Assessing the Performance of Different Behavior Selection Architectures in a Large and Complex Virtual Environment

Wan Ching Ho, Orlando Avila-Garcia, Chrystopher L. Nehaniv
Adaptive Systems Research Group
School of Computer Science, University of Hertfordshire
College Lane, Hatfield, Hertfordshire, AL10 9AB, UK
W.C.Ho@herts.ac.uk, O.Avila-Garcia@herts.ac.uk, C.L.Nehaniv@herts.ac.uk

Abstract- We compare the performance of autonomous agents with three different behavior selection architectures (Static-Threshold, Winner-Takes-All and Voting-Based) in terms of survival in a large and complex dynamic virtual environment. Experiment results indicate both advantages and disadvantages when applying each architecture in such environmental conditions, and also shows that the performance of Voting-Based architecture is significantly sensitive to the balance of rates at which various environmental resources can serve to satisfy an agent's physiological needs. Some problems with behavior selection architectures are identified and possible solutions are proposed.

1 Introduction

Behavior selection is one of the critical issues in designing autonomous agents and robots [Maes, 1995]. In the perspective of Artificial Life, an artificial animal must make appropriate decisions to guarantee its survival in a given environment using some sort of behavior selection architecture – separately or in combination with other mechanisms such as learning or evolution [Meyer and Guillot, 1990].

Many behavior selection architectures have been proposed (see [Tyrrell, 1993] and [Guillot and Meyer, 1994] for an overview). As pointed out by the behavior selection literatures [Tyrrell, 1993, Bryson, 2000], agents which are implemented with architectures such as winner-takes-all and voting-based – or *compromise candidates* – require different levels of necessary information from the environment in order to effectively satisfy their internal goals. Therefore, voting-based agents performing in a large and complex dynamic environment with a limited sensory range and no memory may not be able to cope with conflicting situations to make good enough behavior selection.

In this paper we perform a comparison between Voting-Based (VB), Winner-Take-All (WTA) and Static-Threshold (ST) architectures in a large and complex environment. Previous research [Avila-García and Cañamero, 2002] indicated that assessing the performance of behavior selection architectures is a complex task and generally requires different indicators to allow comparison and understanding of their essential properties. In this paper we compare architectures in terms of three viability indicators [Avila-García and Cañamero, 2002]: Lifespan, Overall Comport and Physiological Balance (Section 5). A large and complex nature-like environment has been created (Section 2) to test and compare three different behavior selection architectures (Section 3) respectively defined in our previous research [Ho et al., 2003, Ho et al., 2004, Avila-García et al., 2003].

2 Agent Physiology, Environment and Basic Behaviors

In order to test agents with three different behavior selection algorithms, a large, dynamic and complex "nature-like" virtual environment has been created by using VRML and Java programming languages. This environment is fairly different from other simple and flat agent test-beds since it has various types of landforms and concomitant resource differences, including areas of oasis, desert, mountains, cave, river, lake and waterfall. Figure 1 illustrated the virtual environment model from two different perspectives. Each area has its unique features that could be used by the agents to meeting their physiological needs (see below), illustrated as follows:

- *Oasis* - this is generally a warm and flat area, which has three *Apple Trees* in the summer.
- *Desert* - a hot and flat area where efficiently provides body heat to the agents and has *Stones* and *Cactuses*. Cactus is the only resource for agents to increase their *moisture* in the winter. To crush the Cactus, agents need to pick up a Stone. Therefore agents are able to change the Stone distribution in the environment by randomly carrying or laying down the Stone after they have consumed a Cactus.
- Mountain - located between the desert and oasis areas; some edible *Mushrooms* exist permanently on the top of the mountain, however, climbing up the mountain takes an extra amount of internal energy from the agents.
- River - in the summer, it provides water resource to the agents and locates next to the oasis. Agents are able to swim in the river, but they cannot swim toward the north since it is against the current.
- Lake and Waterfall - these provide another source of moisture and environmental complexity. The waterfall connects to the upper river and the lake. Once agents enter the waterfall area, they will be picked up by the downstream current and then fall into the lake area. An agent is not able to either go back to the river from the waterfall or go back to the waterfall from the lake.

