Intention Reconsideration in Complex Environments

Martijn Schut and **Michael Wooldridge**

Department of Computer Science
The University of Liverpool, Liverpool L69 7ZF,UK

{m.c.schut, m.j.wooldridge}@csc.liv.ac.uk

Abstract

One of the key problems in the design of belief-desire-intention (BDI) agents is that of finding an appropriate policy for intention reconsideration. In previous work, Kinny and Georgeff investigated the effectiveness of several such reconsideration policies, and demonstrated that in general, there is no one best approach – different environments demand different intention reconsideration strategies. In this paper, we further investigate the relationship between the effectiveness of an agent and its intention reconsideration policy in different environments. We empirically evaluate the performance of different reconsideration strategies in environments that are to varying degrees dynamic, inaccessible, and non-deterministic. In addition to our empirical results, we are able to give preliminary analytical results to explain some of our findings.

1 Introduction

Computation is a valuable resource for autonomous agents that are required to act in complex environments [13]. Such agents cannot reason indefinitely, either about which goals to achieve, or what actions to perform in furtherance of these goals [3]. Any implemented agent will operate under very real resource bounds – in terms of computation power, memory, and the time available to make decisions. It follows that the effective control of reasoning is a key factor in the success (or otherwise) of an agent system. Research on resource-bounded decision making and the control of reasoning originated in economics and the decision sciences [16, 8]; in AI, such research falls under the banner of meta-level reasoning [15]; and in the agent literature, it falls under work on bounded optimality [14].

In this paper, we examine the relationship between the properties of the environment in which an agent must operate, and the requirements for effective control of reasoning in that environment. Our chosen agent architecture for this study is the belief-desire-intention (BDI) model [3, 7]. In BDI agents, decision-making is composed of two main activities: *deliberation* (deciding *what intentions* to achieve) and *means-ends reasoning* (deciding *how* to achieve these intentions) [3]. Deliberation is a computationally costly process, and in order for a BDI agent to operate effectively, it should choose to deliberate only when necessary; this requires an appropriate *intention reconsideration* policy [3, 9, 18].

In this paper, we investigate the relationship between the success (or otherwise) of intention reconsideration policies in BDI agents, and the characteristics of environments that these agents inhabit. The intuition is that different intention reconsideration policies will be better suited to different environmental niches. Our work builds on, and considerably extends that of Kinny and Georgeff, who studied the performance of different intention reconsideration policies in environments with varying degrees of *dynamism* [9]. In our work, we investigate the performance of intention reconsideration policies in environments where we vary the following parameters (cf. [13, p46]):

- *dynamism* – the rate of change of the environment, independent of the activities of the agent;
- *accessibility* – the extent to which an agent has access to the state of the environment;
- *determinism* – degree of predictability of the system behaviour for identical system inputs.

The remainder of this paper is organised as follows: section 1.1 provides some background to our experiments; section 2 lays out the methodology of the experiments; in section 3 we present and analyse our experimental results; and section 4 concludes by placing the investigation into context and discussing further work.

1.1 Background

Research in the design of autonomous agents throughout the 1970s and early 1980s was dominated by STRIPS-style classical planning approaches [1]. These approaches focussed on algorithms for automatic plan generation, that would take as input a specification of the current world state, a goal to be achieved, and the actions available to an agent, and would produce as output a plan to achieve the goal state. This style of planning, it was believed, is a central component in rational action. By the mid 1980s, a number of researchers, (of whom Rodney Brooks is probably the best known [4]), began to claim that such approaches were fundamentally flawed, for both pragmatic and philosophical reasons. From a pragmatic point of view, STRIPS-style planning algorithms tend to be computationally intractable, rendering them of limited value to agents that must operate in anything like real-time environments [6, 5]. From a philosophical point of view, it was argued that much of what we regard as everyday intelligence does not arise from abstract deliberation of the kind involved in STRIPS-style planning, but from the interaction between comparatively simple agent behaviours and the agent's environment.

The challenge posed by behaviour-based AI research has arguably led to some fundamental changes in the agenda of the AI community. First, it has become widely accepted that intelligent behaviour in an autonomous agent is more closely coupled to the environment occupied by the agent than was perhaps hitherto acknowledged. As a consequence, there has been renewed interest in the use of more realistic environmental settings for the evaluation of agent control architectures. Second, it has become accepted that while reasoning is an important resource for intelligent decision-making, it is not the only such resource. As a consequence, there has been much interest in *hybrid* approaches to agent design, which attempt to combine reasoning and behavioural decision-making [17, 11].

One popular approach to the design of autonomous agents that emerged in the late 1980s is the *belief-desire-intention* (BDI) model [3, 7]. The BDI model gets its name from the fact that it recognises the primacy of beliefs, desires, and intentions in rational action. Intuitively, an agent's *beliefs* correspond to information the agent has about the world. These beliefs may be incomplete or incorrect. An agent's *desires* are states of affairs that the agent would, in an ideal world, wish to bring about. Finally, an agent's *intentions* represent desires that it has *committed* to achieving. The idea is that an agent will not be able to deliberate indefinitely over which states of affairs to bring about; ultimately, it must fix upon some subset of its desires and commit to achieving them. These chosen desires are *intentions*.

