Planning the Whisky Problem

Louise A. Dennis[1] and Alan Bundy[2]

[1] School of Computer Science and Information Technology, University of Nottingham, Nottingham, NG8 1BB lad@cs.nott.ac.uk
[2] Division of Informatics, University of Edinburgh, 80 South Bridge, Edinburgh, EH1 1HN bundy@dai.ed.ac.uk

Proof critics are a technology from the proof planning paradigm. They examine failed proof attempts in order to extract information that can be used to generate a patch which will allow the proof to go through.

We examine the proof of the "whisky problem", a challenge problem from the domain of temporal logic. The proof requires a generalisation of the original conjecture and we examine two proof critics which can be used to create this generalisation. Using these critics we believe we have produced the first fully automatic proofs of this challenge problem.

We chose to use the proof planning system λ*Clam* [2] to attempt an automation of the proof of the whisky problem. The whisky problem is represented in λ*Clam* as three axioms and a goal to be proved. The axioms are:

$$\begin{aligned} &p(a,0), \\ \forall x.\, p(x,0) &\Rightarrow p(h(x),0), \\ \forall x,y.\, p(h(x),y) &\Rightarrow p(x,s(y)). \end{aligned}$$

These are treated as the following rewrite rules in λ*Clam*:

$$p(a,0), \tag{1}$$
$$p(h(X),0) \quad :\Rightarrow \quad p(X,0), \tag{2}$$
$$p(X,s(Y)) \quad :\Rightarrow \quad p(h(X),Y). \tag{3}$$

The goal to be proved is

$$\forall x.\, p(a,x).$$

The challenge is that for a proof to go through the goal must be generalised to $\forall x,n.\, p(h^n(a),x)$ where h^n signifies n applications of h. A λ*Clam* inductive proof of this fails to apply the rewrite rule (3) in the induction to the goal $p(a,n) \vdash p(a,s(n))$ because the resulting embedding is unsinkable. The embedding is part of the rippling heuristic ([3]) used to control rewriting during induction. This can be used to trigger a critic.

The existing λ*Clam* critic for this situation, based on work in [1], generalises the initial goal to

$$\forall x,y.\, p(F(y),x).$$

Where F is gradually instantiated during the course of the proof to the desired function. This proof proved sensitive to the rewrite rules available in the default context and required the provision of a new induction scheme:

$$\frac{P(B) \qquad P(G^n(B)) \vdash P(G(G^n(B)))}{\forall n. P((\lambda x. G^x(B))(n))}, \tag{6}$$

We also investigated a more specialised critic based on work in [4] which speculated the “target” generalisation of

$$\forall x, n.\ p(h^n(a), x).$$

directly based on the contents of the wave front in the unsinkable embedding. The proof then went through in the default $\lambda Clam$ context.

As far as we are aware these represent the first fully automatic proofs of the whisky problem.

Acknowledgments

This research was funded by EPSRC grants Gr/M45030 and Gr/M46624. We thank the members of the Logics Group at Liverpool University and Michael Fisher and Anatoli Degtyarev in particular for drawing our attention to this challenge problem.

References

1. A. Ireland and A. Bundy. Productive use of failure in inductive proof. *Journal of Automated Reasoning*, 16(1–2):79–111, 1996. Also available as DAI Research Paper No 716, Dept. of Artificial Intelligence, Edinburgh.
2. J.D.C. Richardson, A. Smaill, and I. Green. System description: Proof planning in higher-order logic with lambda-clam. In C. Kirchner and H. Kirchner, editors, *Conference on Automated Deduction (CADE'98)*, volume 1421 of *Lecture Notes in Computer Science*, pages 129–133. Springer-Verlag, 1998.
3. A. Smaill and I. Green. Higher-order annotated terms for proof search. In J. von Wright, J. Grundy, and J. Harrison, editors, *Theorem Proving in Higher Order Logics: 9th International Conference, TPHOLs'96*, volume 1275 of *Lecture Notes in Computer Science*, pages 399–414, Turku, Finland, 1996. Springer-Verlag. Also available from Edinburgh as DAI Research Paper 799.
4. T. Walsh. A divergence critic for inductive proof. *Journal of Artificial Intelligence Research*, 4:209–235, 1996.