Broad Agents

Joseph Bates
joseph.bates@cs.cmu.edu

Bryan Loyall
bryan.loyall@cs.cmu.edu

W. Scott Reilly
william.reilly@cs.cmu.edu

School of Computer Science
Carnegie Mellon University
Pittsburgh, PA 15213

Abstract

The Oz project at Carnegie Mellon is developing technology for dramatic virtual worlds. One requirement of such worlds is the presence of broad, though perhaps shallow, agents. To support our needs, we are developing an agent architecture that provides goals and goal directed reactive behavior, emotional state and its effects on behavior, some natural language abilities (especially pragmatics based language generation), and some memory and inference abilities. We are limiting each of these capacities whenever necessary to allow us to build a broadly capable, integrated agent.

In attempting to construct a broad agent, constraints seem to arise between components of the architecture. In this brief note, we discuss some of these constraints.

1 Introduction

The Oz project at the Carnegie Mellon School of Computer Science is developing technology for high quality interactive fiction and virtual realities. Our goal is to provide users with the experience of living in dramatically interesting micro-worlds that include moderately competent, emotional agents [1].

An Oz world is composed of (1) a simulated physical environment, (2) the agents which populate the environment, (3) a user interface to allow one or more humans to participate in the environment, and (4) a computational theory of drama. The drama theory plans and gently controls the overall flow of events in the world to provide long term structure to the user's experiences. The user interface includes a theory of presentation, which continually adjusts the style of presentation to suit the varying dramatic content of the experience. We hope our work will be the basis for one of the first sophisticated knowledge based art forms.

One of the keys to an effective virtual world is for the user to be able to "suspend disbelief". That is, the user must be able to imagine that the world portrayed is real, without being jarred out of this belief by the world's behavior.

In order to foster the illusion of reality, we believe our agents must have broad, though perhaps shallow, capabilities. To this end, we are attempting to produce an agent architecture that includes goals and goal directed reactive behavior, emotional state and its effects on behavior, some natural language abilities (especially pragmatics based language generation), and some memory and inference abilities. Each of these capacities can be as shallow as is necessary to allow us to build a broadly capable, integrated agent [1].

Building broad agents is a little studied area. Even in this proceedings, the majority of efforts are to integrate relatively few capabilities, such as action and learning [2]. We are hopeful that moderate effort in integration may yield good results, such as believable hints of thought and emotion in our limited micro-world domains.

In thinking about the nature of broad agents, and in developing our architecture, called Tok, we have felt constraints arise between the components of the design. In the remainder of this note, we sketch the architecture and mention some of the constraints.

2 The Architecture

As mentioned above, we wish our agents to sufficiently suggest human behavior that the true human participants of Oz worlds can project depth into the agents. We believe this means that agents must provide some signs of internal goals, reactivity, emotion, natural language ability, and knowledge of agents (self and other) as well as of the simulated physical world.

Tok divides these behavioral features among three main components: a goal based reactive engine called HAP, an emotion component called Em, and a natural language system called Glinda.

Reactivity is of key importance for us, because non-reactive behavior immediately breaks the illusion of intelligence that we wish to project. Even animals are reactive, if not particularly smart. Initially we intended to provide reactivity by using a planner in the background feeding a reactive front-end that would execute plans. However, as we pursued this idea, the reactivity seemed to spread through the system, so that we see it now as fundamental to the entire architecture.

The ideas of Agre and Chapman, Brooks, and others reacting against the classical planning/execution paradigm have impressed us. Nonetheless, we believe that at least the appearance of goal directed behavior is key to providing our illusion. Thus, we want a notion of goal and subgoaling in our architecture.

We felt that the notion of a goal as a structured object, which could be inspected and analyzed by a planner, was perhaps unhelpful in a complex environment. Thus, HAP, the central component of Tok, is a "reactive planner", somewhat along the lines of Firby's RAP system and Simmons' TCA. HAP uses goals and perceived world state as indices into mem-

[1] In the context of Oz, instead of demanding that our agents be especially active and smart, we require only that they not be clearly stupid or unreal. An agent that keeps quiet may appear wise, while one that oversteps its abilities may destroy the suspension of disbelief. Broad agents will let us take advantage of the "Eliza effect" [9], in which people see subtlety, understanding, and emotion in an agent as long as the agent does not actively destroy the illusion. There is some evidence that this approach also works in politics.