A key problem in the design of BDI agents is that of *intention reconsideration* [3, 9, 18]. This problem arises when we consider that an agent should not, in general, maintain an intention indefinitely – either the intention should be achieved or it should be dropped. This implies that, from time-to-time, agents should pause to deliberate over their intentions and *reconsider* them. But reconsideration is itself a computationally costly process. As a rule of thumb, therefore, an agent should only reconsider intentions when such reconsideration would lead to a change in intentions – otherwise the effort invested in reconsideration is wasted.

Developing an appropriate *intention reconsideration* policy – which keeps an agent committed to its intentions just as long as it would be rational to do so – is thus a critical issue in the design of any BDI agent. In a series of experiments, Kinny and Georgeff [9] investigated the relative performance of intention reconsideration strategies for BDI agents in different environmental settings. The experimental framework they used involved a PRS BDI system [7] that was situated in Pollack and Ringuette's TILEWORLD domain [12].

In essence, the TILEWORLD is a grid environment on which there are agents, tiles, obstacles, and holes. An agent can move up, down, left, or right, and can move tiles towards holes. An obstacle is a group of immovable grid cells. Holes have to be filled up with tiles by the agent. An agent scores points by filling holes with tiles, with the aim being to score as many points as possible. The TILEWORLD is inherently *dynamic*: starting in some randomly generated world state, based on parameters set by the experimenter, it changes over time in discrete steps, with the appearance and disappearance of holes. The experimenter can set a number of TILEWORLD parameters, including: the frequency of appearance and disappearance of tiles, obstacles, and holes; the shape of distributions of scores associated with holes; and the choice between hard bounds (instantaneous) or soft bounds (slow decrease in value) for the disappearance of holes. In the TILEWORLD, holes appear randomly and exist for as long as their *life-expectancy*, unless they disappear because of the agent's actions. The interval between the appearance of successive holes is called the *hole gestation time*.

The aims of Kinny and Georgeff's investigation were to "(1) assess the feasibility of experimentally measuring agent effectiveness in a simulated environment, (2) investigate how commitment to goals contributes to effective agent behaviour and (3) compare the properties of different strategies for reacting to change" [9, p82]. The full TILEWORLD domain was considered too complex for the experiment, and the testbed was therefore simplified in several ways. First, tiles were omitted: an agent scores points simply by moving to holes. In addition, the agent was assumed to have perfect, zero-cost knowledge of the state of the world. Finally, it was assumed that agents only form correct and complete plans, and only generate plans for visiting a single hole (rather than planning multiple-hole tours).

In Kinny and Georgeff's experiments, two different types of reconsideration strategy were used: *bold* agents, which never pause to reconsider their intentions before their current plan is fully executed, and *cautious* agents, which stop to reconsider after the execution of every action. These characteristics are defined by a *degree of boldness*, which specifies the maximum number of plan steps the agent executes before reconsidering its intentions. Dynamism in the environment is represented by the *rate of world change* and is manipulated by changing the ratio of the clock rates of the TILEWORLD and the agent. The effectiveness of the agent is represented by its score (the sum of values of holes filled) divided by the maximum score it could in principle have achieved (the sum of the scores of all holes appearing in the TILEWORLD during a trial). The results of the experiments show that a cautious agent outperforms a bold agent in highly dynamic environments; intuitively, because in dynamic environments, which change frequently, it pays to reconsider intentions frequently.

In Kinny and Georgeff's investigation, as mentioned previously, the agent has perfect zero-cost knowledge of the world. In later work by Kinny, Georgeff, and Hendler [10] a *sensing cost* was introduced, that represents the time cost of processing sensor information. The aim of this work was to show that an optimal sensing rate exists, depending on the degree of world dynamism and the sensing cost. A model was presented that captures the trade-off between time saved by early detection of change and time wasted by too frequent sensing. Applying a cost to sensing is different from varying the accessibility of the world. Varying accessibility essentially means varying the amount of information accessible to the agent, which implicates that it does not matter how much the agent attempts to obtain information. If a cost is applied to sensing, the information is available, but for a higher price.

Note that intention reconsideration is a kind of *meta-level control* [15, 18]. Other researchers have investigated meta-level control issues in autonomous agents. For example, Boddy and Dean [2] developed anytime algorithms for optimal scheduling of reasoning in dynamic environments.

2 Experimental Methodology

The aim of our work is to experimentally investigate the performance of a range of intention reconsideration policies in environments with different properties. To do this, we make use of a simulation of a single agent inhabiting an adapted TILEWORLD environment [12] – see the preceding section for a more detailed description of the TILEWORLD. The task of the agent is to visit holes in order to gain as many points as possible. The agent decides which hole to visit based on hole distances – it always chooses to visit the nearest hole. We adopted the same simplifications to the original TILEWORLD as Kinny and Georgeff (as discussed in the previous section) and adapted the system in two further ways: (i) we omitted obstacles from the TILEWORLD; and (ii) we allowed the agent to move diagonally over the grid (in addition to moving horizontally and vertically). Omitting obstacles simplifies the problem domain without trivializing it; allowing diagonal movement is an obvious extension.

