Steps Towards a Systematic Investigation of Possible Evolutionary Trajectories from Reactive to Deliberative Control Systems

Matthias Scheutz and **Paul Schermerhorn**
University of Notre Dame, Notre Dame, IN 46556, USA

Abstract

Although they are vastly outnumbered by simpler control systems, complex, "deliberative" control systems have evolved on our planet. Hence, agents with deliberative capabilities must have adaptive advantages over agents with simpler control systems in some environments. This paper examines the tradeoffs between the costs of control systems and the benefits they offer in a variety of environments and discusses the implications of these tradeoffs on evolutionary trajectories from "reactive" control systems, through somewhat more complex "affective" systems, to genuine deliberative systems.

Introduction

Most of the Earth's biomass consists of simple organisms (like bacteria, insects, etc.), and the vast majority of evolved species have simple control systems. This seems to indicate that such organisms are very good survivors and must be well adapted to their environments. Furthermore, it suggests that most evolutionary trajectories lead to such simple organisms (at least starting from the Earth's initial conditions) in that there may be many niches in "niche space" (Sloman 2000) that can be occupied by simple creatures. Yet, not *all* evolutionary trajectories stop there. Complex organisms with complex control systems have evolved as well. The natural question then is: under what circumstances did these complex organisms evolve?

A complete answer to this question will obviously have to involve multifaceted arguments comprising considerations of physical, physiological, and control tradeoffs. There are several dimensions along which such tradeoffs can be studied: genotype vs. phenotype, nature vs. nurture, body size vs. brain size, r-selection vs. K-selection, altricious vs. precocious development, innate vs. acquired strategies, reactive control vs. deliberative control, etc. In particular, the argument will have to show how the transition from simple, expendable, fast-growing, etc. organisms with many offspring to complex, inexpendable, slow growing, etc. organisms with few offspring could work.

In this paper we will focus on tradeoffs at the level of architectures of agent control systems. This is not the level of brain and neural organization *per se*, but rather a higher level, implemented in neural brain tissue, at which we can study control systems consisting of functional components and their interconnections.

The architectural level allows us to study functional components and their effects on an agent's behavior, and *a fortiori* its chances of survival. We will compare different kinds of control systems (as defined by their architectures) under different environmental conditions in order to assess their advantages and disadvantages. Components in a control system will have a cost associated with them, reflecting the energy expenditure to build and maintain them in an operational state. By varying the associated cost under identical environmental conditions we can derive empirically a cost function, which points to the tradeoffs between the functionality (and the behavioral consequences) a particular component may add to the control mechanism of an agent, on one hand, and the increase in energy expenditure caused by adding and using the respective component, on the other. Systematic experiments with various agent architectures allow us to study possible trajectories from low-level reactive to high-level deliberative control architectures (two notions to be defined below), which we hope will contribute to some of the open questions in artificial life research (in particular, questions (5), (6), (7), (10), and (11) in (Bedau *et al.* 2000)).

In this paper, we do not attempt to answer the question of how deliberative agent architectures evolved directly (for the problem is by far too complex to be tackled at once). Rather, we approach the set of possible trajectories from reactive to deliberative systems by ruling out a large class of initial conditions under which deliberative systems would most likely not have evolved, because their cost outweighs their benefits.

The paper is organized as follows: first, we define what we mean by "reactive", "affective", and "deliberative" control architectures. Then, we introduce the employed agents, sketch their control architectures and briefly discuss the resultant behavioral dispositions. After pointing to previous experimental findings regarding the relation of reactive, affective, and deliberative agents with respect to their relative fitness [1] in different kinds of environments, we present our experimental setup and report new findings regarding the relative cost of affective and deliberative extensions of reactive agent architectures. Finally, we analyze the experimental outcome and discuss its significance in the light of evolutionary trajectories from reactive to affective, to deliberative control systems.

Reactive, Affective, and Deliberative Architectures

Agent architectures play an important role in the understanding of natural and development of artificial systems (Sloman & Scheutz 2002). They can be thought of as blueprints of control systems, where different functional components and their interconnections are depicted (e.g., see (Russell & Norvig 1995) for a more detailed definition of "agent architecture"). Since we would like to understand (1) what kinds of components (and arrangements thereof) are required to produce particular kinds of behaviors, and (2) what the relative tradeoffs of different control systems (and their implementations) are, we first need to define what we mean by "reactive", "affective", and "deliberative" control, or more to the point, what "reactive", "affective", and "deliberative architectures" are.

Reactive Architectures

Unfortunately, there seems to be a wide range of definitions of "reactive" that differ in substance (e.g., "reactive" as "stateless" versus "reactive" as "tight sensor-motor coupling"). Hence, it seems that "reactive" is best defined in opposition to "deliberative", i.e., as "not deliberative", which puts the burden on a definition of "deliberative". Since we are interested in demarcating an intellectually interesting difference, rather than trying to say what "deliberative" *really means*, we will construe "deliberative" as "being able to produce and use representations of hypothetical past or future states or as yet unexecuted actions (or sequences of such actions)". Note that according to this (negative) definition of "reactive", reactive architectures may make use of simple representations of the state of the world and/or the agent. But these representations will not explicitly encode goals, hypothetical states of the world or sequences of possible actions. And while we may be able to ascribe intentional states such as beliefs and desires to a reactive agent, the agent architecture contains no explicit representation of these states. For example, an agent which exhibits a behavior that could be described as "avoiding obstacles" can be said to have a goal of "avoiding collisions", even though this goal is not explicitly represented in the agent's control system.

