Using explicit averaging fitness for studying the behaviour of rats in a maze

Ariadne de A. Costa¹, Patrícia A. Vargas² and Renato Tinós³

¹ Department of Physics, FFCLRP, University of São Paulo, Ribeirão Preto, SP, Brazil
² School of Mathematical and Computer Science, Heriot-Watt University, Edinburgh, Scotland, UK
³ Department of Computing and Mathematics, FFCLRP, University of São Paulo, Ribeirão Preto, SP, Brazil
ariadne.ad@pg.ffclrp.usp.br

Abstract

In this paper we study the performance of different evolutionary strategies based on explicit averaging. On a previous study, (Costa et al., 2012) proposed a probabilistic fitness function for an agent model based on neural networks and genetic algorithms employed to investigate the behaviour of rats in an elevated plus-maze (EPM). Differently from other computational models, the virtual rat proposed in (Costa et al., 2012) is not built based on experimental data comparisons with real rats, but, instead, is based on a behavioural model exploring the conflict between fear and anxiety. Despite the good results of the proposed agent, the effects of the uncertain fitness functions in the evolutionary learning process were not studied in the previous study. In our experiments we found significant differences in the performance of the genetic algorithm when the fitness of the individuals is sampled different times thus enabling us to define the best strategy for the studied problem.

Genetic algorithm, Uncertainty, Explicit averaging fitness, Elevated plus-maze, Rat

Introduction

Uncertainty has been widely studied in different areas, like politics (Cioffi-Revilla, 1998; Carmignani, 2003), economics (Naceur Jabnoun and Yusuf, 2003; Pindyck, 2007; Baker et al., 2013), movies (Miller and Shamsie, 1999; Vany, 2004; Gil, 2008), sports (Jennet, 1984; Peel and Dennis, 1992; Knowles et al., 1992; Forrest and Simmons, 2002; Buraimo and Simmons, 2008), criminology (Lattimore and Witte, 1986; Goulas and Zervoyianni, 2013) and sociology (Marris, 1996). By reviewing the uncertainty literature, we find many concepts for this word. Early studies claim that uncertainty is product of unpredictability (Cyert and March, 1963) or environmental turbulence (Emery and Trist., 1965). In (Lawrence and Lorsch, 1967), uncertainty is described as the lack of knowledge in the decision making process and similarly, in (Duncan, 1972), uncertainty is the absence of information for decision making. Uncertainty is also faced as result of the complexity of influential variables (Galbraith, 1973).

Working with evolutionary computation, there are various kinds of uncertainties involved. According to (Jin and Branke, 2005), these uncertainties may be classified into four groups:

1. the fitness function is subject to noise, which also may be comprehend as a problem of a partially observable environment;
2. the variables and/or the environmental parameters may change after simulation;
3. the fitness function is an approximation, which may cause errors;
4. the optimum of the problem is dynamic, so the optimizer has to seek the optimum continuously.

(Campi and Calafiore, 2004) adds another class to this list:

5. the optimization is based on sampling finite, small, number of instances (scenarios), normally because of the cost of obtaining each instance.

One of the classes of uncertainty problems that has been most studied is related to problems where the fitness function is subject to noise (Fitzpatrick and Grefenstette, 1988; Rattray and Shapiro, 1996; Sondahl and Stonedahl, 2010). Many of the solutions adopted to this class of problems investigate the influence of different selection schemes (Miller and Goldberg, 1996; Sano, 2002; Goschin et al., 2011) in the performance of the evolutionary algorithms. The problem studied here belongs to the first class of uncertain problems.

Uncertainty is also one of the greatest difficulties during decision making processes. The process of decision in rats navigating in mazes is widely investigated in animal behavioural studies (Salum et al., 2003; Walf and Frye, 2007). One of the mazes extensively used for this purpose is the elevated plus-maze (EPM). The EPM is an elevated plus-shaped maze composed of two opposed open arms and two opposed enclosed arms. The experiments using EPM are mostly derived from studies realized by Montgomery, in 1955, for the investigation of the conflict between fear and anxiety in rats during their exposition to a new environment (Montgomery, 1955). The conflict in the animal is
caused by sensations of curiosity and fear occurring simultaneously and understood as defence mechanisms (Graeff, 1990). Most of the studies of rats in the EPM are based on the results of Montgomery (Montgomery, 1955), which show that rats spend more time in the enclosed arms of the maze than in the open arms.