Following Kinny and Georgeff [9], we define the *effectiveness* ϵ of an agent as the ratio of the actual score achieved by the agent to the score that could in principle have been achieved. This measurement is thus independent of randomly distributed parameters in a trial. It also avoids problems such as machine-dependency and prevention of repetition of experiments on different machines, which would occur if the effectiveness of an agent was based on such measures as CPU-time or elapsed time [12].

There are three main environmental attributes that we vary in our experiments:

- *Dynamism* (an integer in the range 1 to 80 denoted by γ) represents the ratio between the world clock rate and the agent clock rate [9]. If $\gamma = 1$, then the world executes one cycle for every cycle executed by the agent. Larger values of γ mean that the environment is executing more cycles for every agent cycle; if $\gamma > 1$ then the information the agent has about its environment may not necessarily be up to date.

- *Accessibility* (a real value in the range 0 to 1 denoted by α) represents the proportion of the environment that is visible to the agent. If $\alpha = 1$, then the agent can see the entire TILEWORLD, and thus has complete, perfect information about its environment; if $\alpha = 0$, then the agent can see nothing of its environment but the grid point it currently occupies.

- *Determinism* (an integer in the range 0 to 100 denoted by δ) represents how certain it is that an action has the expected outcome. The idea is that an agent performs actions in order to bring about certain states of affairs. However, in most realistic environments, actions are non-deterministic, in that they can have a number of possible outcomes. Thus δ represents the probability that an action will have its intended outcome, expressed as a percentage. If $\delta = 100$, then the agent can be certain that every action it performs will have the desired effect; as $\delta \to 0$ the probability that an action will have an undesirable outcome increases. In our scenario, actions are movements that can be made by an agent, either north, south, east, west, or diagonally. We model non-determinism by allowing actions to move the agent in an unintended direction – for example, in attempting to move north, the agent may actually end up moving east. This represents the situation in mobile robotics, where a robot attempting to move in some direction can never be sure that it will succeed in moving in that direction.

Parameter	Value/Range
world dimension	20
hole score	10
hole life-expectancy	[240,960]
hole gestation time	[60,240]
dynamism (γ)	(1,80)
accessibility (α)	(0,1)
determinism (δ)	(0,100)
number of time-steps	15,000
number of trials	50
replanning rate	0 or ∞
planning cost (p)	0, 1, 2, or 4

Table 1: Overview of the experiment parameters

The experiments we conducted are divided into two series: the *single parameter variation* series, in which we varied one parameter per experiment; and the *combined parameters variation* series, in which we systematically varied two parameters per experiment. In the single parameter variation we respectively minimized or maximized the parameters other than the one varied: in the dynamism experiment, we maximized accessibility ($\alpha = 1$) and maximized determinism ($\delta = 100$); in the accessibility experiment, we minimized dynamism ($\gamma = 1$) and maximized determinism ($\delta = 100$); finally, in the determinism experiment, we minimized dynamism ($\gamma = 1$) and maximized accessibility ($\alpha = 1$).

With respect to agent properties, we varied the *replanning rate* and the *planning cost*. The replanning rate represents the *boldness* of the agent. We set it for each experimental condition to 1 (the agent replans every time before performing an action – a cautious agent) and ∞ (the agent never replans while executing a plan – a bold agent). The planning cost represents the time cost of planning: the number of time-steps required to form a plan. We set it for each experimental condition to 0, 1, 2, and 4. In what follows, we denote planning cost by p. In table 1 we give an overview of the values of relevant parameters that we used in the experiments ($[x, y]$ denotes a uniform distribution from x to y and (x, y) denotes the range from x to y). Note that each TILEWORLD was run for 15,000 time steps, and each run was repeated 50 times, in order to eliminate experimental error.

3 Results and Analysis

In this section, we present the results of our experiments. The experiments with single parameter variation resulted in the graphs shown in figure 1. The experiments with combined parameter variation resulted in the graphs shown in figures 4, 5, and 6. The graphs for the combined parameter series generalise those of the single parameter series, and so in principle it would suffice to give the graphs of the combined parameter series only. However, in the interest of clarity, we included graphs for both series[1]. We refer to a plot of effectiveness ϵ as in figure 1 as an *effectiveness curve* and to a plot of ϵ as in figures 4, 5, and 6 as an *effectiveness surface*.

[1]To save space, we omitted the graphs from the combined parameter variation series for planning cost equal to 1 and 2, as although we conducted these experiments, the results were consistent with those of the single parameter variation experiments.

