Referee Report for a Doctoral Thesis

Candidate: Štefan Gurský
Title: Special Classes of Boolean Functions with Respect to the Complexity of their Minimization
Supervisor: doc. RNDr. Ondřej Čepek, Ph.D.
Referee: Mikoláš Janota, Ph.D., INESC-ID, Lisbon, Portugal

The thesis studies the computational complexity of certain problems related to Boolean functions in normal forms. Its contributions are principally in two areas: 1) *minimization* of formulas in conjunctive/disjunctive normal form and 2) *unit-propagation completeness.* The thesis is well written with only minor grammatical errors. The material is presented clearly.

What I was missing was certain global context for the studied problems. In particular, minimization is important in areas like circuit manufacturing and significant effort was invested into providing practical solutions to the problem [2]. Similarly, the motivation for knowledge compilation could be expanded. Further, it should be stressed that unit-propagation complete formulas represent only one of many ways of how Boolean functions can be compiled [3]. Chapter 3 extensively studies the class of matched formulas; it would be interesting to have some understanding of where and when they appear in practice.

Chapter 2 gives basic definitions of concepts and problems that are studied in the thesis. Here I would point out that it is *not* true that *most* SAT solvers use the pure literal elimination rule during search (even though it is true that the current trend in SAT solving is to employ various in-processing techniques during search). A citation to SAT solving literature is missing. Also [4] should have been cited for 1-UIP in Sec 5.3 rather than the later work.

Chapter 3 studies the complexity of minimization of Boolean functions. It is recalled that minimization is in Σ_2^p-complete for general CNF. However, for special classes of CNF the problem becomes easier (it is argued that it cannot be easier than SAT). It is shown that for a set of *tractable classes of CNFs* minimization always drops to NP. This is an important characterization of "easily minimizeable" classes. Here I would comment on the paragraph at the beginning of Sec. 3.7. As such, the paragraph is somewhat misleading because it claims that minimization drops to NP if the equivalence test drops to P. We might envision a class $\mathcal{C}$ of formulas for which an equivalence test is in P as long as the two formulas are in $\mathcal{C}$ but is not in P if one of the formulas is not in that class. Then, minimization is not necessarily in P because a minimal form of a formula $\phi \in \mathcal{C}$ might go outside of $\mathcal{C}$. Nevertheless, this suggests that the tractability condition could be weakened to the condition: equivalence being in P, and, a minimal formula within the same class always exists.

Chapter 4 studies the class of *matched formulas.* A very elegant proof is given that shows that while SAT is in P, minimization stays in Σ_2^p. Two rather interesting results are given for matched formulas. Firstly, that for every matched ϕ there is an equivalent matched ϕ' that is irredundant and prime. Secondly, any clause-minimum representation of a matched formula is also matched. Proofs of these two theorems are nontrivial and provide a good insight into matched formulas. I would note that it could be pointed out that a consequence of the second result is that equivalence of matched formulas cannot be in P as otherwise minimization would be in NP.

Chapter 5 studies unit propagation completeness of CNF formulas. A number of interesting results are shown for empowering implicates. In particular, that determining whether there exists an empowering implicate for a given formula is NP-complete. The existing results on empowering implicates are nicely used for the proofs. Connections are also made to CDCL SAT solving. However, it seems to me that there is a missed opportunity to make a connection to "conflict finding" in a CDCL solver. Since it is shown that finding an empowering implicate is NP-hard and at the same time, it is known that generating an empowering implicate from a conflict can be done in P, the thesis implicitly shows that steering a CDCL solver into a conflict (on purpose) is in fact hard. While this chapter is inspired by [1], it is omitted that there it was shown that if empowering implicates are generated in increasing size they are also irredundant. In this context the fact that generating *one* empowering implicate is in NP, does not help. Lem 5.3.6 and the related discussion seems to suggest that this might be more difficult but it is unclear to me.

I believe that the thesis provides a number of novel and interesting insights into the area of CNF/DNF minimization as well as unit propagation completeness. The results regarding empowering implicates are particulary interesting as they might be availed of in knowledge compilation in further research.

The results are sound, well presented, and non-trivial. The candidate has shown a good grasp of the subject. Therefore I recommend the thesis to be defended.

24 August 2014 Mikoláš Janota

References

[1] L. Bordeaux and J. Marques-Silva. Knowledge compilation with empowerment. In *SOFSEM*, 2012.

[2] O. Coudert. Two-level logic minimization: an overview. *Integration*, 17(2), 1994.

[3] A. Darwiche and P. Marquis. A knowledge compilation map. *JAIR*, 17, 2002.

[4] J. Marques-Silva and K. A. Sakallah. GRASP — a new search algorithm for satisfiability. In *ICCAD*, 1996.