In the last years, some researches proposed computational models to simulate the rats behavior in the EPM (Salum et al., 2000; Giddings, 2002; Miranda et al., 2009; Shimo et al., 2010; Costa et al., 2012). In (Salum et al., 2000), artificial neural networks (ANNs), trained by competitive learning, control a virtual rat in a virtual EPM. The inputs of the ANN are: need for exploration, aversion repulsive stimuli, and spontaneous motor activity. In the paper, the authors also studied a completely open and a completely closed maze. (Miranda et al., 2009) proposes a very similar model. (Giddings, 2002) implements a computational model based on the fact that the rat generally remains in the direction that it already is. The probabilities of following the previous direction, or moving to other direction, are defined a priori for different parts of the maze.

Other approach is used in (Shimo et al., 2010; Costa et al., 2012), in which the virtual agent is obtained from a optimisation process. In this case, the agent is a mobile robot controlled by an ANN. The weights of the ANN are defined by a genetic algorithm (GA), which uses a fitness function based on the comparison of the trajectories of the individuals in a replica of the EPM and the trajectories of real rats in an EPM.

According to the authors’ knowledge, (Costa et al., 2012) proposed the first model in which the virtual agent is obtained independently from experimental data obtained with real rats. In the same way of the model proposed in (Salum et al., 2010; Costa et al., 2012), the computational agent (a virtual robot) is represented by an artificial neural network, whose weights are evolved by a genetic algorithm. The outputs of the ANN determine the next action of the virtual robot. The main contribution of (Costa et al., 2012) is a probabilistic fitness function that is not dependent of experimental data obtained with experiments using real rats, but is based on the relation of fear and anxiety related by Montgomery. The results of the experiments with real rats are employed only in the validation of the model. The experimental results showed that the virtual model was capable of reproducing the rats behaviour, when some parameters of the trajectories like number of entrances and time spent in each arm are considered. However, the effects of the uncertainty inherent to the fitness function were not studied on our previous study (Costa et al., 2012). The problem by using a probabilistic function, is that the comparison of candidate solutions can be unfair, causing bias in the selection of the individuals, and affecting the evolutionary process.

As mentioned in (Jin and Branke, 2005), one of the classes of uncertainty in evolutionary optimisation occurs when the fitness function is noisy, and there are many types of strategies to deal with that uncertainty. One of them is explicit averaging, where the fitness of an individual is sampled a number of times. In this paper, we test different evolutionary strategies based on explicit averaging for the model proposed in (Costa et al., 2012), with the intention of reducing the effects of the uncertainty in the fitness function employed by the virtual robot. One problem with explicit averaging is the additional computational costs for the optimisation algorithm. This is particularly true for problems where the evaluation process is costly, like in robots. If the evaluation is costly, an option is to use a less accurate fitness function and another option is to calculate the average based on fewer independent individual evaluations (Sondahl and Stonedahl, 2010). However, this is not the case of our problem, as the fitness evaluation can be split in two parts: one to obtain the trajectory of the agent, which is costly, and another to compute the fitness based on this trajectory. Since the last one is not costly, we are able to make many independent evaluations to calculate the mean fitness for each individual without significantly changing the time required to evaluate an individual.

In next section we explain the computational methods for the virtual rat and the four strategies studied in this paper to deal with uncertainty. In Section III, we present experiments comparing the fitness and the time spent in the enclosed and open arms for the virtual rats obtained with these strategies. Conclusions and future directions are presented in Section IV.

Methodology

In this paper, a computational agent (virtual robot) is employed to simulate the behaviour of a rat in the EPM. The virtual robot is controlled by a recurrent multilayer perceptron (Elman’s network) with ten inputs (six sensors and four recurrent signals from the hidden layer), four neurons in the hidden layer, and four neurons in the output layer. The recurrence is important because it allows that previous inputs be stored in the internal neurons, acting like a memory. The sensors are placed around the robot in order to detect the walls of the EPM. The outputs of the ANN indicate the next robot’s action (stay, turn left, turn right or go forward to the next position).

A genetic algorithm optimises the weights of the artificial neural network. This way, each individual of the GA corresponds to a chromosome composed by an array of integers, representing a possible solution in the fitness landscape. The initial population is randomly chosen.

The same virtual EPM proposed in (Salum et al., 2000) is employed in our work. Each arm of the plus-maze is divided in five positions plus the central position that links the four arms of the EPM, totalling 21 positions. The agent (rat or virtual robot) is evaluated by its navigation in this virtual EPM during a period of time corresponding to 5 minutes for
the real rat, or 300 time steps for the virtual rat.