0-7803-9363-5/05/$20.00 ©2005 IEEE.

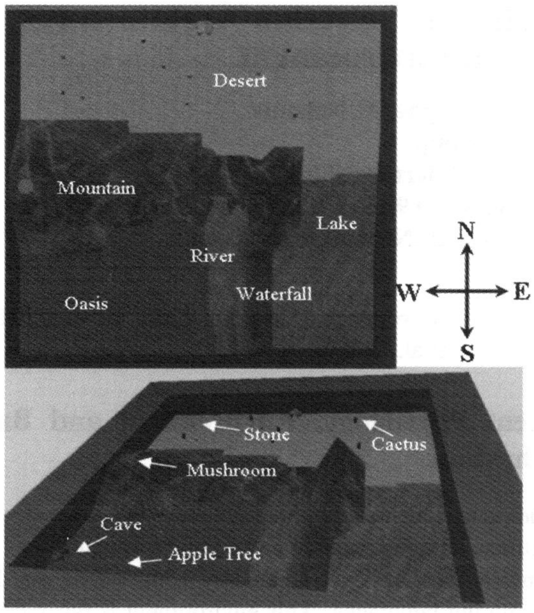

Figure 1: The simulated virtual environment viewed from two different perspectives.

	Level of cold in Oasis	Level of heat in Desert	River
Summer	Cool	Hot	Flowing (Agents cannot pass)
Winter	Cold	Warm	Frozen (Agents can pass)

Table 1: The level of heat and cold in different areas of the environment, and the accessibility of the river in the seasons of summer and winter

- Cave - there are two caves in the environment for agents to regain their energy, one located in the oasis area and the other one located in the desert area.

Alternatively, two seasons *Summer* and *Winter*, have been simulated in the environment to have a higher level of environmental dynamics (Table 1). Each season has the same duration but different effects on a) the level of heat and cold in different areas of the environment, b) dynamic resources allocation and c) the accessibility of the river.

2.1 Physiology and Basic Behaviors

2.1.1 Sensing and acting

Agents are equipped with nine external sensors: seven Hit-Ray sensors [Blaxxun, 2004] formed a 90 degree fan-shaped for detecting the objects, landforms, as well as the environment heat from different types of landforms; agent body has a landform sensor and also a time sensor for sensing the current season of the environment; Figure 2 shows the distributions of these sensors.

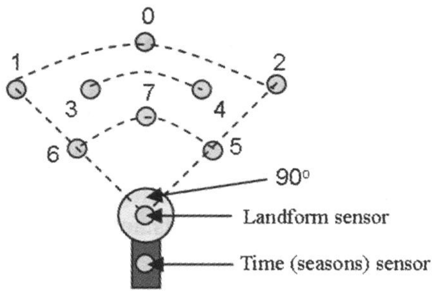

Figure 2: Hit-Ray sensors 0 - 7 for sensing both objects and Landforms, agent body has a landform sensor and a time sensor.

2.1.2 Physiology

There are four essential internal variables [Meyer and Guillot, 1990] implementing the agent's synthetic physiology. The agent must keep those variables within certain viable ranges in order to survive: (1) *Energy*, *Moisture* and *Glucose* must be maintain higher than the minimum value, and (2) *Temperature* within a range between a maximum and the minimum values. If one of these values exceeds the boundary the agent 'dies'. There are conflicts between the different behavioral alternatives to satisfy those internal needs, so the agent must decide at every step of its life what to do next in order to stay 'alive'.

2.1.3 External stimuli

Besides the internal variables, the presence of external stimuli is also a factor to influence the outcome of the behavior selection. There are five types of external stimuli in the environment including *Apple Trees*, *Water*, *Mushrooms*, *Environmental Heat* and *Cold*, some of them are static objects, such as *Mushroom* on the top of the mountain area, *Apple Trees* in the oasis area; some of them are the spread through a huge area, for example, the agent's body can be warmed up in the desert area and cooled down in other areas. Furthermore, the river provides *Water* only in the summer, which is a dynamic external stimulus to the agent.

