]; N; \{P \vee Q\}; 1; \top).$$

Most importantly, the new clause $P \vee Q$ is *non-redundant*, i.e., it is not implied by smaller clauses from N , where the ordering is a lifting of the literal ordering generated by the trail [1, 14]. Propositional non-redundancy is itself an NP-complete property. Hence, the CDCL polynomial time model generation procedure either finds a model or eventually leads to learned clause that enjoys an NP-complete property. This has several immediate consequences: (i) termination (ii) the approach of forgetting clauses works. Note that since only non-redundant clauses are learned by CDCL, a forgotten clause that has become redundant will not be generated a second time. Hence, learning plus forgetting can be seen as an efficient way to getting rid of redundant clauses.

Please recall, the classical first-order notion of redundancy means “not needed to find a proof or model”. Therefore, a redundant clause can be deleted. In the superposition or ordered resolution context this abstract definition is instantiated with “not implied by smaller clauses” [3, 20]. However, in the context of SAT, redundancy is often defined as a synonym for “satisfiability preserving” [16]. So in the context of many SAT related papers redundant clauses must not be removed, in general, or completeness is lost. These two notions of redundancy must not be confused and I stick here to the classical first-order notion.

The prerequisites for learning non-redundant clauses are twofold:

1. Propagation needs to be exhaustive.
2. Conflict detection needs to be eager.

The number of literals of a SAT problem is typically small compared to the respective clause set size. Therefore, exhaustive propagation in SAT is not an efficiency issue, as well as the detection of conflicts, i.e., false clauses. Surprisingly, what holds for SAT, does not hold for NP-complete problems, in general. It seems that the language of propositional logic is a nice compromise between expressivity, succinctness, and the efficiency of propagation.

Consider the NP-complete problem of testing satisfiability of a system of linear arithmetic inequations over the integers [17, 21]. The language of linear integer arithmetic is more succinct than propositional logic, i.e., the encoding of an integer variable x requires linearly many propositional variables, because the a priori bounds for solvability are simply exponential. Now consider a CDCL style procedure testing solvability of a LIA (Linear Integer Arithmetic) system of inequations [11]. Consider the example system

$$N = \{1 - x - y \leq 0, x - y \leq 0, \dots\}$$

and a CDCL-style run via CATSAT+ [11], where the partial model assumption is represented by simple bounds, i.e., inequations of the form $x \# c$, $c \in \mathbb{Z}$, $\# \in \{<, >, \leq, \geq\}$.

$$\begin{aligned} &(\varepsilon; N; \emptyset; 0; \top) \\ &\Rightarrow_{\text{CATSAT}^+}^* ([x \geq 5^1 y \geq 6^{1-x-y \leq 0} x \geq 6^{x-y \leq 0} \dots]; N; \emptyset; 1; \top) \end{aligned}$$

Obviously, the propagation does not terminate, except for a priori simply exponential bounds: if m is the number of inequations in N , n the number of different variables in N and a the maximal coefficient in N , then the problem has a solution iff $-n \cdot (m \cdot a)^{2m+1} \leq x \leq n \cdot (m \cdot a)^{2m+1}$ for every variable x in

N [21]. Hence, there is currently no CDCL-style procedure known that does exhaustive propagation. Therefore, the learned clauses [11] are not non-redundant. The current state-of-the-art procedures don't use a CDCL-style approach but apply a relaxation of LIA to LRA (linear rational arithmetic) and use LRA solutions for a branch-and-bound approach. This general idea is complemented via simplifications and fast, sufficient tests for the existence of a solution [9, 18, 12, 10].

3 NEXPTIME-Complete Problems

The same phenomenon that occurs at LIA, see Section 2, also shows up if the complexity is “slightly” increased from SAT to the satisfiability of the Bernays-Schoenfinkel (BS) fragment of first-order logic [5], which is NEXPTIME-Complete [25]. For example, consider the following clause set [22]

$$N = \left\{ \begin{array}{l} 1 : P(0,0,0,0), \\ 2 : \neg P(x_1, x_2, x_3, 0) \vee P(x_1, x_2, x_3, 1), \\ 3 : \neg P(x_1, x_2, 0, 1) \vee P(x_1, x_2, 1, 0), \\ 4 : \neg P(x_1, 0, 1, 1) \vee P(x_1, 1, 0, 0), \\ 5 : \neg P(0, 1, 1, 1) \vee P(1, 0, 0, 0), \\ 6 : \neg P(1, 1, 1, 1) \end{array} \right\}$$

where a CDCL-style run, using ground literals for the partial model assumptions [14], will propagate all values of the 4-bit counter represented by P :