The fitness function, proposed in (Costa et al., 2012) and employed here, is based on the conflict of fear and anxiety model (Montgomery, 1955). Two terms compose the fitness function: one for reward and other for punishment. The reward represents the curiosity of the rat in exploring not recently visited positions of the maze, while the punishment represents the exposure to damage. This last term is probabilistic, i.e., the same trajectory of the agent in the EPM can generate different values for this term. The fitness of individual $x$ is computed based on the trajectory of the agent controlled by the ANN with weights given by the chromosome of the individual in the virtual EPM. In this way, after selection and reproduction, each individual of the GA generates a trajectory in the virtual EPM and this trajectory generates the fitness of the individual according to the following function:

$$f(x) = \sum_{t=1}^{n} r(x, p_t) + s(x, p_t) \cdot \beta,$$  \hspace{1cm} (1)

where $p_t$ is the position of the virtual rat at time step $t$. The term of reward $(r(x, p_t))$ may increase the fitness of the individual, following the rule:

$$r(x, p_t) = \begin{cases} 
1, & \text{if } p_t \text{ was not visited for the agent in the last } \gamma \text{ time steps} \\
0, & \text{otherwise},
\end{cases}$$  \hspace{1cm} (2)

where $\gamma$ is a parameter of the model that is related to the memory of the virtual rat.

On the other hand, the punishment $(s(x, p_t))$ in Eq. 1 represents the exposure to the damage and decreases the fitness of the individual. It is known that the rat avoids the damage by spending more time in the enclosed arms, i.e., the level of damage is different for different positions of the maze. This way, the punishment is given by $s(x, p_t) \cdot \beta$ in Eq.(1), in which $\beta$ is the weight of the punishment and

$$s(x, p_t) = \begin{cases} 
-1, & \text{if } z_t < \alpha(p_t) \\
0, & \text{otherwise},
\end{cases}$$  \hspace{1cm} (3)

where $z$ is a random number, and $\alpha(p_t)$ is the level of damage of the position $p_t$ occupied in the time step $t$:

$$\alpha(p_t) = \begin{cases} 
\alpha_o \in [0, 1], & \text{if } p_t \text{ is an open arm} \\
\alpha_e \in [0, 1], & \text{if } p_t \text{ is an enclosed arm} \\
\alpha_c \in [0, 1], & \text{otherwise},
\end{cases}$$  \hspace{1cm} (4)

The more the rat is exposed to danger, the higher is the chance of it being punished, which means that the punishment is probabilistic. That is, the fitness function is noisy and may be compared to a robot walking around an uncertain environment. The strategies described in the next section were studied to deal with this kind of uncertainty.

Evaluation

As mentioned previously, we study four different evolutionary strategies for the GA. The individual’s selection and fitness depend on the employed strategy. In the experiments, for each strategy, we calculate the mean over 30 executions of the GA and the population size is constant throughout the simulation (500 individuals).

The fitness of the individual is calculated based on the trajectory of the corresponding agent. After navigating in the EPM, the corresponding trajectory of the virtual robot is recorded. Then, the fitness based on this trajectory is sampled $n$ times using Eq. (1) and the individual’s fitness is given by the mean fitness of these $n$ independent results. That is, the trajectory of a individual, which is the more costly part of the evaluation process, is performed only once. The mean of various fitness computations are calculated because of the uncertainties of the fitness function, as seen in the last section. The evaluation process may be seen in Figure 1. It is important to highlight that, in (Costa et al., 2012), the fitness is sampled only one time, i.e., $n = 1$.

Figure 1: Process of individual evaluation. First, the individual navigates in a simulated five-minute test. The fitness based on this trajectory is then sampled $n$ times. The resultant individual fitness is the mean of these independent fitness calculations.

Strategies

The four strategies studied here are presented below.

Strategy 1 The individuals are selected by elitism and tournament. By elitism, the two best individuals of the population are selected to the next population. In tournament selection, the best of two random individuals is selected, with probability 0.75, and the selected individual is then submitted to the reproduction operators. The tournament occurs until the population is completed. One-point crossover (with rate of 0.6) and mutation with uniform distribution (with rate of 0.05) are employed. The population evolves during 500 or 1500 generations (depending on the experiment), concluding an execution of the GA. Then (after GA simulation), the best individual of each of the 30 executions is evaluated with $n$ partial fitness calculations and the mean fitness
represent the fitness of these individuals. It is important to observe that during the GA evaluation, each individual is sampled with \( n = 1 \) on each generation of the evolutionary process.