2.1.4 Motivations

Motivations constitute urges to action based on bodily needs related to self-sufficiency and survival. They implement a homeostatic control to maintain the essential physiological variables within certain ranges. Agents' motivations are characterized by: a controlled (essential) physiological variable, a drive to increase or decrease the level of the controlled variable and an (external) incentive stimulus that can increase the motivation's intensity. Table 2 shows agents' motivations with their drives and incentive stimuli.

2.1.5 Behaviors

Following the usual distinction in ethology, our agents have consummatory (goal-achieving) and appetitive (goal-directed) behaviors [McFarland, 1999]. A consummatory

Motivation	Drive	Incentive stimulus (Cue)
Fatigue	Energy ↑	Cave
Thirst - Water	Moisture ↑	River (in summer only), Lake
Thirst - Apple	Moisture ↑	Apple Tree
Hunger - Mushroom	Glucose ↑	Mushroom
Hunger - Apple	Glucose ↑	Apple Tree
Body Heat	Body Temperature ↑	Desert area
Body Cold - Oasis	Body Temperature ↓	Oasis area
Body Cold - Mountain	Body Temperature ↓	Mountain area
Body Cold - Water	Body Temperature ↓	River, Waterfall, Lake

Table 2: Agents' motivations with their drives and incentive stimuli.

behavior is executed only if it has been selected by the motivational state of the agent and its incentive stimulus is being observed. Appetitive behaviors help the agent to reach incentive stimuli so as to execute consummatory behaviors. The execution of a behavior has an impact on (increasing or decreasing) the level of specific physiological variables. Table 3 shows the agent's behaviors.

The different resource and environment types, the agent physiology and behaviors, as well as multiple possible behaviors and situations where conflicts for action selection may arise all contribute to the complexity of potential agent-environment dynamics.[1]

3 Architectures for Behavior Selection

We have tested 3 types of agent architectures which have been previously used in agent simulations and robotic experiments: Static-Threshold [Ho et al., 2003, Ho et al., 2004, Avila-García et al., 2003], and Winner-Takes-All and Voting-Based [Avila-García and Cañamero, 2002, Avila-García et al., 2003]. All of these three behavior selection architectures have (1) the same appetitive – randomly searching – and reflexive – obstacle avoiding – behaviors and (2) the same elements introduced in the next section. They are based on a subsumption control architecture [Brooks, 1986], as illustrated on Figure 3.

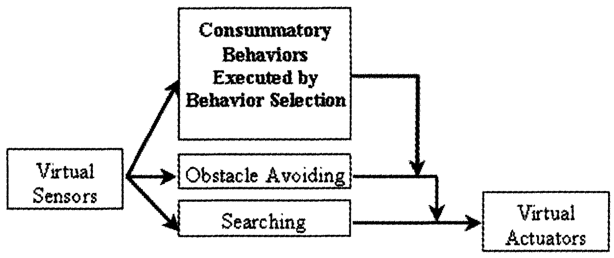

Figure 3: Behavior hierarchy which is based on the subsumption architecture for a Purely Reactive agent.

3.1 Behavior Selection Policies

3.1.1 Static Threshold (ST)

Agents with ST behavior selection architecture have fixed points as behavior performing thresholds. A ST agent executes the consummatory behavior only if the specific internal variable is lower than the static threshold and the agent encounters the corresponding external stimulus. At that moment, the agent will stay and consume the resource until the specific internal variable reaches its upper limit. For *Body Temperature*, the agent will stay in the area where can provide heat or cold to the agent to adjust its own *Body Temperature* value to the ideal range before it leaves this area. When the agent is not able to sense any kind of external stimuli, or its physiological variables are all higher than the static threshold or in the ideal range, it will be just wandering around and avoiding obstacles.

3.1.2 Winner-Takes-All (WTA)

In order to make the decision of which behavior is going to be executed in the next time step, WTA behavior selection architecture calculates and updates all the motivation values in each time step by using the following formula:

$$Motivation = Deficit + (Deficit * Cue)$$

In this formula, *Deficit* of a specific physiological variable is calculated by the ideal value of that variable minus the current value. The value of *Cue* increases when the agent is getting closer to the corresponding stimulus, and vice versa. After all motivations' intensity have been calculated, the motivation with the highest value triggers the relevant behavior to execute in the next time step.

