



*Citation for published version:*

Rohlfshagen, P & Bryson, JJ 2010, 'Flexible latching: a biologically-inspired mechanism for improving the management of homeostatic goals', *Cognitive Computation*, vol. 2, no. 3, pp. 230-241.  
<https://doi.org/10.1007/s12559-010-9057-0>

*DOI:*

[10.1007/s12559-010-9057-0](https://doi.org/10.1007/s12559-010-9057-0)

*Publication date:*

2010

[Link to publication](#)

The original publication is available at [www.springerlink.com](http://www.springerlink.com)

## University of Bath

### Alternative formats

If you require this document in an alternative format, please contact:  
[openaccess@bath.ac.uk](mailto:openaccess@bath.ac.uk)

#### General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

#### Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

# Improved Animal-Like Maintenance of Homeostatic Goals via Flexible Latching\*

Philipp Rohlfshagen

School of Computer Science, University of Birmingham  
Edgbaston, Birmingham B15 2TT, United Kingdom  
p.rohlfshagen@cs.bham.ac.uk

Joanna J. Bryson †

Department of Computer Science, University of Bath  
Bath BA2 7AY, United Kingdom  
jjb@cs.bath.ac.uk

## Abstract

Controlling cognitive systems like domestic robots or intelligent assistive environments requires striking an appropriate balance between responsiveness and persistence. Basic goal arbitration is an essential element of low-level *action selection* for cognitive systems, necessarily preceding even deliberate control in the direction of attention. In

---

\*Research funded by the British Engineering and Physical Sciences Research Council (EPSRC) grants GR/S79299/01 and EP/E058884/1. A preliminary version of this paper appeared in the AAAI Fall Symposium on Biologically Inspired Cognitive Architectures [P. Rohlfshagen, J.J. Bryson, Improved Animal-Like Maintenance of Homeostatic Goals via Flexible Latching, in: Proceedings of BICA-08, 2008]. This version contains results omitted from the workshop paper, as well as further and improved expositions.

†The majority of this work was carried out when PR was in the Department of Computer Science, University of Bath, Bath, BA2 7AY, United Kingdom, and JJB was on sabbatical at The Konrad Lorenz Institute for Evolution and Cognition Research, Adolf Lorenz Gasse 2, A-3422 Altenberg, Austria. JJB is the corresponding author.

natural intelligence, chemically-regulated motivation systems focus an agent's behavioural attention on one problem at a time. Such simple durative decision state can improve the efficiency of artificial action selection by avoiding dithering, but taken to extremes such systems can be inefficient and produce cognitively-implausible results. This article describes and demonstrates an easy-to-implement, general-purpose latching method that allows for a balance between persistence and flexibility in the presence of interruptions. This appraisal-based system facilitates automatic reassessment of the current focus of attention by existing action-selection mechanisms. We propose a mechanism, *flexible latching*, and demonstrate that it drastically improves efficiency in handling multiple competing goals at the cost of a surprisingly small amount of extra code (or cognitive) complexity. We briefly discuss implications of these results to understanding natural cognitive systems.

**Keywords:** Action selection; drives; modularity; cognitive architectures

## 1 Introduction

The term *action selection* might seem to imply cognition, but this is merely due to anthropomorphic labelling. If we take *cognition* to be a process requiring time (probably a form of on-line search; [42]), and *action selection* to be any mechanism for determining the present course of action [11], then

6 much of action selection is really non-cognitive. Action choices in animals  
7 are limited both by evolution and individual skill learning; for adult animals  
8 many actions may be essentially reflexive [5, 7]. Such limiting is necessary if  
9 action selection is to be achieved in a timely manner [37, 15, 21]. However,  
10 there is no question that animals (including humans) do engage in cognition  
11 in some contexts. This article examines one such context: the arbitration  
12 between different goals. Even here, basic arbitration must necessarily be  
13 automatic. However, functional and efficient behaviour requires that the  
14 automated system can in some situations be interrupted and controlled cog-  
15 nitively [39]. Here we present a way to efficiently facilitate this capacity in  
16 artificial cognitive systems.

17 Budgeting time and pursuing multiple conflicting goals is a key aspect  
18 of any cognitive system [17, 22]. In the simulation of real-time animal-like  
19 intelligence considered in this paper, artificial agents must carry out a set  
20 of tasks, essential to their survival, while also interacting with dynamic sur-  
21 roundings, including other agents. Other-agent interactions in particular  
22 may include activities that are potentially essential to the species as a whole  
23 but not necessarily in the interest of the performing individual's viability.  
24 This characterisation might suggest rather dramatic activities, e.g. fending  
25 off attack, but it can also apply to ordinary duties. In some sense, the tasks  
26 that the system was originally designed to carry out (e.g. mating in nature,  
27 or perhaps tea making for an office robot) are of lower immediate priority  
28 than making certain that the system maintains working order, since working

29 order (e.g. the ability to move and manipulate) is a precondition of any  
30 other activity. Nevertheless, it is clear that we require an agent to devote  
31 considerable time to the goals that motivated its construction. Such critical  
32 but non-urgent goals are common amongst animals, such as maintaining a  
33 social network, reproducing or keeping clean. All these behaviours require  
34 both time and energy, and it follows that agents possessing more efficient be-  
35 haviour management should, in general, fare better than other agents with  
36 less efficient behaviour selection.

37 In this article, we demonstrate our goal-arbitration system using a simple  
38 artificial life task environment. Our agents must ensure they have the ability  
39 to store excess energy in order to pursue auxiliary behaviours. We discov-  
40 ered the need for an improved arbitration mechanism during the course of  
41 research on the evolution of primate social structures, so our examples de-  
42 rive from these models. The immediately urgent goals concern feeding, while  
43 the ultimately-important goals are social networking and exploration. Note  
44 that in nature such goals could also be considered survival-oriented, since  
45 socialising promotes long-term survival by facilitating group living [17, 25].  
46 However, their payoff is more diffuse — it is seldom knowable when addi-  
47 tional goodwill or information gathered may become critical, in contrast to  
48 starvation which has clear endogenous indicators. Thus we place essential  
49 behaviours at a high priority, but design an action-selection mechanism to  
50 ensure they are executed as efficiently as possible.

51 In this article we present a comparative study of three variants of a simple

52 action-selection mechanism designed to improve the agent’s capacity for goal  
53 arbitration. Our primary motivation is a potential inefficiency that may  
54 occur when an agent attempts to acquire a buffer of excess satisfaction before  
55 pursuing its next goal. We propose that if an agent is interrupted at any stage  
56 during this period, a choice needs to be made concerning whether to continue  
57 with the current goal or whether to attend to other, possibly more relevant  
58 behaviours. Persistence avoids the inefficiency of *dithering* between multiple  
59 goals. Dithering is inefficient because there is typically a significant start-  
60 up cost to pursuing new goals before consummatory actions can take place.  
61 However, some degree of flexibility avoids the inefficiency of pursuing a goal  
62 which is no longer urgent and has locally become excessively costly.

63 We look to biological motivation systems for inspiration because these  
64 have presumably evolved to manage this trade off. However, here we do not  
65 attempt a perfect or neurological model nature. Rather, our emphasis in this  
66 article is engineering. We present and evaluate a simple control mechanism  
67 that achieves the requisite level of flexibility at minimal cost. In fact, two  
68 types of costs are kept minimal: both the advance, coding-time costs for  
69 the agent’s designers and the real-time, cognitive-processing costs for the  
70 agents. We use a basic latching system augmented with the ability to detect  
71 potentially relevant interruptions. This threshold-based addition triggers a  
72 reevaluation of priorities already present in the agents’ overall action-selection  
73 system.

