

Evolving a circadian rhythm

Sjoerd Lagarde (0513431)

December 4, 2009

Bachelor Thesis
Author: Sjoerd Lagarde
Email: S.Lagarde@student.ru.nl
Student Number: 0513431
Supervisors: Pim Haselager and Ida Sprinkhuizen-Kuyper
Radboud University Nijmegen
December 4, 2009

Evolving a circadian rhythm

Sjoerd Lagarde

Department of Artificial Intelligence, Radboud University
Nijmegen

December 4, 2009

Abstract

The theory of Embodied Embedded Cognition (EEC) proposes that cognition and behavior emerge out of the interaction between body and world. The brain can be envisioned as a traffic regulator, assisting in the selection of appropriate behavior based on cues from the body and environment. In this paper the theory of EEC with respect to the traffic regulator hypothesis is tested by simulation. Agents are placed in an environment with a day and night rhythm. In order to survive, these agents have to search for food and avoid obstacles. At day this is easy, but at night the perception is troubled by the darkness and agents may hit obstacles. Two types of agents are tested: purely reactive ones (based on the subsumption architecture) and reactive ones with a control structure on top. This control structure is a neural network with inhibitory output links to the behavioral layers of the agent. The control structure takes input from the body and environment.

Both types of agents are compared for their effectiveness in the environment. Results show that the control system agents are more effective in surviving than the reactive agents. Further investigation also showed that this effectiveness is due to the development of a circadian rhythm. The results support the theory of EEC with the brain as traffic regulator. The control structures act as traffic regulators, inhibiting behavioral responses when necessary.

1 Introduction

Circadian rhythms are found in many species, ranging from fruit flies to humans. The behavior of these living beings is synchronized with periodic factors in the environment, like the daily cycle of light and dark (Ishida, 1995; Shaw, Cirelli, Greenspan, & Tononi, 2000). A well known example of a behavior that demonstrates the circadian rhythm is the alternation of periods of activity with periods of inactivity (e.g. sleep). This wake and sleep cycle is synchronized with the light and dark cycle. Although this circadian rhythm, but also circadian rhythms in general, seems to rely on external cues (Tokura & Aschoff, 1983), experiments have shown that circadian rhythms are also produced when these cues are absent (Toh, 2008; Schaffer et al., 1998).

It has been argued that the circadian rhythm of sleep serves not only a psychological function (Kavanau, 2002), but also confers a survival advantage (Pittendrigh & Minis, 1972; Berger & Phillips, 1995; Ouyang, Andersson, Kondo, Golden, & Johnson, 1998). As Pinel (2006) notes, sleep may conserve the energy resources and reduce the risk of mishap. When the chances of finding food are very low or the risks of searching for food are very high, it is probably better to conserve energy resources for better times. This can also be seen in the hibernation behavior of some animals. During winter these animals sleep to escape the cold weather and food shortage (Carey, Andrews, & Martin, 2003).

Given the fact that sleep-waking cycles are often synchronized with the day-night cycles of the environment and the fact that so many species show this sleep-waking rhythm, it is interesting to look at the theory of embodied embedded cognition (EEC). The theory of embodied embedded cognition proposes that cognition and behavior emerge out of the interaction between body and world. The brain is just one of the equally important parts of the body that interacts with the environment. The brain can be envisioned as a traffic regulator, assisting in the selection of appropriate behavior based on cues from the body and environment (Haselager, Van Dijk, & Van Rooij, 2008; Van Dijk, Kerkhofs, Van Rooij, & Haselager, 2008). One of the simplest forms of a traffic regulator would be an inhibitory brain, inhibiting behavioral responses in certain circumstances depending on the state of the body and the environment.

In this paper one of the hypotheses of the theory of EEC, the brain as a traffic regulator, will be tested by using a computer simulation of an environment in which different types of agents have to survive. The environment shows a day and night rhythm and the agents are supposed to survive as long as possible in this environment. The first type of agent to be tested is a reactive agent (Brooks, 1986; Murphy, 2000). This agent responds directly to stimuli from the environment and has no higher control structures. The second type of agent is also a reactive agent, but with the difference that this agent does have a higher control structure. This structure, much like the prefrontal cortex in humans, can influence the different behavioral responses by means of cancellation.

By using the environment and agents described above, the following questions will be addressed:

1. Will reactive agents with an inhibitory control system be more effective than purely reactive agents in an environment displaying a day and night rhythm?
2. If so, is this effectiveness due to the development of an interaction between the control system agents and environments?
3. To what extent do these results support EEC?

The first two questions will be addressed by experiments; the third question is more philosophical in nature and will be discussed after the experiments.

The remainder of this paper is structured as follows. First the simulation environment will be discussed, followed by the parameters used in the simulation. Next the results of the simulations will be discussed together with their implications. This paper will end with the conclusion and ideas for future research.

2 Method

Different types of agents have to survive as long as possible in a given environment. The time the agents survive averaged over a number of different environments is the measure by which the agents are compared. Agents are able to survive in the environment by eating food and avoiding obstacles when searching for food. The aspects of the environment and the types of agents will be discussed in more detail below.

2.1 Environment

The world of the agents is rather simple. It consists of a two-dimensional grid. There is also a day and night rhythm in this world. A number of cells of the grid are randomly filled with food and obstacles. Cells can only be occupied by one object at a time, which means that each cell can be classified as food, obstacle or ground. An example of an environments can be seen in Figure 1. Obstacles are like quicksand: low to the ground and, with a lot of effort, one can go through these obstacles. The feeding places and obstacles reflect their own unique pattern of light, while ground reflects no light at all. Agents are able to classify each cell by looking at the reflections. At daylight this reflection is perfectly clear and so an agent does not make mistakes in classifying. However, at night the reflections become fuzzy and agents will not be able to classify each of the cells correctly. One other characteristic of the reflection is that obstacles have a weaker reflection than food has. This means that agents are able to detect food two cells away, using Manhattan distance, and obstacles one cell away from the current position. Since obstacles are low to the ground, these differences in reflection imply that food can be detected behind obstacles. The comparison with scent might be intuitive; one is able to smell a good meal being prepared, even when standing behind a closed door.

At the start of each simulation one agent is placed at a ground cell in the environment. Agents are able to move to one of the adjacent cells of their current position. They are allowed to move through the four boundaries of the environment and to reappear at the opposite side of the environment. Time passes in discrete steps and agents may act at each of these time steps.

Agents need food in order to survive. When the agent hits a feeding place, it will consume the food at that location. At this point, the food is no longer available at the current cell and the simulation will create a new food source at a random location in the environment. This will keep agents from staying at one food source. The obstacles are a bit different. When an agent hits an obstacle, some of its energy will be lost. In contrast to food, obstacles do not move and remain static during one run of the simulation.

