Discovering the competitors

Luc Steels Artificial Intelligence Laboratory Vrije Universiteit Brussel E-mail: steels@arti.vub.ac.be (Journal of Adaptive Behavior 4(2) 1996)

January 11, 1996

Abstract

This paper reports on an experiment testing whether a particular representation of robotic control processes is adequate for capturing significant variations in robot behavior. These variations can then be explored by a selectionist mechanism that generates and tests variations. An ecosystem modeled after a physical robotic ecosystem is introduced. The ecosystem contains a robot that occasionally has to recharge, as well as competitors that take away energy from the total system. The robot has to discover that its viability requires combatting the competitors.

1 Introduction

Behavioral development implies the incremental acquisition of new behavioral competences by an agent while remaining viable in an unknown dynamically changing environment. Several approaches to behavioral development have been discussed in the literature, including supervised neural networks that generalise or form associations based on a series of examples [14], reinforcement learning algorithms [6] that do the same based on a reinforcement signal, and genetic algorithms that perform a parallel selectionist search through a series of alternative behavioral networks [5]. Selectionist experiments have also been reported by Edelman for categorisation competence [3] and Holland for rule-based decision making as implemented by classifier systems [4].

This paper contributes to the selectionist paradigm. It focuses on the acquisition of competence for behavior regulation, i.e. the control of which behavior will be performed, rather than the acquisition of new behavior (as in [11]). We seek a mechanism whereby such competence can be acquired on-line and while the robot stays viable. This requires that a representation must be found such that a variation in this representation captures a significant difference in behavior. Second a search mechanism must be found that performs the exploration of alternatives without endangering the viability of the agent. This paper focuses on the first aspect. It proposes representations capturing dynamic couplings between quantities and tests whether these representations adequately capture significant variations needed to handle a concrete challenging example ecosystem. Another paper [13] focuses on a possible search mechanism and shows for the same example that an exploration while remaining viable is possible.

The rest of the paper has the following parts. The first part introduces the ecosystem. The second part discusses the control system on the robot as well as the selectionist mechanism. Then variations of behavioral programs linking alignment with perceived energy level are studied. One of these variations turns out to yield the most adapted behavior. Some conclusions end the paper.

2 The ecosystem

The experiments discussed in the paper involve an abstract ecosystem, modeled after a physical ecosystem built in the laboratory (fig.1). The ecosystem contains a robot (fig.2) and a charging station. The robot has a battery which is depleted as it operates and moves around in the environment. The robot can recharge itself by sliding into the charging station. The ecosystem also features competitors. The competitors take away energy from the global



Figure 1: Robotic ecosystem that is assumed for the experiments. The robots have a battery, a battery sensor, and photosensors. They can control their own forward movement and slide into the charging station on which a lamp is mounted.



Figure 2: Physical robot modeled in the experiments. The robot has a charging rod on top which is linked to a contact switch. When the robot slides in the charging station, this switch is closed.

energy flowing into the system and thus from the charging station. In the physical ecosystem the competitors take the form of lamps mounted in boxes (fig.3). When the robots run against the boxes, the competitors diminish in strength but they regrow later. Thus the robot has to work continuously to get sufficient energy in the charging station. This ecosystem was developed in collaboration with David McFarland and has been designed to reflect many essential characteristics of natural ecosystems.

The internal control processes on the robot use a behavior-oriented approach [9], which means that behavior systems as opposed to actions are the primary building blocks. Different behavior systems operate in parallel and the observed behavior is the sum of all their influences. The intensity with which a behavior system impacts global behavior is controlled by a motivational parameter which is itself regulated using a dynamical system [12].



Figure 3: Competitors take the form of lamps mounted in a box. When the robot pushes against a box, the lamp diminishes and thus takes away less energy from the global energy flowing into the ecosystem.

2.1 Conventions

The definition of the ecosystem uses the following conventions. We will need a set of quantities $Q = \{q_1, ..., q_n\}$. The value of a quantity is bounded. $0.0 \le q_i \le 1.0$ with some exceptions. The modeling and simulation assumes a discretisation of time. The value of a quantity at time t is denoted as q(t, i). t is left out if the value intended is the current time. i is left out when we just want to talk about a quantity in general.

