N94-34053

FILTERING AS A REASONING-CONTROL STRATEGY: AN EXPERIMENTAL ASSESSMENT

Martha E. Pollack
Department of Computer Science
and
Intelligent Systems Program
University of Pittsburgh
Pittsburgh, PA 15260

ABSTRACT

In dynamic environments, optimal deliberation about what actions to perform is impossible. Instead, it is sometimes necessary to trade potential decision quality for decision timeliness. One approach to achieving this trade-off is to endow intelligent agents with meta-level strategies that provide them guidance about when to reason (and what to reason about) and when to act. We describe our investigations of a particular meta-level reasoning strategy, *filtering*, in which an agent commits to the goals it has already adopted, and then filters from consideration new options that would conflict with the successful completion of existing goals [1]. To investigate the utility of filtering, we conducted a series of experiments using the Tileworld testbed [12]. Previous experiments conducted by Kinny and Georgeff used an earlier version of the Tileworld to demonstrate the feasibility of filtering [5]. We present results that replicate and extend those of Kinny and Georgeff, and demonstrate some significant environmental influences on the value of filtering.

INTRODUCTION

Many existing and potential AI applications involve systems that are situated in dynamic environments: Laffey *et al.* list examples from aerospace, communications, medical, process control, and robotics applications [6]. Optimal deliberation about what actions to perform is impossible in such environments. This is because all systems have computational resource-limits: their deliberations take time. During the time in which a system in a dynamic environment is deliberating about what actions to perform, the environment may change—and it may change in ways that undermine the assumptions underlying the deliberation. A system may begin a deliberation process with a particular set of available options for action, but new options may arise and formerly existing options may disappear during the course of the deliberation. Moreover, the utilities associated with each option are subject to change during the deliberation. A system that blindly pushes forward with its original deliberation process, without regard to the amount of time it is taking or the changes meanwhile going on, is not likely to make rational decisions about what to do. It is thus sometimes necessary to trade potential decision quality for decision timeliness [14, 10, 13].

One approach is to endow intelligent systems, or agents, with meta-level strategies that provide them guidance about when to reason (and what to reason about) and when to act. In previous work, we have proposed two such strategies: *filtering* [1], and *overloading* [9]. In the present paper, we focus on filtering, a strategy in which an agent commits to the goals it has already adopted, and tends to bypass, or filter from consideration, new options that would conflict with the successful completion of existing goals.

To investigate the utility of filtering, we conducted a series of experiments using a simple, abstract testbed: the Tileworld. Our use of the Tileworld is part of an experimental research methodology that we discuss in detail elsewhere [3, especially Section 5.2]. We first described the Tileworld several years ago [12]. Since then, we have made a number of enhancements to the original system, so that it can support a wider range of experiments. A simplified version of the original Tileworld was used by Kinny and Georgeff [5] in a series of experiments that demonstrated the utility of filtering. The experiments we report on in this paper replicate and extend those of Kinny and Georgeff, and demonstrate some significant environmental influences on the value of filtering.

THE TILEWORLD TESTBED

The Tileworld testbed is a tool that we developed to support controlled experimentation with agents in dynamic environments. It is designed to run under Unix, using Lucid Common Lisp and CLX (the Common Lisp X Interface). We first described the Tileworld several years ago [12]; since then, we have made a number of enhancements to the system, so that it now supports a wider range of experiments. We briefly describe the current state of the system, focusing on those aspects of it that are most pertinent to our experimental investigations of filtering. Details about the system implementation, along with information about how to obtain a copy of it, can be found in the Tileworld User's Guide [4].

The Tileworld consists of an abstract, dynamic, simulated environment with an embedded agent. It is built around the idea of an agent carrying "tiles" around a two-dimensional grid, delivering them to "holes", and avoiding obstacles. The environment is dynamic; during the course of a simulation, objects appear and disappear at rates specified by the researcher. The Tileworld is obviously, and intentionally, a highly artificial environment. In keeping the environment divorced from any particular application, our goal has been to provide a tool that allows researchers concerned with *any* application to focus on what they consider to be key features of that application's environment, without the confounding effects of the actual, complex environment itself. We have, in other words, traded realism—in the short run, at least—for sufficient control to allow for systematic experimentation. This methodological decision is one that has also been made in several other testbeds for studying AI planning, for example, the independently developed NASA Tileworld [8] and the MICE system [2, 7], both of which are also organized around the theme of agents situated on two-dimensional grids, pushing tiles. See Hanks, Pollack, and Cohen [3] for a discussion of the methodological issues surrounding the use of simplified testbeds.

