Master Thesis Proposal

Interpolation in First Order Logic with Equality

Bernhard Mallinger

Advisor: Ass.Prof. Stefan Hetzl

March 6, 2014

1 Motivation and problem statement

After decades of continued research, the area of software verification still lacks effective methods for reasoning about real world programs, which is necessary to prove vital safety or liveness properties. Major advances in the form the emergence of symbolic model checking and bounded model checking have ameliorated the situation. But in dealing with loops, the highly non-trivial but at the same time essential problem of discovering loop invariants still remains a challenging task. In recent years, the approach of applying Craig interpolation in order to calculate invariants enjoyed increasing popularity, especially after successful applications for instance in [McM03].

The Interpolation theorem is a long known basic result of mathematical logic. Interpolants lay bare certain relations between formulas or sets of formulas by describing it in a concise way. Given a logical specification of transitions in a program, they can be used to describe reachable states. This process is fully analytic in the sense that interpolants can efficiently be calculated from proofs. Leveraging the tremendous progress of automatic deduction systems in the last decades, obtaining the required proofs is feasible.

For practical applicabilty, often relatively weak formalisms such as propositional logic or equational logic with uninterpreted function symbols are employed. However for first-order logic with equality, no efficient algorithms for computing interpolants are known, even though a basic procedure is already provided in [Cra57a].

2 Aim of the work

This thesis aims to work towards finding an algorithm to calculate interpolants in first-order logic in the presence of equality. Currently no procedures of practical applicability are known for this logic, This should be accomplished by either improving existing solutions such as the aforementioned procedure by Craig or exploring a nouveau approach.

Furthermore, a comprehensive account of existing techniques and results will be presented. This includes different proofs of interpolation results with a focus on constructive proofs which give rise to concrete algorithms. Non-constructive methods, especially of a model theoretic nature, will be treated to meet theoretical curiosity and to put the algorithm in a perspective and give a broader picture. In this spirit, further corollaries and also applications of the interpolation theorem will be presented.

3 Methodology and approach

As the problem at hand is a well-defined mathematical task, standard mathematical methodology applies.

Determined by the results in the development of an algorithm, an implementation is deemed scientifically valuable but most likely beyond the scope of this thesis.

4 State of the art

Current research and application is based on the fundumental result by Craig [Cra57a], here given in a formulation for resolution calculus:

Theorem 1 (Interpolation)**.** Let A and B be sets of first-order sentences such that $A \cup B$ is refutable. Then there exists an interpolant I such that

1. $A \supset I$ is valid
2. $I \wedge B$ is unsatisfiable
3. the non-logical symbols of I are only those that appear in both A and B.

This result has been proven in different formalisms using different syntactic methods (cf. [Cra57a]; [Tak87]; [Kra97]; [Pud97]), but also via semantic, model theoretic means (cf. [Sho67], section 5.2; [CK90], theorem 2.2.20). To this end, the interpolation theorem can be seen as a corollary Robinson's joint consistency theorem, but even further, latter can also be proven from the former. This suggest a close relation between one the one hand the proof-theoretic and on the other hand the model-theoretic view.

Another major corollary of the interpolation theorem is given by Beth in form of the definability theorem, which shows the the notions of implicit and explicit definition coincide. In

sloppy terms, an implicit definitions refers to a definition by usage in a formula. This can be made explicit by means of interpolation.

4.1 Interpolation algorithms

Constructive proofs of the interpolation theorem directly give rise to algorithms for computing interpolants. For instance in [Tak87], the well known Maehara lemma is used in the proof, which forms an efficient procedure for extracting interpolants from first-order LK proofs, but fails in the presence of equality.

5 Relevance to the curriculum of Computational Intelligence

Logic as core machinery of computer science is featured prominently in the curriculum of Computational Intelligence in the mandatory module "Logic and Computability" as well as the module "Logic, Mathematics, and Theoretical Computer Science", but also frequently appears in theoretic foundations of other areas.

The logic used in this thesis, first-order logic with equality, is without any doubt one of the most common and useful ones; the interpolation theorem is hereby a celebrated result. As argued in section 1, advancements in this field have direct consequences for the area of formal verification, which is also featured in the curriculum.

The following courses possess a direct relation to the topic of this thesis:

- Formal Methods in Computer Science
- Proof Theory 1
- Logic and Computability
- Advanced Mathematical Logic

References

[CK90] C.C. Chang and H.J. Keisler. *Model Theory.* Studies in Logic and the Foundations of Mathematics. Elsevier Science, 1990.

[Cra57a] William Craig. Linear Reasoning. A New Form of the Herbrand-Gentzen Theorem. *The Journal of Symbolic Logic*, 22(3):250–268, September 1957.

[Cra57b] William Craig. Three uses of the herbrand-gentzen theorem in relating model theory and proof theory. *The Journal of Symbolic Logic*, 22(3):269–285, September 1957.

[Hua95] G. Huang. Constructing craig interpolation formulas. In *Proc. of the First Annual International Conference on Computing and Combinatorics*, pages 181–190, Xian, China, 1995.

[Kra97] Jan Krajíček. Interpolation theorems, lower bounds for proof systems, and independence results for bounded arithmetic. *Journal of Symbolic Logic*, pages 457–486, 1997.

[McM03] Kenneth L. McMillan. Interpolation and sat-based model checking. In Jr. Hunt, Warren A. and Fabio Somenzi, editors, *Computer Aided Verification*, volume 2725 of *Lecture Notes in Computer Science*, pages 1–13. Springer Berlin Heidelberg, 2003.

[Pud97] Pavel Pudlák. Lower bounds for resolution and cutting plane proofs and monotone computations. *J. Symb. Log.*, 62(3):981–998, 1997.

[Sho67] J.R. Shoenfield. *Mathematical logic*. Addison-Wesley series in logic. Addison-Wesley Pub. Co., 1967.

[Tak87] G. Takeuti. *Proof Theory*. Studies in logic and the foundations of mathematics. North-Holland, 1987.

[Wei10] Georg Weissenbacher. *Program Analysis with Interpolants*. PhD thesis, 2010.