A process is written as:

$$q_j \longleftarrow v_k \quad if \ p_k \tag{1}$$

which is to be read as " q_j increases with v_k when p_k " where v_k is a number or a numerical expression and p_k is a proposition whose truth value determines whether the increase or decrease should happen. Processes form a process network because the influenced quantity q_j of one process may itself be the source of another process. Note that there is an implicit parallelism. All processes run at the same time. The quantities are shared variables and changed through side-effect. A further notational shorthand is the following:

$$q \Longleftarrow v \quad if \ p \tag{2}$$

which is to be read as "q becomes v when p" where v is a number or a numerical expression. If the proposition p is true at time t, the value of q will be attempted to be bridged at the next time instant by the difference between v and the current value of q. The above expression is therefore equivalent to:

$$q(t+1) \longleftarrow (v - q(t)) \quad if \ p \tag{3}$$

There is no guarantee that v is reached because another process may also perform a change to another value. In general, no single process has ever control, neither over another process nor over the final state of a quantity.

One special kind of proposition is

$$q \ll v \tag{4}$$

which is to be read as "q changes into v", where v is a number or a numerical expression denoting a number. This proposition is true if there is a change from any kind of value into the value v.

2.2 Initial situation

The initial ecosystem (without competitors) contains a single robot and a charging station. The robot starts out with two behaviors: one for forward movement and one for alignment towards a light source using photosensors. The operation of the ecosystem and these behaviors is governed by the following laws:

[1] When the agent is moving forward it uses energy at a certain rate r_1 . Let f be the speed of forward movement and e the energy level then this is captured in the following process. By default $r_1 = 0.1$.

$$e \leftarrow -r_1 f \qquad [P1] \tag{5}$$

[2] There is a constant energy loss at rate r_2 . By default, $r_2 = 0.01$.

$$e \leftarrow -r_2 \quad [P2] \tag{6}$$

[3] When the agent has access to the charging station there is an increase in energy at the rate r_3 . Let c be 0.0 if there is contact and 1.0 if there is no contact.

$$e \leftarrow r_3 c \qquad [P3] \tag{7}$$

[4] The motivation of the forward movement behavior is regulated by the quantity b_f . There is a rate r_4 which determines the success with which this motivation is translated into actual forward movement. The success is also dependent on the amount of energy available. This leads to the following process definition. By default $r_4 = 1.0$.

$$f \Longleftarrow r_4 \ b_f \ e \qquad [P4] \tag{8}$$

[5] The forward movement behavioral motivation decreases towards zero (the resting state of the robot), meaning that if there are no forces pushing towards forward movement, there should be a tendency towards halting. This is captured in the following process definition, where r_5 is the rate at which the behavioral motivation decreases. By default $r_5 = 0.2$.



Figure 4: Main parameters determining the geometrical layout of the charging station and its surrounding area. d determines the depth of the charging station, m the minimal area which needs to be traversed after leaving the station, and n the typical distance to be traversed before the station can be entered again. j,k, and l model how far these quantities have been traversed.

$$b_f \longleftarrow -r_5 \ b_f \qquad [P5] \tag{9}$$

[6] The charging station is assumed to have a certain depth d (fig.4). The robot slides into the charging station and will slide through it depending on its forward speed. This means that we must model the robot's position in the charging station or more precisely the distance of joined overlap j between the robot's charging apparatus and the charging station. When the robot enters the charging station j is set equal to d. Then j evolves according to the next process definition. By default, $r_6 = 0.2$.

$$j \leftarrow -r_6 f j \qquad [P6] \tag{10}$$

[7] As long as j is positive, contact remains established:

$$c \Longleftarrow \begin{cases} 1.0 & \text{if } j > 0.0 & [P7] \\ 0.0 & \text{otherwise} \end{cases}$$

(11)

[8] r_6 is the travelling rate also used for sliding through the charging station. k = m when the robot leaves the charging station.