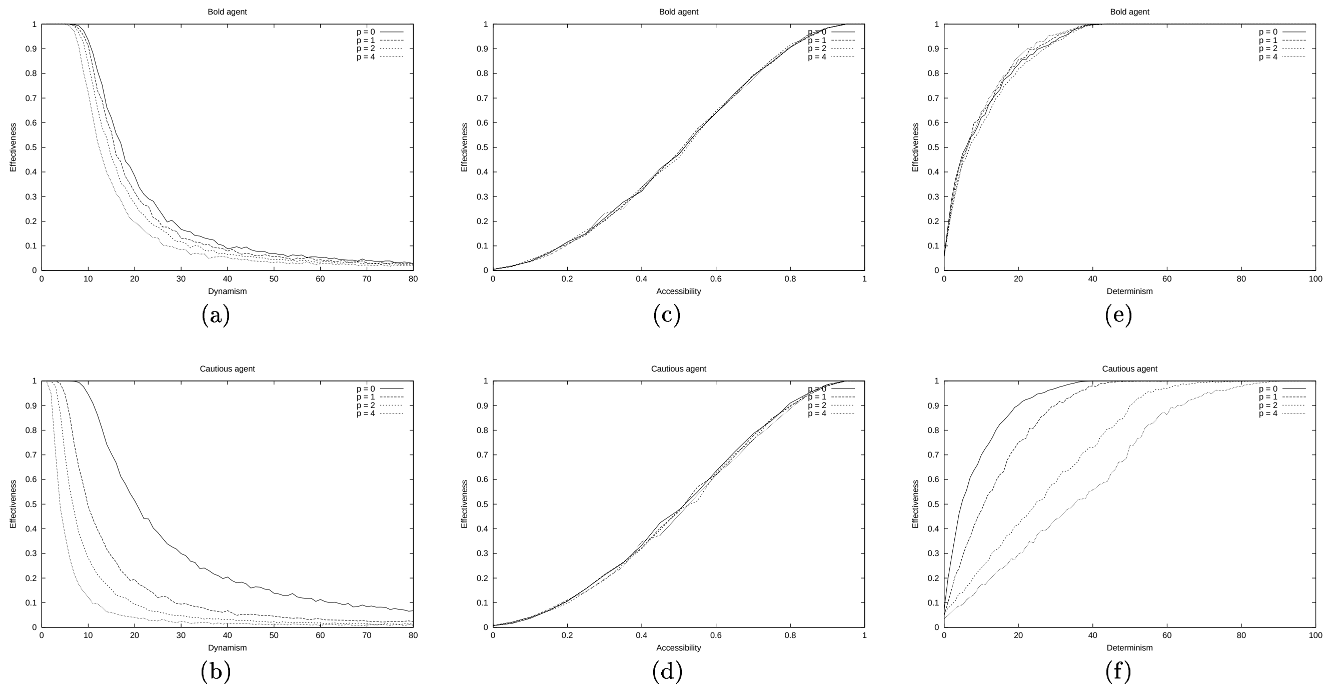

Figure 1: Experimental results (single parameter variation).

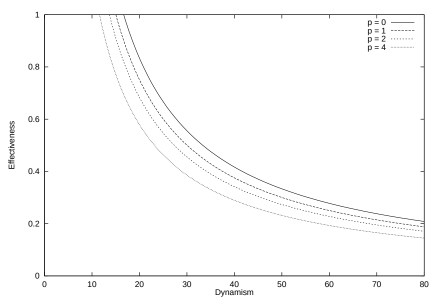

Figure 2: Theoretical effectiveness for a bold agent when dynamism is varied.

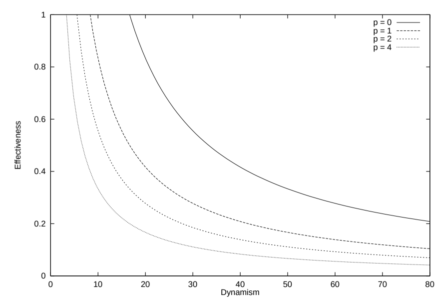

Figure 3: Theoretical effectiveness for a cautious agent when dynamism is varied.

3.1 Single parameter variation

Dynamism

From the results of the dynamism experiment, as plotted in figures 1.a and 1.b, we observe that the shapes of the effectiveness curves are similar, but not the curves themselves. We can explain the shape of the effectiveness curves and the differences between the curves as follows. If the dynamism of the world is at a minimum ($\gamma = 1$), then holes appear and disappear sufficiently slowly that the agent can visit each hole before it disappears, which results in a perfect score ($\epsilon = 1$) of the agent. As γ increases, then at some point, holes start to disappear before the agent has visited them, and ϵ starts to drop below 1. The effectiveness curve first declines steeply, later more gradually and eventually asymptotically approaches zero.

Some observations on the differences in the curves can be made directly. First, it is clear that varying the cost of planning has much more influence on the effectiveness of a cautious agent than on the effectiveness of a bold agent. Second, if planning is free ($p = 0$), then a cautious agent performs better than a bold agent if $\gamma > 7$. Third, if $p = 1$, then a cautious agent performs worse than a bold agent, independent of the dynamism of the world.