2.2 Agents

Agents are supposed to survive as long as possible in the environment. Each agent has a certain amount of energy available which is used to perform actions in the environment. The agent dies when all of its energy is depleted and the agent therefore needs to eat the food found in the environment. However, to move around in the environment, the agent also uses some of its energy. And even worse, when the agent hits an obstacle a lot of its energy is taken away

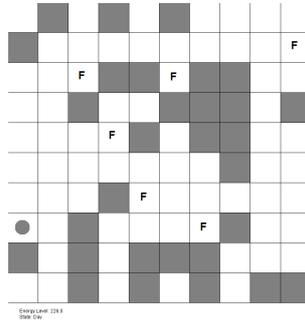


Figure 1: A typical environment. The squares with the letter ‘F’ indicate food, dark colored squares are obstacles. The agent is indicated with the gray circle.

at once. To make life not too hard, the agent is equipped with a light sensor that can be turned up to 360 degrees so it can detect obstacles and food in the surrounding. It also has a motor system with which the agent can turn up to 360 degrees and move forward.

Four different agents are tested in the environment and these agents can be divided into two categories: reactive agents and control system agents. The two categories of agents are a lot alike. The main difference between the two is that control system agents have a higher control structure, like the prefrontal cortex in humans. This higher control structure, however, also uses some energy. Therefore an additional cost is calculated. This extra energy usage is not totally artificial. In humans, for example, the brain also uses a rather larger portion of the energy available (Shulman, Rothman, Behar, & Fahmeed, 2004; Oz et al., 2007).

The four agents that will be used in the simulation are:

- Reactive. A reactive agent with no higher control structures.
- Reactive-DN. A reactive agent with a built-in day and night rhythm. DN stands for Day and Night.
- Control. An agent with a higher control structure. The control structure may inhibit behavioral responses of the agent. The cost of having this control structure is some extra energy usage.
- Control-2. An agent with a higher control structure. The Control-2 agent is the same as a Control agent with the only difference being the extra energy consumption of the control structure. The Control-2 agent will base its additional energy usage on the number of behavioral responses inhibited by the control structure.

A more in-depth description of these agents is given below. Table 1 summarizes the parameter settings used in the experiments.

2.2.1 Reactive

The first type of agent is the reactive agent. This agent is designed according to the reactive paradigm (see Murphy, 2000 for a detailed description of this

Parameter	Value
Agent Types	Reactive, Reactive-DN, Control, Control-2
Costs of Sleeping	Reactive: 1; Reactive-DN: 2; Control: 2; Control-2: 1+extra costs
Costs of Moving	Reactive: 2; Reactive-DN: 3; Control: 3; Control-2: 2+extra costs
Initial Energy Level	250
Hunger Threshold	240
Extreme Hunger Threshold	20
Environment Size	10x10 cells
Energy Increment by Food	10
Energy Decrement by Obstacle	15
Number of Feeding Places	6
Number of Obstacles	30
Maximum number of time steps	750
One full day cycle	30 time steps
Number of steps in day light	15 time steps

Table 1: Parameter settings used in experiments

paradigm). The design of the behavioral layers (Brooks, 1986) are displayed in Figure 2. The layers will be discussed in more detail.

- The first layer (from bottom to top) is the Wander layer. It takes no inputs from the body or environment. Its output is the motor system. The wander behavior turns the agent to a random direction and then moves the agent in that direction. Note that having the agent remain at its position is also an output of this layer. Also, the result of the wander layer is likely to be overwritten by output from higher layers.
- The Food Direction layer takes the readings from the light sensor as its input. Based on these readings it finds food in the surrounding of the agent. The output of this layer is an excitatory link to the first layer: the direction of movement can be changed towards the food. If no food is found, there will be no output.
- Evaluate Hunger is the third layer and takes as input the energy level of the agent. Based on this energy level, the agent decides whether or not it will look for food. If the agent is not hungry, the inhibitory link to the second layer, Food Direction, will be activated.
- The fourth layer is the Obstacle Avoidance layer. This is a rather important layer. It takes input from the light sensor again which is then used to detect if there are any obstacles in the surrounding. It then makes sure that the agent does not walk into an obstacle by having a direction not occupied by obstacles as output.
- The final layer is the Evaluate Extreme Hunger layer which takes the energy level of the agent as input. If the energy level is very low (i.e. the agent is starving to death), the obstacle avoidance layer will be deactivated

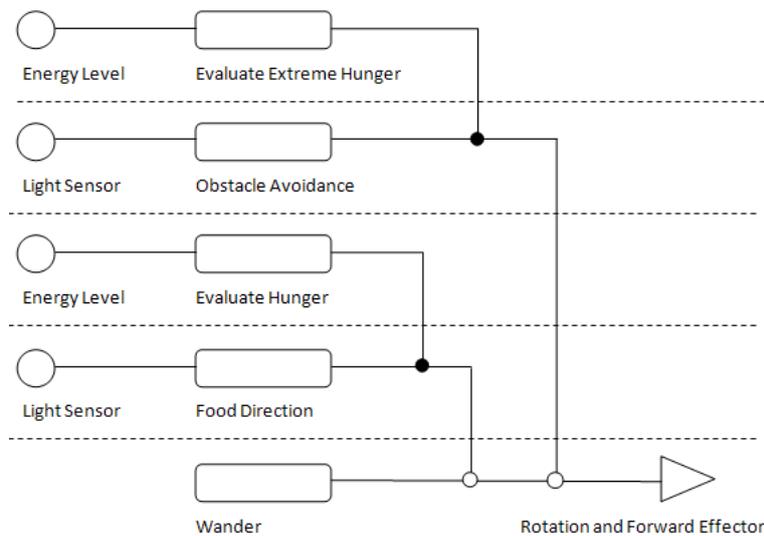


Figure 2: Behavioral layers of the Reactive Agent

by an inhibitory link to that layer. The agent may then still have a chance of finding food, although it has to move over an obstacle to reach it.

These layers together should produce rather complex behavior in the agent. At daylight the agent will function normally and search for food when needed. However, when night falls, the sensor readings from the light sensor become unreliable and the agent may begin to make mistakes.

2.2.2 Reactive-DN

The second type of agent is the reactive day-night agent. As the name implies, this agent is also a reactive agent and it has a day and night rhythm. The reactive part is a copy of the one used in the Reactive agent. Having a day and night rhythm, means that the agent knows when to go to sleep and when to wake up again. The agent therefore provides a good measure to compare the control system agents to. Since it is of interest whether or not these control system agents develop a day and night rhythm the results of Reactive-DN are very useful. It is expected that if control system agents develop day and night rhythms, their results should be closer to the results obtained from the Reactive-DN agents than to the results of the Reactive agents. Also, to make the results comparable, the costs of performing actions and going to sleep for this type of agent are equal to those of the control system agents.

When the agent goes to sleep, it will remain at its current position until it wakes up again. Energy consumption is reduced to a minimum in this state.

2.2.3 Control

The Control System Agent is the third type of agents. The agent is the same as the reactive agent, but now with a control system on top. This control system provides inhibitory links to all of the behavioral layers. With this brain, the

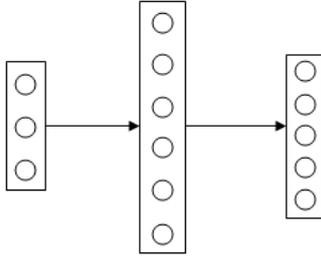


Figure 4: Topology of the multilayer perceptron. Three units take inputs (left layer) and propagate it to a hidden layer (middle layer, 6 units) which propagates it to the output layer (right layer, 5 units). The layers are fully connected.