A researcher using the Tileworld can manipulate and monitor characteristics of the simulated environment (such as how quickly it changes) and of the embedded agent (such as what kind of meta-level reasoning principles it employs). These characteristics can be defined either interactively using a menu-based interface, or by storing parameter settings in files that are then used to control batch-style experiments.

In originally developing the Tileworld, we adopted a minimalist philosophy: our policy was to keep the environment as abstract and simple as possible, in order to provide the experimenter with maximal control over the environment and to ensure that the system's performance is not tied to the particulars of any given domain. Each of the parameters in the original Tileworld was introduced because it represented an abstraction of what we believed to be a potentially important and interesting environmental characteristic. Thus, the original Tileworld allowed us to manipulate a number of environmental characteristics, including the degree of dynamism in the environment, the degree of uniformity of task difficulty, and the degree of uniformity of task reward.

Our early experiences with the Tileworld led us to conclude that, while this was a good set of parameters with which to begin, some extensions were necessary to support the range of experiments we hoped to conduct. In particular, in the original system, agents had only a single type of top-level goal, hole-filling, and no matter how they achieved such a goal, they were always awarded the same score (i.e., the score associated with the hole in question). This made the original Tileworld environment one in which there was very little about which to deliberate, and it was thus difficult to study the trade-offs involved in extra deliberation. We thus extended the system in several ways:

- We added the requirement that agents maintain fuel level: we can thus now study goals of maintenance.

- To enable agents to maintain their fuel levels, we added a "gas station" where they can go to get more fuel. We also added a top-level goal of building stockpiles of tiles having particular shapes at strategic locations on the grid. Thus, where for the original Tileworld agent all top-level goals were of the same type (fill a hole), in the new version there are several different top-level goals.

- We assigned "shapes" to tiles and holes, and changed the reward structure associated with successfully filling a hole. The agent may fill a hole with any tiles, but it gets more points if it uses tiles whose shapes match the shape associated with the hole. As a result, there is now the possibility of investigating trade-offs between the value of alternative plans to achieve a goal. An additional complication is that the agent can carry more than one tile—in the original version it only pushed a tile—but the more tiles it carries, the more rapidly it burns fuel. Again, this means that the quality of alternative alternative solutions to some goal may vary.

In implementing these extensions, we adhered to our original minimalist philosophy: we only introduced those extensions that were needed to support the experiments of interest to us. However, we think that one of the strengths of the Tileworld is its conceptual flexibility: we have found that it is relatively easy to design Tileworld modifications that support experiments that investigate environmental and agent-design issues other than those for which it was originally designed.

EXPERIMENTAL RESULTS

We now present the experiments that we conducted using the Tileworld system, to investigate the properties of filtering as a strategy for controlling reasoning in dynamic environments. Due to space limitations, we do not describe either the motivation or details of the mechanism for filtering here, but see [1, 10, 11]. Our central hypothesis, predicated on the earlier work on IRMA, was that that, in a dynamic environment, a tendency to commit to one's plans can result in overall improved performance, despite the fact that the resulting behavior will sometimes be suboptimal. This hypothesis had previously been explored by Kinny and Georgeff, using a simplified version of the original Tileworld system, along with a somewhat modified notion of filtering [5]. Our first goal was to attempt to replicate the Kinny/Georgeff results in the more-complex environment provided by the enhanced Tileworld system, using the original, better-motivated notion of filtering. We were successful in this: like Kinny and Georgeff, we showed that filtering is an effective control strategy. In addition, we generalized their results: we found that the influence of commitment is bounded, i.e., beyond a certain point, additional commitment does not lead to improved behavior, nor does increased lack of commitment lead to poorer performance. We also observed a relation between the rate of change in the environment and the value of commitment. Here we focus on this final observation.