In an attempt to explain the shape of the graph in figure 1.a, we used brute force computation to calculate the mean distance an agent has to travel to any hole in our TILEWORLD – as it turns out, the mean distance to any hole in our experiments is approximately 9. As previously stated, the effectiveness of an agent is the ratio of its actual score to the maximum score. This can be denoted

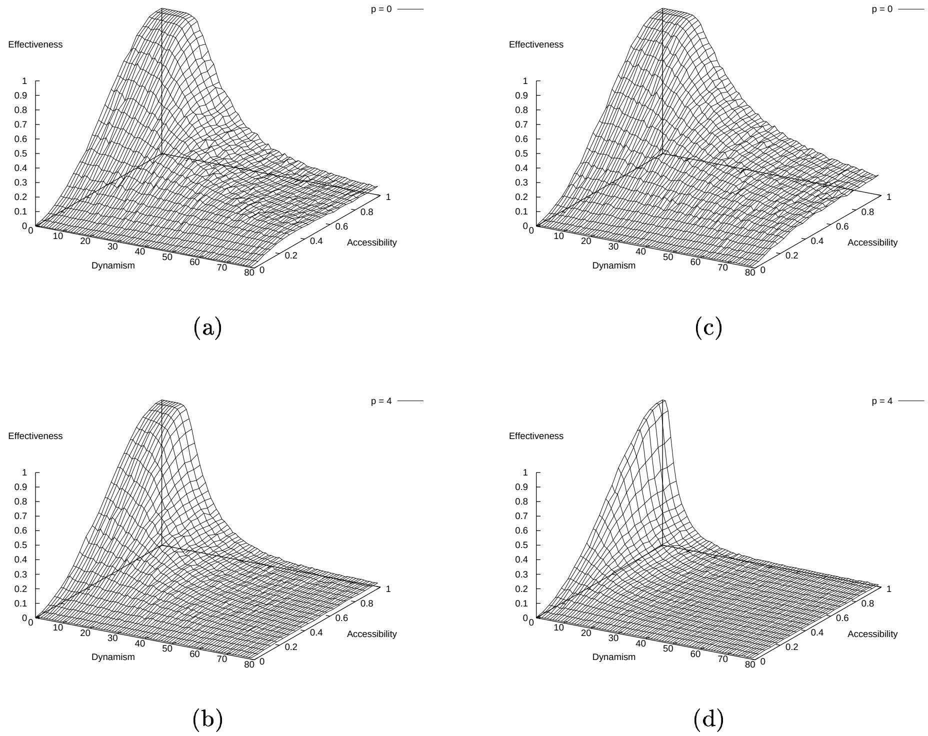

Figure 4: Dynamism and Accessibility – the results for a bold agent are in (a) and (b), for a cautious agent in (c) and (d).

by $\epsilon = score_{agent}/score_{max}$. We can easily calculate the maximum score, namely $score_{max} = T/g$, where T denotes the number of time-steps and g denotes the hole gestation time. The agent's actual score can be calculated by $score_{agent} = T/f$, where f denotes the total time the agent takes to fill a hole. Similar to Kinny and Georgeff, we define f to be given by $f = d \times (p/k + m)$, where d is the hole distance, p is the planning cost, k is the reconsideration frequency, and m is the time to move a single step (here always 1). If we set $k = d$, we have a bold agent, and when we set $k = 1$, we have a cautious agent. Now we can define the effectiveness of the agent as $\epsilon = g/(\gamma \times f)$. The curves in figures 1.a and 1.b can be approximated by this function[2], using the values from table 1 and a mean hole distance of 9. This approximation is shown in figure 2 and 3 for a bold agent and cautious agent, respectively.

Accessibility

The shape of the effectiveness curves in the graphs 1.c and 1.d can be explained from the way we implemented the accessibility of the agent. If the accessibility is minimal ($\alpha = 0$), the agent can only see the point where it is currently located. With the exception of a hole appearing coincidentally on that location, the agent cannot score any points, and its effectiveness is minimal ($\epsilon = 0$). If the accessibility is maximal ($\alpha = 1$), the agent can see all points in the world, and has sufficient time to reach holes before they disappear, in which case its effectiveness is perfect ($\epsilon = 1$). If $\alpha < 0.5$, then the curve is concave; if $\alpha > 0.5$, the curve is convex. This value can be explained from the fact that if $\alpha > 0.5$ and the agent is on an optimal location, (the middle of the grid), then it can see all the points in the world.

From figures 1.c and 1.d it appears that there is no great difference between the results for the bold agent if planning cost is varied and between the curves for the cautious agent if planning cost is varied. Neither is there much difference between the curves for the bold agent and the curves for the cautious agent. A variance analysis on the experimental data confirms that the differences between the curves, within the bold and cautious agent effectiveness curves as well as between them, are not significant. An explanation for this might be that when accessibility is varied, the amount of deliberation an agent engages in does not influence the effectiveness of the agent. Intuitively, there is not enough information for the agent to deliberate over in order to increase its effectiveness.

Note that in addition to giving agents "limited vision", we conducted a series of experiments in which we simulated agents with *noisy sensors*. The idea was that there would be a probability η that any given piece of information (percept) received by the agent was incorrect. If $\eta = 0$, then the agent's sensors would be perfect: all information available to the agent would be correct. If $\eta = 1$, then every piece of information available to the agent would be incorrect. We systematically varied the value of η from 0 to 1, and investigated the performance of bold and cautious agents for each, with different planning costs. These experiments yielded a linear relationship between effectiveness and η.

The shape of the graphs in figures 1.c and 1.d can easily

[2]The function however assumes that the life-expectancy of holes is infinite.