Deliberative Architectures

As mentioned before, a *deliberative* architecture is one in which there is some consideration of alternative courses of action before an action is taken. Hence, there is need for the capacity to represent counterfactual states referring to hypothetical past or future states or as yet unexecuted actions (or sequences of such actions), in which at least some of the basic operations of the architecture is to produce/read/write such counterfactual states. Such states include goals (descriptions of states to be achieved), plans (sequences of unexecuted actions), states describing the imagined consequences of performing an action in the current state or some hypothetical state, partial solutions generated during planning or problem solving, the hypothetical states of the agent's beliefs generated during belief revision and many others. We further require that such states should be influential in the production of actions, in the counterfactual sense that, had the (counterfactual) state not been generated, the agent would have chosen a different action to execute.[2]

To represent counterfactual states, a deliberative agent requires a reusable working memory for the construction and comparison of hypothetical states and some means of deriving the consequences of actions performed in these states. At its simplest, this might be a set of memories of the consequences of performing the action in similar states in the past. The use of a common working memory limits the number of alternative courses of action that can be considered in parallel, and hence the degree of parallelism possible within a deliberative architecture.

All other things being equal, a deliberative architecture must be slower and require more resources than a reactive architecture which encodes a solution to any specific goal solvable by the deliberative architecture, since the generation of alternatives will take time. However, a deliberative architecture will typically be more space efficient than an equivalent reactive architecture, even though it will often require more space than a reactive solution to any given problem instance, since it can solve

[1]'Fitness' is here defined as the agent's ability to survive to reproductive age. It is measured indirectly by counting the number of survivors at the end of an experimental run; agents that generally fail to survive to reproductive age will be poorly represented in the population, whereas agents what generally succeed in surviving to reproductive age will be well represented.

[2]Note that this definition implies no commitments as to whether the states and operations are fine grained, e.g., dealing with partial plans or alternative solutions and their generation and comparison, or whether the states and operations are "coarse grained", e.g., a single "plan" operator which takes a goal and a description of the current state and returns a plan with the rest of the fine-grained states and operators buried in the implementation of the architecture and invisible to the agent program and the agent state. Both cases have at least one counterfactual state and one operator that takes a non-counterfactual state and returns a counterfactual state.

a *class* of problems in a fixed amount of space, whereas a reactive architecture requires space proportional to the number of problems. We can view this as an example of the standard space-time tradeoff, though in this case there is also the time required to code or evolve all the reactive solutions.

Affective Architectures

An *affective* architecture is one in which there are explicit representations of affective control states such as preferences, desires or emotions (Scheutz & Sloman 2001). Such states are directly encoded within the overall state of an agent and implemented in components of the architecture (e.g., in a connectionist unit, a real-valued variable in a C program, etc.) rather than being *supervenient* on the architecture (as in the case of purely reactive agent). Note that this does not mean that *all* affective states that could be *ascribed* to the agent or *emerge* from interactions of various components of the architecture are directly represented in the architecture, only that some are.

Affective architectures are orthogonal to the reactive-deliberative distinction in that they can be combined with both (i.e., both reactive and deliberative architectures can be affective in the sense specified). The fact that some affective states are explicitly represented within the architecture (and do not merely supervene on it) means that the architecture could be extended to monitor the achievement or non-achievement of such states. Hence, affective architectures allow affective states to take a role in learning, deliberation, the modification of reactive behaviors, etc. In other words, the difference between reactive and affective architecture may not so much lie in the actual behavioral dispositions of agents instantiating the architectures, but rather in the counterfactual modifications (e.g., extensions) of the architectures. A deliberative mechanism could monitor an affective control state like hunger to create a representation of the agent's need for food. Such an extension may be easy to achieve (e.g., using evolutionary methods of duplication and specialization (Maynard Smith & Szathmáry 1999)) if the state is already *represented* in the architecture, while it will typically be more difficult (and require more additions) if no such representation is present.

Agents: Architectures and Behavioral Dispositions

In the experiments reported in this paper, we employ three different kinds of agents, *reactive*, *affective*, and *deliberative* agents, where affective and deliberative agents are extensions of reactive agents in that their architectures extend the architecture of reactive agents in different ways, leading to different behavioral dispositions and, hence, different behavior.

All agents are standardly equipped with exteroceptive "sonar", "smell", and "touch" sensors. *Sonar* is used to detect obstacles and other agents, *smell* to detect food and water sources, and *touch* to detect (1) impending collisions with agents or obstacles, and (2) consumable food and water sources that are within reach for digestion. In addition, the touch sensor is connected to a global alarm system (Scheutz, Sloman, & Logan 2000), which triggers an automatic reflex-like action pattern, which the agent cannot suppress, to move it away from other agents and obstacles. These movements are somewhat erratic and will slightly reorient the agent (thus helping it to get out of "local minima"). Furthermore, agents have two proprioceptive sensors to measure their energy and water levels, respectively. These sensors are also connected to alarm mechanisms, which will make agents (1) reduce their speed to the lowest possible level (to minimize the energy and water expenditure) and (2) pursue food or water sources exclusively depending on which level drops below a predetermined critical level first.