$$\begin{aligned} & (\varepsilon; N; \emptyset; 0; \top) \\ & \Rightarrow_{\text{SCL}}^* ([P(0,0,0,0)^{C1} P(0,0,0,1)^{C2} P(0,0,1,0)^{C3} \dots P(1,1,1,1)^{C6}], N, \emptyset, 0, \neg P(1,1,1,1)) \end{aligned}$$

before detecting the conflict, where for the propagation justifications the notation $C < \text{clause number} >$ was used. Thus, there are exponentially many propagations, in general. Still, the clause set can be refuted in linearly many steps by starting with resolution steps between the clauses 2 – 4:

$$\begin{array}{ll} 2.2 \text{ Res } 3.1 & 7 : \neg P(x_1, x_2, 0, 0) \vee P(x_1, x_2, 1, 0) \\ 7.2 \text{ Res } 2.1 & 8 : \neg P(x_1, x_2, 0, 0) \vee P(x_1, x_2, 1, 1) \\ 8.2 \text{ Res } 4.1 & 9 : \neg P(x_1, 0, 0, 0) \vee P(x_1, 1, 0, 0) \\ 9.2 \text{ Res } 8.1 & 10 : \neg P(x_1, 0, 0, 0) \vee P(x_1, 1, 1, 1) \\ 10.2 \text{ Res } 5.1 & 11 : \neg P(0, 0, 0, 0) \vee P(1, 0, 0, 0) \\ 11.2 \text{ Res } 10.1 & 12 : \neg P(0, 0, 0, 0) \vee P(1, 1, 1, 1) \\ 12.1 \text{ Res } 6.1 & 13 : \perp. \end{array}$$

where this proof can be implemented by a standard ordering restriction on the P atoms, e.g., via a KBO (Knuth Bendix Ordering), plus a respective selection strategy [2]. For example, for the first step the second literal out of clause 2 is maximal with respect to a KBO instance where $1 >_{\text{KBO}} 0$ and the first literal of clause 3 is selected.

The above clause set N without clause 6 is obviously satisfiable. Still, a CDCL-style calculus using ground literals generates exponentially many ground atoms before it detects satisfiability. An ordered resolution calculus using a standard KBO instance where $1 >_{\text{KBO}} 0$ does not generate any clause at all, because exactly the positive literals are maximal. Extending the model representation language from ground atoms [14] to more a more expressive language including variables [4, 13, 24, 1, 8] solves the

issue with the above example by starting with a model assumption $[P(x_1, x_2, x_3, x_4)]^1$. However, the more expressive model language requires more complex operations in order to guarantee consistency of the model assumption and to detect propagating literals and false clauses. The discrepancy between a compact model representation and complex calculations cannot be resolved in general, because $\text{NP} \neq \text{NEXPTIME}$.

In addition, for “practically relevant” instances of the BS fragment the situation is the same. We have shown [28] that the YAGO [27] fragment of BS can be effectively saturated by a variant of ordered resolution with chaining, whereas all model-guided approaches fail. Again for the reason that propagation with respect to millions of constants cannot be efficiently done, yet.

4 Unification

The two frameworks, syntactic inference restrictions and model-guided inferences cannot be easily combined. Obviously, since ordered resolution may generate redundant clauses, but model-guided inferences with exhaustive propagation and eager conflict detection do not, model-guided inferences cannot simulate resolution inferences. The resolution inferences in Section 3 refuting the clause set encoding a counter are not redundant. Still, they cannot be simulated by a CDCL-style calculus [14] because it will immediately run into the exponentially many propagation steps and find a refutation of exponential size this way. It seems to be non-trivial and non-obvious how the two paradigms can be unified. If eager propagation is dropped, then model-guided inferences can simulate resolution [23, 14], however, in this case non-redundancy of learned clauses is lost, in general.

One way out of this dilemma could be to limit the amount of propagations by limiting the number of literals that may be derived by propagation. For example, the InstGen calculus [15] limits ground instantiation and the generation of new ground literals to using exactly one constant. Then the potential size of a model assumption remains linear in the size of the investigated clause set. However, a terminating model assumption search does not result in an overall model for the clause set anymore. Thus the model generation process itself needs to turn into a “learning” procedure.

