

A New Model of Reflective Introspective Learning

Susan Fox

Department of Mathematics and Computer Science
Macalester College
1600 Grand Avenue
Saint Paul, MN 55105 USA
E-mail: fox@macalester.edu

Abstract

Systems which introspect about their own processes can improve their reasoning behavior in response to experience using "introspective learning" techniques. Many systems which perform introspective learning analyze and change only an underlying domain task's reasoning processes. They do not possess the ability to reflectively introspect about the introspective task itself. We present a model of a truly reflective introspective learning system which uses the same case-based reasoning mechanisms for both its domain and introspective tasks. The reuse of case-based reasoning mechanisms enables the reuse of introspective models and mechanisms developed for a CBR domain reasoner to reflect about the introspective process itself. We suggest an introspective model capable of reflection, and begin to address the issue of balancing introspection and action within a reflective system.

Introduction

Meta-reasoning — reasoning that incorporates knowledge of the task being addressed and methods for achieving it — has been used to predict the behavior of other agents (e.g., (Gurer, des Jardins, & Schlager 1995)), to guide the acquisition and application of domain knowledge (Bradzil & Konolige 1990; Clancy 1988; Davis 1982), and to adjust the system's own reasoning processes in response to feedback about its performance in its domain (Collins *et al.* 1993; Ram & Cox 1994; Cox & Freed 1995; Stroulia & Goel 1995). This last application of meta-reasoning has been referred to as "introspective reasoning" or "introspective learning."

Most work on introspective reasoning focuses on improving the performance of an underlying reasoning system on some domain task. Introspective reasoning is rarely applied to the introspective reasoning process itself; the system is never fully "reflective." A reflective system must be able to manipulate and reason about its own processes, and alter its own processing behavior (Ibrahim 1992). SOAR (Rosenbloom, Laird, & Newell 1993b) can reflectively reason about the constructs it

makes, but does not explicitly represent its introspective model of behavior and cannot necessarily introspect about all of its mechanisms.

Incorporating reflection into an introspective learning framework requires extending the system's knowledge about its reasoning processes to include the learner itself. Unlike a system where introspection is only applied to an underlying task, a reflective introspective learner must control when and to what extent introspection occurs, so that it does not choose to spend all its time constructing a reflective tower.

In this paper we present a preliminary model for reflective introspective learning, using case-based reasoning (CBR) techniques to implement both domain and introspective processes. We explore the needs of such a system for representing meta-cognitive reasoning processes, and begin to address the balance between taking action and introspecting, and the costs associated with unrestricted reflection.

Our initial results suggest that the use of case-based reasoning for introspective learning carries little overhead compared to other methods. Monitoring of the introspective process is feasible and requires little additions to an existing model of case-based processes. The correct balance between reflecting and acting on the domain task is still an open question, but we suggest that reflective introspection be driven by necessity only. By default, introspection should occur only at the basic level, and reflections to higher levels should occur only in failure situations when lower level introspection has failed to resolve the problem.

Reflection

A system is reflective if it can shift the focus of its processes from its basic "domain" task to the problem-solving task itself (Smith 1982). Such a system can construct an unbounded "reflective tower" of reasoning processes where each analyzes the process beneath it. The idea of reflection has been applied to programming language designs as well as artificial intelligence reasoning systems (Ibrahim 1992).

Ideally, a reflective system should use the same reasoning mechanisms at all levels of operation. It can reason about its own processes, including the introspec-

but still represents declaratively the entire reasoning process at both domain and introspective levels.

Our work on RILS is still preliminary. We can see already that the costs of reflection make it a dangerous tool to use. It seems clear that the only real solution at this point is to reflect only on an "as-needed" basis: RILS will only reflect to a higher level when it cannot solve the problem with a lower level of analysis. Future work will examine the issue of controlling reflection in more detail by comparing different control methods and their resulting costs.