## 74 **2 Methods**

75 In this section we first describe the particular agent architectures we use  
76 to test our new goal arbitration system. Although we use a single system  
77 here, it is an example of a common type of action-selection system, and we  
78 describe the augmentation in general terms so that it may be applied on  
79 other systems as well. We then describe the specific goals to be manipu-  
80 lated in the experiments, and define the metrics of success in terms of these.  
81 Next, we describe the various latching mechanisms we have implemented for  
82 comparison. Finally, we describe the testing scenarios, including the agents'  
83 operating environment, followed by the presentation and discussion of our  
84 results.

### 85 **2.1 Basic Action Selection**

86 The agents are specified using the behaviour-oriented design (BOD) method-  
87 ology [12], a system that produces complete, complex agents consisting of (a)  
88 modules that specify details of their behaviour and (b) dynamic plans that  
89 specify agent-wide, cross-modular priorities. Actions are produced by the  
90 modules; action selection (where there is contention) is carried out using  
91 the Parallel-rooted, Ordered Slip-stack Hierarchical (POSH) dynamic plan  
92 system [10].

93 We chose BOD as a fairly simple example of an architectural consensus  
94 achieved in the late 1990s for real-time, situated systems: That AI is best

95 constructed using a combination of modularity, for providing intelligent prim-  
96 itives, and structured hierarchical plans, for encoding priorities [24, 26, 8].  
97 Even mainstream cognitive architectures such as Soar and ACT-R can be  
98 described in this way [28, 38]. Such approaches have been somewhat ne-  
99 glected in the academic literature in the last decade due to an emphasis on  
100 machine learning approaches to action selection. However, in applied human-  
101 like AI such as games programming and cognitive robotics, such modular,  
102 hand-coded approaches are still very much the norm [23, 31].

103 The details of the structured action-selection system are unimportant to  
104 the mechanism presented in this paper. All that is assumed is

- 105 • some mechanism for storing temporary values of long-term state  
106 (e.g. learning),
- 107 • some mechanism of expressing a variety of goals and their associated  
108 actions, and
- 109 • the notion of a trigger or precondition as part of the mechanism for  
110 choosing between goals and actions.

111 A single POSH plan was used to specify the priorities of all the agents  
112 tested here. That is, all the agents have the same priorities and therefore  
113 the same dynamic plan, though of course their expressed behaviour will vary  
114 due to their environment and their previous experience. What differs between  
115 conditions in the experiments described below are only the action-selection  
116 mechanisms and the testing environments.



117 The plan, shown in Figure 1, assumes four basic behaviours (drives):  $B_1$   
118 to  $B_4$ . In POSH, the top level of a plan hierarchy (the *drive collection*) is  
119 checked on every cycle of the controller. Control is passed to the highest-  
120 priority drive element whose trigger (line-labels in Figure 1) is true. All  
121 but behaviour  $B_4$  further contain a sub-plan, in POSH called a *competence*.  
122 Competences also contain elements each with their own trigger, but these  
123 are plans for the purpose of pursuing a single goal, and as such require less  
124 sophisticated scheduling than the drive collection. Competences maintain  
125 decision memory and control behaviour until they either terminate, pass  
126 control to a child competence of their own, or the main drive collection takes  
127 control back for a higher-priority problem. Their execution is similar to teleo-  
128 reactive plans [32] or indeed to the generalised plans created by STRIPS [18].

129 The first two behaviours, which are of the highest (and equal) priority,  
130 fulfil consumption-related needs, such as eating or drinking, the neglect of  
131 which would cause the agent to die. Behaviours  $B_3$  and  $B_4$  are of lower  
132 priority and are only considered for potential execution if  $B_1$  and  $B_2$  are  
133 not triggered. It should be noted that these behaviours are of lower priority  
134 simply because behaviours  $B_1$  and  $B_2$  are essential to the agent’s immediate  
135 survival. This does not imply, however, that lower-priority behaviours are  
136 not important, they could be critical to the agent’s mission. Since our experi-  
137 mental environment represents primate social behaviour, these behaviours in  
138 fact relate to increasing the probability of longer life. As such, behaviour  $B_3$   
139 represents social networking through grooming, which requires two agents to

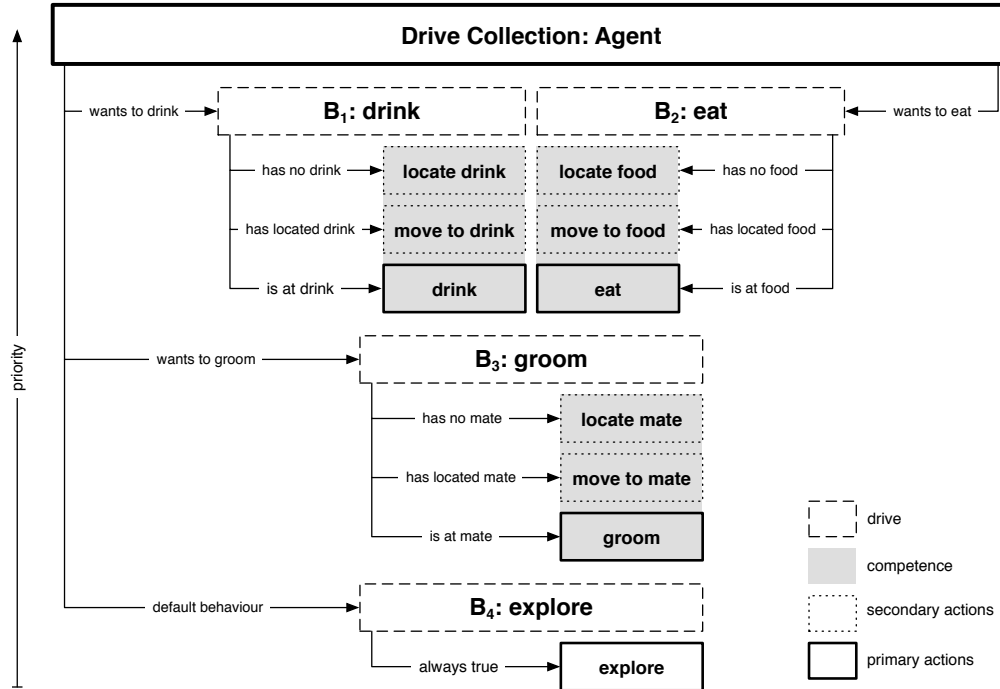


Figure 1: The POSH plan that determines priorities for the agents: the *drive collection* (SDC) is called at every time step and its elements checked in order:  $\{B_1=\text{eat}, B_2=\text{drink}\}$ ,  $\{B_3=\text{groom}\}$ ,  $\{B_4=\text{explore}\}$ . The highest-priority element whose trigger is true is executed. Equal priority elements (i.e.,  $B_1$  and  $B_2$ ) are checked in random order.

140 interact with one another. The final behaviour ( $B_4$ ) is exploration, possibly  
 141 to find new food sources. In a POSH plan, the lowest-priority goal serves as  
 142 a default behaviour and should always be triggerable. Thus if an agent with  
 143 this plan is efficiently arbitrating between goals, it should be able to spend  
 144 most of its time exploring new space.