2.2.4 Control-2

The fourth and last type of agent to be discussed is the Control-2 agent. This type of agent is basically the same as the control system agent. The difference is that the additional cost of the higher control structure is not a fixed value, but is dependent on the output of the brain. Each active output link, or deactivated layer, adds some small value to the cost of having a higher control structure. If no links are active, there is also no extra cost involved. Each active output link adds an amount of 0.20 (one divided by the total number of output links) to the energy consumption. When all behavioral layers are inhibited and thus all output links are active, the extra energy consumption is one. Compared to the Control agent, the Control-2 agent will only use the same amount of energy when all behavioral layers are inhibited. In all other cases, the Control-2 agent will use less energy.

2.3 Evolution

An Evolutionary Algorithm is used to optimize the weights of the neural network of the control system agents. The representation used here is inspired by Van Dartel, Sprinkhuizen-Kuyper, Postma, & Van Den Herik, 2005. In the evolutionary algorithm a genome represents the neural network by having all the weights of the network at specific positions in a vector. A population of 125 individuals is created in which each individual has 48 genes (3 inputs, 6 hidden nodes and 5 output nodes¹). Since using real valued numbers would result in a large search space, even when the values are restricted to a given interval, the genes are initialized with random integer values in the interval $[-300, 300]$. Weights for the neural network are then calculated by dividing each gene by 100. Weights of the network thus lie in and are constrained to the interval $[-3, 3]$. Fitness is calculated as follows. The genome is first converted to a neural network which is then given to an agent. Next this agent goes through one simulation in a random environment. The number of time steps the agent survived in this environment is taken as the fitness. Since there is a constraint

¹The three input nodes are fully connected to the 6 hidden nodes, resulting in $3 * 6 = 18$ links. The 6 hidden nodes are also fully connected to the 5 output nodes: $6 * 5 = 30$. This results in $18 + 30 = 48$ links for which weights have to be set. The weights are represented in each individual with 48 genes. No bias nodes are used in the network. Tests showed no change in performance when bias nodes were used.

on the number of time steps the agent may live, 750 time steps, the fitness value is also bounded: $[0, 750]$.

Based on the fitness values, parents are selected to create new offspring by means of mutation and recombination (Eiben & Smith, 2007). The Evolutionary Algorithm is implemented with the use of JGAP (Rotstan, 2009). The default configuration for mutation rate and selection method provided by this package was used. Details can be found in Appendix A.

After 400 generations the evolutionary algorithm is terminated and the best individual is selected as the one that will be used in the final simulation.

2.4 Simulation

Performance of agents was tested by running a number of simulations. In these simulations, the agents were placed in randomly created environments. The environments are of size 10 by 10. In other words: the agent has 100 cells on which it can stand. Due to randomness some of the environments are particularly nice for the agents, while others are very bad to live in. Therefore, each agent runs in 5000 environments and the results are then averaged. The agent receives an energy level of 250 at the start of each simulation. The maximum number of time steps each agent may live is set to 750. An agent able of reaching this limit is rewarded with the highest fitness available. Each full cycle of day and night consists of 30 time steps: 15 for the day and 15 for the night.

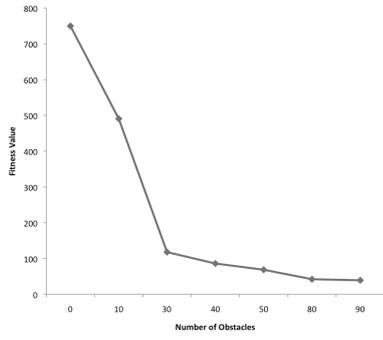
Several variables are tested for their effects on the performance of the agent. Two of these variables are used to answer the research questions. First, the number of food and obstacles is varied. It is likely that this variable will have a large effect on the performance. The second variable is the cost associated with doing nothing and moving around in the environment. The results can be found in Section 3. Other variables were also tested and the results can be found in Appendix B. These results are used to fix the parameters to reasonable values and to verify that agents show the expected behavior.

3 Results

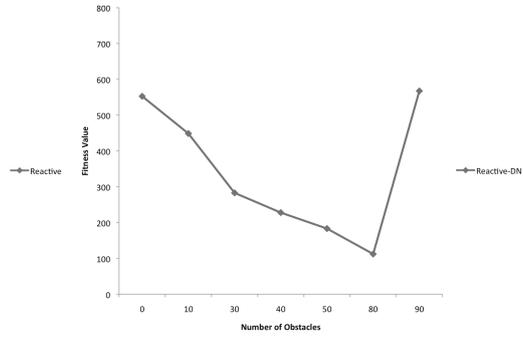
The first part of the research question is: will reactive agents with an inhibitory control system be more effective than purely reactive agents in an environment displaying a day and night rhythm? To answer this question, agents are tested in a number of environments with different numbers of food and obstacles. Varying costs of performing actions are also tested. The results should give an indication of the effectiveness of a control system.

The second part of the question – is this effectiveness due to the development of a circadian rhythm within the control system agents? – can only be answered when the results to the first part of the question are positive. If the results show positive, and indeed they do, an analysis of the simulation will provide the answer. The analyses consist of a visual inspection and a data analysis.

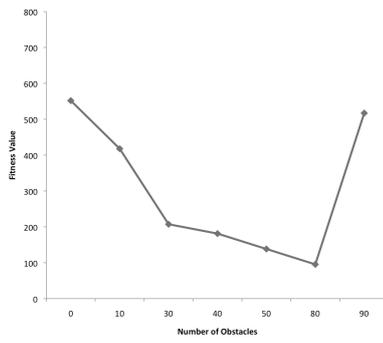
The third question – to what extent do these results support EEC – will be discussed last. The answer to this question depends on the results from the first two questions.