$$k \longleftarrow -r_6 f \qquad [P8] \tag{12}$$

[9] I models the distance to be covered to reach the charging station once the corridor around the station has been left. I decreases at the same rate as k or j:

$$l \longleftarrow -r_6 f \qquad [p9] \tag{13}$$

[10] The different distances j, k, and l are set to their corresponding maximum values by the following processes:

$$j \Longleftarrow d \quad if \ c \ll 1.0 \qquad [p10] \tag{14}$$

$$k \longleftarrow m \quad if \ j \ll 0.0 \qquad [p11] \tag{15}$$

$$l \Leftarrow n \ if \ k \ll 0.0 \qquad [p12] \tag{16}$$

[11] There is a contact sensor mounted on the charging rod which switches on when contact is made with the charging station:

$$s_c \Leftarrow c \quad [P13] \tag{17}$$

[12] The next processes are relevant for the alignment behavior. The first one converts the alignment motivation (b_a) to actual alignment, which will depend on the available energy in the batteries e and on a rate $r_4 = 1.0$:

$$a \Longleftarrow r_4 \ ba \ e \qquad [P14] \tag{18}$$

[12bis] alignment also uses up energy:

$$e \leftarrow -r_1 a \quad [P14bis] \tag{19}$$

[13] The traversal of l (determined by process [p9]) is made to depend on the alignment as well:

$$l \leftarrow -r_6 a \quad [p15] \tag{20}$$

2.3 Ecosystem with competitors

The ecosystem defined in the previous subsection is now extended. Not only are competitors added but several non-linear effects are deliberately introduced to make the ecosystem more interesting. For example, recharging will no longer be linear but dependent on the level of the battery (as is indeed the case with physical batteries).

First a global source of energy g is introduced. The global source is constantly replenished. It is maximally equal to 1.0 and replenishment goes slower as this limit is reached. The charging station takes energy away from this global energy source if the robot is charging and the competitors take from it as well. The robot can only sense indirectly the impact of the competitors when (a) it is in the charging station and (b) the competitors are sufficiently strong to cause a decrease of g, and hence a decrease of the amount of energy available to the robot as it is recharging.

The competitors are distributed in the environment around the charging station and at first the robot will hit a competitor by pure chance. The probability with which a competitor is hit depend on the distance to the charging station. This probability depends also on the number of competitors p. The ecosystem can be made tougher for the robot by making the competitors harder to find.

These considerations are translated in the following set of processes. First, process [P3] which regulates the energy increase when the agent is in the charging station is revised so that it is in relation to the energy in the global energy source. The process is also made non-linear.

$$e \leftarrow r_3 \ c \ g \ (1.0 - e) \qquad [P3] \tag{21}$$

There is a corresponding loss in the global energy source:

$$g \longleftarrow -r_3 \ c \ g \ (1.0 - e) \qquad [P3bis] \tag{22}$$

[13] The competitors take away energy from the global energy source g at the rate r_7 .

$$g \leftarrow -r_7 p g \qquad [P16] \tag{23}$$

[14] The competitors grow non-linearly at the rate r_8 :

$$p \leftarrow -r_8 (1.0 - p) \qquad [P17] \tag{24}$$

[15] The competitors non-linearly diminish at the rate r_9 when the robot touches them. The more competitors there are the more effect touching will have. Touching is captured in the binary quantity c2. c2 is not a sensor but a property in the world (just like c).

$$p \leftarrow -r_9 c_2 p \quad [P18] \tag{25}$$

[16] Hitting a competitor slows the robot down due to the obstacle avoidance necessary for getting away from a competitor:

$$f \longleftarrow -r_{10} c_2 p \quad [P19] \tag{26}$$

[17] Hitting a competitor depends probabilistically on the distance away from the charging station l, the speed of forward movement f, and a probability factor r_{11} .

$$c2 \longleftarrow prob(r_{11}, l \ f) \quad [P19] \tag{27}$$

where

$$prob(\theta, n) = \begin{cases} 1.0 & \text{if } random(n) < n/\theta \\ 0.0 & \text{otherwise} \end{cases} [P7]$$

(28)

The robot must have the physical means to detect the competitors at a distance so that a second type of alignment can be used. In the physical implementation this has been accomplished by using infrared light strobed

at a certain frequency as 'lamp' and two infrared sensors mounted on the left and right side of the robot's body which filter for the same frequency. These sensors make a second alignment behavior a_2 possible whose behavioral motivation is controllable with the parameter b_{a2} . The following processes (which are analogous to those for the first kind of alignment) need to be added:

[18] The first process converts the alignment motivation (b_{a2}) to actual alignment, depending on the available energy in the batteries e and on a rate $r_4 = 1.0$:

$$a_2 \Longleftarrow r_4 \ b_{a2} \ e \qquad [P20] \tag{29}$$

[19] The second process models the energy used by alignment:

$$e \longleftarrow -r_1 \ a_2 \qquad [P21] \tag{30}$$

[20] The alignment motivation also moves to a quiescent state, just like forward movement. By default $r_5 = 0.2$.

$$b_a \longleftarrow -r_5 \ b_a \qquad [P22] \tag{31}$$

3 The selectionist mechanism

3.1 Representation of control processes

The control processes on the robot can be viewed as dynamically coupled maps. A dynamically coupled map $DCM = \langle C, F, U \rangle$ is an extension of coupled maps which are a well researched formalism in complex systems theory [7]. It consists of a set of quantities C with associated values, a set of functions (maps) F, and a function assignment U mapping a pair of quantities [i, j] to a function at a particular time: $U: C \times C \times T \to F$. The functions define processes. F includes Z, a constant function $Z(x_i, x_j) = 0.0$ which is assigned by default. Note that a function always takes two arguments involving a target quantity i and a source quantity j. Updating is governed by the following equation:

$$x_i(t+1) = \sum_{j=1}^{N} f(i, j, t)(x_i(t), x_j(t))$$
(32)

The kinds of couplings that may occur between quantities is restricted. In this paper the following couplings are used:

The similarity-coupling **S** couples the values associated with the quantities i and j in such a way that the value of i approaches that of j. It is a kind of goal-seeking coupling. The goal for i is given by the value of the other quantity j. r determines the rate at which j should be approached.

$$\mathbf{S}(i,j) \equiv x_i(t+1) = r[x_j(t)) - x_i(t)]$$

$$(33)$$

The reverse-coupling \mathbf{R} couples the values associated with the quantities i and j in such a way that they attain opposite values with respect to the maximum value 1.0. It is another kind of generalisation of a goal-seeking coupling.

$$\mathbf{R}(i,j) \equiv x_i(t+1) = r[(1.0 - x_j(t)) - x_i(t)]$$
(34)

Couplings can be made dependent on the evolution of the source quantity. Two such conditions will be further explored in this paper:

• The coupling can be dependendent on an increase I in the source quantity:

$$I(j) \iff x_j(t-1) < x_j(t) \tag{35}$$

• The coupling can be dependent on a decrease **D** in the source quantity:

$$D(j) \iff x_j(t-1) > x_j(t) \tag{36}$$

The possible combinations of S, R, I and D are illustrated in fig. 5.

I call the representation of a coupling in the memory of the robot a *bene*, to suggest an analogy with genes. Indeed, genes in some sense also represent couplings because they co-determine the presence of enzymes which regulate the nature and rate of processes in a metabolic process network. Sets of benes are grouped in *strands*. One strand represents a set of couplings and thus a process network.



Figure 5: q_1 will follow q_2 in similar motion when a similarity coupling (S) exists between q_1 and q_2 or in reverse motion (R) when a reverse coupling exists. The figure shows the different possibilities depending on whether the source quantity is increasing (I) or decreasing (D).

A bene is represented in the following way:

$$q_1 - d_1 - d_2 - q_2 - r \tag{37}$$

where q_1 and q_2 are quantities (often called target and source quantities respectively), $d_1 \in \{S, R\}$ expresses how q_1 will follow the change in q_2 . $d_2 \in \{D, I\}$ determines the required direction of change of q_2 . r is the rate. The default value of r is r=1.0, in which case r is not written. The above expression is equal to

$$q_1 \Longleftarrow rw \quad if \ dq_2/dt < 0 \tag{38}$$

if d_2 is D (decreasing) or

$$q_1 \longleftarrow rw \quad if \ dq_2/dt > 0 \tag{39}$$

if d_2 is I (increasing).

$$w = \begin{cases} q_2 & \text{if } d_2 = S \\ 1.0 - q_2 & \text{if } d_2 = R \end{cases}$$

(40)

Assuming two quantities: b_f for the motivation of forward movement and s_e for the sensed energy level, then the coupling $b_f - R - D - s_e - r$ [b1] reads as "forward movement reversely follows the decreasing sensed energy level with rate r". It is equivalent to the following process:

$$b_f \Longleftarrow r(1.0 - s_e) \quad if \ d \ s_e/dt0 \tag{41}$$

by default r = 1.0.