3.1.3 Voting-Based (VB)

VB can be seen as another calculating process building on the top of WTA, since it takes the results from the motivation calculation process, and then performs another process of calculations for behavior executions in the next time step, as shown in the following formula:

[1] This complexity is reflected in situations that may appear in an action selection problem. For instance, our environment presents situations with:

- The same behavior affecting more than one physiological variable: e.g., *Searching*, *Avoiding Obstacle* or *Eating Apple*.
- The same behavior correcting more than one physiological variable: e.g., *Eating Apple* raises the level of both *Moisture* and *Glucose*.
- The same physiological variable being corrected by more than one behavior: e.g., decreasing *Body Temperature* can be obtained going in a mountain, oasis, river or lake.
- The same physiological variable being corrected in both directions: e.g., both low and high *Body Temperature* must be corrected.
- The same physiological variable being corrected using more than one external stimulus: e.g., glucose can be obtained either consuming *Apples* or *Mushrooms*
- The same external stimulus helping correct different physiological variables: e.g., apples help to obtain both *Moisture* and *Glucose*.

Behavior	Stimulus	Effects	Motivations
Searching	NIL	Energy ↓, Glucose ↓, Moisture ↓	NIL
Obstacle Avoiding	Obstacle	Energy ↓, Glucose ↓, Moisture ↓	NIL
Resting	Cave	Energy ↑	Fatigue
Drinking Water	River, Lake	Moisture ↑	Thirst - Water
Eating Apple	Apple Tree	Moisture ↑, Glucose ↑	Thirst - Apple, Hunger - Apple
Eating Mushroom	Mushroom	Glucose ↑	Hunger - Mushroom
Move to Desert	Desert	Body Temperature ↑	Body Heat
Move to Oasis	Oasis	Body Temperature ↓	Body Cold - Oasis
Move to Mountain	Mountain	Body Temperature ↓	Body Cold - Mountain
Move to Water	River, Lake	Body Temperature ↓	Body Cold - Water

Table 3: Agents' behaviors with their corresponding stimuli, effects and motivations.

$$behavior = Motivation_1 * Rate_1 + Motivation_2 * Rate_2$$

As described in the previous subsection, there are behaviors which can satisfy more than one motivation. For example *Eating Apple* behavior increases the level of two physiological variables: *Glucose* and *Moisture*; the changes of these two variables will influence the motivation calculation in the next time step of the simulation. In this equation, *Rate* is the speed of consuming a specific external stimulus. Similar to WTA architecture, the behavior with the highest value will be executed by the agent in the next time step.

Setting (Apple Nutritional Rate)	ST	WTA	VB
(1) 1 : 2	10 Runs	10 Runs	10 Runs
(2) 1 : 1	10 Runs	10 Runs	10 Runs
(3) 2 : 3	10 Runs	10 Runs	10 Runs
(4) 4 : 7	10 Runs	10 Runs	10 Runs

Table 4: Apple Nutritional Rate is the amount of *Moisture* and *Glucose* the Apple provides to the agent in each time step it is consumed, in contrast with that provided by Water and Mushroom. For example, setting (1) shows the rate as 1:2, which means eating Apple provide half of the *Moisture* that Water does per time-step when the agent is consuming it, and also provide half of the *Glucose* that Mushroom does.

4 Experiments

4.1 Method

For testing the efficiency of three behavior selection architectures in terms of agents' survival in the proposed environment, and particularly investigating the hypothesis we made for VB architecture – i.e. it is likely to be unable to perform as well in large and complex environments – we designed four different experimental settings (see Table 4). All experimental settings were set up to testing and comparing the performance of ST, WTA and VB architectures; the reason to have different *Apple Nutritional Rate* is that we would like to confirm our hypothesis of the weaknesses of VB architecture and attempt to rectify this problem by slightly changing the parameters of environmental conditions.

Each experiment run takes about 15 minutes in a Pentium 4 2.0GHz PC with 512MB RAM.