**Strategy 2** This strategy is similar to Strategy 1. The difference is that each individual of the population is evaluated based on \( n \) samples of the fitness function. Then, at the end of each execution, the best individual is selected and it is not evaluated again.

**Strategy 3** Strategy 3 is the same of Strategy 1, except that the only type of selection is tournament. There is not elitism.

**Strategy 4** In this strategy we do the same of Strategy 2, but without elitism.

**Results**

In previous experiments, we have tested several sets of parameters and the selected for the virtual robot simulations are: \( \gamma(p_t) = 3 \), \( \beta = 5 \), \( \alpha_o = 0.015 \), \( \alpha_e = 0.012 \) and \( \alpha_c = 0.011 \), which are used in all simulations presented here.

Table 1 exhibits the mean fitness obtained with the Strategies 1, 2, 3 and 4, based on 30 executions of 500 generations each one. In the evaluation, the fitness is computed 10 times for the best individuals of each run. By the table, it is clear that the best results are achieved with Strategy 1, in which the fitness exhibited corresponds to the mean of the best individuals of 30 executions of the GA evaluated after simulations. Strategies without elitism (Strategy 3 and 4) are worse than their similar with elitism. It shows that elitism is important to the model.

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Strategy 1</th>
<th>Strategy 2</th>
<th>Strategy 3</th>
<th>Strategy 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>-10.94</td>
<td>-15.60</td>
<td>-11.29</td>
<td>-16.08</td>
</tr>
<tr>
<td>Sd</td>
<td>3.98</td>
<td>1.31</td>
<td>4.52</td>
<td>0.01</td>
</tr>
<tr>
<td>Min</td>
<td>-18.84</td>
<td>-16.12</td>
<td>-22.48</td>
<td>-16.11</td>
</tr>
<tr>
<td>Median</td>
<td>-11.15</td>
<td>-16.072</td>
<td>-11.60</td>
<td>-16.08</td>
</tr>
<tr>
<td>Max</td>
<td>-2.88</td>
<td>-11.41</td>
<td>-1.98</td>
<td>-16.06</td>
</tr>
</tbody>
</table>

Table 1: Simulated results of the fitness in 30 runs of 500 generations for Strategies 1, 2, 3 and 4, with \( n = 10 \).

To clarify how close our results are to the experiments with real rats, Figure 2 contains mean and standard deviation of the time spent in each arm and central position obtained by experiments with real and virtual robots. The experimental results with real rats are based on the trajectory of 47 rodents (Costa et al., 2012).

Virtual robots spend similar proportions of time in each arm of the EPM in relation to real rats, remaining substantially longer in the enclosed arms. The qualities of the strategies are in agreement with their respective fitness presented in Table 1.

Since we conclude that elitism is fundamental to keep best individuals in the population, the next results are all obtained with elitism. Our intention is to better understand the model and the influence of Strategies 1 and 2. For this purpose, we found the best individuals of 30 executions of both 500 and 1500 generations with Strategy 3 for \( n = 1 \), \( n = 10 \) and \( n = 100 \). Then, we evaluated them again after the runs, like is done in Strategy 1, calculating 100 partial fitness for each one of best individuals obtained in the 30 runs. We calculated the mean, standard deviation, median, maximum and minimum of the 30 fitness. The result is shown in Table 2.

<table>
<thead>
<tr>
<th>Statistic</th>
<th>1ft 500ger</th>
<th>1ft 1500ger</th>
<th>10ft 500ger</th>
<th>10ft 1500ger</th>
<th>100ft 500ger</th>
<th>100ft 1500ger</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>-9.79</td>
<td>-9.46</td>
<td>-11.50</td>
<td>-7.97</td>
<td>-0.92</td>
<td>1.60</td>
</tr>
<tr>
<td>Sd</td>
<td>3.16</td>
<td>3.77</td>
<td>8.04</td>
<td>6.08</td>
<td>9.76</td>
<td>4.50</td>
</tr>
<tr>
<td>Min</td>
<td>-17.91</td>
<td>-16.75</td>
<td>-22.03</td>
<td>-17.63</td>
<td>-21.21</td>
<td>-6.45</td>
</tr>
<tr>
<td>Max</td>
<td>3.57</td>
<td>3.31</td>
<td>2.70</td>
<td>-7.37</td>
<td>6.95</td>
<td>9.09</td>
</tr>
</tbody>
</table>

Table 2: Mean, standard deviation, minimum, median and maximum for the best individual of 30 executions of 500 and 1500 generations of Strategy 2 with \( n = 1 \), \( n = 10 \) and \( n = 100 \), associated with Strategy 1 with \( n = 100 \).

