JOHANNES KEPLER
UNIVERSITY LINZ | JKU

Eingang 09.07.15

Prof. Dr. ARMIN BIERE
Institute for Formal Models and Verification
http://fmv.jku.at

+43 732 2468 4541
biere@jku.at

Review Doctorial Thesis

Resolution-Based Methods for Linear Temporal Reasoning

by Martin Suda

Linz, Tuesday, 30. June 2015

This thesis works out a close connection between imported related problems from (at least) four different communities: *(i)* saturation methods from the first-order theorem proving community, *(ii)* temporal deduction respectively LTL satisfiability from the logic (programming) community, *(iii)* symbolic model checking from the computer aided verification community and *(iv)* planning from the AI community. In essence, they are all PSPACE complete.

Maybe the most canonical problem in this context is symbolic reachability [Savitch'70]. With my background in HW verification I would just say this is just reachability for sequential bit-level circuits, e.g., whether a certain flip-flop can ever be set to a fixed value. These are the problems considered as benchmarks in the single property track of our Hardware Model Checking Competition (HWMCC), which are also used in Chapter 5. I still think it is fair to argue that LTL satisfiability should be considered to be a similar generic problem, particular if you have a background of logic. Thus the title of the thesis still fits well.

Even though the connection between these problems is kind of obvious, in the sense, that they are all PSPACE complete, techniques developed in the listed various communities differ substantially. It is quite impressive how the candidate was able to bridge the gap between these communities, extract the core ideas of these, in general, rather involved methods and provide a deep analysis of their relation. It further makes a substantial contribution in the domain of applying saturation to LTL satisfiability and IC3 to planning.

The thesis does not stop at a theoretical level. In at least two domains, planning and in particular for LTL satisfiability, the tools developed by the candidate, based on these insights, are able to achieve a substantial improvement of the state-of-the-art. Also for the HWMCC problems, a model checker was developed and evaluated. These extensive experimental evaluations do provide novel interesting insights, particularly in relation to IC3/PDR.

It is clear that such an effort to bridge areas needs a substantial amount of formalism and trying to completely understand all aspects of the thesis is pretty challenging. Nevertheless, the chosen level of detail and amount of concepts is as expected and further shows that the candidate clearly masters the art of formalizing and proving important facts in this domain with the necessary mathematical rigor.

Regarding the saturation based method in Chapter 2, which lead to the new algorithm LPSup, an important contribution of the thesis is to show that clausal temporal resolution is very close to the presented saturation based view. In some sense the author shows how to take ideas temporal resolution, combine it with saturation to obtain a complete LTL satisfiability procedure. The exposition is crystal clear, while I had more issues, when I was exposed to the clausal temporal resolution work in the past by other authors. The prototypical implementation shows good results, even though no actual times were provided. The worst-case complexity is as far I can tell double exponential, which is horrible, but also shared with other approaches, particularly with the method of the next Chapter.

In Chapter 3 another view on temporal reasoning is used, motivated by a connection to CDCL SAT solving. Both use partial models to guide the search. This LTL satisfiability procedure called LS4 also uses a partial temporal model to achieve a similar effect. Another motivation is to make use of SAT technology, in particular incremental SAT solving, instead of an explicit resolution procedure as used for LPSup. The new LS4 algorithm works forward in contrast to other algorithms and again needs a global leap inference for termination. The presentation of the algorithm comes with full proofs. As mentioned above space complexity is pretty bad. However, the details are rather involved and I can not claim to understand every detail, but for the expert there is no problem to follow the overall picture. The experimental results are very impressive compared to approaches taken from recent publication on the same subject (and compared to the methods presented in Chapter 2).

The most common canonical problem for PSPACE is usually considered to be checking satisfiability of quantified boolean formulas (QBF). This problem is not really touched in the thesis. This is justified by the fact that QBF has not really found much practical application in reachability problems despite many serious attempts. However, in all these communities SAT solving is used at one point or the other, and the thesis proves detailed knowledge of the candidate on how to apply SAT solvers to solve such problems and further Chapter 4 actually lifts the currently in practice most important preprocessing technique from SAT to temporal deduction / LTL satisfiability. The performance improvement in this context is not as dramatic as in SAT but the experiments clearly show the usefulness of the contribution. On the technical side, lifting this variable elimination technique to the temporal deduction framework requires the introduction of labeled clauses, similar to the temporal labels discussed earlier in thesis and similar restrictions to resolving such clauses had to be introduced.

An important recent breakthrough in the model checking community was the invention of the already mentioned IC3 algorithm by Aaron Bradley in 2010 (sometimes also called PDR). This pushed model checking (at least for HW) to a new level. The original algorithm is rather involved and was first considered to be unnecessarily complex. Todays understanding is that most of the complexity of the original description of IC3 can not be avoided and is required to achieve its efficiency.

The author of the thesis gives a close connection between IC3 and algorithms discussed earlier in the thesis for temporal satisfiability. This leads to the Reach algorithm, which then is turned into IC3 by adapting some of the insights from IC3, including obligation rescheduling, which was proven to be important empirically, particularly in order to find deep error traces, but has not been fully explained from the theoretical side. Here the thesis provides another angle to look at this important aspect of IC3, and the experiments confirm that this technique is particularly useful for satisfiable instances. Clause propagation to improve convergence as well as minimizing relative inductive clauses are considered to be important and also much easier to motivate from an algorithmic perspective. The former, as expected, proves useful for unsatisfiable instances in the experiments.

The last Chapter is about using similar technology in the context of planning. The author succeeded to develop a novel competitive planning algorithm, motivated by IC3, which due the application domain does not require a symbolic decision procedure at all (after proper encoding and preprocessing more general STRIPS problems), an impressive achievement.

This is an outstanding thesis, which of course I mark with the highest grade "summa cum laude". Consequently I recommend to continue with the graduation procedure.

[redacted]

Prof. Dr. Armin Biere

Institut für Formale Modelle und Verifikation • Universität Linz