# An Evolutionary Approach to Balancing Deliberation and Reactiveness in a Multi-Agent Scenario

2 E − 3

**Joachim Baczewski**
Department of Administration Engineering
Faculty of Science and Technology
Keio University

**Jun Tani**
Sony Computer Science Laboratory Inc.

## 1 Introduction

Regarding planning behavior in an ecosystem a trade-off between cost and time exists. If an agent has to plan he will tend to minimize the costs of the plan execution which will lead him to longer deliberation in order to find a suitable low cost plan. On the other hand the agent is situated in a competitive environment which gives him the highest guarantee of plan execution if his planning behavior is strongly reactive. Thus, a dilemma-like situation emerges which is due to the *bounded rationality* [1] which characterizes living beings as well as intelligent agents in complex environments: optimizing its outcome with its limited abilities.

Evolution can be a source for diversity and complexity, but also for stability in a dynamical system such as a system of multiple, goal-driven agents which have to do planning in an ever changing dynamic environment in order to fulfill their goals. We assume such agents to be selfishly motivated in the first place and to compete for resources. Executing a plan is connected with costs, e.g. energy consumed when moving a robot. Fulfilling a goal gives a payoff to an agent, e.g. reaching a loading-station for a robot. Regarding agents as dissipative systems makes it even costly to take no actions. Agents which reason deliberately can reduce arising costs by choosing a carefully selected low-cost plan. On the other hand, such committing agents might be confronted with a changed environmental situation when executing their plan after time-consuming deliberation in a dynamic environment. This might result in failure of execution. Reactive agents will not spend much time on reasoning but will execute the fastest plan, which is normally not the cheapest, in order to fulfill their goal. Because reactive agents do not spend much or no time on planning they are likely to fulfill most of their goals as long as these goals are within their cost range.

In this paper we model this tradeoff problem and will show some simulation results.

## 2 The Model

A number of agents is randomly placed in a 2-dimensional limited discrete world. Each agent has a limited sensing ability in all directions. Objects are placed in this world. It is each agent's goal to reach food. This is done by *mentally* moving within his sensing range. When food is found by this mental searching process, the agent moves *physically* to the food, the food disappears and the agent gets a payoff in form of energy increase. New food is created in fixed temporal steps which models natural growth processes.

In order to model the attraction towards the food and the avoidance of objects in the world, we introduce *forces* which are influencing each agent. These forces are determined in the following way in a discrete world where the agent can move in four directions. *Attracting forces* from food-points on an agent which are just mentally attracting focus his interest on food. The closer the food is, the more it attracts the agent. We introduce a limit distance beyond which the food cannot be sensed by the agent. In analogy, *repulsion forces* are determined as real physical force on the agent. For example, the closer a robot moves to a wall the stronger is the repulsion, the more difficult it gets for him to navigate. The total forces on an agent are listed in Equation 1. The factor $k$ is our *evolutionary determined biasing factor* which makes the agent more deliberative for small $k$ and more reactive for high $k$. The factor $k$ is determined by mutation when creating offspring.

$$\vec{F}_{agent} = \sum_j \vec{F}_{repulsion_j} - k \sum_i \vec{F}_{attraction_i} \qquad (1)$$

Because we are using a discrete space model it is difficult to deal with force vectors directly. Thus, we integrate the forces into a *potential field*. On this potential field the agent's decisions are based. The bias factor $k$ determines the steepness of the attracting food gradient. If an agent moves the costs for him are the integration of the *absolute value* of the potential over the way he moves. The absolute value of the potential represents a environmental difficulty. We refer to this cost function as energy-consumption function. The initial fitness of an agent increases by

- consuming food

and decreases

- by moving,
- as a linear function of time,
- when creating offspring.

The *mental focus-point* of an agent searches within the sensing range of the agent. The search is determined according to the potential of the adjacent fields of the position of the mental search process. The search process moves always to lower potentials. The search is a step by step process with the emergence of backtracking if an agent gets stuck in local minima [2]. Loops in agent plans are cut. After finding food the *agent* moves physically to the food position. The time required for planning steps is longer then the moving time. If two agents have the same food as goal, the faster agent gets the food and the other agent stops the planning process.

In order to create offspring an agent has to reach a certain energy level. Offspring inherits its parent's characteristics which undergo mutations. Our *evolutionary algorithm* mutates the bias factor $k$ in small steps which ascertains a smooth drift to more optimal parameters. The offspring is situated randomly in the environment. Existing elite in the environment is preserved because the parent still exists after reproducing. If an agent's fitness is lower than a certain minimum value the agent dies and disappears.

