Varieties of Problem Solving in a Unified Cognitive Architecture

Pat Langley (langley@asu.edu)
School of Computing and Informatics
Arizona State University, Tempe, AZ 85287 USA

The Icarus Architecture

In previous publications, we have described Icarus, a unified theory of the human cognitive architecture in the spirit of Newell (1990) and Anderson (1993). Like its predecessors, the framework makes strong assumptions about the representation and organization of cognitive structures in memory, the performance mechanisms that operate on these structures, and the learning processes that acquire them from experience. Our aim has been not to model details like timing and error rates, but to explain the broad range of abilities found in humans.

Icarus incorporates some assumptions that distinguish it from earlier frameworks (Langley, 2006). These include separate long-term memories for concepts and skills, hierarchical organization of these memories, grounding of cognition in action and perception, and a strong correspondence between long-term and short-term structures. Recent versions of Icarus have introduced a variant of means-ends problem solving that operates in psychologically realistic ways (Langley & Rogers, 2005), along with a mechanism for learning new skills from traces of successful problem solving. We have demonstrated the architecture on a number of domains.

Forward Seach in Problem Solving

Despite this progress, Icarus remains in its early stages as a theory and does not account for important aspects of human cognition. One issue concerns people's ability to carry out heuristic search either backward from the goal, as in means-ends analysis, or forward from the current state, as in progressive deepening. The current theory models only the former, which appears reliably in novice behavior on puzzles like the Tower of Hanoi, but not the latter, which is common in board games like chess.

We are extending Icarus to support the second capacity as a variation on its existing mechanism for executing skills. Rather than carrying out its skills in the environment, the system will instead simulate them internally, producing mental encodings of the expected state after each step. This forward search will not be random, but rather constrained by the hierarchical structure of skills and the expected value of states.

Moreover, memory limitations will require the architecture to pursue a single path at a time, remembering only one simulated state in addition to the current physical state available from perceptions. Upon reaching an informative state, the module will store its value with the first skill that led to this result, then continue to search down other paths. After each pass, the system will compare the stored value to the newer one, retain the better option, and continue until the time available for search ends, when it will execute the best-scoring alternative.

Unifying Forward and Backward Search

A key question, unaddressed in the cognitive science literature, concerns when humans carry out means-ends analysis and when they invoke forward strategies like progressive deepening. One hypothesis is that they prefer backward chaining except when the branching factor from the goal or subgoal becomes too large, as occurs in most board games. A related possibility is that they favor backward search when goal descriptions are quite specific, as in many puzzles, but prefer forward chaining when goals are abstract, as in typical games.

We plan to incorporate both strategies into the next version of Icarus and compare the behavior they produce to traces of human problem solving on puzzles and games. Note that neither scheme predicts a uniform style of search; the agent may shift from backward to forward chaining, and vice versa, as it attempts to solve a problem, depending on the specific situation in which it finds itself. Human problem solving is a complex activity, and we hope that the extended architecture will improve our understanding of its distinctive varieties.

Acknowledgements

This research was funded in part by Grant FA8750-05-2-0283 from DARPA IPTO. Discussions with Dongkyu Choi, Fernand Gobet, Ugur Kuter, Dana Nau, David Nicholas, Stellan Ohlsson, Daniel Shapiro, and others contributed to many of the ideas presented here.

References

Anderson, J. R. (1993). *Rules of the mind*. Hillsdale, NJ: Lawrence Erlbaum.

Langley, P. (2006). Cognitive architectures and general intelligent systems. *AI Magazine*, *27*, 33–44.

Langley, P., & Rogers, S. (2005). An extended theory of human problem solving. *Proceedings of the Twenty-seventh Annual Meeting of the Cognitive Science Society*. Stresa, Italy.

Newell, A. (1990). *Unified theories of cognition*. Cambridge, MA: Harvard University Press.