## 145 **2.2 Metrics of Efficient Behaviour**

146 The primary focus of our investigations then is on behaviours  $B_3$  and  $B_4$ .  
147 Lower priority behaviours may only be executed if all higher priority be-  
148 haviours are managed efficiently and for artificial agents, the ‘lower’ be-  
149 haviours are typically the ones that define and justify the agent’s mission.  
150 Despite their significance these behaviours are necessarily of lower priority  
151 than those that facilitate the survival of the agent so it can perform these  
152 tasks. It is therefore paramount that these higher-level behaviours are man-  
153 aged efficiently enough to allow agents to pursue other behaviours as well.

154 Each behaviour is composed of numerous elements, some of which may be  
155 classified as *secondary actions*. In the case of feeding, the secondary actions  
156 would be ‘locating food source’ and ‘move towards food source’. The *primary*  
157 *action* would correspond to ‘eat’. For all behaviours, executing the primary  
158 action with a high frequency relative to the secondary actions determines  
159 the degree of efficiency with which the behaviour is executed. *Dithering*, the  
160 rapid switching between goals, results in secondary actions being performed  
161 excessively in proportion to primary ones. In our example, each behaviour  
162  $B_i$  has one such primary action which will be denoted as  $B_i^\alpha$ . The frequency  
163 at which primary actions are executed determines the degree to which all  
164 behaviours may be executed and thus defines the metric of success at the  
165 centre of our investigation.

## 166 2.3 Agents and State

167 Each behaviour  $B_i$  is associated a single-valued internal state  $E_i$ . Here, for  
168 the sake of clarity and without loss of generality, we use the concept of energy  
169 to denote the internal state of the agent: each behaviour  $B_i$  has a current  
170 level of energy  $E_i$ . The agents live in a toroidal, discrete-time world with  
171 dimensions of  $600 \times 600$  pixels. Time is considered to be discrete and at  
172 every time step, all agents in the environment are updated simultaneously.  
173 In particular, at every time-step, all energy states  $E_i$  are decreased by  $e_i^-$ .  
174 If a given behaviour is vital to the agent's survival, death is imminent once  
175  $E_i \leq 0$ . For each behaviour, we define a threshold  $\delta_i$  such that  $B_i$  is trig-  
176 gered once  $E_i < \delta_i$ . Once  $B_i$  is triggered, the agent will execute the actions  
177 associated with that particular behaviour. The behaviours  $B_1$  and  $B_2$  in our  
178 example correspond to sustenance activities (eating or drinking): The agent  
179 first locates an energy source, moves towards the energy source (at a speed  
180 of 2 pixels/time step) and consumes the source once in close proximity. This  
181 consumption raises the agent's internal state by  $e_i^+$ . Clearly we must ensure  
182 that  $e_i^+ \gg e_i^-$ ,  $\forall_i$  as otherwise an agent would never be able to satisfy a  
183 need (and in the case of essential behaviours, the agent would eventually  
184 die). Here we have chosen the same values for all behaviours:  $e^+ = 1.1$   
185 and  $e^- = 0.1$  and hence drop the behaviour-dependent subscript  $i$  from here  
186 on. Since we are interested in the execution of lower-priority behaviours, an  
187 individual choice of energy gain/loss across the different behaviours would  
188 require the adjustment of the individual thresholds (which are tightly related

189 to the net energy gain), unnecessarily complicating the model. Overall, this  
190 gives a net energy gain of  $e^\pm = 1$  for any primary action.

191 Lower-priority behaviours (i.e.  $B_3$  and  $B_4$ ) may only be executed if  $B_1$   
192 and  $B_2$  are satisfied. What it means for a behaviour to be ‘satisfied’ depends  
193 upon the implementation of the agents’ action selection — the basis of this  
194 article which we describe next.

## 195 **2.4 Conditions**

196 We use three different action selection mechanisms and evaluate their impact  
197 on the efficiency of the agent: unlatched, strict latch and flexible latch.

### 198 **2.4.1 Unlatched**

199 As mentioned in the previous section, a behaviour  $B_i$  is triggered if  $E_i < \delta_i$ .  
200 In the basic unlatched model, the drive terminates as soon as  $E_i \geq \delta_i$  and the  
201 time spent at the energy source is expected to be relatively short (although  
202 this depends strictly on  $\delta_i - E_i$  which may vary depending on the number of  
203 equal-priority behaviours). Furthermore, no excess energy is stored and the  
204 behaviour is triggered again very shortly after it is satisfied<sup>1</sup>. When there  
205 are multiple such behaviours, the agent will continue to oscillate between  
206 them (dithering). Even if there is only a single top-priority behaviour, the  
207 agent will spend its entire time in close proximity to the energy source as the

---

<sup>1</sup>The theoretical maximum possible excess energy in this case given the values of  $e^+$  and  $e^-$  is 0.9 which will last for 9 time steps.

208 acquired energy is always insufficient to pursue anything else.

### 209 **2.4.2 Strict latch**

210 In the latched models, the agent only terminates the drive once  $E_i \geq \phi_i$   
211 where  $\phi_i \geq \delta_i$ . Now the agent has an energy reserve of  $(\phi_i - \delta_i)/e^-$  time  
212 steps before the behaviour is triggered again. If all high-priority drives are  
213 latched in this way and the latch is sufficiently large (see next section), the  
214 agent is able to eventually follow lower-priority drives. This form of latching  
215 is very inefficient, however, if the agent inhabits a world where unexpected  
216 interruptions may occur. If an agent is almost finished with one activity but  
217 gets interrupted, the agent will continue to pursue this activity independent  
218 of other, lower-or-same priority needs. For example, an agent that is groom-  
219 ing and whose partner has left, might pursue another partner for five minutes  
220 when only another five seconds of grooming would have satiated it. This is  
221 true even if  $E_i = \phi_i - \epsilon$  where  $\epsilon \ll \phi_i - \delta_i$  and hence this form of latching is  
222 referred to as strict.

### 223 **2.4.3 Flexible latch**

224 If the agent is able to detect interruptions, the interruption could trigger  
225 a decision that determines its subsequent activities. Such a decision might  
226 be conscious, but here we simply relax the latching by using yet another  
227 threshold,  $\psi_i$ , that is situated in-between the previously two established ones,  
228  $\delta_i \leq \psi_i \leq \phi_i$ . This gives rise to two different scenarios. If the interruption

229 occurs when:

- 230 1.  $\delta_i < E_i < \psi_i$ , the drive remains ‘unsatisfied’
- 231 2.  $\psi_i < E_i < \phi_i$ , then the drive is considered ‘satisfied’

232 Note that for  $\delta_i < E_i < \phi_i$  the status of any latch is path or history dependent  
233 — if  $E_i$  was more recently below  $\delta$  the drive is now unsatisfied, if it was more  
234 recently satiated (about  $\phi$ ) than it is not. What is new for the flexible latch is  
235 that if an interruption occurs in the third scenario, where  $E_i$  had been below  
236  $\delta$  but has now been raised above  $\psi_i$ , this path dependency is dismissed.