3.2 The search process

The following mechanism is proposed for regulating the search in the space of possible variations [13]. There is a *current strand* C which represents the best process network so far. At regular intervals, benes and corresponding strands are constructed for each of the possible combinations of two quantities which are randomly selected. Only four possibilities are considered in this paper: follow in similar motion if increasing (S - I), follow in reverse motion if increasing (S - D), follow in similar motion if decreasing (S - D) and follow in reverse motion if decreasing (R - D). Each of these is tried very briefly in combination with the current strand C giving an initial ranking. The possible combinations might not be better than what already exists on the robot in which case they are ignored and other combinations are tried. If not, the ranking is retested for a longer time period and the best strand is combined with the current strand to form the new current strand. This current strand is tested again with the others to see whether a further combination is beneficial, and so on until no better performance from recombination is seen. The procedure then repeats itself for other possible combinations of quantities.

The notion of *viability* [1] is used as selection criterion. A behavior is viable when it keeps the organism within a region of its physical parameter space that guarantees continued existence [8]. During a specific time period, a region is carved out by actual behavior. When this region moves towards the critical boundaries of the viability region, then the corresponding behavior should undergo negative selection pressure.

For example, one of the most important critical parameters for a robotic agent is the level of its internal energy resources. The ideal situation occurs when this level is at its maximum. During a particular time period T the robot's energy level fluctuates between a minimum min and a maximum level max. The center of mass is defined as center(T) = min(T) + (max(T) - min(T))/2. If this center moves up, the robot becomes more viable and there should be a positive selection for this behavior. In real world robots, many other dimensions are to be considered as well. The experiments in this paper however focus on energy maintenance only.

4 Experimental results

The goal of this paper is to investigate the adequacy of benes as a representation of the dynamics controling motivational variables and thus behavior regulation. The representation is adequate if it captures competing alternatives. A number of experiments have been done testing variations. Each time the impact on viability was studied to see whether one variation would stand out. The results for handling the initial situation are reported only briefly. The results for handling the competitors, which is the most challenging task, are discussed in more detail.

4.1 Initial Situation

Forward movement and energy level

There are four possibilities relating forward movement behavior to energy level:

```
[b1] bf - R - I - se ; reversely follow increasing energy level
[b2] bf - R - D - se ; reversely follow decreasing energy level
[b3] bf - S - I - se ; follow increasing energy level
[b4] bf - S - D - se ; follow decreasing energy level
```

The only combination improving the viability of the robot is one where the motivation of the forward movement behavior follows the decreasing perceived energy level: $b_f - S - D - s_e$ [b4]



Figure 6: The different combinations between the forward movement motivation and the energy level sensor are tried for 100 time steps each. b4 (the last one tried) clearly gives the best performance. The robot slows down when going to the charging station. The forward movement shown is actual forward movement and not motivation for forward movement. Actual forward movement depends on available energy and will therefore always decrease when energy decreases.

This way the robot decreases its forward speed as the energy level decreases. Other combinations have no or only a slight effect (fig.6). [b4] is the most beneficial because this way the robot slows down while approaching and thus sliding through the charging station. It also optimises energy use when it is most scarce.

Alignment and energy level

The following couplings are possible between motivation for alignment (b_a) and sensed energy level:

 $b5 = [b_a - R - I - s_e]$; reversely follow increasing energy $b6 = [b_a - R - D - s_e]$; reversely follow decreasing energy



Figure 7: The different combinations between the align motivation and the energy level sensor are shown for 100 time steps each. Actual energy level and alignment are shown. Success of alignment depends on forward movement as well and on the available energy. [b6] is the winner although the others except the first are also beneficial.

 $b7 = [b_a - S - I - s_e]$; follow increasing energy $b8 = [b_a - S - D - s_e]$; follow decreasing energy

Experimenting with these couplings leads to a preference for the following: b_a - S - D - s_e [b6]

The motivation of alignment follows the decreasing energy level. This is the most appropriate behavior because the motivation to perform alignment to the charging station goes up when the energy is getting low (fig. 7).

Forward movement and the contact sensor

The following couplings are possible between forward movement motivation (b_f) and the contact sensor

 $b9 = [b_f \mbox{ - R - I - } s_c]$; reversely follow increasing contact



Figure 8: The different combinations of the forward movement motivation and the contact sensor are tried for 100 time steps each. Actual energy level and forward movement are shown. Only b9 (the first one) produces a significant improvement in performance.