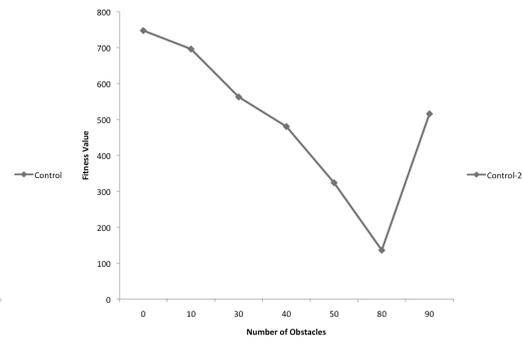
(a) Reactive Agent



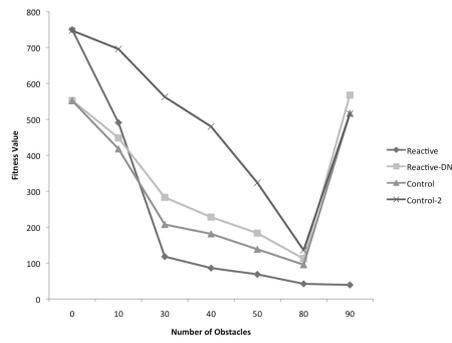
(b) Reactive-DN Agent



(c) Control Agent

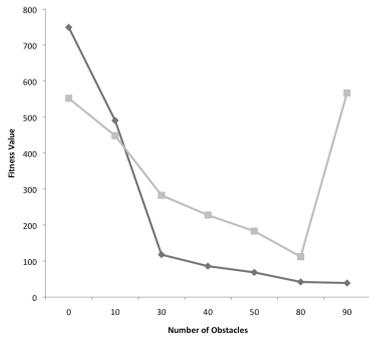


(d) Control-2 Agent

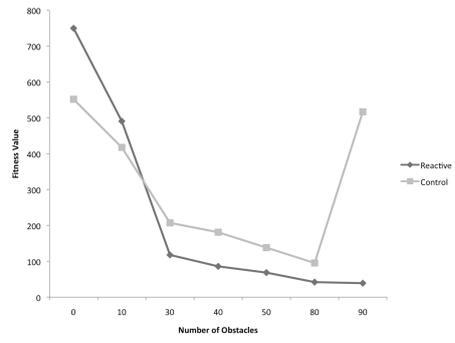


(e) All Agents

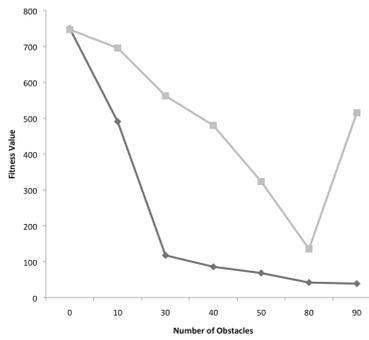
Figure 5: Visualization of the fitness (vertical axis) of different types of agents with a varying number of obstacles (the horizontal axis). The number of feeding places is held constant at the value of 10.



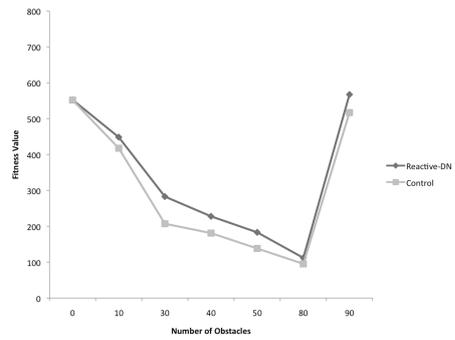
(a) Reactive and Reactive-DN



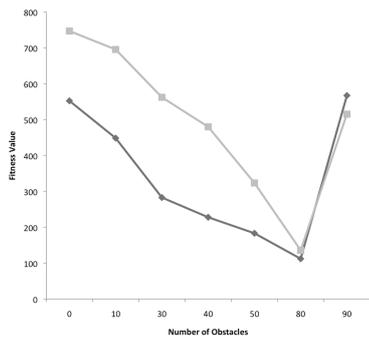
(b) Reactive and Control



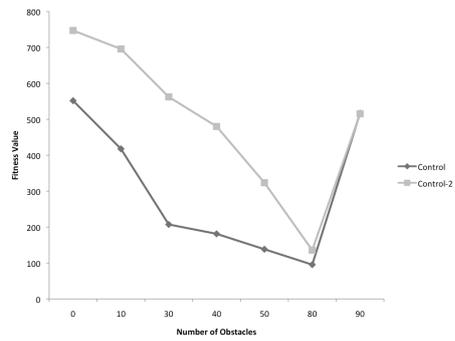
(c) Reactive and Control-2



(d) Reactive-DN and Control

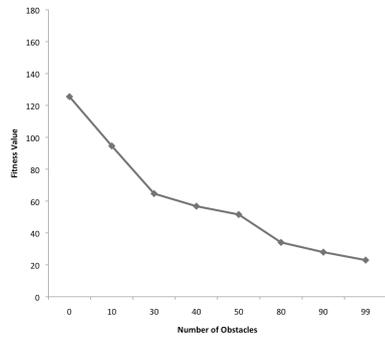


(e) Reactive-DN and Control-2

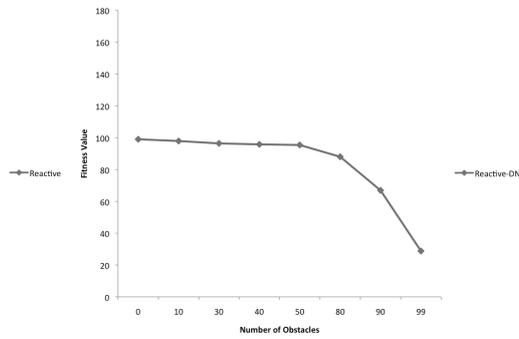


(f) Control and Control-2

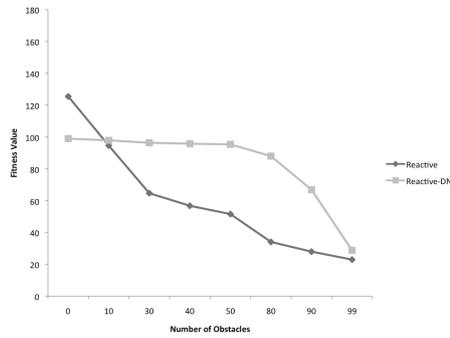
Figure 6: Visualization of the fitness (vertical axis) of different types of agents compared to each other with a varying number of obstacles (the horizontal axis). The number of feeding places is held constant at the value of 10.



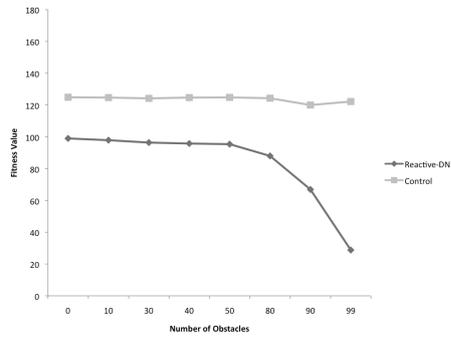
(a) Reactive Agent



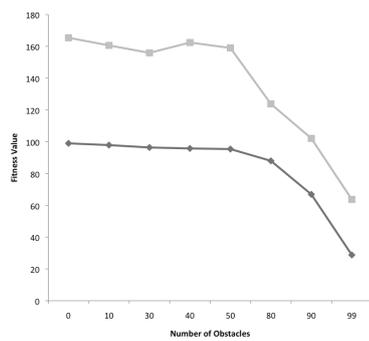
(b) Reactive-DN Agent



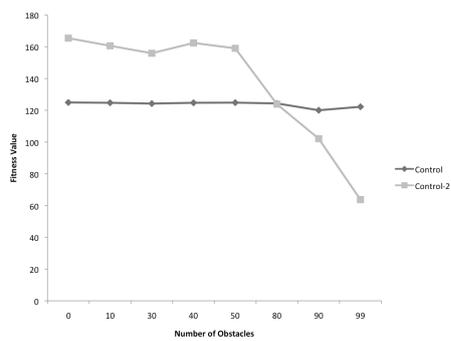
(c) Reactive and Reactive-DN



(d) Reactive-DN and Control



(e) Reactive-DN and Control-2



(f) Control and Control-2

Figure 7: Visualization of the fitness (vertical axis) of different types of agents with a varying number of obstacles (the horizontal axis). The number of feeding places is held constant at the value of 0. The results indicate that control system agents use different behavioral strategies than the reactive agents.