## 237 **2.5 Threshold Selection**

238 The previous section has discussed different thresholds that require initialisa-  
239 tion and the choice of parameters is crucial to the outcome of the simulation.  
240 First, it should be noted that the flexible latch is simply a generalisation of  
241 the strict latch, which in turn is a generalisation of the unlatched technique:

$$\text{Flexible latch} \quad \delta \leq \psi \leq \phi$$

$$\text{Strict latch} \quad \delta \leq \psi = \phi$$

$$\text{Unlatched} \quad \delta = \psi = \phi$$

242 In this investigation, we have two primary points of interest, which are closely  
243 related: Survival and efficiency. The survival of the agent crucially depends  
244 on the choice of  $\delta$ . Efficiency, on the other hand, refers to the agent’s ability

245 to pursue all its behaviours, not just high-priority ones, and depends on the  
 246 choice of  $\phi$  and  $\psi$ . In order for an agent to survive, any vital behaviour must  
 247 be triggered such that the agent has enough energy to approach the energy  
 248 source (locating an energy source can be done in a single time-step and is  
 249 subsequently excluded from the following discussion):

$$\delta_i \geq \mathbb{E}_i^r \quad (1)$$

250 where  $\mathbb{E}_i^r$  is the energy required to reach the source:  $(d_{max}/d_{mov}) \times e^-$ , where  
 251  $d_{mov}$  is the distance an agent can move in a single time step and  $d_{max}$  is the  
 252 maximum possible distance an agent can travel<sup>2</sup>. If there are  $n$  equally vital  
 253 behaviours,  $\delta_i$  has to be adjusted accordingly:

$$\delta_i \geq \sum_{j=1}^{n-1} (\mathbb{E}_j^r + \mathbb{E}_j^c) + \mathbb{E}_j^r \quad (2)$$

254 where  $\mathbb{E}_i^c$  is the energy required to raise the energy level to the appropriate  
 255 level:

$$\mathbb{E}_i^c = \frac{\delta_i - E_i}{e^\pm} \quad (3)$$

---

<sup>2</sup>The theoretical maximum in this case is simply  $\sqrt{(width/2)^2 + (height/2)^2} \approx 424$  and it would take the agent a maximum of  $424/2=212$  time steps to reach the target, consuming  $212 \times 0.1 = 21.2$  units of energy.



256 The value of  $\phi$ , on the other hand, has to be set such that enough energy is  
 257 stored to pursue all vital needs:

$$\phi_i \geq \delta_i + \sum_{j=1}^n (\mathbb{E}_j^r + \mathbb{E}_j^c) \quad (4)$$

258 Any excess energy is subsequently devoted to the other, lower-priority be-  
 259 haviours. This choice of  $\phi_i$  necessarily affects  $\mathbb{E}_c$  as now more time is spent at  
 260 the energy source (a difference of  $\phi_i - \delta_i$ ). Interruptions drastically alter  $\mathbb{E}_c$   
 261 and the energy required to satisfy a latched behaviour given  $m$  interruptions  
 262 is simply:

$$\mathbb{E}_i^c = \sum_{j=1}^m (\mathbb{E}_{ij}^r + \mathbb{E}_{ij}^c) \quad (5)$$

263 At each interruption, the agent should, in theory, decide whether it is worth  
 264 pursuing the currently executed behaviour (i.e. if there is a positive or neg-  
 265 ative energy ratio). Usually there is insufficient knowledge available to make  
 266 an informed decision due of the complexity or indeterminacy of the environ-  
 267 ment. Consequently, heuristic values must be used. Nature selects for agents  
 268 with appropriate or at least adequate thresholds; here we test a range of  
 269 values for  $\psi$  to find which is appropriate for our particular simulations.

## 270 **2.6 Experiment and Simulation Details**

271 Our experiments are organised into two sets. The first set uses *sim1*, a very  
 272 well defined setup that allows a great degree of control over all aspects in-

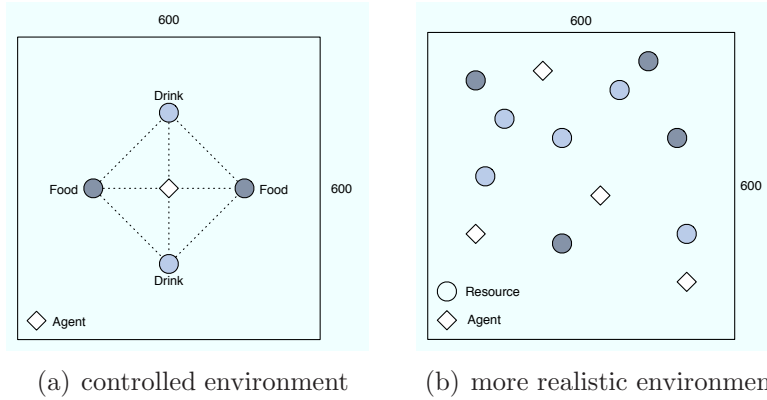


Figure 2: The two simulation environments used to test the overall efficiency of the agents: a completely controlled scenario (a) where energy sources are maximum distance apart, all agents are initially grouped at the centre and interruptions are externally induced, and a more realistic scenario (b) where agents and energy sources are placed randomly.

273 vestigated, particularly the frequency of interruption (see Figure 2(a)). The  
 274 second set use *sim2* (Figure 2(b)), a more realistic simulator where inter-  
 275 rptions are caused by the dynamics of the environment itself. For our  
 276 experiments we consider two types of interrupts. The first type occurs when  
 277 the source of satisfaction is depleted or otherwise removed (e.g., an agent  
 278 loses his current grooming partner). The second type of interrupt is caused  
 279 by higher priority drives that are triggered.

280 In both simulations, there are 5 identical agents. Furthermore, *sim1*  
 281 positions the energy sources such that they are maximum distance from one  
 282 another<sup>3</sup>. In this simulation, we exactly control the number of interruptions  
 283 an agent is exposed to throughout the execution of a single behaviour. Once

---

<sup>3</sup>The simulation is toroidal and agents are able to move, for example, from the far left to the far right in one move.

284 an agent is interrupted, it is forced to consider an alternative energy source  
285 (it is not allowed to remain at the current one). The second simulation is  
286 somewhat more realistic and is used to verify the results obtained from the  
287 first set of experiments. In *sim2*, energy sources are scattered randomly  
288 across the world. Each energy source has a certain load that depletes as an  
289 agent consumes it. Once depleted, the energy source vanishes, but, at the  
290 same time, a new energy source appears elsewhere in the world. The load  
291 of any energy source has a maximum of 50 units and depletes by 2 units if  
292 consumed. All energy sources gain 1 unit per time step.

293 The experiments are executed over 15 distinct trials. Each trial executes  
294 the simulation for 5000 time steps. All internal states are initialised such that  
295  $E_i = \delta_i$ , thus all behaviours are triggered immediately once the simulation  
296 begins. At each time step, the agent may execute a single action. The  
297 results are simply the number of times each primary action has been executed,  
298 averaged over all agents and trials. In all cases, a two-tailed t-test is used to  
299 test for significance with a confidence of 0.995. We chose the same threshold  
300 settings across all behaviours and again, we drop the subscripts from here  
301 on. Furthermore, we set  $\delta = 200$  in all experiments, giving an agent sufficient  
302 energy for  $200/e^- = 2000$  time steps before  $E$  falls to zero after a behaviour  
303 has been triggered.

action	no latch $\phi = \delta$	latched			significance		
		10	50	100	0-10	10-50	50-100
$B_1^\alpha$	443	452	478	494	*	*	*
$B_2^\alpha$	443	452	479	498	*	*	*
$B_3^\alpha$	0	0	454	468		*	
$B_4^\alpha$	0	0	1414	2037		*	*
total	886	903	2824	3498			

Table 1: Comparing latched and unlatched behaviours. The latches are chosen to be  $\phi - \delta \in \{0, 10, 50, 100\}$ .