 $\begin{array}{l} b10 = [b_f - \mathbf{R} - \mathbf{D} - s_c] \text{ ; reversely follow decreasing contact} \\ b11 = [b_f - \mathbf{S} - \mathbf{I} - s_c] \text{ ; follow increasing contact} \\ b12 = [b_f - \mathbf{S} - \mathbf{D} - s_c] \text{ ; follow decreasing contact} \end{array}$

[b9] is indeed a significant improvement because the robot now halts in the charging station based on a contact sensor (fig. 8). The last coupling [b12] leads to death because the robot decreases its forward movement after leaving the charging station and never regains speed in time. This illustrates that restoration periods between trials are required for the robot to remain viable despite negative experiences and that a new behavior should initially not be pursued for too long.

The overall behavior with [b4], [b6], [b9] and a positive growth rate for the competitors is shown in fig.9. As energy decreases, the robot's forward speed increases. This will bring it in the charging station where it stops (based on [b12]) at time step 45. But as the competitors keep on growing, they



Figure 9: The robot shows a cyclic behavior of going to the charging station and then recharging.

eventually cause the energy level of the robot to decrease, even though it is located in the charging station. This happens around time step 60, causing the forward speed to go up and the robot to leave the charging station. We see that on its way back to the charging station, the robot hits a competitor (around time step 80) which causes an increase in the global energy g. Then the same scenario repeats itself. The second round the robot is more lucky and hits several competitors causing a substantial increase in global energy.

The story continues in fig.10. A cyclic behavior continues where the robot alternates between recharging and seeking the charging station. On its way to the charging station the robot occasionally hits one or more competitors. This will cause it to slow down but it has no other effect.

The robot does not appear to be harmed by the competitors. However, this scenario relies on the fact that the robot hits sufficiently often a competitor by mere luck. The robot would quickly die if the pressures in this ecosystem would increase, for example by decreasing the probability factor r_{11} influencing the chance of hitting a competitor, or by increasing the growth rate of the



Figure 10: Continuation of scenario from fig.9. The cyclic behavior continues.

competitors r_8 , or by increasing the rate with which competitors consume energy r_9 . Let us therefore now turn to the problem how the robot might improve its chances of hitting the competitors.

4.2 Handling the competitors

Alignment towards the competitors, controlled by the motivational variable b_{a2} can be coupled to the energy level sensor as follows: [b13] b_{a2} - R - D - s_e ; reversely follow decreasing energy level [b14] b_{a2} - S - D - s_e ; follow decreasing energy level [b15] b_{a2} - R - I - s_e ; reversely follow increasing energy level [b16] b_{a2} - S - I - s_e ; follow increasing energy level [b16] b_{a2} - S - I - s_e ; follow increasing energy level

The following sections consider these variations one by one.



Figure 11: Behavior resulting when [b13] is added after 100 time steps. Alignment motivation follows reversely a decrease in the energy level. Its introduction is not detrimental to the agent.

4.3 The first variation: b_{a2} - R - D - s_e

The first variation is: $b_{a2} - R - D - s_e$ [b13] This coupling prescribes that the competitor alignment b_{a2} has to follow reversily a decrease in the energy level. The results of introducing this coupling can be seen in fig.11. The results are catastrophic. The robot does not survive. This means that the ecosystem is too tough to allow an exploration of this variation or that the overall regulatory mechanism should stop in time the effect of this coupling.

The alignment motivation change occurs the right time (namely when going back to the charging station). However it comes too late and increases when it should decrease. At the critical moment it distracts the robot away from the charging station and takes badly needed energy away. This coupling is clearly not beneficial for the agent.



Figure 12: Behavior when [b14] is added after 100 time steps. Alignment motivation follows a decrease in energy level. We see that a few competitors are hit every time the robot has left the charging station.

4.4 The second variation: b_{a2} - S - D - s_e

The second variation is: b_{a2} - S - D - s_e [b14]

This coupling prescribes that alignment follows the evolution of the energy level when this is perceived to decrease. The effect can be seen in fig.12. The motivation shoots up as soon as the energy decreases. The actual observed alignment is much lower as it depends also on the available energy. It is obvious that this coupling has a positive influence. There is a continued regular cycle where the robot recharges, seeks and hits the competitors and then seeks the charging station. The alignment behavior decreases appropriately so that the robot goes back to the charging station in time.