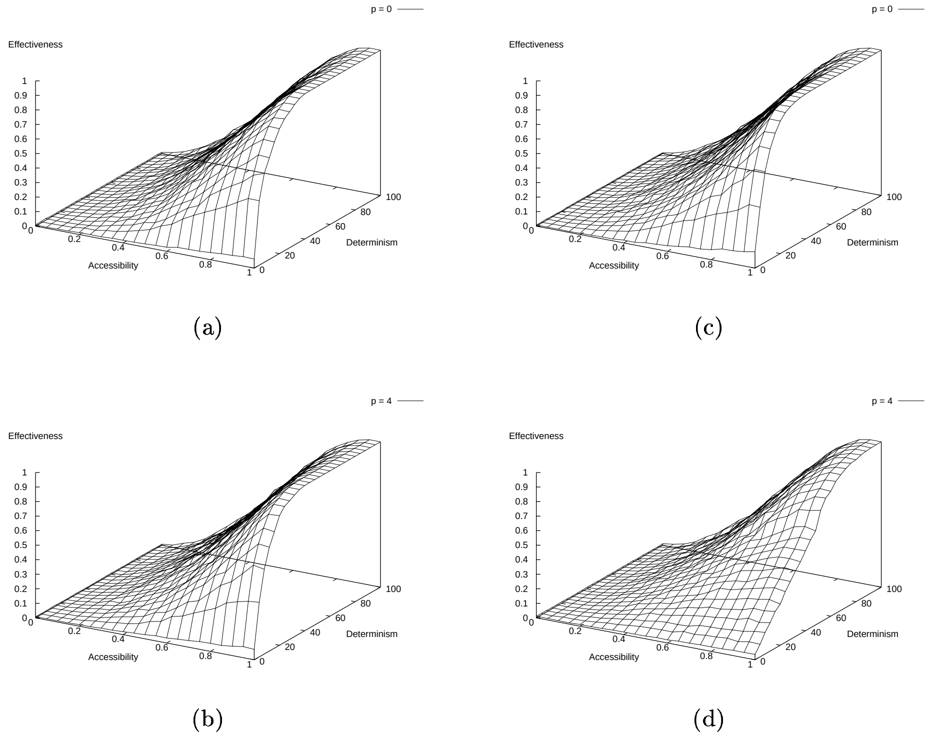

Figure 5: Accessibility and Determinism – the results for a bold agent are in (a) and (b), for a cautious agent in (c) and (d).

be put on a theoretical footing. Because the world changes slow enough for the agent to reach a hole when observed ($\gamma = 1$), the agent's effectiveness corresponds with its *visibility* – the number of grid points the agent can see around itself. Calculating this visibility by brute force computation resulted in a curve identical to the effectiveness curve as in 1.c or 1.d. Using a curve-fitting method, this visibility curve can be approximated by a biquadratic function[3].

Determinism

The effectiveness curves for the determinism experiment are plotted in figures 1.e and 1.f. If the determinism of the world is minimal ($\delta = 0$), the outcomes of the agent's actions are never as intended by the agent. But because the agent can still encounter a hole by accident, it achieves a higher score than minimal ($\epsilon > 0$). If determinism is maximal ($\delta = 1$), the outcomes of the agent's actions are always the outcomes as intended by the agent, and the agent achieves a perfect score ($\epsilon = 1$). The reason for this is that determinism is defined as the chance that the outcome of an agent's action is the outcome intended by the agent. If $\delta = 0$, the agent never arrives at the location it intends. If $\delta = 1$, the agent always arrives at the intended location. As δ increases, the agent slowly starts to arrive at the intended holes and thus increases its score. The curve inclines slowly at first and later steeper, until $\delta > 40$, from where the effectiveness stays approximately perfect ($\epsilon \simeq 1$). We speculate that the agent can achieve a perfect score when $\delta > 40$ for the following reason. If δ exceeds a certain threshold (here: $\delta > 40$), the agent can compensate for failed actions by replanning. As long as the intended hole does not disappear, the agent can replan and in the end will reach the hole. This means an increase in deliberation, but a justified one, because it increases the effectiveness of the agent considerably.

When one considers the effectiveness curves for a bold agent, it is clear there is not much difference between them. As the planning cost p is increased, ϵ decreases. This decline is slight because the agent must replan completely after executing a plan, rather than because the agent does not need to reconsider its plans. This is also the reason why, with the exception of when planning is free ($p = 0$), a bold agent performs better than a cautious agent. A cautious agent has to replan after every step, whereas a bold agent does not do this and therefore a bold agent can perform more effectively. However, when planning is free, the cautious agent outperforms the bold agent, because it does not need to execute its complete plan before replanning. In this case, a cautious agent's plans are more flexible and thus shorter. With reference to figure 1.f, it is immediately obvious that planning cost has a significant impact on effectiveness for cautious agents in non-deterministic environments.

Before we leave this section, we note that the effectiveness of the agent depends on other characteristics of the environment, such as the life-expectancy of holes. If the life-expectancy of a hole is too short, then the agent cannot reach the hole by planning again. In this case, δ must be

[3] For example, for a 5×5 world the agent's visibility can be described by $(-a^2 + 9a + 5)^2/d^2$, where a denotes accessibility of the world before normalization ($a = \alpha \times d$) and d denotes world dimension. The constant values in this function depend on the world dimension d.