### 304 **3 Results**

#### 305 **3.1 Controlled Environment: *Sim1***

306 The first experiment compares the unlatched version with the strictly latched  
307 one. The results are shown in Table 1. The data confirms that in the  
308 unlatched case, dithering prevents the agent from pursuing any of the lower  
309 priority behaviours. The latch effectively solves this problem, although only  
310 if the latch is sufficiently large. A latch of size 10 does increase the activity of  
311 the primary actions for behaviours  $B_1$  and  $B_2$  but still does not allow for the  
312 lower-priority behaviours  $B_3$  and  $B_4$  to be executed. Once the latch increases  
313 sufficiently in size, so does the activity of the lower-priority behaviours. This  
314 result is not surprising. Note though that too large a latch might also lead to  
315 neglect of lower-priority behaviours, since the highest-level goals might never  
316 be satisfied.

317 The next experiment investigates the efficiency of strict latching once  
318 an agent is confronted with interruptions. The data for this experiment

action	10			50			100			significance				
	1	3	5	1	3	5	1	3	5	0-1	0-3	0-5	1-3	3-5
$B_1^\alpha$	458	442	420	478	481	462	519	504	508	*		*	*	
$B_2^\alpha$	454	441	429	474	481	455	521	512	519	*		*		
$B_3^\alpha$	0	0	0	277	1	0	468	421	1		*	*	*	*
$B_4^\alpha$	0	0	0	95	0	0	1119	57	0	*	*	*	*	*
total	912	882	850	1324	962	917	2627	1493	1028					

Table 2: The performance of the agents given  $\phi - \delta \in \{10, 50, 100\}$  and 1, 3 or 5 interruptions. Significance is checked for  $\phi = 100$ . Cases without interruptions (0) are taken from the results shown in table 1 (not shown in this table).

319 is summarised in Table 2. Even in the case of a single interruption, the  
320 frequency of primary actions executed drops significantly. The right-most  
321 column in the table compares the performance of a latch of size 100 with 0,  
322 1, 3 and 5 interruptions and the differences for the lower-priority actions are  
323 almost always significant.

324 The final experiment using *sim1* determines the performance of the flexi-  
325 ble latch using the same settings as in the experiment before. Here, different  
326 values for the intermediate threshold  $\psi$  are tested. The value of  $\psi$  is denoted  
327 as the percentage of the latch itself. If, for example,  $\delta = 100$  and  $\phi = 120$ , a  
328 value of 25% would indicate that  $\psi = 105$ . The results are shown in Table 3  
329 and a setting of  $\psi = \delta$  seems most successful. However, as shown in Table 4,  
330 the differences are usually not significant. In the absence of significant differ-  
331 ence, the zero setting is still to be preferred as it also allows us to simplify the  
332 action-selection mechanism. We can effectively eliminate  $\psi$  altogether but  
333 *always* reconsider priorities when interrupted. Comparing the flexible latch

action	1				3				5			
	0%	25%	50%	75%	0%	25%	50%	75%	0%	25%	50%	75%
$B_1^\alpha$	499	491	489	501	490	491	496	496	482	487	482	495
$B_2^\alpha$	492	490	496	503	483	487	491	496	488	485	493	497
$B_3^\alpha$	481	476	479	481	475	479	469	455	474	470	462	437
$B_4^\alpha$	1723	1689	1528	1312	1458	1342	1059	651	1222	1150	880	495
total	3195	3146	2991	2797	2906	2799	2516	2098	2666	2592	2318	1923

Table 3: The performance of the agents with flexible latching.  $\psi = \delta + p(\phi - \delta)$  where  $p \in \{0, 0.25, 0.5, 0.75\}$ ,  $\delta = 200$ ,  $\phi = 300$  and frequency of interruptions equal to 1, 3 and 5. Significance of results shown in table 4.

	1			3			5			<i>vs. strict</i>		
	0-25	25-50	50-75	0-25	25-50	50-75	0-25	25-50	50-75	1-1	3-3	5-5
$B_1^\alpha$										*		*
$B_2^\alpha$										*	*	*
$B_3^\alpha$											*	*
$B_4^\alpha$			*		*	*		*	*	*	*	*

Table 4: Significance results for table 3. Increasing  $p$  has the most impact on the lowest-priority behaviour. The right-most column compares the strictly and flexibly latched implementation for the different frequencies of interruptions.

334 to the strict latch shows a significant improvement in at least one behaviour's  
335 primary action for any number of interruptions tested (compare Table 2 with  
336 Table 3; significance is indicated in the right-most column of Table 4).

337 Figure 3 shows graphically how the ability to detect interruptions im-  
338 proves the agent's overall efficiency. The graph plots the number of time  
339 steps spent executing the actions of interest given different frequencies of  
340 interruption. Furthermore, as a reference value, the unlatched and uninter-  
341 rupted latched cases are also shown. It is evident that the performance of  
342 the strict latch degrades very quickly while the flexible latch substantially

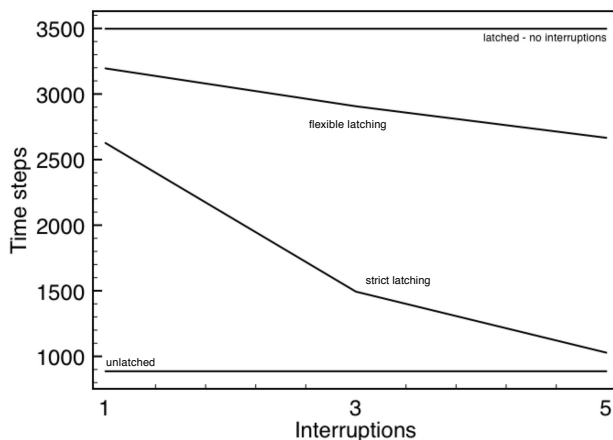


Figure 3: A graphical comparison of strict and flexible latching ( $\sum_{i=1}^4 B_i^\alpha$ ). The top and bottom lines are shown for reference, indicating the latched but uninterrupted and unlatched cases. For uninterrupted latches, the strict and flexible cases are indistinguishable.

343 reduces the impact of interruptions.

### 344 3.1.1 Death Rates

345 In the previous experiments, efficiency was judged by the capacity to devote  
 346 time to all behaviours. For these experiments, the value of  $\delta$  has been set  
 347 such that agents would always survive. In nature, such a threshold would  
 348 evolve in species like primates that invest a great deal in individual survival  
 349 and life histories. Nevertheless, exceptionally extreme environments or other  
 350 unusual circumstances may cause a threshold setting to become (temporarily)  
 351 insufficient.

352 In the present experiment, we set  $\delta$  such that survival in an uncertain  
 353 environment is no longer guaranteed ( $\delta = 40$ ). We then compare death

action	0	strict latch			flexible latch			significance		
		1	2	3	1	2	3	1-1	2-2	3-3
$B_1^a$	478	423	34	34	475	387	360	*	*	*
$B_2^a$	477	415	37	32	475	390	360	*	*	*
$B_3^a$	460	256	0	0	444	324	255	*	*	*
$B_4^a$	1402	90	0	0	750	295	140	*	*	*
Total	2816	1185	71	66	2144	1397	1115			
dead	0	601	4551	4551	0	861	1143	*	*	*

Table 5: A comparison of death rates for agents with lower values of  $\delta$  than are entirely sustainable in the environmental context. Tests are run with strict or flexible latching and with from 0–3 interruptions. Note again that without interruptions, whether the latch is flexible is irrelevant.