The stabilisation in the global energy is seen in the next figure (fig.13.) which shows the evolution of g for the same period as fig.12. The robot does roughly as much work as it needs to do to maintain the supplies that it will take out later. Note that the robot has no sensor to measure the global energy and



Figure 13: The global energy g stabilises because the robot takes as much out as it puts in (indirectly) by combatting competitors. These data are for the same experiments as shown in fig.9.

could therefore not receive a reinforcement signal based on this information.

Fig.14 shows what happens when the challenges in the ecosystem are increased. The probability of finding by chance a competitor has been lowered by half (r11 has been changed from 0.2 to 0.1). This would be mortal if the coupling was not present. But now it has no effect whatsoever. The agent has been able to wrestle itself away from chance and exploit the available resources in the ecosystem.

4.5 The third variation: b_{a2} - R - I - s_e

The next variation is: $b_{a2} - R - I - s_e$ [b15] The alignment motivation now reversely follows an increase in the energy level. The results are seen in fig.15. Predictably, the coupling has no impact on behavior. The reason is that the increased alignment motivation happens at a time when it is not needed, namely when the robot is in the charging station.



Figure 14: The probability of hitting the competitors by chance has been lowered but due to [b14], introduced after 100 time steps, no effect is seen on performance.



Figure 15: Variation [b15] where alignment reversely follows increases in energy level is introduced after 100 time steps. It has no impact on behavior.



Figure 16: Failure of [b15] to counteract the competitors causes the robot to die after a longer time, namely around time step 500.

Although fig. 16 appears to show that the robot is viable, the lack of counterforce to the competitors eventually causes them to increase so much that the robot does not survive as is seen in fig.17.

4.6 The fourth variation: b_{a2} - S - I - s_e

Finally the last variation is attempted: b_{a2} - S - I - s_e [b16]

Alignment towards the competitors now follows an increase in the energy level. Again there is no impact on behavior because alignment happens at the time it is not needed. So, the end conclusion is clearly that only the second variation, namely [b14] has to be retained.

The overall result with all the selected benes active is given in fig. 18.



Figure 17: Variation [b16] where alignment follows increases in energy level is introduced after 100 time steps. It has no impact on behavior.



Figure 18: The final process network causes the desired activity cycle alternating between charging, combatting the competitors, and foraging to find back the charging station. Performance improves and then stabilises.

5 Conclusions

The paper has explored a selectionist approach to behavioral development in which a (possibly complex) behavior is generated and tried out. If the behavior is beneficial with respect to the viability of the robot, then its likelihood of remaining in the behavioral repertoire increases otherwise it decreases. A selectionist approach has the advantage that completely novel situations can be created and tried out, whereas a purely inductive approach has to wait for situations to arise before something can be learned.

The paper has contributed to the selectionist approach to behavioral development by proposing a particular representation of behavioral control processes in the form of couplings between quantities represented as 'benes' and testing the adequacy of this representation for capturing significant variations. It was shown that a surprisingly small number of couplings is enough to capture the basic behavioral regulation needed for dealing with the ecosystem. It was also shown that the viability criterion is adequate for selecting which variation is most adapted.

The difficulty of the test experiment should not be underestimated. The ecosystem is unknown to the robot and has a variety of non-linear dynamical aspects. There are no prior examples on which some form of inductive learning can take place. The robot must first generate a particular behavior before its impact can be seen. There is no direct way to give credit to a specific action, as required by most reinforcement learning methods. The benefit of combatting the competitors becomes clear only much later. The robot does not get immediate feedback from crucial aspects of the environment such as the global energy level. In addition, there are no discrete prior categories, making it difficult to use if-then rules as the representation over which selectionist operations take place (as used in genetic algorithms). Sensory quantities are continuous and stream into the robot in real time and behavior regulation must be continuous and take place within the real-time constraints imposed by the environment.

There are still many open issues. Other types of couplings need to be added to the repertoire of possible couplings, and there will be processes that link more than two quantities. The presented experiments need to be carried out in the physical ecosystem. I have not discussed in detail the overall selectionist procedure that regulates when variations are introduced or the time frames in which these criteria are applied. This is discussed in another paper [13]. Nevertheless the results reported in this paper are encouraging and surprising, partly because one would expect that much more complex mechanisms would have been needed to handle the competitors.