References

- Arcos, J., and Plaza, E. 1993. A reflective architecture for integrated memory-based learning and reasoning. In Wess, S.; Altoff, K.; and Richter, M., eds., *Topics in Case-Based Reasoning*. Kaiserslautern, Germany: Springer-Verlag.
- Birnbaum, L.; Collins, G.; Brand, M.; Freed, M.; Krulwich, B.; and Pryor, L. 1991. A model-based approach to the construction of adaptive case-based planning systems. In Bareiss, R., ed., *Proceedings of the Case-Based Reasoning Workshop*, 215-224. San Mateo: DARPA.
- Bradzil, P., and Konolige, K., eds. 1990. *Machine Learning, Meta-Reasoning and Logics*. Boston: Kluwer Academic Publishers.
- Clancy, W. 1988. Acquiring, representing, and evaluating a competence model of diagnostic strategy. In Chi, M.; Glaser, R.; and Farr, M., eds., *The Nature of Expertise*. Hillsdale, NJ: Lawrence Erlbaum Associates. chapter 12, 343-418.
- Collins, G.; Birnbaum, L.; Krulwich, B.; and Freed, M. 1993. The role of self-models in learning to plan. In *Foundations of Knowledge Acquisition: Machine Learning*. Kluwer Academic Publishers. 83-116.
- Cox, M., and Freed, M. 1995. Using knowledge of cognitive behavior to learn from failure. In *Proceedings of the 1995 AAAI Spring Symposium on Representing Mental States and Mechanisms*.
- Cox, M. 1996. *Introspective multistrategy learning: Constructing a learning strategy under reasoning failure*. Ph.D. Dissertation, College of Computing, Georgia Institute of Technology. Technical Report GIT-CC-96-06.
- Davis, R. 1982. Application of meta level knowledge to the construction maintenance and use of large knowledge bases. In Davis, R., and Lenat, D., eds., *Knowledge-Based Systems in Artificial Intelligence*. McGraw-Hill. 227-484.
- Fox, S., and Leake, D. 1995. Modeling case-based planning for repairing reasoning failures. In *Proceedings of the 1995 AAAI Spring Symposium on Representing Mental States and Mechanisms*. Stanford, CA: AAAI.
- Fox, S. 1995. *Introspective Reasoning for Case-Based Planning*. Ph.D. Dissertation, Indiana University, Computer Science Department. IUCS: Technical Report 462.
- Freed, M., and Collins, G. 1994. Adapting routines to improve task coordination. In *Proceedings of the 1994 Conference on AI Planning Systems*, 255-259.
- Gurer, D.; des Jardins, M.; and Schlager, M. 1995. Representing a student's learning states and transitions. In *Proceedings of the 1995 AAAI Spring Symposium on Representing Mental States and Mechanisms*. Stanford, CA: AAAI.
- Ibrahim, M. 1992. Reflection in object-oriented programming. *International Journal on Artificial Intelligence Tools* 1(1):117-136.
- Krulwich, B.; Birnbaum, L.; and Collins, G. 1992. Learning several lessons from one experience. In *Proceedings of the Fourteenth Annual Conference of the Cognitive Science Society*, 242-247. Bloomington, IN: Cognitive Science Society.
- Oehlmann, R.; Edwards, P.; and Sleeman, D. 1994. Changing the viewpoint: Re-indexing by introspective questioning. In *Proceedings of the Sixteenth Annual Conference of the Cognitive Science Society*, 675-680. Lawrence Erlbaum Associates.
- Ram, A., and Cox, M. 1994. Introspective reasoning using meta-explanations for multistrategy learning. In Michalski, R., and Tecuci, G., eds., *Machine Learning: A multistrategy approach Vol. IV*. Morgan Kaufmann. 349-377.
- Rosenbloom, P.; Laird, J.; and Newell, A. 1993a. *Meta Levels in Soar*, volume I. The MIT Press. chapter 26.
- Rosenbloom, P.; Laird, J.; and Newell, A. 1993b. *R1-SOAR: An Experiment in Knowledge-Intensive Programming in a Problem Solving Architecture*, volume I. The MIT Press. chapter 9.
- Smith, B. 1982. *Reflection and Semantics in a Procedural Language*. Ph.D. Dissertation, Massachusetts Institute of Technology, Cambridge, MA. MIT-LCS-TR-272.
- Stroulia, E., and Goel, A. 1995. Functional representation and reasoning in reflective systems. *Applied Artificial Intelligence: An International Journal, Special Issue on Functional Reasoning* 9(1):101-124.
- Stroulia, E. 1994. *Failure-Driven Learning as Model-Based Self-Redesign*. Ph.D. Dissertation, College of Computing, Georgia Institute of Technology. Technical Report GIT-CC-96106.