Our primary experiment used a factorial design with two factors: degree of commitment, for which we had 14 levels, and degree of dynamism, for which we had 11 levels. "Degree of commitment" refers to the strength of the filtering strategy: the most committed agent seldom reconsidered its options until it had completed its current plan, while the least committed agent always interrupted its actions to weigh the significance of perceived changes in the environment. "Degree of dynamism" refers to the average rate of change in the environment: how frequently, on average, do exogenous events occur? The independent parameter was effectiveness, which is a normalized measure of the agent's score. There were a total of 51 trials conducted per experimental condition, where the length of each trial was 80,000 clock ticks. (A clock tick is the amount of time it takes the agent to move one unit of distance in the simulated environment.) Pre-tests were performed to establish the duration of a trial needed to ensure quiescence of effectiveness and to establish the number of trials needed to ensure quiescence of the mean effectiveness across trials.

The data we collected showed that there was a strong tendency for agents that committed more strongly to their plans to achieve higher degrees of effectiveness. This is most strongly evidenced

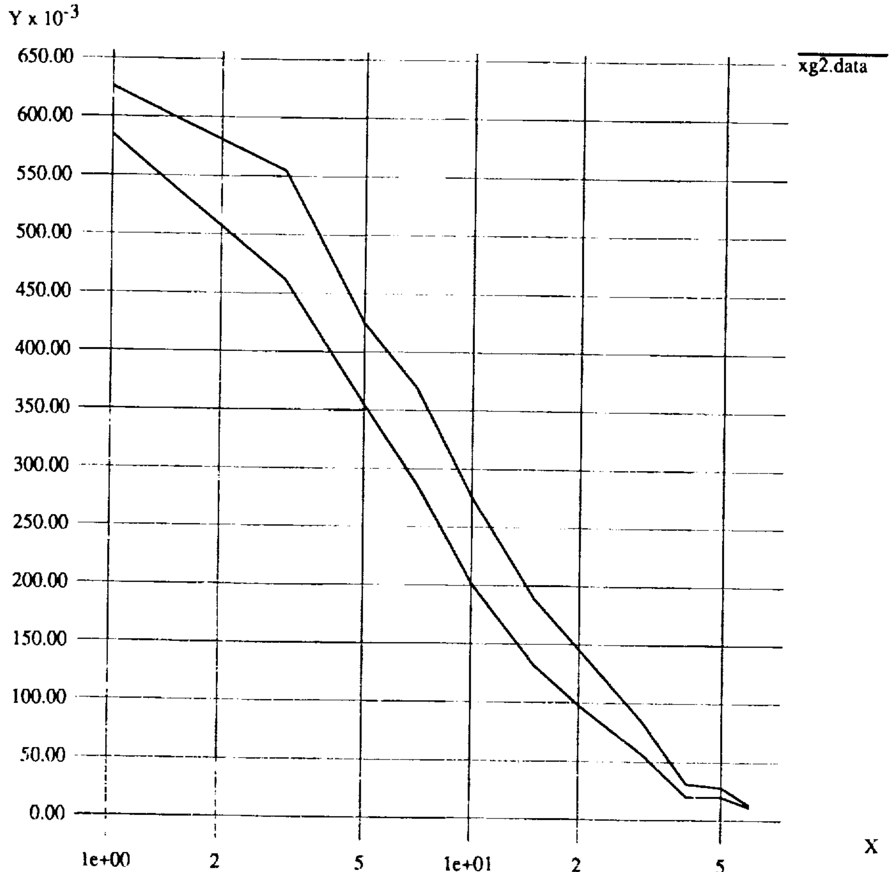

Figure 1: Comparison of Degree of Commitment in the Tileworld

by a comparison of the effectiveness of the most and least committed agents, shown in Figure 1. (When all fourteen levels of committment are plotted, there are some line crossing, but the trend relating effectiveness and degree of commitment is still clear; see [11].)

Table 1 summarizes the significance of the difference in performance between the most committed agent we ran and the least committed one. It shows that the difference between their performance, although not enormous, is statistically significant everywhere except at the endpoints. Further analysis reveals the reason for the collapse at the endpoints. In the slowest environment we studied, there was a great deal of variation in the agent's performance, because it was possible for the environment sometimes to evolve in a way that enabled the agent to succeed at all the tasks it was presented. Because of the high degree of variation in the scores, there was no statistical significance between the agents' performance in these slowly changing environments. At the other endpoint—the most quickly changing environment—the situation is different. In this environment there was very little variation in the scores: both agents scored very poorly, because they were unable to succeed at all but a few of the tasks they were presented. This bottoming effect resulted in a lack of significance between the agents' scores in this environment.

