

THESIS REPORT

Lukasz Stafiniak, GADTs for Reconstruction of Invariants and Postconditions

Martin Sulzmann

SUMMARY

The thesis proposes a type inference method for GADTs based on a combination of abduction and constraint solving. Chapter 2 reviews earlier work in the area. Chapter 3 introduces the GADT type system including a mapping of the type inference problem to implication constraints. Chapter 4 introduces a solving method for implication constraints based on constraint abduction. The system has been implemented and several examples are discussed in Chapter 5.

I have only skimmed through the appendix.

EVALUATION

Contribution to the literature The thesis makes some novel contributions and improves the state of the art of GADT type inference. The main novelty lies in an abductive constraint solving method for implication constraints. See Chapter 4. More detailed discussion below.

Visibility and impact on the field As far as I can tell, most of this work hasn't been published yet. The results are novel, interesting and have the potential to inspire new research in the area of GADT type inference.

All in all, I consider this work well worth of a PhD.

COMMENTS

There are a few points I'd like to discuss further.

Notation/presentation Some notation is used but then only formally introduced later. There are also several important pieces of discussion spread out through the thesis instead of giving a comprehensive discussion at one place.

I'd wish for more helpful examples. Only expert readers will be able to master the material. Some of the algorithms, e.g. see Table 4.3 and 4.4, without any worked out examples.

Connection to earlier work I have only figured out the main contributions of the thesis after reading through some of the related work. As it appears, the main contribution is an alternative abductive constraint solving method in style of earlier work by Sulzmann, Stuckey and Schrijvers.

Problem statement I am missing a clear problem statement. There is some discussion of earlier work, see Chapter 2, but only expert readers will be able to figure what is the problem this thesis is trying to solve.

More on the above points below.

FURTHER DETAILS

Reference to specific sections/pages.

- Section 2.2.2 HMG(X)

The purpose of implication constraints need to be stressed. Thanks to implication constraints type inference becomes almost trivial. The downside is a non-elementary complexity of checking satisfiability of implications.

Aside, to the best of my knowledge, Zenger was the first to employ implication constraints for GADT-like systems.

- Section 2.2.4 Relevance

"There are three major differences between HMG(X) and InvarGenT ... Logically complex formulas, especially involving implications, are likely to not be sufficiently self-explanatory."

While not explicitly said, this paragraph implies the following type inference approach:

- Reduce type inference to implication constraints, e.g. $C \Rightarrow D$
- Unlike earlier work by Zenger and Pottier et al, the approach taken, following work by Sulzmann, Stuckey and Schrijvers, is to find a solution/answer A in the form of conjunctions of primitive constraints such that A solves $C \Rightarrow D$. That is, $A \Rightarrow (C \Rightarrow D)$.

This has important consequences:

Lack of principal types due to (possibly) infinitely many, incomparable solutions etc.

- Section 2.5.3 Relevance

Most of the discussion here is to high-level. Anyway, we haven't seen any details yet. So why not defer the discussion till later?

- Section 2.6 Herbrand Constraint Abduction

See my comment about the 'approach taken'. Would be helpful to discuss HMG(X) and the abduction approach together.

BTW, 'fully maximally answers' are only introduced much later. So much of this discussion here is meaningless (for non-expert readers).

- p45 "We identify clauses p:e of MMG(X) with p when .e."
I couldn't find typing rules for the form "p when .e"?

- Chapter 4

This is where the main contribution of this thesis lies. I still don't fully grok all of the technical details. What I gathered is the following.

Earlier work by Sulzmann, Stuckey, Schrijvers show that the full constraint abduction problem for GADTs is not feasible as there are too many solutions.

Their idea is to restrict the attention to 'intuitive' solutions based on the concept of fully maximally answers introduced by Maher.

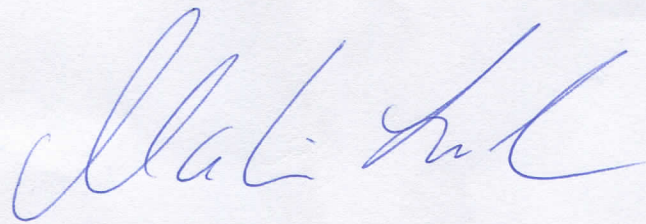
This thesis argues that restricting ourselves to fully maximally answers may rule out some interesting programs.

See the 'eval' example on p64. BTW, I'm wondering about the implication constraint

$$a = \text{Term}((a', b')) \Rightarrow b = (a'', b'')$$

Aren't we missing some constraints on 'eval'?

It's also puzzling that only much later on p82 (top) it is said that "The corresponding implications do not have fully maximal answers" Why not give a comprehensive discussion here?



17.07.2015 Karlsruhe, Germany