[2] Vere's work [8] is a significant exception.

ory to recall relevant sets of actions. These actions can be subgoals to achieve or actions to take in the world.

HAP cannot inspect the structure of goals and sets of actions. Thus, it does not reason about preconditions, postconditions, or add-delete lists. Generally, we see this as a strength of the system. However, people do seem to attach some information to even low level goals and plans, such as when a goal has been fortuitously satisfied and when a plan is no longer remotely applicable. Thus, goals and actions sets include information that let HAP react appropriately when these conditions arise.

Tok presently operates only in the world. However, it is clearly useful to be able to imagine the effects of actions before or instead of doing them. Thus, while we will not represent the known primitive actions in a decomposable fashion (such as via pre and post conditions), we are working to provide an envisionment system for Tok. This will allow a bit of forward search using a simulated world (which is particularly easy to do in Oz).

The natural language component of Tok is presently an independent subsystem. It is used by agents to generate textual "speech". Unlike most earlier systems, Glinda does text planning and surface realization using a single integrated framework. With this approach, we can vary the style of output at many levels, such as concept choice, syntactic form, and lexical choice, based on pragmatic information provided by the agent, such as their emotional state. We believe this will help us project the illusion of humanity when agents speak.

Michael Dyer's work in story understanding showed a way for attaching meaning to emotion words [3]. His approach was heavily based on "goal situations", patterns of goals, actions, and achievement (of one or more agents) that gave rise to certain emotions. For instance, "grateful" could arise when someone else helped an agent achieve an otherwise hard to achieve goal. "Afraid" could arise if an important goal was judged unlikely to be achieved. Tok's emotion subsystem, Em, uses Dyer's general approach, which was further explored by Ortony et al [7].

Further detail on Tok is provided in several technical reports [5, 6, 2].

In organizing the workshop, John Laird asked us to specify whether our system is operational, what earlier work inspired our approach, and how our architecture compares to other integrated architectures.

Tok is presently under development. HAP and Glinda are running, Em is still being developed, and the integration remains incomplete in the implementation, though we have done it on paper.

Agre and Chapman's various work at MIT, especially the paper "What are plans for?", Firby's Ph.D. thesis at Yale on Reactive Action Packets, Lucy Suchman's book on "Plans and Situated Actions", and Bates' earlier work at Cornell on automated reasoning in mathematics (the PRL project), strongly affected our thoughts on action.

The Soar work at CMU and elsewhere is often in our thoughts, though the immediate influence on our efforts is not clear.

Dyer's thesis at Yale, "In Depth Understanding", suggested that it might be possible to build agents that would provide adequate illusions of competence and emotion. However, his work did not try to explain how to produce "in depth behavior", and that has turned out to be non-trivial, as far as we are concerned. Nonetheless we have directly benefited from the BORIS representations and certain aspects of processing.

Ed Hovy's thesis at Yale on natural language generation under pragmatic constraints has influenced our efforts in NL generation. Also influential is work on systemic grammar, the work of the CMU Center for Machine Translation, and general ideas from the Yale NLP school.

In comparing Tok to others systems, such as Soar, we find that Tok has no unified framework, other than our general concern about reactivity. To an extent this troubles us, but we take some comfort from the view that homogeneity is in the eye of the beholder, depending on the knowledge one brings to understanding a system. In addition, we believe that minds must be heterogeneous in a significant way, ie they must be fundamentally very messy. It seems to us that the idea of information as entropy leads to the conclusion that only weak agents can have regular and well structured minds. Another way to say this is that architecture, if taken to be the fundamental organizing principles of the mind, is necessarily spread rather uniformly throughout the agent's mind.

3 Problems in Broad Architectures

In developing Tok, we have come across integration problems that seem to be inherent in the task and which suggest constraints and difficulties for any architecture. Here we mention five such problems.