354 rates between strict and flexible latches. The latch is also set at a relatively  
355 low level of  $\phi = 45$ . The results are shown in Table 5. The flexible latch  
356 shows a significantly reduced death rate in all three relevant conditions (as  
357 determined by the number of interruptions). Furthermore, it is interesting  
358 to note that now, even with the smaller latch, the flexible implementation  
359 performs significantly better in almost all cases when compared to the strictly  
360 latched version.

361 Finally, it is possible to reduce the death rate even further. In another sce-  
362 nario we utilise the agents’ ability to deal with interruptions: Equal-priority  
363 behaviours are allowed to interrupt one another if they reach a critical thresh-  
364 old  $\psi$ . We set  $\psi = 20$ , as per the calculations described in Section 2.5 above.  
365 This critical threshold essentially corresponds to the minimum energy re-  
366 quired to satisfy a single need. The addition of the threshold changes the  
367 death rates from 0, 861, 1143 to 60, 417, 472. Interestingly, the death rate



368 is actually slightly higher in the first case but noticeable lower in the other  
369 two cases. The differences are relatively weakly significant for this  $N$ , with  
370 a confidence of  $p < 0.05$  for both the two- and three-interrupt conditions.

### 371 **3.2 Random Environment: *Sim2***

372 The previous results showed that in *sim1*, latching is necessary to allow  
373 an agent to execute lower-priority behaviours, and that it is best to abort  
374 a latched behaviour immediately upon interruption. We now examine these  
375 results in a system with a more “natural” setup using *sim2*, where the timing  
376 and frequency of interruption depends on the dynamics of the environment  
377 itself.

378 Table 6 compares all three implementation on *sim2*. The overall results  
379 are similar to before although there are some striking differences. Now, a  
380 latch of size 10 is sufficient to generate at least some frequency of execution  
381 for behaviours  $B_3$  and  $B_4$  whether or not it is flexible and indeed the flexi-  
382 bility makes no significant difference at this size latch. The change is due to  
383 the random environment providing more opportunities, which either imple-  
384 mentation is able to exploit. Once the size of the latch increases, flexibility  
385 creates a noticeable (as well as significant) difference for behaviour  $B_4$ , but  
386 no difference for  $B_3$ . This indicates  $B_3$ 's primary action is already executed  
387 sufficiently even without the flexibility in the latch — the flexibility in the  
388 environment provides sufficient opportunities for it to satiate at the threshold  
389 levels we've specified. Nevertheless, the massive increase of opportunity for

action	unlatched	strict latched			flexible latched			significance		
	0	10	50	100	10	50	100	10-10	50-50	100-100
$B_1^\alpha$	451	454	470	500	454	466	468			*
$B_2^\alpha$	452	454	475	490	455	466	469		*	*
$B_3^\alpha$	0	178	365	452	154	423	471			
$B_4^\alpha$	0	71	264	689	22	704	1289		*	*
total	903	1156	1574	2131	1084	2058	2697			
dead	0	0	0	0	0	0	0			

Table 6: Comparing the unlatched, strictly and flexibly latched implementations in *sim2* using latch sizes of  $\phi - \delta \in \{10, 50, 100\}$  and  $\psi = \phi$ . All cases have frequent interruptions (see main text).

390 expressing the exploratory behaviour shows the power of flexible latching.

## 391 4 Discussion

392 We have considered three variants of a simple threshold-based action selec-  
393 tion mechanisms. The completely unlatched condition may seem unrealistic,  
394 but several well-known reactive architectures have added latching only as an  
395 afterthought, handled with rather inelegant exception mechanisms [35, 16].  
396 Others assume latching can be handled by intelligent planning [6, 39]. This,  
397 however, requires a high cognitive load and in general, reasoning about time  
398 and distant rewards is difficult even for cognitive, symbolic systems [1].

399 The basic latched approach is inspired by theories of affect and action  
400 selection, as well as basic control theory. LeDoux [29] for example promotes  
401 the theory that emotions place the brain in a cognitive context appropriate  
402 for a particular course of action. Neuroscience tells us that interrupting such  
403 emotional responses is a cognitive capacity requiring frontal-lobe inhibition

404 of the emotional response [14]. Of course, the frontal-lobe inhibition system  
405 must itself be a fairly automatic gating mechanism. But this mechanism  
406 provides an opportunity for an alternative plan to become most salient [34].

407 Our system for determining appropriate thresholds for the flexible latches  
408 is also inspired by animal mechanisms through ethology. In particular, Dun-  
409 bar’s time-budget theory [17, 25] suggests that animal drives have evolved  
410 to ensure individuals are likely to spend the appropriate amount of time in  
411 behaviours, where *appropriate* is determined by what is adaptive. Our work  
412 here can be seen both as support for this theory and possibly as an elabora-  
413 tion, to the extent that our mechanism helps connect the time budget to the  
414 underlying neuroscience others have proposed (e.g. [34].)

415 In AI in contrast, there have been surprisingly few recent attempts to pro-  
416 pose general-purpose architectural features for homeostatic control. Those  
417 that exist tend to create detailed biomimetic representations of hormone lev-  
418 els [41, 27]. Gadanho [20] has a similar perspective to our work, using emo-  
419 tions to control the temporal expression of behaviour. However, she focuses  
420 on modelling specific emotions and their impact on reinforcement learning  
421 systems, rather than focusing directly on control mechanisms. In contrast,  
422 our flexible latch is simple to implement and incorporate into any standard  
423 module-based agent architecture. Also, she uses rising levels of emotions as  
424 the *source* of interruptions, rather than dealing with inefficiencies caused by  
425 interruptions generated by the external environment.

426 Interestingly, several established models of consciousness are similar to

427 our new model of flexibly-latched drives. Norman and Shallice [33] describe  
428 consciousness as a higher-cost attentional system which is brought on line  
429 whenever the more basic, reliable, low-cost action-sequencing mechanism is  
430 unable to proceed. Our system of flexible latching also operates by recognizing  
431 interruptions. It would be plausible in a system with modules capable  
432 of deliberation to have interruptions trigger these rather than the simple re-  
433 assesment of existing goals demonstrated above. More recently, Shanahan  
434 [36] proposes a model of mutually-inhibiting motives in a global workspace.  
435 We do not agree with Shanahan that such models can account for all of  
436 action selection. Tyrrell [40] provides provides an extensive critique of a  
437 very similar spreading-activation architecture, The Adaptive Neural Archi-  
438 tecture [30] (more commonly referred to as Maes' Nets [19]), explaining why  
439 spreading-activation models cannot scale to a full action-selection mecha-  
440 nism. The problem is simple combinatorics — a problem that architectures  
441 like ACT-R and IDA address by focussing on just one plan subset of the full  
442 network [19, 2]. This focussing makes these architectures functionally simi-  
443 lar to script-based dynamic-planning systems, although their actual action-  
444 selection mechanisms are far more complex. However, as this paper makes  
445 clear, we do think that a system like Shanahan's or Maes' could well account  
446 for high-level goal arbitration.

