Available online at www.sciencedirect.com

SCIENCE DIRECT®

Electronic Notes in Theoretical Computer Science

ELSEVIER Electronic Notes in Theoretical Computer Science 125 (2005) 87–101

www.elsevier.com/locate/entcs

The TM System for Repairing Non-Theorems

Simon Colton

Department of Computing
Imperial College, London
United Kingdom
sgc@doc.ic.ac.uk

Alison Pease

School of Informatics
University of Edinburgh
United Kingdom
alisonp@dai.ed.ac.uk

Abstract

We describe a flexible approach to automated reasoning, where non-theorems can be automatically altered to produce proved results which are related to the original. This is achieved in the TM system through an interaction of the HR machine learning program, the Otter theorem prover and the Mace model generator. Given a non-theorem, Mace is used to generate examples which support the non-theorem, and examples which falsify it. HR then invents concepts which categorise these examples and TM uses these concepts to modify the original non-theorem into specialised theorems which Otter can prove. The methods employed by TM are inspired by the piecemeal exclusion, strategic withdrawal and counterexample barring methods described in Lakatos's philosophy of mathematics. In addition, TM can also determine which modified theorems are likely to be interesting and which are not. We demonstrate the effectiveness of this approach by modifying non-theorems taken from the TPTP library of first order theorems. We show that, for 98 non-theorems, TM produced meaningful modifications for 81 of them. This work forms part of two larger projects. Firstly, we are working towards a full implementation both of the reasoning and the social interaction notions described by Lakatos. Secondly, we are aiming to show that the combination of reasoning systems such as those used in TM will lead to a new generation of more powerful AI systems.

Keywords: Automated theorem modification, automated reasoning, model generation, machine learning, automated theory formation, philosophy of mathematics.

doi:10.1016/j.entcs.2004.07.021

1 Introduction

Mathematics has developed in a much more organic way than its rigid textbook presentation of definition-theorem-proof would suggest. Automated theorem proving systems more closely reflect the textbook notion of mathematics than a developmental approach. In particular, most deduction systems are designed either to prove results if they are true, or find counterexamples if they are false, but not both. System designers also assume that the concepts mentioned in the conjecture are correctly defined and actually relate to the mathematical notions the user is interested in. Clearly, the adoption of these assumptions is not conducive to the kind of exploration more common in research mathematics, in which concept definitions change and become more sophisticated, and flawed conjectures and proofs are gradually refined. Hence, it is time to increase the flexibility of reasoning systems to better handle ill-specified problems.

We describe here the development of the Theorem Modifier (TM) system. This takes a set of axioms and a conjecture in first order logic and tries to prove it. If this fails, TM attempts to modify the conjecture into a set of theorems which it can prove. To achieve this flexibility, TM combines the power of three automated reasoning systems, namely the HR machine learning program [1], the Otter theorem prover [11] and the Mace model generator [12]. As described in §3, TM uses these systems in ways prescribed in the philosophy of mathematics developed by Lakatos [9]. In particular, TM performs counterexample-barring, piecemeal exclusion and strategic withdrawal. These techniques are further explained in §2. As a simple example of TM working, given the non-theorem that all groups are Abelian, it states that it cannot prove the original result, but it has discovered that self-inverse groups are Abelian. To evaluate this approach, in §4, we describe how TM successfully found meaningful modifications to 81 of 98 non-theorems derived from the TPTP library of first order theorems [16].

The development of the TM system forms part of two larger projects. Firstly, we are working towards a full implementation both of the reasoning and the social interaction notions described by Lakatos in [9]. Secondly, we are aiming to show that the combination of reasoning systems such as those used in TM will lead to a new generation of AI systems which are able to solve problems which individual techniques cannot.