From Table 2, it is clear that increasing from 500 to 1,500 the number of generations in the GA, the quality of the re-
sults is improved in any studied situation. It also may be seen that increasing the number of samples of the fitness function is beneficial to the model. To improve the comprehension of what is happening, we analysed histograms of fitness coming from Strategy 2 with \( n = 1 \) and \( n = 100 \). The best individual of the 30 executions is evaluated again with 1000 samples (partial fitness). These samples are in the histograms of Fig. 3, which also exposes the original fitness for \( n = 1 \) and \( n = 100 \), i.e., the fitness of the best individual of the simulations whose trajectory was selected to be evaluated a thousand times (\( f_1 \) and \( f_{100} \), for the case with \( n = 1 \) and \( n = 100 \) respectively). We observe that \( f_{100} \) is a little higher than the average of 1,000 evaluations, while \( f_1 \) is much higher, close to the maximum fitness obtained in the 1,000 evaluations. The value for \( f_1 \) is high because the individual’s fitness obtained with the calculation of only one sample is more susceptible to randomness. Since various individuals in the population have similar genome, that one with the maximum fitness is selected with higher probability, causing the bias in the histogram. On the other hand, \( f_{100} \) is the mean of a thousand samples, which ensures that the fitness will be near the mean fitness allowed for the trajectory evaluated for the various individuals with the same genome in the population. As this individual is selected with higher probability, it is progressively better along the executions. But this process occurs in a very slow way, compared to the case for \( f_1 \).

The Figure 3 enables to observe how noisy is the fitness function studied. One can observe, in the histograms, regions without samples. This occurs because the punishment weight is equal to 5.8 (\( \beta = 5.8 \)) in the runs, whereas the reward is equal to 1. So, there are values that are not covered by the possible combinations of rewards and punishments in Eq. 1 (one can remember that the reward is deterministic while the the punishment is probabilistic for a given trajectory).

Figure 4 shows the maximum, mean and minimum fitness of each generation for the best of 30 executions of the GA in simulations of Strategy 2 with \( n = 1 \), \( n = 10 \) and \( n = 100 \). As it was already presented, increasing the number of samples improves the fitness achieved by the optimisation process. The figure demonstrate that there is a jump in fitness around the 500th generation in strategy with \( n = 100 \), revealing that the virtual robot learns a new skill at this point of evolution. It is interesting to observe that it only happens when 100 samples of the fitness are considered. By analysing the amount of punishments and rewards received by the best individual of each generation, we can note that this jump is due to fact that the virtual robot have learned how to be more rewarded in his trajectory.

Conclusions

The results show that explicit averaging is important to the studied problem, as our fitness function is probabilistic. The

---

Figures:

- Figure 3: Histogram of 1,000 samples for the best individual obtained in 30 runs for Strategy 2 with \( n = 1 \) (up) and \( n = 100 \) (bottom). The tables show the mean, standard deviation, minimum, median, maximum and the original fitness.

- Figure 4: The up three lines are the maximum fitness; the central three lines correspond to the mean fitness; and the bottom lines are the minimum fitness of each generation for a run of Strategy 2 with \( n = 1 \), \( n = 10 \) and \( n = 100 \).
fitness function studied is very noisy, influenced by the reward and punishment weight values in the fitness function.

It is relevant to highlight that the navigation in the EPM is the most costly part of the evaluation process. It takes a robot a long time to perform its trajectory in an EPM's replica (as it did in (Shimo et al., 2010)). Hence, it is possible and fair to calculate a high number of samples of the fitness function for each individual evaluation.

Among the four strategies studied for the problem of the rat navigating in an elevated plus-maze, the most effective one is to evaluate the best individual with 100 samples of the fitness functions during the optimisation process. This is caused by the fact the explicit averaging smooths the effect of randomness in the fitness function. We also noticed that selection by elitism plays an important role in the evolution of the evolutionary robots for the problem studied here.

As a future work, we intend to study the influence performed by the ANN and each neuron from the hidden layer. Moreover, we intend to test the methods also in real robots.

Acknowledgements

This research is supported by FAPESP, CNPq and CAPES. The authors are grateful to Helder Ken Shimo and to Julián Tejada for their contributions.

References