3.1 Effectiveness of Control Systems

3.1.1 Environments

Figures 5 and 6 show the results of the different types of agents. Agents are placed in environments with a fixed amount of 10 feeding places. The number of obstacles are varied (the horizontal axis) and the fitness of agents is measured in different environments (the vertical axis). The reason for using fixing the number of feeding places to 10 is that it will give the most interesting results. Environments with more feeding places tend to be too friendly; agents may constantly eat. Environments with very few feeding places are very hard to survive in. The number of obstacles shows this effect too, although in reverse order, as can be seen in Figure 5. Environments with no obstacles are very friendly, while environments with a lot of obstacles are very hard to survive in. A point of interest is that the graphs in Figure 5, with exception of the Reactive agent, show a large peak at the end when the environment is filled with obstacles (right hand side of the graphs. Figure 5e makes it even more clear). This is rather strange since this environment should be very hard to survive in. However, by filling the whole environment with obstacles and food, the agents will stay at their current position most of the time and will not use any energy for moving around. In case the agents do start looking for food, when they have found a feeding place, it is likely (i.e. 50%) that this feeding place becomes a feeding place again, since only two squares are available to put the new feeding place in². The reason that the Reactive agent does not have this peak in performance is because of the fact that this agent will continue to search for food at night. The reflections of obstacles and food become fuzzy and the agent will most likely mistake an obstacle for food and lose much of its energy as a results.

One other point of interest is that in environments with no obstacles sleeping behavior is suboptimal. Figure 5e shows that both Reactive-DN and Control perform worse than Reactive and Control-2. The results of Reactive-DN and Reactive are more or less intuitive: why would you go to sleep when there is no risk of running into obstacles? It is more effective to just keep on searching for food. The Control-2 agent acts like the Reactive agent. Since the control structure of this agent is not inhibiting any behavioral layers, the costs of having this control structure are zero (see Section 2.2.4 for the calculation of extra costs). This makes the Control-2 agent the same as a Reactive agent and hence the same performance is achieved. The Control agent performs worse than the Reactive and Control-2 agents, but the same as Reactive-DN agent. The reason for this is the cost associated with having a control structure. In contrast to the Control-2 agent, the Control agent will always use one extra energy unit even if the higher control structure is not used (Table 1).

²Cells in the grid of the environment can be occupied by one type of object at a time: food or obstacles. Cells can also be empty in which case they considered as ground. When the agent eats food and a new feeding place is added to the environment, the location of this new feeding place will be a random location in the empty cells. This also means that a new feeding place may be dropped on the agents current location. In this case the agent is lucky and may have its free lunch.

3.1.2 Food and Obstacles

Varying the number of food and obstacles provides a measure to indicate whether or not the control system works in different environments. Figures 5 and 6 show some interesting results which require a bit more explanation. The raw data used in the graphs can be found in Appendix B.1.

The graph of the Reactive-DN agent (Figure 5b) shows the fitness values for an agent with a day and night rhythm. It is interesting to compare this graph with the graphs of Control and Control-2 agents to see how similar they are. When compared with the Control agent (Figures 5c and 6d), it is found that Reactive-DN performs better than the Control agent in all cases. However, the performance of both agents is qualitatively similar. The line of the Control agent follows the same trend as the Reactive-DN agent (Figure 6d). The performance issue can be attributed to the fact that the Reactive-DN is programmed to show the most optimal behavior in the given environment. The Reactive-DN agent results are therefore the best results an agent can achieve. The Control agent is able to learn the optimal behavior, but will most likely, due to early stopping of evolution, only approach the optimal behavior. The rather large deviation in performance around the points of 30 to 50 obstacles might also be due to the fact that Control agents are given a limited number of evolutions. When comparing the results of the Reactive and Reactive-DN (Figure 6a) it can be seen that both types of agents perform about the same with 20 obstacles and that the Reactive agent performs worse than the Reactive-DN agent from that point on. It might be that the Control agent averages over both strategies. In order to test this hypothesis, a larger number of generations and a larger population size in evolution were tested. Results indicate that when the number of generations is increased from 400 to 1000 and the population size is doubled to 250 individuals, performance increases to the expected levels.

The effectiveness of control systems also becomes clear when comparing the results of the previous environment, with 10 feeding places, to the results of a new environment with no feeding places at all. The results of different agents are visualized in Figure 7. The results are a bit counterintuitive, since it would be expected that all agents perform very badly in these environments because there is no way to get any food. However, results indicate that only the reactive agents, Figures 7a and 7b, perform badly, while the control system agents are able to learn some strategy that makes them perform better than expected as can be seen in Figures 7d and 7e. Further inspection showed that control agents learned to sleep constantly, not using any energy for moving around.

When comparing the Reactive-DN agent to the Control-2 agent, see Figure 6e, two surprisingly different graphs are found. The Control-2 agent performs a lot better compared to the Reactive-DN agent in most of the cases. Apparently the Control-2 agent has learned a strategy to save energy. The cost in energy for a Control-2 agent to inhibit one behavioral layer is low. Only when all layers are inhibited at once, a full energy unit is used and in that particular case the Control-2 agent would have the same energy usage as Reactive-DN and Control agents. The results show that the higher control structure has learned to only inhibit the necessary layers. For example if sleeping would be the appropriate behavior given the inputs, only the wander, food direction and obstacle avoidance layers need to be inhibited. The other two layers, evaluate hunger and evaluate extreme hunger, are inhibited implicitly by inhibiting the

Costs (sleep/action)	Reactive	Reactive-DN	Control	Control-2
1/1	132.2	732.0	720.7	743.8
1/2	89.8	429.2	431.2	414.7
2/3	68.4	159.0	149.6	218.6
3/4	56.3	97.5	82.9	86.4
3/5	47.6	81.6	75.6	70.2

Table 2: The fitness values of agents with varying costs values. Higher values are better. The parameter settings used in the experiments are in bold. Note that the Control-2 agent will calculate additional costs based on the amount of active output links of its control structure. The results are based on simulations with 6 feeding links and 30 obstacles. The full table with results can be found in Appendix B.2

other layers (see Figure 3). The strategy is also not a reactive one, because the results of Reactive and Control-2 agents differ too much (Figure 6c).

3.1.3 Costs

The costs associated with the use of a control system are also of importance to the question whether or not a Control System agent performs better than a Reactive agent. When the costs of having a control system are too high, it may not be beneficial to have such a control system. This is also what is found in the results from the experiments. Table 2 shows that when costs are very high, the control system agents will perform more or less equal to or worse than the reactive agents with low costs (e.g. row 1/2 for reactive agents and row 3/4 for control agents). Also note that the costs of Control-2 agents are higher in simulations than the costs displayed in the table. Control-2 agents will use additional energy based on the amount of active output links of the control structure (at most 1 extra energy unit will be used. See Section 2.2.4 and Table 1 for more details on costs of Control-2 agents). This means that the Control-2 agent in the row of 1/2 in Table 2 uses 1 energy unit plus some additional cost for using the brain to sleep and 2 energy units plus some additional cost for using the brain to act.

3.1.4 Conclusion

The results provide a clear answer to the question whether or not agents with a control system perform better than purely reactive agents. It is found that the control system agents perform better than purely reactive agents in most circumstances. The results also indicate the conditions in which this is true. The additional costs of a control structure may be set to relatively high values, but values too high will make control structures useless. The number of obstacles and feeding places can be set to a wide range of values with the restriction that in order to have a useful day and night rhythm, the number of obstacles should be relatively high (e.g. 10% or more of the total number of cells in the environment). This also shows the robustness of the control system agents. The performance of the control system agents is not dependent on very specific settings of the environment.

