

Joined-Up Reasoning for Automated Scientific Discovery

A Position Statement and Research Agenda

Simon Colton

Combined Reasoning Group
Department of Computing
Imperial College, London
<http://www.doc.ic.ac.uk/crg>
sgc@doc.ic.ac.uk

We use the phrase “joined-up” here with a double meaning: to convey two aspects of scientific discovery which we believe are essential, yet under-researched with respect to automating scientific discovery processes. Firstly, from an Artificial Intelligence perspective, the majority of approaches to using AI techniques involve a disjointed sequential application of different problem solving methods, with the user providing the glue in various ways. These include routine logistical aspects such as the pre-processing of data and knowledge, translating outputs into input, choosing parameter settings for running AI methods, etc. More importantly, however, the user performs various aspects of meta-level reasoning, including asking the most pertinent questions, determining what it means if a process terminates with success and identifying – and investigating – anomalies. This approach tends to lead to auto-assisted discoveries where the user knows what they are looking for, but not what it looks like, rather than the deeper discoveries of examples/concepts/hypotheses/explanations that the user didn’t even know he or she was looking for. While AI methods promise the discovery of such surprising and novel scientific artefacts, they rarely deliver on this promise, as their application is too regimented within the problem solving paradigm of AI.

We therefore advocate (and actively pursue) investigations into how to build systems which combine reasoning activities in such a way that the whole is more than a sum of the parts. While automating the logistical aspects mentioned above is tiresome but straightforward, we believe that automating the meta-level reasoning skills employed by the users of AI tools for discovery tasks is a fascinating problem. To this end, we have looked at various ad-hoc combinations for discovery tasks in pure mathematics (e.g., (Colton & Pease 2005), (Charnley, Colton, & Miguel 2006), with a summary in (Colton & Muggleton 2006)), but more recently we have started to investigate the value of more generic approaches based on proof-planning from automated theorem proving (Sorge *et al.* 2007), and global workspace architectures from cognitive science (Charnley & Colton 2008). We are not the first to suggest studying combined reasoning it

its own right¹. However, we perhaps go further than most in suggesting that a pipelining approach for knowledge flow through reasoning systems similar to those employed for data through calculations may not be sophisticated enough to facilitate the kinds of deeper discoveries described above. We are particularly keen on building ontologies – or perhaps more expressive knowledge bases – pertaining to reasoning. In particular, we hope to formally define the input to; output from; processing of; and glueing mechanisms required between reasoning systems such as constraint solvers, theorem provers, machine learners, planners, and so on. In this way, we hope to build discovery systems which can be given only scientific background knowledge and data, and no notion of particular problems to solve or investigations to make. These systems will automatically configure and reconfigure combinations of reasoning systems, and will be driven by aesthetic and utilitarian notions of interestingness (as discussed below) and by the discovery and investigation of anomalies.

Secondly, from a computational creativity point of view, we would like to join up the emerging consensus of notions about automating artistic creativity with notions of automating scientific creativity. An important factor here is the notion of subjective aesthetic preferences. We believe it is often too simplistic to think of scientific discovery as an entirely objective, platonic pursuit, driven by utilitarian measures of value, the most important of which is truth. While subjectivity is largely omitted from published scientific results, personal preferences, hunches, individual experiences and prejudices often drive scientific investigations. Moreover, in computer-aided scientific discovery, the notion of value is sometimes straightforward, but in other situations, it may be more complicated. For instance, a drug designed via constraint solving may optimise some physical characteristic, but might not be easy to synthesise; one machine learned classifier may achieve a higher predictive accuracy for some physical phenomena than another, but the latter may give more insight into the phenomena (which is why logic-based learning approaches have an important role to play in scientific discovery applications). Such contrasting and compet-

¹See, for instance, (Bundy 2007) for a similar call to arms as this one.

ing assessment criteria are commonplace in studies of computational creativity in the arts, which has led to the determination of concrete value measures based on the artefacts produced by creative systems, in particular, Ritchie's criteria, described in (Ritchie 2007).

We therefore advocate (and actively pursue) investigations into modelling individual or group aesthetics which can be used to drive scientific exploration. Such aesthetics may alter during the course of an investigation, they may be inconsistent, and pinpointing the notion of value in a particular domain may itself represent a discovery. We believe that, for computational approaches to gain more autonomy and higher potential for finding the kinds of deeper discoveries described above, software has to be taught not only some received principals about interest and value, but also has to be taught how to invent – and defend – its own heuristics for the value of scientific artefacts. While there is a glut of data in most sciences (although many argue that getting the *right kind of data* is still the most important and difficult aspect of a scientific investigation), and a growing number of scientific knowledge bases, there is very little heuristic information available about what a researcher might find interesting or anomalous. In the application domain of graphic design, we have begun investigations both into how software can learn someone's aesthetic preferences from their actions and choices (Colton 2008c), and how software can invent its own fitness function (Colton 2008a) in order to appear to be more creative.² We hope to show that such learning and invention of notions of interestingness can also be fruitfully applied in scientific discovery applications.

In summary, our position is that it is time for researchers to investigate software which controls other, third party, scientific discovery software, and our research agenda is to build generic mechanisms for the combination of reasoning systems which can undertake large-scale, multi-faceted investigations autonomously, driven by complex and evolving models of value in scientific artefacts. Joining up reasoning systems in practice and joining up our understanding of creative activity across the arts and sciences, will – we hope – lead to greatly improved automation in scientific discovery systems.

Acknowledgements

We would like to thank the organisers of the AAAI Fall symposium on Automated Scientific Discovery for the opportunity to present this position statement and research agenda. We would also like to thank the numerous colleagues who have collaborated with us to conceive, construct and experiment with combined reasoning systems for scientific and artistic creative purposes.

²See (Colton 2008b) for a discussion of the perception of creativity in computational systems.

References

- Bundy, A. 2007. Cooperating reasoning processes: More than just the sum of their parts. In Veloso, M., ed., *Proceedings of the 20th International Joint Conference on Artificial Intelligence*, 2007, 2–11.
- Charnley, J., and Colton, S. 2008. A global workspace framework for combining reasoning systems. In *Proceedings of the CICM workshop on Empirically Successful Automated Reasoning Systems for Mathematics*.
- Charnley, J.; Colton, S.; and Miguel, I. 2006. Automatic generation of implied constraints. In *Proceedings of the 17th European Conference on AI*.
- Colton, S., and Muggleton, S. 2006. Mathematical applications of Inductive Logic Programming. *Machine Learning* 64:25–64.
- Colton, S., and Pease, A. 2005. The TM system for repairing non-theorems. In *Selected papers from the IJ-CAR'04 disproving workshop, Electronic Notes in Theoretical Computer Science*, volume 125(3). Elsevier.
- Colton, S. 2008a. Automatic invention of fitness functions, with application to scene generation. In *Proceedings of the EvoMusArt Workshop*.
- Colton, S. 2008b. Creativity versus the perception of creativity in computational systems. In *Proceedings of the AAAI Spring Symposium on Creative Systems*.
- Colton, S. 2008c. Experiments in constraint-based automated scene generation. In *Proceedings of the 5th International Joint Workshop on Computational Creativity*.
- Ritchie, G. 2007. Some empirical criteria for attributing creativity to a computer program. *Minds and Machines* 17:67–99.
- Sorge, V.; Meier, A.; McCasland, R.; and Colton, S. 2007. Automatic construction and verification of isotopy invariants. *Journal of Automated Reasoning* 40(2-3):221–243.