Figure 2 plots the difference in these two agents' performance. The graph shows that the value of commitment, while always positive, is a function of degree of dynamism in the environment. As dynamism increases, the marginal value of commitment first increases, then peaks, and subsequently drops off, although it does not become negative within the bounds of the experiment. This result can be explained as follows. In slower worlds, there are fewer options presented to the agent, and, hence, fewer opportunities for filtering to result in a savings in reasoning cost. Moreover, the advantages of reducing reasoning are minimal, since there is generally enough time to deal with

T-Test Results:
Significance of Difference between Mean Effectiveness
of Most and Least Committed Agents

Dynamism	t	Significance
1	1.121692	P < .15
3	3.129425	P < .0025
5	3.148238	P < .0025
7	4.130018	P < .0005
10	5.727610	P < .0005
15	6.686329	P < .0005
20	7.000076	P < .0005
30	4.909135	P < .0005
40	3.164884	P < .0005
50	2.260098	P < .02
60	0.709967	P < .25

Table 1: Analysis of Value of Commitment

options. As the world becomes more dynamic, there are more options for consideration, and the penalty for extra reasoning increases, because there is less time to respond to those options. This explains why filtering increasingly pays as dynamism increases. However, another influence comes into play as the rate of change in the environment increases: the missed-opportunity cost grows. As the world changes more rapidly, it becomes increasingly important for the agent to succeed at each individual task, since it will fail to complete a larger proportion of the potential tasks. The shape of the graph in Figure 2 is thus explained by the tension between the increased benefits of reduced reasoning and the increased penalties of missed opportunity, both of which vary directly with rate of change in the world. We expect to see a similar pattern of competing influences on the usefulness of filtering in other domains, and we will pay particular attention to the shape and peak of of the filtering-value curve in other domains, as it reveals useful information about the relative significance of reasoning overhead and missed-opportunity costs.

CONCLUSION

We provided a brief description of a set of experiments aimed at assessing the value of a strategy that may be incorporated in intelligent agents to help focus their reasoning in dynamic environments. The strategy, *filtering*, involves screening from consideration options for action that are incompatible