3.2 Circadian Rhythm

Since the results described above showed that control system agents perform better in most environments, an analysis can be made of the simulations to see whether or not this effectiveness is due to the development of a circadian rhythm. When agents develop a circadian rhythm they should be in a sleeping state at night, while searching for food at day. The analyses performed on the simulations are a visual inspection and a data analysis. Both are discussed below.

3.2.1 Visual Inspection

First, a visual inspection was done of the simulations. The last ten environments of the simulation of control system agents are visualized and inspected (see Figure 1 for an example of one step in a visualization). Reactive agents and Reactive-DN agents are not inspected in detail since their behavior is already known. It is found that the control system agents develop a day and night rhythm. At day they are actively searching for food, but at night they remain mostly in their current position.

The Control-2 agents show, next to the day and night rhythm, some other, unexpected behavior as well. When the energy level of the Control-2 agent is high (circa 100 units or more), the agent behaves as if it is night: it does not move around in the environment to look for food. Or to put it in a wake and sleep context: the agent takes a siesta. Only when the energy level is low enough (i.e. it comes in a range of [100, 150]), the agent starts moving (Figure 8). This is interesting, because, compared to the Control agent, the Control-2 agent shows more complex behavior than just a day and night rhythm. There is no difference in the design of the higher control structure, so the difference in behavior needs to be attributed to other factors. The defining characteristic of Control-2 agents is that the additional costs for using the higher control structures is based on how much this structure is used (i.e. how much of the behavioral layers it inhibits). It is highly likely that inhibiting the wander layer (see Figure 3), and thus effectively putting the agent to sleep, is cheaper in energy usage than to move around in the environment. However, at some point, looking for food will be necessary in order to survive and the agent wakes up again.

3.2.2 Data Analysis

The number of times agents sleep and move in the two phases of the day-night cycle are counted and can be found in Table 3 as percentages of the average fitness of the agents. These results indicate that control system agents are close to the Reactive-DN agents with respect to their sleeping behavior.

An interesting thing to note is that the unexpected behavior of Control-2 agents – taking a siesta when having lots of energy – can also be seen in the results. The average time of sleeping during the day is still very small, but, in comparison to the other types of agents, a lot higher. This suggests that sleeping at day when having enough energy is not just fluke behavior that only happens in a few runs, but that it is in fact systematically occurring in the Control-2 agents.

Another surprising result is that the Control agent also seems to take a siesta once in a while. The percentage of sleeping at day is much lower than

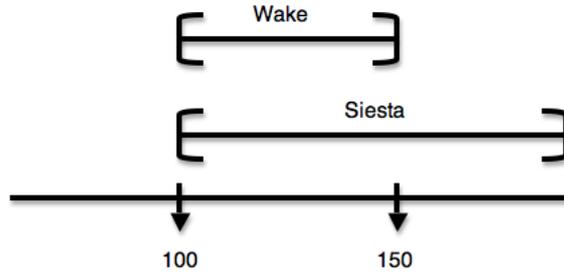


Figure 8: The ranges in which the Control-2 agent will take a siesta and wake up again. From circa 100 energy units and up the agent will take a siesta. Somewhere in the range of [100, 150] energy units the agent will wake up again to search for food.

	Reactive	Reactive-DN	Control	Control-2
Day Actions	52.8%	52.2%	51.5%	49.9%
Night Actions	47.1%	1.4%	0.7%	0.5%
Day Sleep	0.1%	0.1%	0.6%	1.6%
Night Sleep	0.0%	46.3%	47.2%	48.0%
AVG Fitness	89.4	158.4	149.6	215.2

Table 3: Average number of steps as percentages of the average fitness in different situations.

in Control-2 agents, but higher than in Reactive and Reactive-DN agents. The reason for this behavior is not as clear as in the Control-2 agent case. Visual inspection of the simulations with the Control agent shows that agents with very low energy levels and not located near food sources, tend to sleep more at day. By doing so, the agent may live a few steps longer, since it is not using energy for movement. This behavior can be seen as suicide through inertia. Note that this happens only in a few cases and that no research was done to investigate this behavior further.

3.2.3 Conclusion

Based on the analyses, it can be concluded that the effectiveness of the control system agents is due to sleeping behavior. Agents are able to learn when it is most beneficial to them to go to sleep and when to wake up again. In the case of the Control agent, this takes the form of the expected sleeping behavior: sleep at night and eat at day. The Control-2 agents might have even developed more complex behavior, since it also sleeps at day in some cases. However, this might need some further investigation.

3.3 Embodied Embedded Cognition

The results described above show that a day and night rhythm is developed by agents with a control system and that this day and night rhythm is advantageous to the agents. To what extent do these results support the traffic regulator

hypothesis of EEC? To answer this question the embodiment and embeddedness of agents will be discussed first. Next, the relation to traffic regulators will be explained in more detail.

3.3.1 Embodiment and Embeddedness

The embodied part of EEC is supported by the fact that the control system agents are able to use their sensors and internal states to act in the environment. All agents use a light sensor that does not work very well at night. Both control system agents are able to adapt to this light sensor and use the information to come up with appropriate behavior. The motor system is used to achieve a day and night rhythm by not moving at night. With respect to internal states, the energy level of agents is used to guide the agent in surviving. The agents learn that their energy level, which is an internal state, needs to have a high enough value in order to survive. In order to keep the energy level high enough, the agents will have to consume food. However, at night the risks of hitting obstacles are too high and the will agents will start to sleep. An even stronger example of the energy level as internal state is given by the Control-2 agents, as also explained in Section 3.2.1. These agents do not only sleep at night, but they also sleep at day when they have enough energy available. Here the internal state directly influences the higher control structure of the agent.

The results discussed in Section 3.1.2 also clearly show that the agents are embedded. Different environments result in different behavioral responses from both the reactive agents and the control system agents. Figure 7 shows the results for agents in environments with no feeding places at all. The differences with the results given in Figure 5 are clear. When there is no way of finding food, control system agents deploy different behavioral strategies than when food is available. The results of simulations with environments in which no obstacles were placed (see peaks at left hand side in Figure 5e) show that a control structure is not always necessary. Reactive agents perform very well in these environments. The control structure agents deploy the same behavioral strategies as reactive agents to maximize their fitness by shutting down their control systems.

3.3.2 Traffic Regulator

The higher control structures in the control structure agents can be seen as traffic regulators. Behavior of these agents is not generated by the control structure, but is merely directed in appropriate ways based on cues from the agent's body and its environment. The embodiment of the agent provides the inputs to the control structure. When it is dark, the behavioral responses will be limited to sleeping. Inputs that otherwise would result in some action are now ignored (e.g. the food searching and eating). The embeddedness of the agents provides cues to the control structure on what behavior is appropriate in the given environment. When no food is available in the environment, for example, there is no need to move around looking for food. Even when it is daylight and everything is perfectly visible, a search for food would result in nothing but an energy loss. The control structure learns, by trial and error, that inputs should be ignored and sleeping is the best behavior given the environment.

By combining the information from both the embodiment and embeddedness of the agent, the higher control structure learns to inhibit the behavioral responses in appropriate ways given the agent's situation.