On the effector side, they have motors for locomotion (forward and backward), motors for turning (left and right in degrees) and a mechanism for consuming food and water (which can only be activated when the agent is not moving). When agents come to a halt on top of a food or water source, their ingestion mechanism suppresses the motors for locomotion until the item is consumed, which will take a time proportional to the amount of energy or water stored in the food or water source (depending on the maximum amount of food or water an agent can take in at any given time).

While different agents may have different short-term goals at any given time (e.g., getting around obstacles, consuming food, reaching a water source faster than another agent, or having offspring), there are two long-term goals that are common to all of them: (1) *survival* (i.e., to get enough food and/or water to maintain all bodily functions, and to avoid running into obstacles or other agents), and (2) *procreation* (i.e., to live long enough to have offspring). In the following, we will briefly describe the architectures and behavioral dispositions of each agent kind.

The Reactive Agents

All agents process sensory information and produce behavioral responses using a motor schema-based approach (Arkin 1989). Let $Ent_k = \{f, w, o, k, \overline{k}\}$ be an index set of the five types of objects *food*, *water*, *obstacle*, *agents of kind k* and *agents of a kind different from k* relative to a given agent kind k–all subscript variables will range over this set unless stated otherwise. For each object type in Ent_k, a force vector $F_{i,k}$ is computed, which is the sum, scaled by $1/|v|^2$, of all vectors v from the agent to the objects of type i within the respective sensory range, where '$|v|$' is the length of vector v. These five *perceptual schemas* are mapped into motor space by the transformation function

$$T_k(x) = \sum_{i \in Ent_k} g_{i,k} \cdot F_{i,k}(x) \qquad (1)$$

where the $g_{i,k}$ are the respective gain values of the perceptual schemes. The gain values simply scale the effect of sensory input, providing a means by which to prioritize certain inputs (e.g., if food is especially important, its gain value could be higher than the other gain values, so that sensing food has a greater impact on the direction chosen than sensing other entities). These gain values are initialized to values determined to be reasonable via a series of experiments, and are kept constant throughout the life of a reactive agent.

Reactive agents always behave in the same way, given that their gain values are *constants*: their positive $g_{f,k} = g_{w,k}$ make them employ a "consume nearest" strategy (Spier & McFarland 1998), whereas their negative $g_{o,k} = g_{k,k} = g_{\overline{k},k}$ values make them avoid obstacles and other agents. Consequently, the behavior of reactive agents is completely determined by their input–hence their name–and can be described as "greedy".

The Affective Agents

Affective agents have in addition to architectural components of reactive agents a three-layer *interactive activation and competition* (IAC) neural network with five input units *in*, five hidden units *hid*, and five output units *out* (Rumelhart & McClelland 1986).[3]

The input units receive their activations (via appropriate scaling functions) from the internal water (in_w) and energy level sensors (in_f) as well as the global alarm mechanisms (which send impulses to in_o, in_k or $in_{\overline{k}}$ units depending on whether the alarm was triggered by an impending collision with an obstacle or an agent of the same or of a different kind).

The output units are connected to the gain values in the motor scheme via individual scaling functions $f_i(x) = x \cdot c_i + b_i$ (where b_i is the *base gain value* and c_i the scaling factor for the activation of out_i).

The activation value $act_i(t)$ of an IAC unit i at time t is defined by

$$act_i(t) = \begin{cases} (max - act_i(t-1)) \cdot net_i(t) - decay, & net_i(t) \geq 0 \\ (act_i(t-1) - min) \cdot net_i(t) - decay, & net_i(t) < 0 \end{cases}$$

where *min* and *max* are the minimum and maximum activation level, respectively, *decay* is a decay factor defined by $d \cdot (act_i(t) - rest)$ (where d is a constant), *rest* the rest level, and $net_i(t)$ the weighted sum of all inputs to unit i at time t.

The choice of IAC units over standard perceptrons is based on their update rule, which is particularly suited to implement important temporal features of affective states in that it (1) takes into account the *previous activation* (hence, can be used to implement "inner states"), and (2) incorporates a *decay term* to raise or lower the activation to a predetermined *base level* (both features that seem to be typical of the temporal development of certain affective states, e.g., basic emotional states).

Although fully connected IAC networks are possible, we will focus on a subset of networks to reduce complexity, where weights between in_i and hid_i are always non-zero and some (possibly all) of the weights between hid_i and out_i, call them "gain weights" ow_i, are non-zero, all other weights being zero.

Foraging in affective agents, then, is not solely determined by their sensory inputs, but also by their "inner states" as defined by the activation of the hidden units in the neural network. These states can implement primitive motivational and emotional states like "hunger", "thirst", "fear", and "aggression" as argued in (Scheutz 2001). Hence, affective agents forage based on their perceptions and needs: they may avoid food if not "hungry", or ignore food if very "thirsty", or skip food if "afraid" of obstacles, etc. By modifying the gain values of the motor schema depending on their internal states, affective agents can influence the combination of the sensory vector fields in the motor schema to their advantage. Note that the integration of this affective system in the reactive system is *conservative* in the sense that no modification of the reactive system is necessary; the affective components are simply "added" to the existing components, which could be relevant for the evolution of force-field based control systems in nature (as adding components without the need to change the existing structure seems to be favorable over additions that require modifications, thereby raising the chances that a previously functional organization may become distorted in the extension process).

