In J. Hendler (Ed.), *Artificial Intelligence planning systems.* San Mateo, CA: Morgan Kaufman. 1992.

Reaction and Reflection in Tetris

David Kirsh **Paul Maglio**
Cognitive Science Department, 0515
University of California, San Diego
La Jolla, CA 92093

Abstract

To discover how to couple reflection with reaction we have been studying how people play the computer game Tetris. Our basic intuition is that the job of the reasoner is to monitor the environment and the agent's behavior over time to discover trends or deviations from the agent's normative policy, and tune the priorities of the attentional system accordingly.

Introduction

How can a high level planner interact with a reactive system? The question is of interest because it is likely that a system able to cope in real-time with the complexities characteristic of many human activities will require an architecture with both reactive and reasoning components. The question is hard because truly reactive systems tend to be immune to top-down interference.

In mobot architectures, for instance, activity control systems are linked together in a layered fashion that permits communication only at lines of input or output (Brooks, 1990). Input to a layer may be suppressed, output inhibited or augmented. But the processes occurring inside each activity layer are well-insulated and modular, sealed off from the computations occuring in other layers. Cross talk can occur only at the periphery.

A similar information encapsulation is thought to obtain for human subjects during skilled behavior. After extensive training, typists learn to carry out long sequences of movements with little attention and seem unable to interrupt particular chunks of their behavior on cue (Gentner, 1988). Apparently subjects cannot directly control automatic processing. The computation is data driven, fast, displays little temporal variability, and is probably parallel (Schneider, 1985). In short, practiced behavior seems the product of specific highly dedicated procedures, often parallel, which are insensitive to outside information.

To discover how humans integrate planning, or rather reflection, with reaction we have been studying subjects at different levels of skill playing the interactive video game Tetris. In Tetris players must choose from three actions: rotate, translate or drop. Tetrazoids enter from the upper boundary of a rectangular playing field at a fixed speed. As the game proceeds the pieces drop faster, leaving the player with less and less time to make the decision as to the column and orientation to place a zoid. We have implemented a computational laboratory that lets us record keystrokes and game situations, and allows us to dynamically create situations. Our goal is to tease apart the contributions a rational reflective component makes to performance from the contributions a highly automatic reactive component makes, and then to explain how these components interact.

Skill-Based Systems

In the literature on human process control it is widely claimed that one consequence of practice is that an agent begins (unconsciously) to form a model of the statistical structure of the inputs it confronts and the effects of the actions it takes (Moray, 1986). These models allow (imperfect) predictions of future inputs and the actions required to compensate for those future inputs. Typically process control theorists study domains that are slowly varying dynamic systems such as steel making, or control of a bakery oven, where non-linearities and discontinuities occur, but which in general can be expected to vary smoothly. They have noted that by attending to locally detectable error signals, practiced technicians are able to select appropriate control actions to minimize a smooth cost function.

We can exploit this idea to define, to a first approximation, a skill as an error reducing control mechanism built on a statistical model of the environment. From our viewpoint it is an implementation detail whether the agent's model is explicit and declarative or implicit in the structure and success conditions of its proce-

dures, though we assume the model is implicit.

Central to this notion of skill is the idea that behavior is *perceptually driven*—since errors are perceptually discernable—and *goal specific*—since the goal of reducing differences is intrinsic to a skill. Thus skills do not set goals, they adaptively carry them out. They rely on their perceptual representation of the current situation and their implicit model of the domain to respond adaptively. Activity layers in mobots qualify as skills according to this definition.

RoboTetris

We have implemented a simple skill-based system that plays Tetris. Although in principle there are 17×2^{270} possible Tetris states, our system represents a board situation by a vector of six features: number of holes present, total board height, mountain height, covered holes, filled rows, and local fitness. This particular feature set is based on experts' verbal protocols and performance data. RoboTetris computes the overall goodness of a possible placement as a weighted sum of the features. All possible placements are considered and the best one is selected.

Though RoboTetris' weights can be adjusted to outperform any human expert, it does not play like a person. For example, it is much less consistent than intermediate players, even though it performs above experts; that is, the standard deviation of its scores is unnaturally high. Similarly, RoboTetris has a bizarre highest peak/lowest valley profile. Of course, these unnatural characteristics may disappear with better weighting functions, better qualitative feature sets, and the addition of specific case knowledge. But at present we take its inability to achieve thoroughly reliable behavior as an indication of limitations of the skill-based approach.

It is in keeping with this view that verbal protocols of players become significant. Experts report that they have a high level strategy which shows itself as a *set of concerns*. They claim to monitor ongoing activity mindful of these concerns. If the protocols are to be taken at face value, these concerns serve to focus attention on certain regions of the board or onto certain board properties. In particular, flatness seems to be a crucial topic of concern.

Exactly what an expert means by flatness, however, is not apparent. We assume that the concept must be perceptually identifiable: that in principle it is a geometric or topological notion that can be applied on the basis of current perceptual input. Hence, in principle, it might be added as a new feature to the feature set of the skill module. But again judging from protocols it would seem that an expert's concern with flatness varies in the course of a game, often with one's current estimation of abilities at the time. If flatness, or rather one's degree of concern with flatness, varies with the internal state of the agent, it cannot be reducible to currently perceivable properties. Pending new evidence to the contrary, we take the varying nature of concern with flatness to be a key empirical indicator of higher level control.

Integrating Reflection and Reaction

We believe that reflective concerns can improve performance. But what is the mechanism by which high level reflection can interact with lower level skills? Most existing accounts of architectures coupling planners to reactive systems begin with the assumption that planners are able, if necessary, to *override* the output of reactive systems, or to *suppress* the input to particular activity layers. Little effort is paid to actively redirecting the sensors of the system to *bias* the input stream.

We are just in the process of implementing a model which works by controlling *input* in a sophisticated manner, sensitive to what we call *concerns* and *policies*. A *policy* is a set of norms, specified in a high level language reflecting a global perspective on the game. Policies identify how things ought to go in the course of a game. The job of the planner is to monitor for signigficant differences between how things ought to be and how things are. These significant differences are concerns—high level descriptions of respects in which the reactive agency is falling short of implementing the system's policies. Once concerns are identified they are translated into directives for changing the focus of attention. Changes in attention, in turn, affect the feature vector presented to the skill module and hence indirectly affects behavior. Accordingly, reasoning affects attention, which has the effect of biasing the perception of certain board regions, or increasing the precision of values being returned by visual routines and the like, which in turn tends to change the behavior with predictable effect over several moves.

References

Brooks, R. (1990). Intelligence without Representation. *Artificial Intelligence*, *47*, 139–160.

Gentner, D. (1988). Expertise in Typewriting. In M. Chi, R. Glaser, & M. Farr (Eds.), *The Nature of Expertise*, pages 23–70. Hillsdale, NJ: Lawrence Erlbaum Inc.

Moray, N. (1986). Monitoring Behavior and Supervisory Control. In *Handbook of Perception and Human Performance, Vol II: Cognitive Processes and Performance*, chapter 40. New York, NY: John Wiley & Sons.

Schneider, W. (1985). Toward a Model of Attention and the Development of Automatic Processing. In M. Posner & O. Marin (Eds.), *Attention and Performance XI*, pages 475–492. Hillsdale, NJ: Lawerence Erlbaum Inc.