4.2 Results and discussions

Performance results in terms of *Life span* (Figure 4), *Overall Comfort* (5) and *Physiological Variance* (6) we obtained from all experiments runs with four different settings. *Life Span* (LS) represents agents' average lifetime in 10 runs in simulation time steps; *Overall Comfort* (OvC) is the average value of all physiological variables and *Physiological Variance* (PhV) is the average value of variance of all physiological variables – it is the inverse of the *Physiological Balance* indicator defined in [Avila-García and Cañamero, 2002].

Agents with ST and WTA architectures obtained approximately the same performance in terms of the lifespan. In terms of OvC, ST outperforms WTA, while in terms of PhV the latter is better than the former. These results can be explained as the difference in persistence when the agent is executing consummatory behaviors. ST has significantly more persistence in consuming resources, which means the agent always consumes a specific external stimulus more than it needs to be able to survive. In WTA the moment at which the agent stops consuming one resource and go to satisfy another motivation is given by the competition between motivations; it can be seen as creating dynamic thresholds for the agent. Therefore WTA produces less variance in the satisfaction of its internal motivations (PhV), while apparently has an impact in the their average level of satisfaction

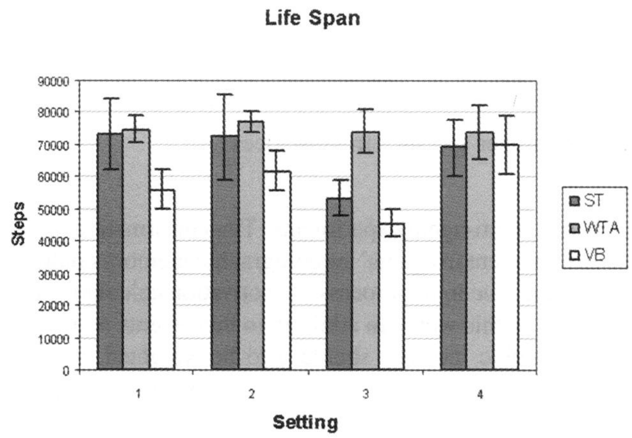

Figure 4: Experiment results of Life Span (LS) with standard errors for three different behavior selection architectures.

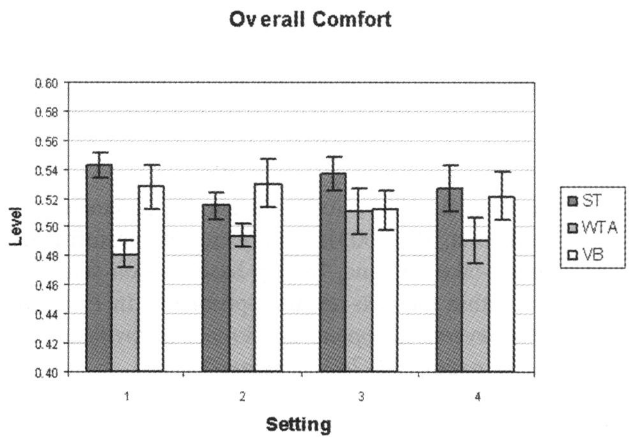

Figure 5: Experiment results of Overall Comfort (OvC) with standard errors.

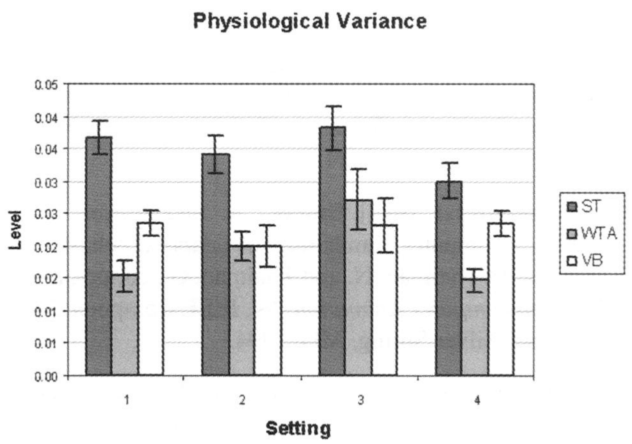

Figure 6: Experiment results of Physiological Variance (PhV) with standard errors.

(OvC). It is also observable from the experiment results that ST and WTA are reasonably stable in terms of the three indicators.

