

Foundations of Automated Deduction and Satisfiability, iDEAS TIN2004-04343

María Luisa Bonet Carbonell *
Universidad Politecnica de Catalunya

Abstract

The project is about searching for new algorithms to show the satisfiability of Boolean formulas. It is also about studying which proof systems are the most efficient to show their unsatisfiability. Both directions have applications in the fields of circuit verification, program verification with model checking, automated deduction, and others.

The satisfiability problem is a particular case of the constraint satisfaction problem. The decisional form of both problems is NP-Complete. Constraint satisfaction problems appear in a large variety of practical situations, such as scheduling, temporal reasoning, machine vision, etc., so their study is important. We will study this topic both from the point of view of verifying satisfiability, and from the point of view of showing the unsatisfiability of constraints via proof systems. For that we will need to define proof systems that will work with domains more general than 0/1. This is a new and promising approach, because it brings to the field of constraint satisfaction problems the techniques from proof complexity and finite model theory that had not been used before in this context.

Keywords: Constraint Satisfaction, Satisfiability, Automatizability, Proof Systems.

1 Project Objectives

Obj. 1 *Proving Lower Bounds for Propositional Proof Systems.*

Obj. 2 *Automated Deduction. Analysis of the Theoretical Limitations*

Obj. 3 *Study and Comparison of Different Measures of Complexity for Propositional Proof Systems*

Obj. 4 *Study of Structural Properties of SAT and CSP Instances*

Task 4.1 Try to find structural properties of boolean formulas that allow us to detect tractable cases of SAT.

*Email: bonet@lsi.upc.edu

Task 4.2 We will study the topic of phase transition of random instances of SAT. The point of phase transition is the ratio number of variables divided by number of clauses that separates the satisfiable formulas from the unsatisfiable ones. We want to study properties of this random formulas related to the point of phase transition.

Task 4.3 We will work on the tractable cases of the problem of morphism. This is part of the general objective of obtaining a complete characterization of the tractable cases of the CSP problem.

Obj. 5 *Definition of Proof Systems for CSPs and Study of the Connection with Finite Model Theory*

Task 5.1 Definition of Proof Systems for CSPs. This proof systems would allow to refute the satisfiability of constraints. Our theoretical results would give insights for future tools to refute CSPs.

Task 5.2 Study the relationship between Finite Model Theory and Constraint Satisfaction Problems (CSPs).

Tar.	Centro	Investigadores	1º año	2º año	3º año
T 1.1	UPC	Atserias, Bonet, Postdoc 1			
T 1.2	UPC	Esteban, Galesi			
T 1.3	UPC	Atserias, Esteban, Galesi			
T 2.1	UPC, CSIC	Bonet, Levy, Villaret			
T 2.2	UPC, U. Toronto	Bonet, Pitassi, Postdoc 1			
T 2.3	UPC	Atserias, Bonet			
T 2.4	UPC, CSIC	Bonet, Levy, Villaret			
T 3.1	UPC	Esteban, Galesi			
T 3.2	UPC	Atserias, Esteban, Bonet			
T 4.1	UPC	Lozano, Galesi			
T 4.2	UPC, U. Toronto	Atserias, Pitassi			
T 4.3	UPF	Dalmau, Postdoc 2			
T 5.1	UPC, Rice U.	Atserias, Dalmau, Vardi			
T 5.2	UPC, UPF, Rice U.	Dalmau, Postdoc 2, Vardi			

2 Level of Success Attained

2.1 Results on Objective 1

We obtained some results on the task 1.2 that appear in the publication [24]. In it the authors prove lower bounds for the size and the space of refutations for the propositional proof system

Res(k) tree-like. Res(k) is an extension of the Resolution proof system where we don't only work with clauses, but also we can use disjunctions of conjunctions of up to k literals. Also in the article the authors show that the proof systems Res(k) for $k \geq 1$ form a strict hierarchy. This means that we can exponentially separate Res(k) from Res($k-1$), and therefore the second system has to refute some classes of formulas with exponentially more clauses than the first.