4 Conclusion and Future Research

As discussed in Section 1, in EEC the brain can be envisioned as a traffic regulator. It assists in the selection of appropriate behavior based on cues from the body and environment. In this paper the hypothesis of the brain as a traffic regulator was tested by placing agents in environments with a day and night rhythm. The agents then had to survive as long as possible by eating food and avoiding obstacles in their search for food. This goal was easy to achieve at day, but at night the agents were not able to clearly perceive the environment due to the darkness. Two types of agents were tested: the reactive agents and the control system agents (Figures 2 and 3). The control system of the control system agents consisted of a neural network with inhibitory output links to the behavioral responses of the agent. The inputs of the network are bodily states and environmental perception. Results showed that these control system agents were able to effectively adapt to the day and night rhythm by developing a day and night rhythm in which agents rested at night and searched for food at day. Control system agents performed, in almost all environments, better than reactive agents. Compared to the idea of traffic regulators, the inhibitory control structures can be said to do exactly this. The control system helps in the selection of appropriate behavior based on cues from the body and environment. The results are thus in favor of the traffic regulator hypothesis.

An interesting finding was the unexpected behavior of Control-2 agents. The Control-2 agents are, as their name implies, control structure agents. It was found that these agents were also sleeping at day when they had enough energy left. The behavior seemed to result from the costs involved with having and using a higher control structure. This complex behavior resulting from the embodiment of the agent is an even stronger argument for EEC.

The control system agents use extra energy for having or using the control system. The Control agents always use one extra energy unit and the extra energy consumption of Control-2 agents is dependent on the number of active output links. While the extra cost for the control system is neurophysiological plausible, the plausibility of the amount of extra energy consumption as used in the experiments described in this paper might be low. Measures of energy usage based on research in human brain metabolism (e.g. Oz et al., 2007) might be an interesting topic for future research.

Future research might also focus on a comparison of different, more complex higher control structures. The control structure described in this paper is a very simple one and it might be interesting to see how other types of control structures perform. The data showed that control system agents are able to learn to survive in the given environments, but these environments are rather simple. More complex environments could, for example, consist of moving objects. The control system might need some form of memory or planning. This is also in line with the proposal for more complex control structures. It would be interesting to see to what extent EEC holds in such complex environments with such complex higher control structures.

References

- Berger, R. J., & Phillips, N. H. (1995). Energy conservation and sleep. *Behavioural Brain Research*, *69*, 65-73.
- Bishop, C. M. (2006). *Pattern recognition and machine learning*. Springer.
- Brooks, R. A. (1986). A robust layered control system for a mobile robot. *IEEE Journal of Robotics and Automation*, *2*(1), 14-23.
- Carey, H. V., Andrews, M. T., & Martin, S. L. (2003). Mammalian hibernation: Cellular and molecular responses to depressed metabolism and low temperature. *Physiological Reviews*, *83*(4), 1153-1181.
- Eiben, A. E., & Smith, J. E. (2007). *Introduction to evolutionary computing*. Springer.
- Haselager, W. F. G., Van Dijk, J., & Van Rooij, I. (2008). A lazy brain? embodied embedded cognition and cognitive neuroscience. In P. Calvo & T. Gomila (Eds.), *Handbook of embodied cognitive science: An embodied approach* (p. 273-290). Oxford: Elsevier.
- Ishida, N. (1995). Molecular biological approach to the circadian clock mechanism. *Neuroscience Research*, *23*, 231-240.
- Kavanau, J. L. (2002). REM and NREM sleep as natural accompaniments of the evolution of warm-bloodedness. *Neuroscience and Biobehavioral Reviews*, *26*, 889-906.
- Murphy, R. R. (2000). *Introduction to AI robotics*. The MIT Press.
- Ouyang, Y., Andersson, C. R., Kondo, T., Golden, S. S., & Johnson, C. H. (1998, July). Resonating circadian clocks enhance fitness in cyanobacteria. *Proceedings of the National Academy of Sciences of the United States of America*, *95*, 8660-8664.
- Oz, G., Seaquist, E. R., Kumar, A., Criego, A. B., Benedict, L. E., Rao, J. P., et al. (2007). Human brain glycogen content and metabolism: implications on its role in brain energy metabolism. *American journal of physiology. Endocrinology and metabolism*, *293*(5), E1378-E1384.
- Pinel, J. P. J. (2006). *Biopsychology*. Pearson Education.
- Pittendrigh, C. S., & Minis, D. H. (1972, June). Circadian systems: Longevity as a function of circadian resonance in *Drosophila melanogaster*. *Proceedings of the National Academy of Sciences of the United States of America*, *69*(6), 1537-1539.
- Rotstan, N. (2009). *JGAP: The Java Genetic Algorithm Package*. Available from <http://jgap.sourceforge.net/> (Version 3.4.3)
- Schaffer, R., Ramsay, N., Samach, A., Corden, S., Putterill, J., Carre, I. A., et al. (1998, June). The *late elongated hypocotyl* mutation of *Arabidopsis* disrupts circadian rhythms and the photoperiodic control of flowering. *Cell*, *93*, 1219-1229.
- Shaw, P. J., Cirelli, C., Greenspan, R. J., & Tononi, G. (2000). Correlates of sleep and waking in *Drosophila melanogaster*. *Science*, *287*, 1834-1837.
- Shulman, R. G., Rothman, D. L., Behar, K. L., & Fahmeed, H. (2004, August). Energetic basis of brain activity: implications for neuroimaging. *TRENDS in Neurosciences*, *27*(9), 489-495.
- Toh, K. L. (2008). Basic science review on circadian rhythm biology and circadian sleep disorders. *Annals of the Academy of Medicine, Singapore*, *37*(8), 662-668.
- Tokura, H., & Aschoff, J. (1983). Effects of temperature on the circadian rhythm

- of pig-tailed macaques *Macaca nemestrina*. *AJP - Regulatory, Integrative and Comparative Physiology*, 245(6), 800-804.
- Van Dartel, M., Sprinkhuizen-Kuyper, I. G., Postma, E., & Van Den Herik, J. (2005). Reactive agents and perceptual ambiguity. *Adaptive Behavior*, 13(3), 227-242.
- Van Dijk, J., Kerkhofs, R., Van Rooij, I., & Haselager, W. F. G. (2008). Can there be such a thing as embodied embedded cognitive neuroscience? *Theory & Psychology*, 18(3).

Appendices

A JGAP Settings

JGAP is an open-source project which provides a robust Java framework for Genetic Algorithms (Rotstan, 2009). The default configuration provided by this package was used to run the evolutionary algorithm. The specific values and methods used in the algorithm are displayed in Table 4.

The evolutionary algorithm is used to find weight settings for the control structure of agents that makes these agents perform well in the given environments. The weights used in the control structure are values in the interval $[-3, 3]$. To reduce the search space in the evolutionary algorithm, an integer vector with values in the interval $[-300, 300]$ is used. The results of the evolutionary algorithm are divided by 100 before they are inserted into the weight settings for the control structure.

Population size	125
Representation	Integer Vector
Mutation	Addition of value in interval $[-100, 100]$
Mutation probability	$\frac{1}{12}$
Recombination	One-point crossover
Recombination probability	0.5
Parent selection	Roulette Wheel
Survivor selection	Generational
Termination condition	400 generations

Table 4: Values and methods used in the evolutionary algorithm.