Since there are four environment settings with different *Apple Nutritional Rate*, they provide actual differences of cost-benefit scheme for the VB agent to execute the Eating Apple behavior:

- With setting (1), *Apple Nutritional Rate* is the lowest one, the agent simply ignores the existence of the apple and looks for the water and mushroom all the time, the reason for this phenomenon is that eating apple is relatively costly.
- With setting (2), it is very beneficial to eat the apple, the agent only looks for apple but ignores the existence of water and mushroom stimuli all the time.
- With setting (3), it is less beneficial to eat the apple, however the agent behaves roughly the same as with (1).
- With setting (4), *Apple Nutritional Rate* has been adjusted to be slight more than (1), which meets a balanced point and the agent consumes Apples, Mushroom and Water stimuli.

In (4) VB presents good opportunistic behavior as the agent 'goes for' all types of resources. This implies that when the agent encounters a resource, even if it was not looking for it, the agent will consume it. The more opportunistic behavior of VB in (4) is reflected in a LS increment, which is statistically higher than in the other three settings and also very close to that obtained by ST and WTA. It is interesting to note that, in terms of LS, WTA performs approximately the same with with all the settings, as the rate of resource consumption does not affect its ability to perform opportunistic behavior.

It can be observed that the VB agent in setting (4) is more 'open-minded' for executing other behaviors, which does not happen in other environment settings; we say that the agent in those settings is 'close-minded'. This also indicates that in the large size environment with more complex agent physiology and multiple resources possibly satisfying multiple needs, agents have to be open-minded – attempting to execute more different types of behavior, and taking the opportunities which offered by the environment. Opportunism has actually been characterized as one of the desiderata for behavior selection [Maes, 1995, Tyrrell, 1993]. In these experiments using a large and complex environment we see that opportunism indeed play an important role in the behavior selection process.

VB also obtained higher average satisfaction of their internal needs (OvC) than WTA in all the settings, equivalent to that of ST. Moreover, it obtained worse PhV than WTA although better than ST. This is because VB makes use of dynamic thresholds – like WTA –, although its more cost-effective policies improve the performance in terms of OvC while at the cost of more variance in the satisfaction of motivations (PhV).

	Searching for Water	Sensed (Move to) Water	Drinking Water	Stop Drinking	Leaving Water area
α	FALSE	FALSE	TRUE	FALSE	FALSE
β	TRUE	TRUE	FALSE	FALSE	TRUE

Table 5: Details of assigning the values to α and β under specific circumstances

5 Conclusions and Future Directions

We have presented an experimental study of the performance of different behavior selection architectures in a large and complex virtual environment with various environmental settings. The different viability indicators that we have used to measure the performance – *Life span*, *Overall Comfort* and *Physiological Variance* – showed fundamental differences between three behavior selection architectures: Static-Threshold (ST), Winner-Take-All (WTA) and Voting-Based (VB).

Our results illustrated that in terms of stimuli consuming cycles, WTA architecture produces the least amplitude thus it has good *Physiological Variances* comparing to ST architecture. They also confirmed that VB architecture is affected by the balance of environmental and physiological parameters for calculating the cost and benefit factors, in comparison to other architectures such as ST and WTA in order to be efficient.

There is an important problem which was created by the persistence of the behavior selection architecture and the conditions of the environment. Water stimuli from the environment includes the *River* and the *Lake* areas, when the WTA and VB agents stay in one of this areas, they carry on the behavior of *Drinking Water* and prefer staying in these areas even when there is a chance to move to another area, such as the *Desert* area which connects to the *Lake*. This problem seems to expose the hidden flaw of WTA and VB behavior selection architectures generated by extra persistence for their consummatory behaviors. We propose a solution for this problem by introducing new boolean parameters α and β to the motivation calculation process, as follow:

$$Motivation = Deficit + (Deficit * Cue) * \alpha + (Deficit * Cue) * \beta$$

The values of parameters α and β are intended to cancel the persistence effects, in the case here, when the agent is staying in the areas of water stimuli and intending to take the opportunity to execute the *Drinking Water* behavior successively in a short period of time. (The details of assigning the values to these two parameters under specific circumstances are shown in the Table 5.)