The Deliberative Agents

Deliberative agents extend reactive agents in various ways. First, they have an additional perceptual mechanism that allows them to extract objects in the environment and to represent their type, distances, and

[3]Note that neural networks employed in other simulations to control the behavior of agents (Menczer & Belew 1966; Seth 2000) usually compute the mapping from sensors to effectors, while the neural network here is intended to implement the affective system, thus adding another layer on top of the input-output mapping of reactive agents (which is accomplished in a schema-based manner; of course, this mapping, in turn, could have been implemented as neural network as well).

directions relative to an agent-centric polar coordinate system. This "vision facility" allows deliberative agents to single out and represent an individual entity (e.g., a food or water item) and store it in memory for later use (in planning, for example). It is first and foremost this ability of being able to represent items in the environment that opens up further possibilities such as storing and retrieving representations, using them in planning and plan execution, etc. None of these possibilities are available to reactive agents, which have access to sensed objects only in a holistic manner (via an agglomerated force vectors).

Second, they have memory components that allow them to store information about food and water items and obstacles. A "comparator" component constantly checks whether a perceived item is present in memory, and if it is not found there, causes it to be stored. In addition to the memory components, they have a mechanism that will update the relative positions of the object stored in memory depending on the agent's movements. Furthermore, a coherence mechanism constantly checks whether a stored item within sensory range is actually perceived and removes an item from memory if it does not seem to exist any more in the environment.

Third, deliberative agents have a simple route planning mechanism which allows them to find a route to the nearest food or water item, avoiding obstacles. The planner is given a list of obstacles, food and water items known to the agent (i.e., stored in the agent's memory), and returns a *plan*, which is a list of headings and distances, of how to get to the nearest reachable item of a given type. The plan is then passed to a *plan execution mechanism*, which ensures that plan steps are executed by overriding the headings generated by the reactive mechanisms in a manner similar to subsumption-style architectures (Brooks 1986).

The planner is based on a simplified version of the A^*_ϵ algorithm (Pearl 1982). A^*_ϵ is a variant of A^* in which the cost of the solution returned is guaranteed to be no greater than $(1+\epsilon)\times$ (the cost of the optimum solution). A^*_ϵ is a good choice for a route planning agent, as all that is needed are "good" rather than optimal plans. The cost of a plan is the distance the agent has to travel to reach the goal, with a penalty for routes which pass through the collision region around an obstacle (any reasonable cost function must be such that no route through a collision region is ever cheaper than a route around the region).

Planning (or "re-planning") can be triggered in various ways: (1) by the obstacle alarm mechanism, (2) by the completion of an existing plan, (3) by the disappearance of the goal item, and (4) by the appearance of a closer item of the same kind as the goal item within sensory range (for details of the planner and a description of some of the difficulties of integrating a discrete planner with the continuous control of a reactive system, see (Scheutz & Logan 2001)).

Decisions about the goal object for the planner (i.e., food or water) are made by explicitly comparing representations of energy and water levels; whichever is needed more (relative to the maximum capacity) will be selected as a goal. If, however, no goal item of the chosen kind can be sensed in the environment or retrieved from memory, deliberative agents will attempt to go for the other kind. If the other kind is not available either, agents will engage in an "search behavior" that makes them sample their environment (without moving) until they sense a food or water item.

Previous Findings

Various previous experiments have confirmed that affective control mechanisms can and will evolve under different environmental conditions. For example, we found that agents with primitive motivational states (i.e., "hunger" and "thirst" drives) can be evolved from reactive agents if reactive agents are allowed to mutate into affective agents, which in turn can mutate their ow_f and ow_w weights (**?**). Such "motivational agents" are likely to evolve from reactive agents independent of many environmental conditions such as the frequency of appearance of new food and water sources, or the numbers and initial distributions of food and water sources, obstacles and agents. Some of them can also be learned during the lifetime of an agent (using associative learning) if the corresponding weights have the right sign (i.e., lead to attractive or repulsive behavioral disposition depending on the affective state to be implemented (Scheutz 2000)). Furthermore, we found that starting with *motivational agents*, different kinds of agents with different combinations of primitive emotional state like "fear" or "anger" will evolve, if the ow_o and ow_a weights are mutated (Scheutz 2001). Such primitive "emotional" agents will also evolve directly from reactive agents (Scheutz 2002). In all of the above cases, we argued in detail that the evolved mechanisms (i.e., positive or negative "gain weights") indeed implement affective processes, based on (1) the functional characterizations of the involved affective processes, (2) the observable behavior of the agents in the environment, and (3) the evolved architectural components (i.e., connection weights).

In other previous experiments we compared the performance of particular kinds of reactive, affective, and deliberative agents in one-resource tasks (i.e., food foraging) and found that in certain environments affective agents perform better than certain kinds of deliberative agents (Scheutz, Sloman, & Logan 2000; Scheutz & Logan 2001).

Common to all these findings is that various kinds of affective agents seemed to perform better than merely reactive or even certain kinds of deliberative agents (in certain environments). However, none of the previous experiments attempted to assess the additional structural and processing costs of the added components in the employed affective or deliberative architectures (relative to the components in the basic reactive architectures). Hence, it was not clear to what extent affective or deliberative control would pay off relative to the energy expenditure required by the additional components. Such tradeoffs, however, are crucial to an understanding of possible evolutionary trajectories from simple to complex creatures, for if the gain in fitness is disproportional to the cost of the added mechanism, it is unlikely that it would have evolved as part of an agent control system in a multi-agent, multi-species environment.

