

# Reflective Reasoning and Learning

Eleni Stroulia \*

College of Computing

Georgia Institute of Technology

Atlanta, GA 30332-0280

eleni@cc.gatech.edu

The capability of learning is a prerequisite for autonomy. Autonomous intelligent agents, who solve problems in a realistic environment need to learn in order to extend the classes of problems they can solve, to improve their performance on these problems, and to improve the quality of the solutions they produce.

One way in which an intelligent agent may effectively use its experiences to learn, is by reflection upon its own problem-solving process. To do that, the agent needs to have an explicit meta-model of its own reasoning and knowledge. This work takes a functional stance towards reflective learning. This stance gives rise to a specific computational model which is based on three major hypotheses: (i) agents can be viewed as abstract devices, (ii) their reasoning can be understood in terms of structure-behavior-function (SBF) models, and (iii) learning can be viewed as a self-redesign task in which the agent uses its understanding of its own reasoning to improve its subsequent performance.

**Modeling Reasoning in terms of SBF Models** Intelligent agents accomplish their problem-solving tasks through the internal information-processing behaviors of their functional architecture. The SBF model describes the agent's reasoning as a non-deterministic sequence of information transformations, through which the output information of the agent's task is produced from its input.

Each information transformation corresponds to some subtask of the agent. The SBF model explicitly specifies the interactions among subtasks in terms of information and control flow. For each subtask, the model also specifies the role it plays to the accomplishment of the overall task in terms of the information it consumes, the information it produces, and a set of "correctness relations" between its input and output. Further, the SBF model specifies the representational and organizational assumptions on which the agent's world knowledge is based, and the ways in which this knowledge is used by the agent to select among its reasoning methods and accomplish its tasks.

**Reflection** If an agent has such a model of its own reasoning, it can use it to

1. *monitor* its own behavior. The SBF model of the problem solver provides a language for the agent to interpret its own reasoning steps in terms of its current goals and to generate expectations regarding their outcomes.

\*This research was conducted under the advice of Dr. Ashok Goel. It has been supported by the NSF (grant IRI-92-10925), the ONR (contract N00014-92-J-1234), the Advanced Projects Research Agency, and an IBM graduate fellowship.

2. *assign blame* when it fails. The SBF model of the problem solver along with the record of its reasoning on the failed problem-solving episode, enables the agent to localize the cause of its failure to some element of its task structure.
3. *redesign* itself appropriately. The semantics of the SBF language enable the the agent to modify itself in a way that maintains the overall consistency of its reasoning.

This reflection process enables problem solvers, to some extent, to redesign themselves and adapt their problem-solving behavior to meet new requirements. Therefore, it is especially useful when the agent's tasks are complex and when the requirements imposed by the the environment upon the agent regarding its processing and the quality of the solutions it produces can change. If, on the other hand, the environment is stable and the task structure is close to optimal, it may be less cost efficient.

**Discussion** AUTOGNOSTIC (Stroulia and Goel 1993) is a system that implements and evaluates this theory of reflective reasoning and learning. Given the SBF model of a problem solver, AUTOGNOSTIC monitors its reasoning, and when it fails, assigns blame to some of its elements and redesigns it appropriately. AUTOGNOSTIC presently operates in two widely different task domains: on top of ROUTER, in the domain of navigational planning, and on top of KRITIK2, in the domain of engineering design. Both ROUTER and KRITIK2 are autonomous multistrategy systems developed independently of AUTOGNOSTIC, and since they solve widely different tasks in widely different domains, the success of AUTOGNOSTIC's reflection process suggests that its SBF models of problem solving and reflection process are quite general. AUTOGNOSTIC is capable of identifying and correcting errors

1. in the representation scheme of the world knowledge,
2. in the organization of the world knowledge,
3. in the content of the world knowledge itself,
4. in the assumptions on the applicability and utility of different reasoning methods, and
5. the role of a subtask in the overall reasoning process.

## References

[Stroulia and Goel 1993] Stroulia, E; and Goel, A. 1993. Functional Representation and Reasoning for. Reflective Systems. *Applied Artificial Intelligence: An International Journal* Forthcoming.