Another direction of future research work is to apply information about the environment gathered by the agents' interaction histories, which can be constructed from agents' significant experiences in their autobiographic memory [Nehaniv and Dautenhahn, 1998]; and this is possible for not only for VB, but also for ST and WTA behavior selection architectures. Our previous research work [Ho et al., 2003, Ho et al., 2004] showed evidences that agents with autobiographic memory can outperform reactive agents with static threshold architecture in both single agent and multi-agent experiments. Thus information about the environment in agents' autobiographic memory could be used as the *Cue* in the process of motivation calculations in regarding to this work. In addition to this, events in agents' autobiographic memory should also be weighted with respect to the significance of a particular event to the agent. We believe that the information available in agents' autobiographic memory would be able to compensate for the weaknesses of VB and enhance the performance of other architectures in complex environments.

Bibliography

[Avila-García and Cañamero, 2002] Avila-García, O. and Cañamero, L. (2002). A comparison of behavior selection architectures using viability indicators. In *International Workshop on Biollogically-Inspired Robotics: The Legacy of W. Grey Walter*, pages 86–93, Bristol HP Labs, UK.

[Avila-García et al., 2003] Avila-García, O., Cañamero, L., and Boekhorst, R. (2003). Analyzing the performance of "winner-take-all" and "voting-base" action selection policies within the two-resource-problem. In *Proceeding of the Seventh European Conference in Artificial Life (ECAL03)*, pages 733–742. Springer-Verlag.

[Blaxxun, 2004] Blaxxun (2004). Blaxxun technologies - products - blaxxun contact 5.1. http://www.blaxxun.com/en/products/contact/index.html.

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

[Bryson, 2000] Bryson, J. (2000). Hierarchy and sequence vs. full parallelism in action selection. In *From Animals to Animats 6: Proceedings of the Sixth Intl.Conf. on Simulation of Adaptive Behavior (SAB00)*, pages 147–156. Cambridge, MA: The MIT Press.

[Guillot and Meyer, 1994] Guillot, A. and Meyer, J.-A. (1994). Computer simulations of adaptive behavior in animats. In Thalman, N. and Thalman, G., editors, *IEEE Proc. on computer Animation '94*. IEEE Computer Society Press. Silver Spring, MD, 1994.

[Ho et al., 2003] Ho, W. C., Dautenhahn, K., and Nehaniv, C. L. (2003). Comparing different control architectures for autobiographic agents in static virtual environments. In *Intelligent Virtual Agents 2003 (IVA 2003)*, pages 182–191. Springer LNAI.

[Ho et al., 2004] Ho, W. C., Dautenhahn, K., Nehaniv, C. L., and te Boekhorst, R. (2004). Sharing memories: An experimental investigation with multiple autonomous autobiographic agents. In *IAS-8, 8th Conference on Intelligent Autonomous Systems*, pages 361–370, Amsterdam, NL. IOS Press.

[Maes, 1995] Maes, P. (1995). *Modeling Adaptive Autonomous Agents*, pages 135–162. Cambridge, MA: The MIT Press.

[McFarland, 1999] McFarland, D. (1999). *Animal Behaviour*. Addison Wesley Longman Limited, 3rd edition.

[Meyer and Guillot, 1990] Meyer, J.-A. and Guillot, A. (1990). Simulation of adaptive behavior in animats: Review and prospect. In *From Animals to Animats: Proceedings of the First Intl.Conf. on Simulation of Adaptive Behavior (SAB90)*, pages 2–14. Cambridge, MA: The MIT Press.

[Nehaniv and Dautenhahn, 1998] Nehaniv, C. L. and Dautenhahn, K. (1998). Embodiment and memories - algebras of time and history for autobiographic agents. In *14th European Meeting on Cybernetics and Systems Research*, Embodied Cognition and AI symposium, pages 651–656.

[Tyrrell, 1993] Tyrrell, T. (1993). *Computational Mechanism for Action Selection*. PhD thesis, Centre for Cognitive Sciences, University of Edinburgh, online at: http://www.cs.bham.ac.uk/~sra/People/Stu/Tyrrell.