In the following, we will remedy this lack by investigating the relative fitness of affective and deliberative agents with respect to different costs associated with their architectural features.

Experiments

The Agent-Based Alife Simulation Environment SimWorld

SimWorld is an agent-based artificial life simulation built on top of the SimAgent toolkit developed by Aaron Sloman and colleagues at the University of Birmingham, England.[4] It consists of a continuous, potentially unlimited two-dimensional surface populated with various kinds of spatially extended objects, in particular, different kinds of agents, static and moving obstacles of varying size, and food and water sources.

Static and moving obstacles are typically generated once at the beginning of a simulation run and either remain in place or move in a predetermined direction at a predetermined speed throughout the run of the simulation.

Food and water sources, on the other hand, pop up at random locations within predefined areas of the simulation environment, at predefined frequencies, and stay for a predefined period of time, after which they disappear unless consumed by agents. They contain a fixed amount of energy and water that can be consumed by agents at their maximum intake capacity per simulation update cycle.

Agents are in constant need of food and water. Every simulation cycle they will spend a certain amount of their stored energy and water to maintain the functionality of their bodies and their control systems (typically at least one unit of each). Moving agents consume even more energy and water as measured in terms of some function (typically quadratic) of their speed. When the energy or water level of an agent drops below a certain threshold ω, agents "die" and are removed from the simulation. They also die and are removed if they run into other agents or obstacles.

After a certain age α (measured in terms of simulation cycles), agents reach maturity and can procreate asexually, if their energy and water levels are above the minimum necessary for procreation.[5] The energy and water necessary for creating the offspring are subtracted from the parent, and a new agent will pop up in the vicinity of the parent in the subsequent simulation cycle. After giving birth, parents cannot have offspring for a predetermined time, which is due to a built-in mechanism that is intended to prevent them from depleting themselves too much (which typically results in their death).[6]

Experimental Setup and Results

We conducted various experiments to study the benefit of the added components in affective and deliberative agents with respect to their potential for survival (i.e., the increase in fitness) relative to the added cost imposed by these components. For all of the following experiments, we limited the world to a squared area of 800 by 800 units, where water and food appear equally distributed in the centered subregion of 720 and 720. The probability of a new food source appearing at any update cycle is 0.25, that for water 0.2.

To determine the effect of introducing additional cost on architecture extensions (i.e., "relative cost" to a given base architecture), we conducted various experiments in which two agent kinds had to compete for survival in various environments with different numbers of static obstacles. Each experiment started with 10 agents of each type, 10 food and 10 water sources, and a varying number of obstacles, and proceeded for 10,000 cycles; each generation (i.e., the average time between the birth of parents and that of their offspring) is slightly longer than ω, so 10,000 cycles allows roughly 35 generations. The performance measure used (and depicted in the tables below) is the average number of surviving agents of a particular kind at the end of an individual experimental run averaged over 20 runs with different initial positions of agents, obstacles, food and water sources. Over a series of experiments, the number of static objects in the

[4] SimWorld is freely available and can be downloaded from http://www.nd.edu/~airolab/simworld. The SimAgent toolkit can be downloaded from http://www.cs.bham.ac.uk/research/poplog/newkit.tar.gz

[5] Note that both parameters, α and ω, can be used to specify whether the simulation is used as an exogenous or as an endogenous fitness model.

[6] SimWorld also features a complex mutation mechanism that can modify the architectural parameters of agents. If *mutation* is turned on, a predefined set of architectural components can be modified with a predefined probability μ, again according to predefined mutation operations. For example, a connection weight in a neural network could be increased or decreased with a certain probability, or a condition-action rule added to the ruleset of an agent. Hence, some offspring will start out with the modified parameters instead of being exact copies of the parent. Furthermore, it is possible to mutate agents of one kind into agents of another kind. In the experiments described below, mutation is not used; agents with predefined parameters are compared.

environment was varied (from 30 to 50), as was the relative cost of the architecture of one agent kind compared to the cost of other (i.e., the cost of the first was set equal to the cost of the second multiplied by C, where C ranged from 1.0 to 5.0). Cost was assessed as a reduction in energy reserves derived from the consumption of food.

Table 1 gives the results for reactive and deliberative agents. When the relative cost of deliberation is 1.0 (i.e., the same as the cost of reactive agents), deliberative agents enjoy a clear advantage across all obstacle environments. Especially high numbers of obstacles make foraging a very difficult (if not impossible) task for reactive agents, for one because being repelled by the proximity of high density obstacle regions causes them to ignore safe routes through these regions to food and water, which deliberative agents are able to exploit. As the cost of deliberation increases (up to five times the cost of reactive agents), however, the reactive agents begin to make inroads, as expected. Still, deliberative agents proved better able to navigate the environment as obstacle count increase (relative to reactive agents).

Comparing affective and deliberative agents, we find that affective agents fare better than their reactive counterparts against deliberative agents (Table 2). However, the same trend appears with regard to the obstacle environment, albeit less pronounced: deliberative agents are better suited for survival in cluttered environments.

Finally, comparing reactive to affective (in Table 3), we notice that the advantages of affective agents over reactive agents are sufficient to overcome only minor increases in the cost function of the affective agents. Reactive agents begin to perform better even before the cost of being affective is double the cost of reactive agents over all obstacle environments examined here.