2.2 Results on Objective 2

In the article number [5] we study different notions of automatizability for various propositional proof systems. Also we determine which of these notions are equivalent and which aren't. Furthermore we see which properties should fulfill the system of Resolution to be weakly automatizable.

2.3 Results on Objective 3

Members of our group (among others) see [15] are writing a manuscript about a new measure of formulas, called the Strahler number. One way of defining this measure for a formula is the minimal complete binary tree that can be embedded into a refutation of the formula. This measure also corresponds with the minimal tree-like space of writing a refutation of a formula. We show how the Strahler number is a measure of how structured a formula is, and a small Strahler means that the formula is easy to refute and viceversa. Small Strahler usually means real life instance while high Strahler is related to random instances and hard combinatorial principles (crafted instances). We have developed an algorithm to obtain refutations based on our new measure that could be competitive for real life instances.

2.4 Results on Objective 4

Paper [25] is about task T.4.1. In this work, the authors combine graph theory and linear algebra to study SAT problems of low "linear algebra complexity", i.e. formulas with bounded hermitian rank. They show polynomial time SAT algorithms for classes of formulas with hermitian rank at most one. They do this by applying methods from hypergraph transversal theory.

Task T.4.2 proposed the study of properties of random instances for distributions with phase transition. The most important result that we obtained in this topic is the article [3]. In this publication the authors show that the standard distribution on 3-CNF formulas is indistinguishable from the planted distribution respect to the properties expressible in first order logic. To obtain this result, the authors needed a detailed study of the Shelah-Spencer results. Adapting these results to the planted distribution is not at all trivial technically. On the contrary, the results are highly difficult.

In task 4.3 great progress has been made. The articles [9, 13, 18, 17, 21, 16, 19, 20, 22, 23] deal with possible tractable problems of constraint satisfaction (CSP). Some of these articles study methods like gaussian elimination that have helped to detect the class of paraprimal algebras that has been shown tractable. Also, we have studied notions of consistency for quantified CSPs and CSPs of bounded arity.

2.5 Results on Objective 5

In task T.5.1 we proposed to define proof systems for CSPs. Our proposal (article [7]) consists in defining instances of the constraint satisfaction problem as a relational database, and as a consequence to use the relational algebra operations or join, projection and weakening to derive new restrictions. In this article, the authors show that a CSP instance is unsatisfiable if and only if one can derive the empty relation with the relational algebra operations. Next the authors propose the concept of width of a CSP refutation and they characterize it using Ehrenfeucht-Fraïssé games. This article manages to connect the area of constraint satisfaction problems with the area of finite model theory, as we also wished to do in this task.

Also in the same article the authors propose a new proof system where the lines of the proof are OBDDs, and the rules are again join, projection and weakening. The article contains proofs of the fact that this system simulates Resolution and Cutting Planes with polynomial coefficients, and also that the system admits monotone interpolation. It would be desirable that this new proof system would have practical applications the same that Resolution has. In fact in the conference SAT 2004 Pan and Vardi published an experimental study.

Recently, Jan Krajíček from the Academy of Sciences of the Czech Republic, communicated us a solution to one of the open problems proposed in the article. Cavignetto and himself have proposed a method to build propositional unsatisfiable formulas for which, for every ordering of the variables, the minimum OBDD refutation is exponential. The result is based on [7] using an additional idea. The fact that such prestigious researcher as Krajíček has worked on open problems that our group formulated is an indication of our success.

Under the task T.5.2 several papers have been written. The first is [4]. In this work the authors work with Tseitin tautologies, well known in the area of proof complexity, to show width lower bounds on Datalog programs to solve certain CSPs. This work presents a simplification and improvement to a former result of Feder and Vardi.

The second paper we want to mention for this task is [2]. This article received the “Best Newcomer Award” of the conference ICDT 2005. In this article the authors propose the adaptation of an algorithm of Beame and Pitassi to show how to automatize tree-like resolution to the context of CSPs. The adaptation is not as direct as it seems, and the article leaves some open problems that we know some foreign researchers are working on.