447 IDA is a cognitive architecture specifically designed to implement a the-  
448 ory of consciousness [3]. IDA is not only a model, but also a working AI  
449 architecture which has been used to create recommender systems for the US

450 Navy. Its newest version, LIDA provides the functionality of flexible latches  
451 through “timekeeper codelets” [4, p. 30] which keep a proposed action salient  
452 long enough for a variety of options to be debated. This system could well  
453 be effective, and is certainly more conducive to human-like meta cognition  
454 than the system proposed here. However, our flexible latches are simpler and  
455 probably sufficient for most autonomous AI applications.

456 The problems Tyrrell identified with spreading activation models are to  
457 some extent addressed by [22], who recommend generating a system of attrac-  
458 tors in the networks. This achieves an effect similar to the latching shown  
459 here. However, again the mechanism and architecture presented here are  
460 much simpler than spreading activation, even without the attractor system  
461 [9].

462 The difficulties in scaling spreading activation networks draw attention  
463 to an important limit of our work. Although we have shown substantial  
464 efficiency improvements, temporal costs still increase linearly with the num-  
465 ber of interruptions. Further, some forms of interruptions will necessarily  
466 increase with the number of potential behaviours — in particular those that  
467 are generated by the action-selection mechanism itself as higher priorities  
468 trigger. What this implies is that agents should have a limited number of  
469 high-level motivations which are contested this way.

470 What we present here is a cognitively-minimal mechanism which makes  
471 substantial improvements to an otherwise reactive action-selection system.  
472 Elsewhere, we explore in more detail the earlier suggestion that due to

473 LeDoux that the psychological entities called *drives* and *emotions* may be  
474 seen as a chemically-based latching system, evolved to provide persistence  
475 and coherence to the otherwise electrically-based action selection provided  
476 by the central nervous system [13]. We hypothesise that in nature, each  
477 drive or emotion — with its associated pattern of hormonal regulators and  
478 species-typical actions — might be viewed as serving one such high-level goal  
479 or need. We recommend that a system such as our flexible latch should simi-  
480 larly be used for each high-level goal an agent has that requires a time budget  
481 in an artificial cognitive system.

## 482 **5 Conclusions**

483 In this paper we have presented a relatively simple way to introduce flexible  
484 latching into an autonomous system and presented an analysis of how to de-  
485 termine appropriate thresholds that govern the execution of lower-priority be-  
486 haviours. The agents we considered have been specified using the behaviour-  
487 oriented design methodology: each agent consists of a set of modules that  
488 specify specific behaviours as well as a dynamic plan that prioritises amongst  
489 these behaviours. We take this as a fairly standard modular architecture us-  
490 ing scripted dynamic plans for action selection, and then demonstrate how  
491 to extend that action selection to improve its efficiency.

492 We demonstrate our system using four behaviours derived from a tool for  
493 modelling primate social behaviour. Two behaviours — eating and drinking

494 — are essential to the immediate survival of the agent and are of highest (and  
495 equal) priority. The third, grooming, represents a mission-critical behaviour  
496 though it is not essential for immediate survival. This and the fourth, default  
497 behaviour (exploring) can only be executed if the higher priority behaviours  
498 are managed efficiently. Each behaviour is composed of a number of indi-  
499 vidual actions and we distinguish between *primary* and *secondary* actions.  
500 Secondary actions are those required to perform the primary action; the pri-  
501 mary action is the core consumatory action of the behaviour and satisfies  
502 the agent’s need that triggers the behavioural module. Efficient execution  
503 of behaviours requires the agents to (a) minimise the execution of secondary  
504 actions, and (b) acquire sufficient satisfaction (energy in our case) to be able  
505 to carry out lower-priority behaviours.

506 The behaviour- (or action-) selection mechanism we have introduced con-  
507 sists of three thresholds: A lower threshold  $\delta$  that triggers the behaviour  
508 depending on the agent’s internal state, an intermediate threshold,  $\psi$ , that  
509 acts in case the agent is interrupted and an upper threshold,  $\phi$ , that causes  
510 the behaviour to terminate. The addition of these thresholds does not al-  
511 ter the priorities of the behaviours (which are still governed by the dynamic  
512 plan) but may delay (or not) the execution of lower-priority behaviours and  
513 may have a significant impact on the ratio of secondary to primary actions  
514 performed by the agent. We demonstrated their efficacy in two experimental  
515 settings. Without latching (i.e., only a lower threshold), the agent dithers  
516 between food sources, leaving no time to execute lower-priority behaviours.

517 Latching (i.e., lower and upper threshold) allows for *persistence* but may be  
518 hugely inefficient in the presence of interruptions. The persistent pursue of  
519 unsatisfied behaviours may lead to an unsustainable frequency of secondary  
520 task executions.

521 The experiments allowed us to determine the most useful setting for the  
522 intermediate threshold, above which an interrupted agent may reconsider its  
523 behaviour priorities. The results show that the utility of latching, as long  
524 as the latch is sufficiently large, where there is a significant cost of switch-  
525 ing between goals. *Flexible* latching addresses a reduction in performance of  
526 latches when there are interruptions. We found however that the interme-  
527 diate threshold is usually not required, or more precisely, can be set to be  
528 equal to the lower threshold. In our experiments, it was optimal for agents  
529 to reconsider priorities *whenever* interrupted. This result may not hold if  
530 interruptions are more frequent and/or the size of the latch is smaller, since  
531 either case would increase the probability that persistence is needed. Finally,  
532 we also explored the case where the agent may die if essential behaviours are  
533 carried out inefficiently. We found that latching significantly improves the  
534 rate of survival of the agent.

535 We have discussed how this mechanism, despite its simplicity, or because  
536 of it, may be relevant to numerous existing artificial cognitive architectures,  
537 and we have drawn parallels to animal-like decision making processes. Al-  
538 though the validation presented here is admittedly limited, these results do  
539 match expectations derived from our observations in nature concerning the



540 life-history strategies for species that tend to be correlated with more cog-  
541 nitive ability. At the same time, the work presented here also allows for  
542 extremely simple implementations such as hand-coding heuristic indicators  
543 of interruption.

544 There are numerous possible avenues to be explored in the near future.  
545 In our experiments, we chose the same thresholds for all behaviours, allowing  
546 a centralised approach that involves little overhead. However, it would be  
547 interesting to highlight potential differences in the efficiency of an agent's  
548 action selection when all behaviours have individual threshold settings. Fur-  
549 thermore, the thresholds may be adjusted dynamically over time (e.g., using  
550 a simple feedback control loop) or in artificial life contexts might be individ-  
551 ually evolved.

## Acknowledgements

Hagen Lehmann first identified that strict latching seemed inappropriate for modelling primate behaviour; his model provided the initial behaviour code that was extended for this research. The research was conducted by PR based on a suggested fix by JJB. This research was funded by the British EPSRC AIBACS programme, grant number GR/S79299/01 and EPSRC grant EP/E058884/1. It was first presented at the AAI 2008 Fall Symposium on Biologically Inspired Cognitive Architectures; we thank the organisers of that meeting, the other participants and the reviewers both for AAI and Cognitive Computation.