6 Acknowledgement

This research has been carried out while visiting the Sony Computer Science Laboratory in Tokyo. I am strongly endebted to Toshi Doi and Mario Tokoro for creating a great research environment and providing the opportunity to work in it. Chisato Numaoka has provided valuable comments on the first drafts of the paper. Further comments by J-A Meyer have greatly improved the text. I am also indebted to the long term collaboration with David Mc-Farland from Oxford University and the members of the VUB AI Lab robotic agents group in Brussels. Special thanks to Dany Vereertbrugghe for robot design and construction and Peter Stuer for programming of the ecosystem and help with the visual material. The agents group is financed by the Belgian Federal government FKFO project on emergent functionality (NFWO contract nr. G.0014.95) and the IUAP project (nr. 20) CONSTRUCT.

References

- [1] Ashby, R. (1952) Design for a brain. Chapman and Hall., London.
- [2] Brooks, R. (1995) Intelligence without reason. In: Steels, L. and R. Brooks (eds.) (1995) The 'artificial life' route to 'artificial intelligence'. Building situated embodied agents. Lawrence Erlbaum Associates, New Haven.
- [3] Edelman, G. (1987) Neural Darwinism: The Theory of Neuronal Group Selection. Basic Books, New York.
- [4] Holland, J.H. (1975) Adaptation in Natural and Artificial Systems. The University of Michigan Press, Ann Arbor, Michigan.

- [5] Husbands, P., I. Harvey and D. Cliff (1995) Circle in the round: State space attractors for evolved sighted robots. Robotics and Autonomous Systems 15 (1995) 83-106.
- [6] Kaelbling, L., Littman and Moore (1995) An overview of reinforcement learning. In: Steels, L. (1995) The Biology and Technology of Intelligent Autonomous Agents. NATO ASI Series. Springer Verlag, Berlin.
- [7] Kaneko, K. (ed.) (1992) Coupled Map Lattices: Theory and Applications. Wiley, New York.
- [8] McFarland, D. and A. Houston (1981) Quantitative Ethology: the state-space approach. Pitman Books, London.
- [9] Steels, L. (1994) The artificial life roots of artificial intelligence. Artificial Life Journal. 1(1).
- [10] Steels, L. (1994a) A case study in the behavior-oriented design of autonomous agents. Proceedings of the Simulation of Adaptive Behavior Conference. Brighton. MIT Press. Cambridge.
- [11] Steels, L. (1994b) Emergent functionality through on-line evolution.
 In: Proceedings of the 4th Artificial Life Conference. MIT Press. Cambridge.
- [12] Steels, L. (1995) Building agents out of autonomous behavior systems. Steels, L. and R. Brooks (ed.) (1995) The 'artificial life' route to 'artificial intelligence'. Building situated embodied agents. Lawrence Erlbaum Associates, New Haven.
- [13] Steels, L. (1996) A selectionist approach to behavior development. Journal of Robotics and Autonomous Systems. Vol 16 (to appear).
- [14] Torras, C. (1995) An overview of neural network learning for robot adaptivity. Journal of Robotics and Autonomous Systems Vol 15 (1-2).

```
Appendix: Overview of quantities.
External quantities
a: alignment to charging station
a2: alignment to competitors
c: contact or not with the station
d: depth of station
e: energy level
f: forward movement
j: percentage of d already traversed
k: percentage of 1 already traversed
1: percentage of n already traversed
m: radius of corridor just outside station
n: minimum distance to station after corridor has been traversed
c2: contact with competitor
g: global energy
p: number of competitors
Internal quantities
bf: motivation of forward movement
ba: motivation of alignment (to charging station)
ba2: motivation of secondary alignment (to competitors)
sc: contact sensor
se: energy level sensor
r1: rate of energy use in forward motion
r2: rate of constant energy use
r3: recharge rate
r4: efficiency of forward movement
r5: rate for reaching rest state
r6: rate of travelling
r7: energy consumption by competitors
r8: competitor growth rate
```

r9: effectiveness of dimming competitors
r10: rate at which obstacle avoidance slows forward movement
r11: probability factor for hitting competitors