Also for the task 5.2, other results have been obtained. In [10, 1] our group has defined refutational systems (like Resolution) that are sound and complete for the optimization versions of SAT and CSP, called Max-SAT and Weighted CSP. These problems can be defined as: given a CNF formula or a constraint satisfaction problem, find an interpretation that satisfies the maximum number of clauses or restrictions. These variants of SAT and CSP are very interesting both from the point of view of theory as well as applications. There are many practical problems that require optimization techniques. At present we are in the process of implementing software based on our results. We are doing this in conjunction with the Lerida group, given that they have a cluster and we have collaborated with them on the theoretical results.

3 Indicators of Success

First we want to mention that a Ph.D. thesis has been written during this period. This is the one of Mateu Villaret, directed by Jordi Levy (both in the project). Their joint work has given

rise to many publications as can be seen in the next section.

Our results have had a big impact in the community. In the previous text we have already given some indication of that, showing how our work opened new directions and how other people have tried or answered questions that we proposed in our work. Also some articles of us like [1], have been invited to be published as book chapters. So an extension of [1] will be published in the book "Current Trends in Constraint Programming". Also, the paper [2] obtained an award for the best paper written by somebody of outside the field of the conference.

But the main indication of the quality of our work are our publications. Most of them appear in first rate conferences and journals. We have published in conferences like LICS, IJCAI, SAT, CP, RTA, etc., and in journals like SIAM journal on Computing, Artificial Intelligence, Journal of Computational Biology, Information and Computation, Journal of Combinatorics, etc.

Below, in the references section, we give the list of publications obtained by members of our group in the last 2 years. We have kept in the document 4 papers that belong to the year 2004 because after writting the proposal of this project we started working on it right away.

Notice that we have obtained 31 publications for a project of 6 EDPs.

References

- [1] C. Ansótegui, M. L. Bonet, J. Levy, and F. Manyà. The logic behind weighted csp. In *Proc. of the Twentieth International Joint Conference on Artificial Intelligence, IJCAI-07*, pages 32–37. AAAI Press, 2007.
- [2] A. Atserias. Conjunctive query evaluation by search-tree revisited. In *10th International Conference on Database Theory (ICDT)*, volume 3363 of *Lecture Notes in Computer Science*, pages 53–67. Springer-Verlag, 2005.
- [3] A. Atserias. Definability on a random 3-cnf formula. In *20th IEEE Symposium on Logic in Computer Science (LICS)*, pages 106–115, 2005. Long version accepted for publication in European Journal of Combinatorics.
- [4] A. Atserias. On digraph coloring problems and treewidth duality. In *20th IEEE Symposium on Logic in Computer Science (LICS)*, pages 106–115, 2005. Long version accepted for publication in European Journal of Combinatorics.
- [5] A. Atserias and M. L. Bonet. On the automatizability of resolution and related propositional proof systems. *Information and Computation*, 189(2):182–201, 2004.
- [6] A. Atserias, A. Bulatov, and V. Dalmau. On the power of k -consistency. in preparation, work announced at Dagstuhl workshop on Complexity of Constraints, 2006.
- [7] A. Atserias, P. G. Kolaitis, and M. Vardi. Constraint propagation as a proof system. In *10th International Conference on Principles and Practice of Constraint Programming*, volume 3258 of *Lecture Notes in Computer Science*, pages 77–91. Springer-Verlag, 2004.
- [8] Balcazar, Bifet, and A. Lozano. Intersection algorithms and a closure operator on unordered trees. Workshop on Mining and Learning with Graphs, MLG 2006.