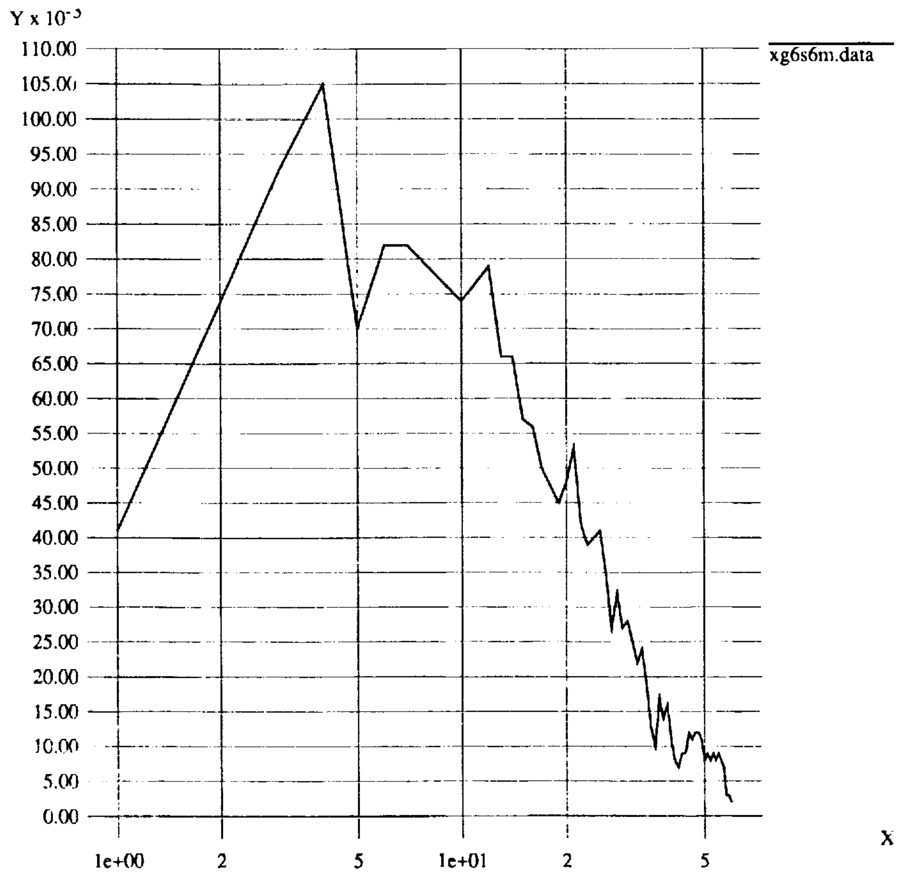

Figure 2: Difference in Effectiveness between Most and Least Committed Agents

with already established plans, except where those options are *prima facie* important enough to trigger a pre-defined override. We relied on a testbed system, the Tileworld, to conduct our experiments. We have made a number of enhancements to the Tileworld since the time it was originally developed, and we described some of the more important of those here. Our experiments demonstrate filtering is a feasible strategy, at least within the Tileworld, a result that suggests to us that it is worth investigating this strategy in more-complex systems. Additionally, our results showed an interesting relationship between the rate of change in the environment and the amount of benefit that one can derive from using a filtering strategy.

Acknowledgements

This work was has been supported been supported by the Air Force Office of Scientific Research (Contracts F49620-91-C-0005 and F49620-92-J-0422), by the Rome Laboratory (RL) of the Air Force Material Command and the Defense Advanced Research Projects Agency (Contract F30602-93-C-0038), and by an NSF Young Investigator's Award (IRI-9258392). Marc Ringuette, Michael Young, Petra Funk, Sigalit Ur, David Joslin, and Arthur Nunes contributed to the development of the Tileworld system and to the experiments described herein.

References

[1] Michael E. Bratman, David J. Israel, and Martha E. Pollack. Plans and resource-bounded practical reasoning. *Computational Intelligence*, 4:349–355, 1988. Also appears in J. Pollock

and R. Cummins, eds., *Philosophy and AI: Essays at the Interface*, MIT Press, Cambridge, MA.

[2] Edward H. Durfee and T. A. Montgomery. MICE: A flexible testbed for intelligent coordination experiments. In Lee Erman, editor, *Intelligent Real-Time Problem Solving: Workshop Report*, pages X1–X16, Palo Alto, CA, 1990. Cimflex Teknowledge Corp.

[3] Steve Hanks, Martha E. Pollack, and Paul R. Cohen. Benchmarks, testbeds, controlled experimentation, and the design of agent architectures. To appear AI Magazine, Winter 1993., 1993.

[4] David Joslin, Arthur Nunes, and Martha E. Pollack. Tileworld users' manual. Technical Report 93-12, University of Pittsburgh Department of Computer Science, 1993.

[5] David N. Kinny and Michael P. Georgeff. Commitment and effectiveness of situated agents. In *Proceedings of the Twelfth International Joint Conference on Artificial Intelligence*, pages 82–88, Sydney, Australia, 1991.

[6] T. J. Laffey, P. A. Cox, J.L. Schmidt, S.M. Kao, and J.Y. Read. Real-time knowledge-based systems. *Artificial Intelligence Magazine*, 9:27–45, 1988.

[7] Thomas A. Montgomery and Edward H. Durfee. Using MICE to study intelligent dynamic coordination. In *Second International Conference on Tools for Artificial Intelligence*, pages 438–444. IEEE, 1990.

[8] Andrew B. Philips and John L. Bresina. Nasa tileworld manual. Technical Report FIA-91-11, NASA Ames Research Center, Moffett Field, CA, 1991.

[9] Martha E. Pollack. Overloading intentions for efficient practical reasoning. *Noûs*, 25(4):513–536, 1991.

[10] Martha E. Pollack. The uses of plans. *Artificial Intelligence*, 57:43–68, 1992.

[11] Martha E. Pollack, David Joslin, Arthur Nunes, and Sigalit Ur. Experimental investigation of an agent commitment strategy. In preparation.

[12] Martha E. Pollack and Marc Ringuette. Introducing the Tileworld: Experimentally evaluating agent architectures. In *Proceedings of the Ninth National Conference on Artificial Intelligence*, pages 183–189, Boston, MA, 1990.

[13] Stuart J. Russell and Eric H. Wefald. *Do the Right Thing*. MIT Press, Cambridge, MA, 1991.

[14] Herbert A. Simon. A behavioral model of rational choice. *Quarterly Journal of Economics*, 69:99–118, 1955.