## References

- [1] G. Ainslie. Emotion as a motivated behavior. In *Agents that Want and Like: Motivational and Emotional Roots of Cognition and Action (AISB 2005)*, Hertsfordshire, 2005. The Society for the Study of Artificial Intelligence and the Simulation of Behaviour.
- [2] J. R. Anderson, M. P. Matessa, and C. Lebiere. ACT-R: A theory of higher level cognition and its relation to visual attention. *Human Computer Interaction*, 12(4):439–462, 1997.
- [3] B. J. Baars and S. Franklin. How conscious experience and working memory interact. *Trends in Cognitive Sciences*, 7(4):166–172, 2003.
- [4] B. J. Baars and S. Franklin. Consciousness is computational: The LIDA model of global workspace theory. *International Journal of Machine Consciousness*, 1(1):23–32, 2009.
- [5] J. A. Bargh and T. L. Chartrand. The unbearable automaticity of being. *American Psychologist*, 54(7):462–479, 1999.
- [6] R. P. Bonasso, R. J. Firby, E. Gat, D. Kortenkamp, D. P. Miller, and M. G. Slack. Experiences with an architecture for intelligent, reactive agents. *Journal of Experimental and Theoretical Artificial Intelligence*, 9(2/3):237–256, 1997.

- [7] R. A. Brooks. Intelligence without representation. *Artificial Intelligence*, 47:139–159, 1991.
- [8] J. J. Bryson. Cross-paradigm analysis of autonomous agent architecture. *Journal of Experimental and Theoretical Artificial Intelligence*, 12(2):165–190, 2000.
- [9] J. J. Bryson. Hierarchy and sequence vs. full parallelism in reactive action selection architectures. In *From Animals to Animats 6 (SAB00)*, pages 147–156, Cambridge, MA, 2000. MIT Press.
- [10] J. J. Bryson. Action selection and individuation in agent based modelling. In D. L. Sallach and C. Macal, editors, *Proceedings of Agent 2003: Challenges in Social Simulation*, pages 317–330, Argonne, IL, 2003. Argonne National Laboratory.
- [11] J. J. Bryson. Mechanisms of action selection: Introduction to the theme issue. *Adaptive Behavior*, 15(1):5–8, 2007.
- [12] J. J. Bryson and L. A. Stein. Modularity and design in reactive intelligence. In *Proceedings of the 17<sup>th</sup> International Joint Conference on Artificial Intelligence*, pages 1115–1120, Seattle, August 2001. Morgan Kaufmann.
- [13] J. J. Bryson and E. A. R. Tanguy. Simplifying the design of human-like behaviour: Emotions as durative dynamic state for action selection. *International Journal of Synthetic Emotions*, 1(1):30–50, January 2010.

- [14] N. R. Carlson. *Physiology of Behavior*. Allyn and Bacon, Boston, seventh edition, 2000.
- [15] D. Chapman. Planning for conjunctive goals. *Artificial Intelligence*, 32:333–378, 1987.
- [16] J. H. Connell. *Minimalist Mobile Robotics: A Colony-Style Architecture for a Mobile Robot*. Academic Press, Cambridge, MA, 1990. also MIT TR-1151.
- [17] R. I. M. Dunbar. Coevolution of neocortical size, group size and language in humans. *Behavioral and Brain Sciences*, 16(4):681–735, 1993.
- [18] R. E. Fikes, P. E. Hart, and N. J. Nilsson. Learning and executing generalized robot plans. *Artificial Intelligence*, 3:251–288, 1972.
- [19] S. Franklin. A “consciousness” based architecture for a functioning mind. *Visions of Mind: Architectures for Cognition and Affect*, pages 149–175, 2005.
- [20] S. C. Gadanho. *Reinforcement Learning in Autonomous Robots: An Empirical Investigation of the Role of Emotions*. PhD thesis, University of Edinburgh, 1999.
- [21] G. Gigerenzer and P. M. Todd, editors. *Simple Heuristics that Make Us Smart*. Oxford University Press, 1999.

- [22] P. Goetz and D. Walters. The dynamics of recurrent behavior networks. *Adaptive Behavior*, 6(2):247–283, September 1997.
- [23] N. Hawes, J. Wyatt, and A. Sloman. Exploring design space for an integrated intelligent system. *Knowledge-Based Systems*, 22(7):509–515, 2009.
- [24] H. Hexmoor, I. Horswill, and D. Kortenkamp. Special issue: Software architectures for hardware agents. *Journal of Experimental and Theoretical Artificial Intelligence*, 9(2/3), 1997.
- [25] A. H. Korstjens, I. L. Verhoeckx, and R. I. M. Dunbar. Time as a constraint on group size in spider monkeys. *Behavioral Ecology and Sociobiology*, 60(5):683–694, 2006.
- [26] D. Kortenkamp, R. P. Bonasso, and R. Murphy, editors. *Artificial Intelligence and Mobile Robots: Case Studies of Successful Robot Systems*. MIT Press, Cambridge, MA, 1998.
- [27] J. L. Krichmar. The neuromodulatory system: A framework for survival and adaptive behavior in a challenging world. *Adaptive Behavior*, 16(6):385–399, December 2008.
- [28] J. E. Laird and P. S. Rosenbloom. The evolution of the Soar cognitive architecture. In D. M. Steier and T. M. Mitchell, editors, *Mind Matters*. Erlbaum, 1996.

- [29] J. LeDoux. *The Emotional Brain: The Mysterious Underpinnings of Emotional Life*. Simon and Schuster, New York, 1996.
- [30] P. Maes. The agent network architecture (ANA). *SIGART Bulletin*, 2(4):115–120, 1991.
- [31] I. Millington and J. Funge. *Artificial Intelligence for Games*. Elsevier, 2009.
- [32] N. J. Nilsson. Teleo-reactive programs for agent control. *Journal of Artificial Intelligence Research*, 1:139–158, 1994.
- [33] D. A. Norman and T. Shallice. Attention to action: Willed and automatic control of behavior. In R. Davidson, G. Schwartz, and D. Shapiro, editors, *Consciousness and Self Regulation: Advances in Research and Theory*, volume 4, pages 1–18. Plenum, New York, 1986.
- [34] P. Redgrave, T. J. Prescott, and K. Gurney. Is the short-latency dopamine response too short to signal reward error? *Trends in Neurosciences*, 22(4):146–151, 1999.
- [35] K. Rosenblatt and D. Payton. A fine-grained alternative to the subsumption architecture for mobile robot control. In *Proceedings of the IEEE/INNS International Joint Conference on Neural Networks*, Montreal, 1989. Springer.
- [36] M. P. Shanahan. Global access, embodiment, and the conscious subject. *Journal of Consciousness Studies*, 12(12):46–66, 2005.

- [37] H. A. Simon. Theories of bounded rationality. In C. B. Radner and R. Radner, editors, *Decision and Organization*, pages 161–176. North Holland, Amsterdam, 1972.
- [38] A. Sloman. What sort of architecture is required for a human-like agent. *Foundations of Rational Agency*, pages 35–52, 1999.
- [39] R. Sun. Motivational representations within a computational cognitive architecture. *Cognitive Computation*, 1(1):91–103, March 2009.
- [40] T. Tyrrell. An evaluation of Maes’s bottom-up mechanism for behavior selection. *Adaptive Behavior*, 2(4):307–348, 1994.
- [41] P. Vargas, R. Moioli, L. N. de Castro, J. Timmis, M. Neal, and F. J. Von Zuben. Artificial Homeostatic System: A Novel Approach. *Eighth European Conference of Advances in Artificial Life, ECAL*, pages 754–764, 2005.
- [42] P. H. Winston. *Artificial Intelligence*. Addison-Wesley, Boston, MA, 1984.