Reset reproduction of a response written December 7, 1978, to the Brachman-Smith Knowledge Representation Survey, appearing in *ACM SIGART Newsletter*, No. 70 (February 1980), p. 41, references on p. 12. Reset in LaTeX by Jon Doyle December 1995. Current address: MIT Laboratory for Computer Science. Available via `http://www.medg.lcs.mit.edu/doyle`.

Why I Don't Play the Piano

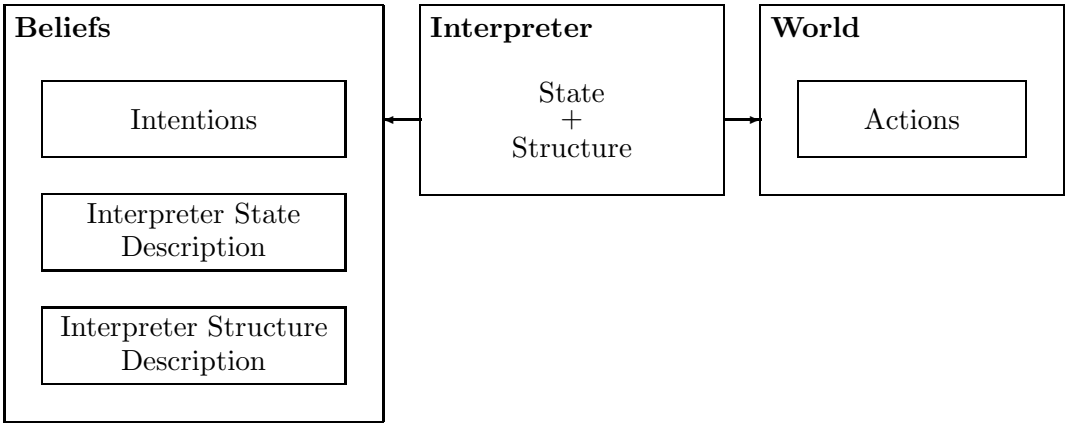
Jon Doyle DOYLE@MIT.EDU
Massachusetts Institute of Technology, Artificial Intelligence Laboratory
Cambridge, Massachusetts 02139, U.S.A.

My work in the past several years has dealt primarily with the study of the processes used in constructing varieties of problem solving behaviors. I have studied processes ranging from the propagation of constraints and non-chronological control schemes to belief revision, introspection, self-consciousness, and learning. While I consider many processes like these to be crucial ingredients of effective problem solvers, and thus important parts of any knowledge representation system purporting to be useful for constructing problem solvers, I find that your questionnaire lacks almost completely any mention of questions concerning problem solving processes. The theoretical bias of the questionnaire recognizes only a sense of representation as mirroring conditions existing in the world, and this leads to many questions traditionally arising in logic, semantics and the philosophy of language. Absent are questions concerning intention, action, purposive communication, and the processes of problem solving. While I am quite interested in some of the problems touched upon in the questionnaire, few, if any, of the questions presented bear directly on the issues I study. I therefore will describe my concerns without regard to most of the questionnaire.

My current work centers on the problem of constructing introspective, self-conscious, self-modifying problem solving interpreters, my thesis being that many of the processes involved in expert problem solving, common sense reasoning, and learning or developmental systems rely to a greater or lesser extent on a reasoner being able to reason about his own structure, beliefs, intentions, and behavior as well as being able to reason about external domains. When applied to the representation of descriptions in the problem solver's data base, self-reference in the small (the ability to refer to descriptions and their parts) is a fundamental aspect of the structured descriptions recently developed by Brachman, Fahlman, Philip Hayes, Hendrix, and Sussman. When applied to the representation of the problem solver's belief system, self-reference in the large (the ability to refer to properties of the entire belief system) is the basis of the non-montonic reasoning and belief revision processes developed by Stallman and Sussman, McDermott, London, and myself. When applied to the representation of the intentions of the problem solver, self-reference underlies the reasoning processes studied by Sussman, Pat Hayes, Sacerdoti, McDermott, de Kleer, Steele, Shrobe, and myself. Finally, applied to representing models of the problem solver and external agents, self-reference supports learning and communication as investigated by McCarthy, Minsky, Sussman, Cohen, and myself.

In overview, my envisioned design of what I call a "reflexive interpreter" has the interpreter consisting of a belief system (in which are represented beliefs about the world, the interpreter's intentions, and descriptions of the interpreter's state and structure) and the interpreter proper (which links the belief system with the world). In a diagram these components are displayed as follows:[1]

[1]Note for the reprinting: The published version omitted this sentence and diagram.



The interpreter is committed to fulfill the intentions when possible by making changes in the world, to update beliefs by making observations about the world, and importantly, to update beliefs by making observations about its own state and that of the belief system, and to modify itself to maintain a correspondence with the interpreter description contained in the belief system.

In somewhat more detail, the belief system uses a sufficiently expressive system of structural descriptions (just about any of the systems mentioned above should do, as the real demands on the descriptive language arise in describing external domains) being maintained by a truth maintenance system. The truth maintenance system supplies the basis for the representation of actions used by the system, and an extended McDermott-like task network links the interpreter state description with the intentions of the system. Finally, the description of the interpreter's structure is a static representation of the interpreter, to be used by learning and compilation processes in which modifications made in the interpreter description lead to changes in the interpreter itself. Again, I emphasize that I focus on the architecture and processes which comprise the interpreter, rather than on the design of a universal language for representing worldly knowledge.

The specific concerns I have addressed or am now addressing include; truth maintenance systems, which support the drawing of conclusions from incomplete information and the subsequent revision of beliefs if these assumptions are later contradicted; non-monotonic logic, which is a mathematical formalization of the logic underlying truth maintenance systems; an action representation, which is based on a certain method for using the non-monotonic arguments of truth maintenance systems; a simple notion of causality appropriate for this action representation; task networks and their interpretation; interpreter self-descriptions; and finally, learning processes by which the interpreter description is modified through experience.

References

Expositions most descriptive of my work are as follows:

1. Jon Doyle, "The Use of Dependencies in the Control of Reasoning"
2. Jon Doyle, "Truth Maintenance Systems for Problem Solving"
3. Jon Doyle, "A Glimpse of Truth Maintenance" (revised version)
4. Jon Doyle, "Reflexive Interpreters" (under revision)
5. Jon Doyle, "Action, Causality, and the Change of Beliefs" (in progress)
6. Johan de Kleer, Jon Doyle, Guy Steele and Gerald Sussman, "Explicit Control of Reasoning"
7. Johan de Kleer, Jon Doyle, Guy Steele and Gerald Sussman, "AMORD: A Deductive Procedure System"
8. Drew McDermott and Jon Doyle, "Non-Monotonic Logic I"
9. Drew McDermott and Jon Doyle, "Non-Monotonic Logic II" (in progress)