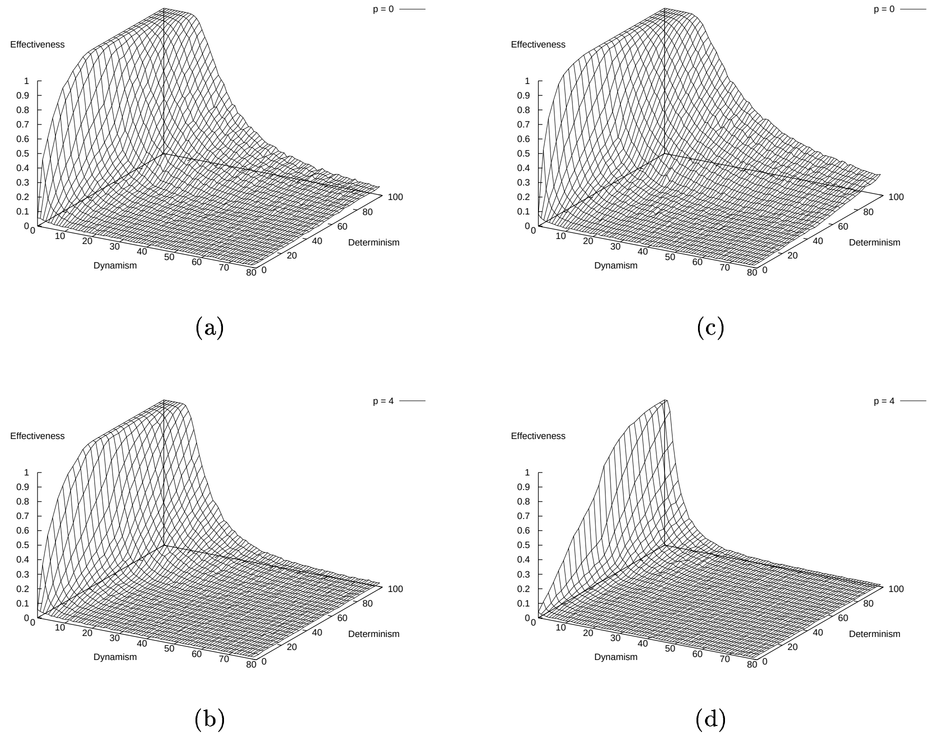

Figure 6: Dynamism and Determinism – the results for a bold agent are in (a) and (b), for a cautious agent in (c) and (d).

very high in order for the agent to score any points. On the other hand, if holes never disappear, the agent would achieve a perfect score, even when δ is very low.

3.2 Combined parameter variation

The experimental results from the combined parameter variation of dynamism and accessibility are shown in figure 4; of accessibility and determinism in figure 5; and of dynamism and determinism in figure 6. All effectiveness surfaces are consistent with the effectiveness curves individually. In for example the variation of dynamism and accessibility, if dynamism is minimal ($\gamma = 1$), the curve corresponds to the individual effectiveness curve for accessibility and if accessibility is maximal ($\alpha = 1$), the curve corresponds to the individual effectiveness curve for dynamism. For these values, the analysis is thus similar to the analysis for the single parameter variations.

With the combined parameter variation experiments we want to show which parameters dominate in complex environments. It is clear from the effectiveness surfaces in figure 4 that dynamism has more influence on the effectiveness of the agent than accessibility. This follows from the fact that the surfaces change more rapidly over the dynamism axis than over the accessibility axis. From maximal effectiveness ($\epsilon = 1$), where dynamism is minimal ($\gamma = 1$) and accessibility is maximal ($\alpha = 1$), the decline in effectiveness is much steeper when dynamism increases than when accessibility decreases.

It is clear from figure 5 that accessibility has more influence on the effectiveness of the agent than determinism. Even in a worst case scenario – a cautious agent where planning cost is 4 – the decrease in effectiveness from maximal effectiveness ($\epsilon = 1$), where accessibility is maximal ($\alpha = 1$) and determinism is maximal ($\delta = 100$), is steeper when accessibility decreases than when determinism decreases. In other cases, effectiveness stays maximal until determinism is approximately 40 ($\delta = 40$). We explained the reason for this in section 3.1: the agent can compensate for non-determinism in the environment by replanning.

Figure 6 shows that the agent's effectiveness changes faster over the dynamism axis than over the determinism axis, from which we conclude that dynamism has more influence on the effectiveness of the agent than determinism.

4 Discussion

In this paper, we examined the effectiveness of bold and cautious intention reconsideration strategies, for a range of planning costs, in environments defined by varying degrees of dynamism, accessibility, and determinism. From our single parameter variation experiment we can derive the following conclusions:

- All tested characteristics influence the effectiveness of the agent.
- We obtained the same results for dynamic environments as Kinny and Georgeff [9]. The intention reconsideration policy (bold or cautious) the agent adopts has a significant impact on the effectiveness of the agent. In a more static environment, a bold agent

performs best, and in a more dynamic environment, a cautious agent performs best. Our experiment also shows that this observation heavily depends on the cost of planning. If planning is expensive, too much deliberation results in sub-optimal performance.