References

- [1] Gábor Alagi & Christoph Weidenbach (2015): *NRCL - A Model Building Approach to the Bernays-Schönfinkel Fragment*. In Carsten Lutz & Silvio Ranise, editors: *Frontiers of Combining Systems - 10th International Symposium, FroCoS 2015, Wroclaw, Poland, September 21-24, 2015. Proceedings, Lecture Notes in Computer Science* 9322, Springer, pp. 69–84, doi:10.1007/978-3-319-24246-0_5.
- [2] Leo Bachmair & Harald Ganzinger (1994): *Rewrite-based Equational Theorem Proving with Selection and Simplification*. *Journal of Logic and Computation* 4(3), pp. 217–247, doi:10.1093/logcom/4.3.217. Revised version of Max-Planck-Institut für Informatik technical report, MPI-I-91-208, 1991.
- [3] Leo Bachmair & Harald Ganzinger (2001): *Resolution Theorem Proving*. In Alan Robinson & Andrei Voronkov, editors: *Handbook of Automated Reasoning*, chapter 2, I, Elsevier, pp. 19–99, doi:10.1016/B978-044450813-3/50004-7.
- [4] Peter Baumgartner, Alexander Fuchs & Cesare Tinelli (2006): *Lemma Learning in the Model Evolution Calculus*. In: *LPAR, Lecture Notes in Computer Science* 4246, Springer, pp. 572–586, doi:10.1016/0004-3702(94)90077-9.
- [5] Paul Bernays & Moses Schönfinkel (1928): *Zum Entscheidungsproblem der mathematischen Logik*. *Mathematische Annalen* 99, pp. 342–372, doi:10.1007/BF014591101.

- [6] Armin Biere, Marijn Heule, Hans van Maaren & Toby Walsh, editors (2009): *Handbook of Satisfiability. Frontiers in Artificial Intelligence and Applications* 185, IOS Press.
- [7] Jasmin Christian Blanchette, Mathias Fleury & Christoph Weidenbach (2016): *A Verified SAT Solver Framework with Learn, Forget, Restart, and Incrementality*. In Nicola Olivetti & Ashish Tiwari, editors: *Automated Reasoning - 8th International Joint Conference, IJCAR 2016, Coimbra, Portugal, June 27 - July 2, 2016, Proceedings, Lecture Notes in Computer Science* 9706, Springer, pp. 25–44, doi:10.1007/978-3-319-43144-4.
- [8] Maria Paola Bonacina & David A. Plaisted (2016): *Semantically-Guided Goal-Sensitive Reasoning: Model Representation*. *Journal of Automated Reasoning* 56(2), pp. 113–141, doi:10.1007/s10817-015-9334-4.
- [9] Aaron R. Bradley & Zohar Manna (2007): *The Calculus of Computation – Decision Procedures with Applications to Verification*. Springer, doi:10.1007/978-3-540-74113-8.
- [10] Martin Bromberger (2018): *A Reduction from Unbounded Linear Mixed Arithmetic Problems into Bounded Problems*. In Didier Galmiche, Stephan Schulz & Roberto Sebastiani, editors: *Automated Reasoning - 9th International Joint Conference, IJCAR 2018, Held as Part of the Federated Logic Conference, FloC 2018, Oxford, UK, July 14-17, 2018, Proceedings, Lecture Notes in Computer Science* 10900, Springer, pp. 329–345, doi:10.1145/322276.322287.
- [11] Martin Bromberger, Thomas Sturm & Christoph Weidenbach (2015): *Linear Integer Arithmetic Revisited*. In Amy P. Felty & Aart Middeldorp, editors: *Automated Deduction - CADE-25 - 25th International Conference on Automated Deduction, Berlin, Germany, August 1-7, 2015, Proceedings, Lecture Notes in Computer Science* 9195, Springer, pp. 623–637, doi:10.1016/S0747-7171(08)80051-X.
- [12] Martin Bromberger & Christoph Weidenbach (2017): *New techniques for linear arithmetic: cubes and equalities*. *Formal Methods in System Design* 51(3), pp. 433–461, doi:10.1007/s10703-017-0278-7.
- [13] Christian G. Fermüller & Reinhard Pichler (2007): *Model Representation over Finite and Infinite Signatures*. *J. Log. Comput.* 17(3), pp. 453–477, doi:10.1093/logcom/exm008.
- [14] Alberto Fiori & Christoph Weidenbach (2019): *SCL – Clause Learning from Simple Models*. In: *Automated Deduction - CADE-27, 27th International Conference on Automated Deduction, Lecture Notes in Artificial Intelligence* 11716, Springer, pp. 233–249, doi:10.1007/978-3-030-29436-6_14.
- [15] Harald Ganzinger & Konstantin Korovin (2003): *New Directions in Instatiation-Based Theorem Proving*. In Samson Abramsky, editor: *18th Annual IEEE Symposium on Logic in Computer Science, LICS'03, LICS'03*, IEEE Computer Society, pp. 55–64, doi:10.1109/LICS.2003.1210045.
- [16] Marijn J. H. Heule, Benjamin Kiesl & Armin Biere (2019): *Encoding Redundancy for Satisfaction-Driven Clause Learning*. In Tomás Vojnar & Lijun Zhang, editors: *Tools and Algorithms for the Construction and Analysis of Systems - 25th International Conference, TACAS 2019, Held as Part of the European Joint Conferences on Theory and Practice of Software, ETAPS 2019, Prague, Czech Republic, April 6-11, 2019, Proceedings, Part I, Lecture Notes in Computer Science* 11427, Springer, pp. 41–58, doi:10.1007/978-3-030-17462-0.
- [17] Richard M. Karp (1972): *Reducibility Among Combinatorial Problems*. In Raymond E. Miller & James W. Thatcher, editors: *Proceedings of a symposium on the Complexity of Computer Computations, held March 20-22, 1972, at the IBM Thomas J. Watson Research Center, Yorktown Heights, New York.*, The IBM Research Symposia Series, Plenum Press, New York, pp. 85–103, doi:10.1007/978-1-4684-2001-2_9.
- [18] Daniel Kroening & Ofer Strichman (2008): *Decision Procedures An Algorithmic Point of View*. Texts in Theoretical Computer Science, Springer. Second Edition 2016.
- [19] Matthew W. Moskewicz, Conor F. Madigan, Ying Zhao, Lintao Zhang & Sharad Malik (2001): *Chaff: Engineering an Efficient SAT Solver*. In: *Design Automation Conference, 2001. Proceedings*, ACM, pp. 530–535, doi:10.1145/378239.379017.
- [20] Robert Nieuwenhuis & Albert Rubio (2001): *Paramodulation-Based Theorem Proving*. In Alan Robinson & Andrei Voronkov, editors: *Handbook of Automated Reasoning*, chapter 7, I, Elsevier, pp. 371–443, doi:10.1016/B978-044450813-3/50009-6.