Analysis

The above results point to a general fitness ordering among the examined agent kinds: deliberative agents are generally better at surviving than affective agents, which, in turn, are better than reactive agents. Moreover, the relative differences in fitness levels increase as the obstacle environment becomes more dense, which is not surprising, since the added deliberative mechanisms will only develop their fullest potential (relative to the other foraging mechanisms) in very cluttered environments, where the planner can detect routes that are not open to force field-based control systems. In fact, experiments using environments without any or with only very few obstacles show that deliberative agents are at a disadvantage against both reactive and affective agents, even without taking cost into account (which is largely due to the fact that they do not pick up items on the way to their goal location if these items are of a different kind than the goal item–a behavior that is a natural consequence of the force field-based control favored by many simple species).

Deliberative agents are able to maintain their advantage over both affective and reactive agents through more dramatic increases in relative cost, especially in more dense obstacle environments. However, further increase will eventually lead to an inversion of this ordering. By the time the cost of being deliberative reaches five times the cost of the competing agents, deliberative agents perform very poorly in comparison.

Comparing the relative cost of the added architectural mechanisms (i.e., memory to hold locations of items in the environment, updating mechanisms, coherency mechanisms for new perceptions, etc.) in terms of actual memory (at least 100 times) and actual computation time (at least 10 times) to the cost of the break-even point in the above experiments (between 3 and 5 if estimated conservatively), it is easy to see that deliberation would not pay off in the scenarios we investigated. Note that the problem here is not simply one of finding a trajectory from reactive/affective agents to deliberative agents composed of adaptations with costs small enough for deliberative agents to compete effectively; the "finished product", i.e., agents with fully developed memory capacity and planning mechanisms, fail to perform well enough even at very modest relative cost levels.

Affective control, on the other hand, despite its modest increase in fitness might well pay off given that the actual additional cost (again measured in terms of computation time and memory) of affective systems is only a fraction of that of the reactive system. Hence, while it seems unlikely that a deliberative control system (with similar capacities to the ones used here) would have evolved, affective control does have a "competitive edge" over mere reactive control within the given cost framework.

Discussion

Agents possessing deliberative capacities do exist in nature. The difficulty is to understand what ecological niche might enable deliberative agents to perform better than reactive or affective agents despite the significant increase in cost caused by deliberative extensions of reactive or affective architectures.

For one, it seems that the more regular and predictable an environment is, while at the same time being dangerous and difficult to explore for agents with simple control systems, the more deliberation will pay off. In the above experiments, this can be seen from the fact that environments with more static obstacles become more predictable (as more items in the environment become predictable) and more dangerous (as there are more chances for agents to accidentally crash into an obstacle).

The defining properties of such "stable" environments, where change occurs only infrequently and essential resources are by and large predictable, square well with

Cost mult.	30 static obstacles				40 static obstacles				50 static obstacles			
	Reactive		Deliberative		Reactive		Deliberative		Reactive		Deliberative	
	Av.	Std.	Av.	Std.	Av.	Std.	Av.	Std.	Av.	Std.	Av.	Std.
1.0	0.3	1.34	24.65	4.0	0.0	0.0	21.25	4.0	0.0	0.0	18.55	5.34
1.4	0.95	3.27	20.6	7.1	0.0	0.0	19.8	2.78	0.0	0.0	16.75	4.88
1.8	0.6	2.26	21.8	4.34	0.0	0.0	15.7	3.61	0.0	0.0	14.35	5.25
2.2	1.7	3.57	15.6	5.13	0.0	0.0	15.05	4.32	0.3	1.34	11.35	5.74
2.3	6.5	9.05	10.4	6.92	0.0	0.0	14.75	5.87	0.3	1.34	12.2	5.27
2.4	8.95	9.75	6.7	7.34	1.4	3.57	12.85	6.2	0.0	0.0	11.2	3.37
2.5	7.4	9.88	8.9	8.94	0.7	2.9	11.7	4.22	0.3	1.34	8.0	6.33
2.6	9.35	9.07	6.1	6.52	3.55	5.6	7.6	6.06	0.35	1.35	9.0	4.81
3.0	9.15	8.5	4.7	6.26	3.05	5.24	8.0	6.97	2.05	3.9	8.15	5.44
3.4	11.35	7.92	2.25	4.7	2.95	5.84	6.6	4.5	1.78	3.74	5.9	5.45
3.8	11.75	6.84	2.2	4.54	2.9	4.42	4.52	4.63	2.0	4.34	4.7	5.79
4.2	12.53	6.71	1.45	3.77	5.6	4.67	3.25	5.5	2.95	4.82	2.8	3.81
4.6	14.05	4.71	0.2	0.89	7.35	5.88	1.85	3.45	2.45	3.91	1.65	3.12
5.0	13.45	6.51	0.3	1.34	8.05	7.85	1.0	2.58	4.9	5.59	0.7	1.98