- [9] M. Bodirsky and V. Dalmau. Datalog and constraint satisfaction with infinite templates. In *STACS*, pages 646–659, 2006.
- [10] M. L. Bonet, J. Levy, and F. Manyà. A complete calculus for max-sat. In *Proc. of the 9th International Conference on Theory and Applications of Satisfiability Testing, SAT’06*, volume 4121 of *Lecture Notes in Computer Science*, pages 240–251. Springer-Verlag, 2006.
- [11] M. L. Bonet, J. Levy, and F. Manyà. Resolution for max-sat. *Journal of Artificial Intelligence*, to appear.
- [12] M. L. Bonet, K. StJohn, R. Mahindru, and N. Amenta. Approximating subtree distances between phylogenies. *Journal of Computational Biology*, 13(8), 2006.
- [13] A. A. Bulatov and V. Dalmau. A simple algorithm for mal’tsev constraints. *SIAM J. Comput.*, 36(1):16–27, 2006.
- [14] J. Buresh-Oppenheim, N. Galesi, S. Hoory, A. Magen, and T. Pitassi. Rank bounds and integrality gaps for cutting planes procedures. *Theory of Computing*, 2:65–90, 2006.
- [15] J. L. Carlos Ansotegui, Maria Luisa Bonet and F. Manyà. Measure of how structured a formula is. Manuscript.
- [16] H. Chen and V. Dalmau. Beyond hypertree width: Decomposition methods without decompositions. In *CP*, pages 167–181, 2005.
- [17] H. Chen and V. Dalmau. From pebble games to tractability: An ambidextrous consistency algorithm for quantified constraint satisfaction. In *CSL*, pages 232–247, 2005.
- [18] V. Dalmau. Generalized majority-minority operations are tractable. In *LICS*, pages 438–447, 2005.
- [19] V. Dalmau. Linear datalog and bounded path duality of relational structures. *Logical Methods in Computer Science*, 1(1), 2005.
- [20] V. Dalmau. A new tractable class of constraint satisfaction problems. *Ann. Math. Artif. Intell.*, 44(1-2):61–85, 2005.
- [21] V. Dalmau, R. Gavaldà, P. Tesson, and D. Thérien. Tractable clones of polynomials over semigroups. In *CP*, pages 196–210, 2005.
- [22] V. Dalmau and A. Krokhin. Majority constraints have bounded pathwidth duality. *European Journal on Combinatorics*, to appear.
- [23] V. Dalmau, A. Krokhin, and B. Larose. Retractions onto series parallel posets. *Discrete Mathematics*, to appear.
- [24] J. L. Esteban, N. Galesi, and J. Messner. On the complexity of resolution with bounded conjunctions. *Theoretical Computer Science*, 321(2-3):347–370, 2004.
- [25] N. Galesi and O. Kullmann. Polynomial time sat decision, hypergraph transversals and the hermitian rank. In *SAT Selected Papers*, pages 89–104, 2004.

- [26] N. Galesi and N. Thapen. Resolution and pebbling games. In *SAT*, LNCS, pages 76–90, 2005.
- [27] J. Levy, J. Niehren, and M. Villaret. Well-nested context unification. In *Proc. of the 20th Int. Conf. on Automated Deduction, CADE-20*, volume 3632 of *Lecture Notes in Artificial Intelligence*, pages 149–163. Springer-Verlag, 2005.
- [28] J. Levy, M. Schmidt-Schauss, and M. Villaret. Bounded second-order unification is np-complete. In *Proc. of the 17th International Conference on Rewriting Techniques and Applications, RTA '06*, volume 4098 of *Lecture Notes in Computer Science*, pages 400–414. Springer-Verlag, 2006.
- [29] J. Levy, M. Schmidt-Schauss, and M. Villaret. Stratified context unification is np-complete. In *Proc. of the 3rd International Joint Conference on Automated Reasoning, IJCAR'06*, volume 4130 of *Lecture Notes in Computer Science*, pages 82–96. Springer-Verlag, 2006.
- [30] J. Niehren and M. Villaret. Describing lambda-terms in context unification. In *Proceedings of the 5th. Int Conf. on Logical Aspects of Computational Linguistics, LACL'05*, volume 3492 of *Lecture Notes in Artificial Intelligence*, pages 221–237, Bordeaux, France, 2005. Springer-Verlag.
- [31] M. Villaret. *On Some Variants of Second-Order Unification*. PhD thesis, UPC, 2005.