

Three Next Generation Approaches to Automated Mathematical Theory Formation

Simon Colton
Computational Creativity Group,
Department of Computing, Imperial College, London
www.doc.ic.ac.uk/~sgc

Around a decade ago, we introduced Automated Theory Formation as a technique for mathematical discovery, implemented in the HR system. Starting with some basic background information such as a set of axioms for an algebraic domain, or fundamental concepts such as addition and multiplication in number theory, HR forms concepts which categorise the examples; makes conjectures which relate the concepts; and generates proofs which explain the conjectures (or counterexamples which disprove them). In addition to inventing concepts and discovering theorems which have been published in the mathematical literature, HR has been applied successfully to AI tasks including constraint solving and machine learning. In the talk, I will describe three PhD projects which have been influenced by the Automated Theory Formation approach.

The first project is the work of Alison Pease, and crosses the boundary between Artificial Intelligence and Philosophy by building and exploring a computational model of notions from Lakatos's philosophy of mathematics as advocated in his book "Proofs and Refutations". Via the HRL system, which simulates a classroom environment consisting of theory formation agents acting as a teacher and a set of students, Pease implemented methods identified by Lakatos such as strategic withdrawal and monster barring, which change the theory in reaction to the discovery of a counterexample to a conjecture under discussion. This led to both an advancement in the state of the art of automated theory formation, and a clarification of some of the issues raised and methods introduced by Lakatos.

The HR system is driven by production rules which turn old concepts into new ones. In his PhD work, Pedro Torres has addressed the meta-level question of automating the generation of production rules. Given a seed theory which contains fundamental background information and exemplar concepts known to be of interest in the domain, with his Suricata system, Torres has implemented methods for analysing the seed theory and deriving production rules which can be used to generate the exemplar concepts from the background material. We are currently investigating the potential of this method to streamline theory formation, so that, in addition to recreating the exemplar concepts, Suricata is able to generate focused theories with a high yield of interesting concepts and conjectures.

The combination of reasoning techniques such as induction, deduction and constraint solving has always been at the heart of automated theory formation. Taking inspiration from the Global Workspace Architecture model of human thought processing advocated by Baars and others in Cognitive Science communities, John Charnley implemented the GC framework for combining AI reasoning systems in a systematic, straightforward way. By removing the communication between processes in favour of a very pared-down blackboard-style system, Charnley has shown that complex combinations of reasoning can be achieved with relatively little effort from the system designer. He has configured his framework to undertake the tasks previously tackled by ad-hoc combined reasoning systems such as HR. This has the benefit of greatly simplifying the construction of combined reasoning systems, and can also take advantage of the natural parallelisation afforded by the Global Workspace approach.