- [21] Christos H. Papadimitriou (1981): *On the complexity of integer programming*. *Journal of the ACM* 28(4), pp. 765–768, doi:10.1145/322276.322287.
- [22] Juan Antonio Navarro Pérez & Andrei Voronkov (2008): *Proof Systems for Effectively Propositional Logic*. In Alessandro Armando, Peter Baumgartner & Gilles Dowek, editors: *Automated Reasoning, 4th International Joint Conference, IJCAR 2008, Sydney, Australia, August 12-15, 2008, Proceedings, LNCS 5195*, Springer, pp. 426–440, doi:10.1023/A:1005806324129.
- [23] Knot Pipatsrisawat & Adnan Darwiche (2009): *On the Power of Clause-Learning SAT Solvers with Restarts*. In: *CP, Lecture Notes in Computer Science 5732*, Springer, pp. 654–668, doi:10.1007/11591191_40.
- [24] Ruzica Piskac, Leonardo Mendonça de Moura & Nikolaj Bjørner (2010): *Deciding Effectively Propositional Logic Using DPLL and Substitution Sets*. *Journal of Automated Reasoning* 44(4), pp. 401–424, doi:10.1007/s10817-009-9161-6.
- [25] David A. Plaisted (1984): *Complete Problems in the First-Order Predicate Calculus*. *Journal of Computer and System Sciences* 29, pp. 8–35, doi:10.1016/0022-0000(84)90010-2.
- [26] João P. Marques Silva & Karem A. Sakallah (1996): *GRASP - a new search algorithm for satisfiability*. In: *International Conference on Computer Aided Design, ICCAD, IEEE Computer Society Press*, pp. 220–227, doi:10.1109/ICCAD.1996.569607.
- [27] Fabian M. Suchanek, Gjergji Kasneci & Gerhard Weikum (2007): *Yago: a core of semantic knowledge*. In Carey L. Williamson, Mary Ellen Zurko, Peter F. Patel-Schneider & Prashant J. Shenoy, editors: *Proceedings of the 16th International Conference on World Wide Web, WWW 2007, Banff, Alberta, Canada, May 8-12, 2007*, ACM, pp. 697–706, doi:10.1145/1242572.1242667.
- [28] Martin Suda, Christoph Weidenbach & Patrick Wischniewski (2010): *On the Saturation of YAGO*. In: *Automated Reasoning, 5th International Joint Conference, IJCAR 2010, LNAI 6173*, Springer, Edinburgh, United Kingdom, pp. 441–456, doi:10.1007/978-3-642-14203-1_38.
- [29] Christoph Weidenbach (2017): *Do Portfolio Solvers Harm?* In Giles Reger & Dmitriy Traytel, editors: *ARCADE 2017, 1st International Workshop on Automated Reasoning: Challenges, Applications, Directions, Exemplary Achievements, Gothenburg, Sweden, 6th August 2017, EPiC Series in Computing 51*, EasyChair, pp. 76–81, doi:10.29007/vpxm.