Table 1: Deliberative agent performance relative to reactive agents

Cost mult.	30 static obstacles				40 static obstacles				50 static obstacles			
	Affective		Deliberative		Affective		Deliberative		Affective		Deliberative	
	Av.	Std.	Av.	Std.	Av.	Std.	Av.	Std.	Av.	Std.	Av.	Std.
1.0	0.85	3.68	22.97	5.72	0.0	0.0	23.1	3.78	0.0	0.0	18.0	4.69
1.4	4.4	8.04	15.75	9.33	0.15	0.67	19.1	3.32	0.85	3.57	16.2	5.63
1.8	6.5	7.61	12.65	8.15	1.4	4.17	14.25	6.32	0.35	1.57	13.9	6.42
2.2	7.55	8.25	8.75	7.38	2.3	4.99	11.45	6.06	1.8	5.03	9.45	5.39
2.3	13.8	10.36	6.0	7.81	3.0	5.52	9.75	7.06	1.4	2.93	10.1	6.69
2.4	9.6	7.84	6.8	8.04	5.55	6.78	8.2	7.74	2.15	4.12	9.35	6.21
2.5	11.2	10.47	5.55	6.96	7.25	8.34	6.1	6.72	2.7	6.41	9.05	6.32
2.6	14.05	8.5	2.95	5.98	9.1	9.01	5.9	6.92	3.32	4.97	8.28	5.79
3.0	19.25	5.07	0.05	0.22	7.9	7.45	4.4	5.67	4.2	6.01	5.8	5.05
3.4	15.55	5.86	0.55	2.46	8.3	7.1	4.35	5.69	3.65	5.53	3.25	4.94
3.8	19.5	4.48	0.0	0.0	12.1	5.88	0.55	1.76	5.55	6.07	3.35	4.82
4.2	19.8	5.08	0.0	0.0	11.9	7.31	0.75	2.31	6.6	5.53	1.85	3.45
4.6	17.9	5.78	0.0	0.0	12.4	5.73	0.45	1.15	9.15	5.43	0.35	1.57
5.0	19.35	5.39	0.0	0.0	13.4	4.16	0.25	1.12	7.27	6.0	0.53	2.08

Table 2: Deliberative agent performance relative to affective agents

deliberative mechanisms. These properties ensure that information regarding the location of a resource will normally be accurate, hence representational capacities enabling information to be stored in memory and used in planning or reasoning are advantageous.

Conversely, in environments where resources are not predictable, the benefits of deliberative mechanisms are likely not sufficient to outweigh the increased cost of those mechanisms. In these “unstable” environments, affective agents will typically perform better than deliberative agents, and memory and planning may be of little use.

Conclusion and Future Work

We conjecture that most evolutionary trajectories of control systems starting with only reactive mechanisms reach a fitness maximum in some of their affective extensions. What is still unclear is exactly under what environmental conditions this fitness peak of affective control systems can be overcome by adding deliberative extensions. Predictability and stability of resources seem to be part of the requirement. This may either involve the prediction of locations where to find food, or in the case of a predator, the ability to predict its prey’s behavior when confronted, chased, etc. Social factors (such as tracking mental states of other group members) will certainly play an important role, although they do not seem to be necessary for the development of deliberative control systems. Most importantly, perhaps, the co-evolution of control system and body may figure in the search for the need of more complex control systems (e.g., energy, size, and other constraints may make mere “reactive” extensions impossible).

Although deliberative extensions might not pay off right away, they need to add some initial benefit to the control system’s function, otherwise they will not evolve in competitive multi-species environments (as their bearers will likely not survive in the long run). Based on

Cost mult.	30 static obstacles				40 static obstacles				50 static obstacles			
	Reactive		Affective		Reactive		Affective		Reactive		Affective	
	Av.	Std.	Av.	Std.	Av.	Std.	Av.	Std.	Av.	Std.	Av.	Std.
1.0	6.85	9.58	11.05	10.04	1.45	4.49	12.6	6.84	2.3	4.51	4.4	5.45
1.1	3.5	7.56	14.25	8.5	3.9	4.89	9.3	8.88	0.9	2.36	5.8	5.55
1.2	6.9	8.78	9.25	9.08	2.45	4.65	8.45	7.93	2.05	4.19	3.5	4.58
1.3	7.23	8.96	9.15	9.07	5.2	6.66	5.4	6.24	1.5	3.32	2.9	3.58
1.4	8.75	8.17	4.65	7.26	6.4	6.55	4.35	5.89	2.45	3.05	1.65	3.12
1.8	9.6	9.28	3.45	5.74	5.0	6.11	3.95	6.08	2.7	4.55	1.15	2.46
2.2	14.2	6.74	0.0	0.0	7.8	6.18	0.95	2.46	3.25	4.71	0.0	0.0
2.6	13.15	6.88	0.45	2.01	8.95	6.14	0.0	0.0	3.9	4.0	0.0	0.0
3.0	14.1	7.24	0.25	1.12	9.2	6.86	0.2	0.89	6.5	6.28	0.0	0.0
3.4	14.2	6.81	0.0	0.0	9.75	4.79	0.0	0.0	4.33	5.16	0.0	0.0
3.8	12.1	6.49	0.0	0.0	9.9	5.39	0.0	0.0	3.5	4.68	0.0	0.0
4.2	11.6	4.45	0.0	0.0	10.0	7.03	0.0	0.0	2.35	3.39	0.0	0.0
4.6	16.55	5.06	0.0	0.0	10.35	6.27	0.0	0.0	3.6	4.63	0.0	0.0
5.0	14.1	7.17	0.0	0.0	11.65	5.78	0.0	0.0	3.25	5.02	0.0	0.0

Table 3: Affective agent performance relative to reactive agents

preliminary evolutionary experiments with agent architectures, we predict that, when there is interspecies competition for resources, it will be difficult to find neighborhoods in design space of architectures where even primitive deliberative extensions are viable (relative to their cost). The main theoretical, explanatory difficulties are not so much connected to a full-fledged, fully functional deliberative system, which already has enormous adaptive advantage over affective control systems (e.g., in highly structured, regular environments). Rather, they are connected to explanations of the intermediary stages, where *some* deliberative capacities have evolved and are combined with affective control. What complicates the pictures is the fact that not all combinations of affective and deliberative control are beneficial, even if both the affective and deliberative subsystems may show satisfactory performance in isolation (e.g., see (Scheutz & Logan 2001)).