## 3 Results

We analyzed the planning behavior of different types of agents ranging from very deliberative to reactive under fixed conditions for a single agent scenario. In this frame we chose agents with bias factor between 0.04 as minimum and 10.24 as maximum. In Figure 1 a) to c) the average values for planning time, path length of the chosen plan and cost of the executed plan can be observed. The exponential falling curve of the planning time for the very deliberative agents (very small $k$) indicates that these agents are too slow or careful for this environment. They frequently get stuck in local minima. But because their planning time is by factor 10 longer than the length of their executed plans it is obvious that they optimize their plans quite well. For higher $k$ the curve of the planning time comes closer to the curve of the plan length which means that the agent's planning characteristic converges towards reactiveness which is reached when both curves meet. In Figure 1 c) it can be seen that the minimum costs for plan execution are achieved for a moderate reactive agent with $k$ about 1.6. Further it has to be kept in mind that the agents plan as feedback to the environmental conditions. In our example we chose a smooth environment which favored reactive agents. In a more rugged landscape it can be expected that very deliberative agents have an advantage.

Various simulations for multiple agents showed us that the ecosystem is very sensitive to parameter-settings like food production rate and agent sensing range. Here too, we placed the agents in a rather smooth environment. For certain parameter-settings the agents die out after a while. For the simulation shown here we chose a parameter range similar to the one we used for the single agent scenario. At the beginning the average $k$ was chosen to 1.6 in order to observe whether a competitive ecosystem would drift away from this optimal parameter. As can be observed in Figure 1 d) the ecosystem drifts away from the optimal parameter by moving slightly towards reactiveness. This might be due to local dominant strategies which develop in some parts of the system and which have a different optimal $k$. This could happen because we chose a large ecosystem and agents with a smaller sensing range. The bigger the world, the agents are set in, and the smaller the sensing range of the agents, the stronger becomes the effect of locality. Thus, an adaptable diversity of *planning methods* can evolve naturally because the agents have mutated offspring.

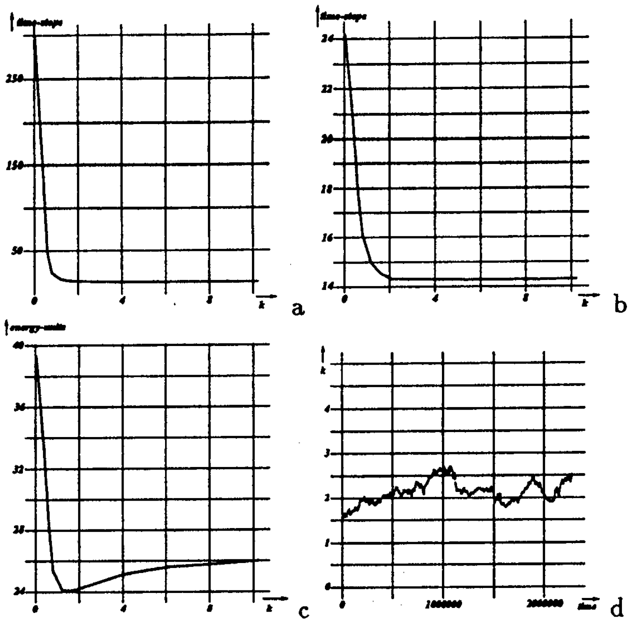


Figure 1: Average a) planning time, b) path length, c) cost for plan execution, and d) development of the factor $k$ averaged over all agents

## 4 Conclusion

We implemented a simulated ecosystem of multiple agents and proposed a potential field method by which agents are adapted between deliberation and reactiveness by a process of evolution. We could show that it depends on the environment to decide whether deliberation or reactiveness is favorable. Adaption is the key to the solution for the tradeoff between deliberation and reactiveness. In one particular example we could see the effect of local dominant solutions keeping the system away from the global optimal solution. For future work we are interested in the detailed analysis of the emergence of a variety of planning strategies.

## References

[1] Stuart Russell, "Rationality and Intelligence," Proceedings of IJCAI 95, pp. 950-957, 1995.

[2] Toru Ishida, "Moving Target Search with Intelligence," Proceedings of AAAI 92, pp. 525-532, 1992.