Automating Science and Finding Patterns in the Data

Ross D. King

Department of Computer Science, University of Wales, Aberystwyth, SY23 3DB, UK
rdk@aber.ac.uk

Abstract. The basis of science is the hypothetico-deductive method and the recording of experiments in sufficient detail to enable reproducibility. We report the development of the Robot Scientist "Adam" which advances the automation of both. Adam has autonomously generated functional genomics hypotheses about the yeast Saccharomyces cerevisiae, and experimentally tested these hypotheses using laboratory automation. We have confirmed Adam's conclusions through manual experiments. To describe Adam's research we have developed an ontology and logical language. The resulting formalization involves over 10,000 different research units in a nested tree-like structure, ten levels deep, that relates the 6.6 million biomass measurements to their logical description. This formalization describes how a machine discovered new scientific knowledge. Describing scientific investigations in this way opens up new opportunities to apply data-mining to discover new knowledge. To test this idea we have applied relational data-mining, guided by ontology, to evaluate the repeatability of Adam's investigations.