1. how to have goals when the external environment is very rich

   When the external world is rich, it appears very difficult to characterize the possible world states and the effects of actions on the world. Nonetheless, to help organize action (and perhaps emotion) we would like a notion of "goal" in our agents. Thus we need some way to represent goals, do subgoaling, and notice achievement and serendipity, while using a limited representation for goals and operators.

2. how to think ahead when the external environment is very rich

   If our goals and operators have no decomposable structure, in particular if there are no analyzable pre or post conditions, we cannot do conventional planning. However, it seems clear that agents can envision some of the effects of their actions, and "plan" accordingly. What sorts of representations permit this behavior?

3. how to make use of chunks when the external environment is very rich

   The initial creator of an agent will supply certain mappings from goals (and other state information) to actions. Additional such mappings may be created by learning: chunking experience, chunking simulated experience, or other means. However, because the rich real world may not react as our chunks suggest, we need to be able to apply them in intelligent ways as we carry them out. What representations and processing techniques support the "intelligent" use of chunked action?

4. how to integrate emotion with reactivity

   Work on cognitive aspects of emotion provides one means for building emotional agents. However, it seems to require an architecture to explicitly represent goals, plans, and actions of self and other agents. Can we retain the essentials of a reactive architecture while incorporating enough explicit representation to support cognition based emotion?

5. how to integrate natural language generation with other action

   As noted earlier, text planning and surface realization are becoming increasingly integrated in modern natural language generators. We have developed one such highly integrated system in our attempt to extend Hovy's work on pragmatics based generation. However, as we connect that system to the rest of our agent architecture, it is clear that generation must dynamically access and modify the goals, beliefs, emotions, sensory memory, and other aspects of the state of the agent. For instance, we would like decisions made during speech generation to affect choice of non-speech actions, such as pointing and crying. Does this mean that architectures for action must equally well support the specialized needs of natural language processing?

4 Conclusion

While our work on Tok is still at an early stage, we feel that it has helped us begin to understand the range of constraints that arise in broad agents. While we could have abstractly speculated on these constraints prior to our effort, the specifics and their impact on the design of individual components of Tok were not apparent to us before we began the work.

By virtue of its size and goals, the Soar community (or at least its leaders) is coming up against many of the issues that arise in building broad agents. However, most efforts at integrated agents in fact focus on relatively narrow kinds of integration. This is useful, just as developing individual facets of architectures is useful. However, from our experience, we would recommend that those researchers who want their work to apply directly to future broadly integrated agents should spend some time studying broad agents now.

Acknowledgements

The Oz project is supported in part by Fujitsu Laboratories, Ltd.

References

[1] Bates, J. *Deep Structure for Virtual Reality.* Submitted to "Presence, The Journal of Teleoperators and Virtual Environments", MIT Press, 1991. Also available as Technical report CMU-CS-91-133, School of Computer Science, Carnegie Mellon University, Pittsburgh, PA, 1991.

[2] Bates, J., Loyall, A., Reilly, S., Kantrowitz, M. *Tok: A Broad Architecture for Intelligent, Emotional Agents.* Technical report, School of Computer Science, Carnegie Mellon University, Pittsburgh, PA, forthcoming summer 1991.

[3] Dyer, M. *In-Depth Understanding.* The MIT Press, Cambridge, MA, 1983.

[4] Hovy, E. *Generating Natural Language under Pragmatic Constraints.* Lawrence Erlbaum Associates, Hillsdale, New Jersey, 1988.

[5] Kantrowitz, M. *Glinda: Natural Language Text Generation in the Oz Interactive Fiction Project.* Technical report CMU-CS-90-158, School of Computer Science, Carnegie Mellon University, Pittsburgh, PA, 1990.

[6] Loyall, B., and Bates, J. *HAP: An Adaptive, Reactive Architecture for Agents.* Technical report, School of Computer Science, Carnegie Mellon University, Pittsburgh, PA, forthcoming summer 1991.

[7] Ortony, A., Clore, G., and Collins, A. *The cognitive structure of emotions.* Cambridge University Press, 1988.

[8] Vere, S., and Bickmore, T. *A basic agent.* Computational Intelligence 6, pp. 41-60, 1990.

[9] Weizenbaum, J. *Eliza.* Communications of the ACM, 9, 1966, pp. 36-45.