It almost seems that individual deliberative subcomponents in natural evolutionary trajectories must have had a sufficient benefit of their own, i.e., independent of other deliberative subcomponents, before they became part of a larger, even more beneficial deliberative system. We are currently investigating this hypothesis by adding various deliberative subcomponents (e.g., memory, "perceptual coherency mechanisms", etc.) to affective agents to see whether they are beneficial in certain environments (e.g., where food and water still appear at random, but are confined to certain subareas in the environment).

We are also working on a more detailed analysis of the notion of "cost of a component in an agent architecture", which will allow us to get a finer-grained break-down of the net benefit of different functional components with respect to the overall agent behavior, and hence on the likelihood that they will evolve in certain environments.

Finally, we see the need for many more experiments with different reactive and affective control systems in different environments to map out the space of possible control mechanisms. It is their relative fitness with which deliberative control systems will have to compete.

Only if we understand the potential and limitations of reactive and affective control, we believe, will we be able to understand the circumstances under which deliberative systems, and consequently minds, have evolved.

References

Arkin, R. C. 1989. Motor schema-based mobile robot navigation. *International Journal of Robotic Research* 8(4):92–112.

Bedau, M. A.; McCaskill, J. S.; Packard, N. H.; Rasmussen, S.; Adami, C.; Green, D. G.; Ikegami, T.; Kaneko, K.; and Ray, T. S. 2000. Open problems in artificial life. *Artificial Life.*

Brooks, R. A. 1986. A robust layered control system for a mobile robot. *IEEE Journal of Robotics and Automation* 2(1):14–23.

Maynard Smith, J., and Szathmáry, E. 1999. *The Origins of Life: From the Birth of Life to the Origin of Language.* Oxford: Oxford University Press.

Menczer, F., and Belew, R. 1966. From complex environments to complex behaviors. *Adaptive Behavior* 4(3/4):317–363.

Pearl, J. 1982. A^*_ϵ—an algorithm using search effort estimates. In *IEEE Transactions on Pattern Analysis and Machine Intelligence*, volume 4, 392–399.

Rumelhart, D., and McClelland, J. 1986. *Parallel distributed processing: Explorations in the microstructure of cognition.* Cambridge, MA: MIT Press.

Russell, S., and Norvig, P. 1995. *Artificial Intelligence, A Modern Approach.* Prentice Hall.

Scheutz, M., and Logan, B. 2001. Affective versus deliberative agent control. In Colton, S., ed., *Proceedings of the AISB'01 Symposium on Emotion, Cogni-*

tion and Affective Computing, 1–10. York: Society for the Study of Artificial Intelligence and the Simulation of Behaviour.

Scheutz, M., and Sloman, A. 2001. Affect and agent control: Experiments with simple affective states. In Zhong, N.; Liu, J.; Ohsuga, S.; and Bradshaw, J., eds., *Intelligent Agent Technology: Research and Development*. New Jersey: World Scientific Publisher. 200–209.

Scheutz, M.; Sloman, A.; and Logan, B. 2000. Emotional states and realistic agent behaviour. In Geril, P., ed., *Proceedings of GameOn 2000, Imperial College London*, 81–88. Delft: Society for Computer Simulation.

Scheutz, M. 2000. Surviving in a hostile multiagent environment: How simple affective states can aid in the competition for resources. In Hamilton, H. J., ed., *Advances in Artificial Intelligence, 13th Biennial Conference of the Canadian Society for Computational Studies of Intelligence, AI 2000, Montréal, Quebec, Canada, May 14-17, 2000, Proceedings*, volume 1822 of *Lecture Notes in Computer Science*, 389–399. Springer.

Scheutz, M. 2001. The evolution of simple affective states in multi-agent environments. In Cañamero, D., ed., *Proceedings of AAAI Fall Symposium*, 123–128. Falmouth, MA: AAAI Press.

Scheutz, M. 2002. The evolution of affective states and social control. In Hemelrijk, C. K., ed., *Proceedings of International Workshop on Self-Organisation and Evolution of Social Behaviour*.

Seth, A. 2000. *On the Relations between Behaviour, Mechanism, and Environment: Explorations in Artificial Evolution*. Ph.D. Dissertation, University of Sussex, UK.

Sloman, A., and Scheutz, M. 2002. A framework for comparing agent architectures. In *UK Workshop on Computational Intelligence*. forthcoming.

Sloman, A. 2000. Interacting trajectories in design space and niche space: A philosopher speculates about evolution. In et al., M., ed., *Parallel Problem Solving from Nature – PPSN VI*, Lecture Notes in Computer Science, No 1917, 3–16. Berlin: Springer-Verlag.

Spier, E., and McFarland, D. 1998. Possibly optimal decision making under self-sufficiency and autonomy. *Journal of Theoretical Biology* 189(317–331).