- The accessibility of the world has a severe impact on the effectiveness of an agent. However, when the world is static and deterministic, accessibility does not influence the effectiveness of the intention reconsideration strategy. Our explanation for this is that, if α is low, there is not enough information for the agent to deliberate over in order to increase its effectiveness. Hence, additional deliberation in such circumstances does not pay off, or, in other words, deliberation is never useful when the environment is not accessible.
- Reconsidering intentions in a highly non-deterministic world only pays off if planning is free. The reason for this is that an agent does not need to reconsider its plans, but just needs to plan more often in order to increase its effectiveness.

We can derive the following conclusions from the experiments with combined parameter variation:

- All results on combined parameter variation are consistent with the results on single parameter variation. This enables us to compare the results of the two series with each other: the conclusions for the single parameter variation series are valid for the combined parameter variation too.
- The effectiveness of agents is minimal in environments that are both dynamic and inaccessible or inaccessible and non-deterministic. In environments that are both dynamic and non-deterministic, an agent can achieve effectiveness only by chance.
- Finally, a more important observation we made concerned which environmental characteristic influenced the agent's effectiveness the most: the dynamism of the environment. Determinism has the least influence on the effectiveness. This might be because the agent can compensate for non-determinism in the environment by replanning instead of reconsidering its plans.

The experiments we conducted can be extended in a number of obvious ways. First, we can extend the system we used by implementing *domain-dependent* decision strategies, as in [9]. Second, the accessibility characteristic of the environment can be implemented differently. We implemented it as how much of the environment the agent can see. But we can also interpret a *space-bounded* variation of accessibility, as how much information the agent can hold about its environment. Third, we can change the way an agent reacts to an non-deterministic environment. The testbed offers a possibility to implement an investigation into decision strategies based solely on the agent's utility function and probability function in a very clean way. The agent only needs to have a utility function defined over its surrounding points and the determinism of the environment delivers the probability function. (The derivation of such utility functions in a non-trivial way is, of course, itself a matter for research.) Finally, we can attempt to derive analytical results to explain the experimental results we obtained.

References

[1] J. F. Allen, J. Hendler, and A. Tate, editors. *Readings in Planning.* Morgan Kaufmann Publishers: San Mateo, CA, 1990.

[2] M. Boddy and T. L Dean. Decision-theoretic deliberation scheduling for problem solving in time-constrained environments. *Artificial Intelligence*, 67(2):245–285, 1994.

[3] M. E. Bratman, D. J. Israel, and M. E. Pollack. Plans and resource-bounded practical reasoning. *Computational Intelligence*, 4:349–355, 1988.

[4] R. A. Brooks. *Cambrian Intelligence.* The MIT Press: Cambridge, MA, 1999.

[5] T. Bylander. The computational complexity of propositional STRIPS planning. *Artificial Intelligence*, 69(1-2):165–204, 1994.

[6] D. Chapman. Planning for conjunctive goals. *Artificial Intelligence*, 32:333–378, 1987.

[7] M. P. Georgeff and A. L. Lansky. Reactive reasoning and planning. In *Proceedings of the Sixth National Conference on Artificial Intelligence (AAAI-87)*, pages 677–682, Seattle, WA, 1987.

[8] I. J. Good. Twenty-seven principles of rationality. In V. P. Godambe and D. A. Sprott, editors, *Foundations of statistical inference*, pages 108–141. Holt Rinehart Wilson, Toronto, 1971.

[9] D. Kinny and M. Georgeff. Commitment and effectiveness of situated agents. In *Proceedings of the Twelfth International Joint Conference on Artificial Intelligence (IJCAI-91)*, pages 82–88, Sydney, Australia, 1991.

[10] D. Kinny, M. Georgeff, and Hendler J. Experiments in optimal sensing for situated agents. In *Proceedings of the Second Pacific Rim International Conference on AI (PRICAI-92)*, 1992.

[11] J. P. Mueller. *The Design of Intelligent Agents (LNAI Volume 1177).* Springer-Verlag: Berlin, Germany, 1997.

[12] M. E. Pollack and M. Ringuette. Introducing the Tileworld: Experimentally evaluating agent architectures. In *Proceedings of the Eighth National Conference on Artificial Intelligence (AAAI-90)*, pages 183–189, Boston, MA, 1990.

[13] S. Russell and P. Norvig. *Artificial Intelligence: A Modern Approach.* Prentice-Hall, 1995.

[14] S. Russell and D. Subramanian. Provably bounded-optimal agents. *Journal of AI Research*, 2:575–609, 1995.

[15] S. J. Russell and E. Wefald. *Do the Right Thing — Studies in Limited Rationality.* The MIT Press: Cambridge, MA, 1991.

[16] H. A. Simon. Rational choice and the structure of the environment. In *Models of bounded rationality, Volume 2*, pages 259–268. The MIT Press: Cambridge, MA, 1957.

[17] M. Wooldridge and N. R. Jennings. Intelligent agents: Theory and practice. *The Knowledge Engineering Review*, 10(2):115–152, 1995.

[18] M. Wooldridge and S. D. Parsons. Intention reconsideration reconsidered. In J. P. Müller, M. P. Singh, and A. S. Rao, editors, *Intelligent Agents V (LNAI Volume 1555)*, pages 63–80. Springer-Verlag: Berlin, Germany, 1999.