B Results

The following section will give more a detailed overview of the results obtained from the experiments. In all experiments the fitness value of the agents was measured. The fitness value an agent could obtain was 750. Results close or equal to this number indicate very fit agents.

B.1 Number of Obstacles and Feeding Places

Different values for the number of feeding places and obstacles were tested. The results can be found in Table 5.

In further experiments the number of obstacles was fixed to 30 and the number of feeding places to 6.

B.2 Costs of Actions

The costs of performing actions and staying in place (resting) were varied. These costs can also be thought of as additional energy consumption for having a higher control structure. The results of different values for the costs can be found in Table 6

The results indicate that cost values of $1/2$ (1 for resting and 2 for action) for reactive agents and $2/3$ for control system agents produce reasonable results. Higher costs for the control system agents (e.g. $2/4$) are closer to the values of the reactive agents with $1/2$ costs, but using these costs would most likely take away the advantage of having a brain, because the costs are too high.

B.3 Extreme Hunger Threshold

The threshold value for extreme hunger is also of influence to the behavior of agents. Different values may thus produce different results. Table 7 shows the results for different values for extreme hunger.

The results indicate that control system agents do not take the extreme hunger threshold into account, or assign very little priority to it. Only when using very high thresholds, there is some effect, but the magnitude of this effect is very low.

B.4 Energy Updates

Using different values for the update in energy when running into obstacles and eating food influences the average fitness of the agents. Different values for both parameters were tested. The results can be found in Table 8. Based on the results it can be concluded that reactive agents are in the advantage when the absolute values of energy updating by obstacles and food is larger for food than for obstacles. Otherwise, it is more advantageous to be a control system agent.

B.5 Day Night Rhythm

The fitness of agents may also be influenced by the length of the day-night rhythm in the environment. Table 9 shows the results of experiments in which the parameters of the day-night rhythm are changed. The first parameter is Duration, which indicates how many time steps one full cycle takes. The second parameter is the ratio of day and night. The values in that column indicate the proportion of the full cycle in which it is night.

From these results it follows that different durations of the day-night cycle do not have much influence on the fitness of agents. Differences in the ratio of day and night in one cycle does change the fitness of the agents.

Obstacles	Food	Reactive	Reactive-DN	Control	Control-2
99	0	23.0	28.8	122.2	63.7
90	0	28.0	66.9	120.0	102.1
	9	39.1	567.4	517.0	515.4
80	0	34.1	88.0	124.3	123.8
	10	42.1	112.2	95.4	135.8
	19	66.4	735.9	703.0	462.7
50	0	51.6	95.4	124.8	159.0
	10	68.6	183.3	138.3	323.6
	20	98.9	553.3	527.8	547.7
	40	347.9	743.5	711.0	657.5
	49	700.9	750.0	718.2	693.4
40	0	56.8	95.8	124.7	162.4
	10	86.0	227.9	181.2	480.1
	20	155.5	707.8	676.2	730.5
	40	614.4	748.5	734.6	731.7
	59	746.1	750.0	731.9	732.1
30	0	64.7	96.4	124.2	155.9
	10	117.9	282.9	207.3	562.4
	20	334.3	743.2	697.9	640.1
	40	737.3	750.0	601.7	642.1
	60	749.0	750.0	734.9	733.0
	69	749.8	750.0	748.5	748.3
10	0	94.6	97.9	124.7	160.6
	10	490.8	448.6	417.8	695.8
	20	749.7	750.0	688.9	743.1
	40	750.0	750.0	716.3	749.7
	60	750.0	750.0	745.8	749.0
0	0	125.5	99.0	124.9	165.4
	10	749.9	552.6	551.8	747.2
	20	750.0	750.0	749.8	745.8
	40	750.0	750.0	750.0	750.0
	60	750.0	750.0	750.0	750.0
	80	750.0	750.0	750.0	749.9
	99	750.0	750.0	750.0	750.0

Table 5: The fitness values of different agents in a variety of environment with respect to the number of obstacles and feeding places.

Costs (rest/action)	Reactive	Reactive-DN	Control	Control-2
1/1	132.2	732.0	720.7	743.8
1/2	89.8	429.2	431.2	414.7
1/3	68.7	231.1	174.8	117.4
2/1	130.8	424.8	423.1	429.6
2/2	89.8	230.8	178.0	217.2
2/3	68.4	159.0	149.6	218.6
2/4	56.1	119.6	95.3	86.1
3/1	130.2	230.4	125.8	148.8
3/2	89.1	159.2	135.4	181.4
3/3	68.3	121.9	113.2	91.7
3/4	56.3	97.5	82.9	86.4
3/5	47.6	81.6	75.6	70.2

Table 6: The fitness values of agents with varying costs values. The number of obstacles is fixed to 30 and the number of feeding places is fixed to 6.

Threshold	Reactive	Reactive-DN	Control	Control-2
0	91.1	163.2	146.7	215.9
10	90.9	163.7	147.6	213.6
20	89.6	158.4	147.7	239.6
50	86.5	144.7	142.0	220.9
100	86.5	144.9	148.1	210.2
150	74.5	96.9	152.6	179.5
200	68.9	73.1	139.5	203.8
250	60.9	49.7	129.2	182.6

Table 7: The fitness values of agents with varying thresholds for extreme hunger. The number of obstacles is fixed to 30 and the number of feeding places is fixed to 6.

Obstacles	Food	Reactive	Reactive-DN	Control	Control-2
-0	0	125.3	99.0	125.0	166.7
	5	184.8	123.5	127.9	183.8
	10	342.6	164.8	179.0	364.3
	15	657.1	233.7	349.2	688.9
-5	0	93.4	97.5	124.6	164.7
	5	122.4	121.1	112.3	157.3
	10	170.7	160.6	123.7	196.5
	15	286.6	226.8	181.2	476.1
-10	0	76.4	96.8	124.9	156.2
	5	93.4	120.5	116.6	157.6
	10	116.0	158.3	143.0	212.9
	15	156.3	227.8	172.2	488.6
-15	0	64.8	96.5	124.5	149.0
	5	76.0	120.1	107.9	145.4
	10	89.6	158.5	154.2	185.7
	15	109.4	225.1	142.7	408.5

Table 8: The fitness values of agents with varying values for the energy updates when hitting obstacles and eating food. The number of obstacles is fixed to 30 and the number of feeding places is fixed to 6.

Duration	Proportion	Reactive	Reactive-DN	Control	Control-2
20	1/3	149.1	180.6	161.9	415.8
	1/2	89.8	158.1	144.8	169.7
	2/3	69.5	145.3	114.2	182.4
30	1/3	132.8	176.7	177.1	385.3
	1/2	89.7	159.0	151.8	198.2
	2/3	67.5	143.6	118.5	199.7
40	1/3	134.3	176.7	171.0	302.1
	1/2	90.1	158.0	155.4	186.2
	2/3	70.6	144.6	127.8	178.4

Table 9: The fitness values of agents with varying values for the duration of the day and the proportion of the night. A proportion of $\frac{1}{3}$ means that $\frac{1}{3}$ of the full day-night cycle is night. The number of obstacles is fixed to 30 and the number of feeding places is fixed to 6.