Single Step Evolution of Robot Controllers for Sequential Tasks

Stéphane Doncieux
ISIR, CNRS UMR 7222
Université Pierre et Marie Curie (UPMC)
4 place Jussieu, F-75005, Paris, France
doncieux@isir.fr

Jean-Baptiste Mouret
ISIR, CNRS UMR 7222
Université Pierre et Marie Curie (UPMC)
4 place Jussieu, F-75005, Paris, France
mouret@isir.fr

ABSTRACT

The generation of robot controllers for a task requiring a sequence of elementary behaviors is still a challenge. If these behaviors are known, intermediate steps can be given to help bootstrap the search, thus leading to task decomposition or incremental approaches. The goal of this paper is to propose an alternative, within which such behaviors do not need to be known. The proposed approach relies on a classical multi-objective evolutionary algorithm and consists in designing objectives dedicated to the enhancement of evolutionary search abilities. These objectives are to be used in addition to performance objectives rewarding the efficiency, robustness, or whatever aspect a robot designer might be interested in. Two different kinds of objectives are proposed, tested and compared on a ball collecting problem. Both rely on states that can be directly extracted from the sensors and are completely independent from the genotype and phenotype. They show promising results, even with a simple direct neural network encoding.

Categories and Subject Descriptors: I.2.6 [Artificial intelligence]: Connectionism and neural nets

General Terms: Algorithms.

Keywords: Evolutionary robotics, multi-objective evolutionary algorithms, behavioral diversity, sequential tasks.

Complex task resolution with Evolutionary Algorithms is often faced with premature convergence to an easy to find local optimum. In a robotics context, several solutions have been proposed. Incremental approaches require a manual decomposition of the task to be solved in simpler sub-tasks. The task is then solved through several steps, each step concentrating on a particular sub-task [4]. Such approaches imply a sequential resolution that is highly dependent on the quality of the task decomposition. Another related possibility consists in using a multi-objective scheme to add as many different selective pressures as necessary. New objectives may thus reward sub-tasks [5], with the advantage of an automatic switch between the different tasks. To avoid premature convergence, another objective may also directly push towards diversity on the genotype [1]. Multi-objective evolutionary algorithms can then be exploited to enhance the search ability of the algorithm. In this *multiobjectivization* approach [3], we will propose two new simple objectives that are adapted to robot tasks requiring a sequence of elementary behaviors.

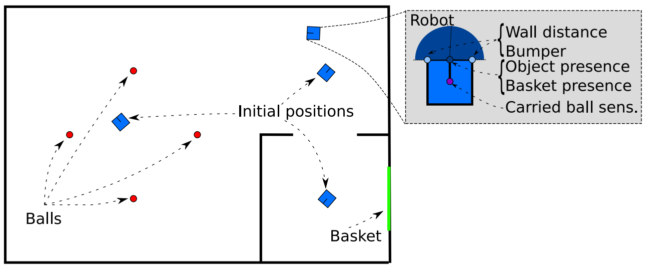

Figure 1: Overview of the arena and of the robot. The initial positions of the three evaluation experiments used for the fitness computation are plotted on the figure.

Both objectives rely on the definition of discrete states on the basis of sensor information. The states are used for fitness design only and are not used at all in the robot controller.

The first objective relies on an available knowledge on the 'desired" transitions between different states and consists in rewarding the robot for following a supposedly right sequence of states. Fitness evaluation requires the definition of two counters for each state: n_t and n_g; n_t evaluates the total number of times the state has been reached and n_g evaluates the number of times that the following state was the desired one. If N is the number of different states, the fitness objective may be expressed as follows:

$$o_{st}(x) = \sum_{i=0}^{N} stc(x,i) * 10^{rg(x,i)}$$

$$\text{with } stc(x,i) = \begin{cases} min(9, 1+n_g[i]) & \text{if } n_t[i] \neq 0 \\ 0 & otherwise \end{cases}$$

and $rg(x,i) = \sum_{j=0}^{i}(1-\delta_0^{n_t[j]})$, δ_i^j being Kronecker delta.

Our second objective defines a selective pressure towards diversity, as evaluated on the basis of the behavior, not on the genotype or phenotype. The behavioral diversity will be evaluated here on the basis of simple statistical information concerning the states we have previously defined. The behavioral diversity objective requires to compute the vector $n_{t,x}$ reflecting the total number of times each state has been reached by the robot described by genotype x. The behavioral diversity objective associated to individual x is then: $o_{bd}(x) = \frac{1}{size(P)} \sum_{y \in P} d(x,y)$ with P the current pop-

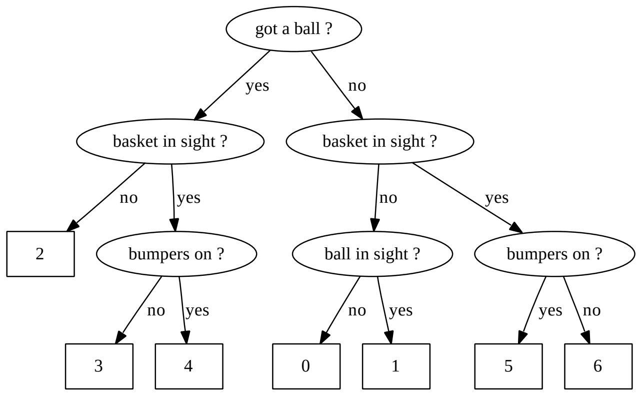

Figure 2: State definition tree used with the arena experiment described on figure 1.

ulation and $d(x,y) = \sum_{i=0}^{n_s-1} d_i(x,y)$ with n_s the number of states. $d_i(x,y)$ is a square euclidian distance normalized by the maximum number of times state i has been reached by x and y. This avoids individuals reaching a huge number of times a particular state to be over-rewarded.

To test the proposed fitness objectives, we have defined a task consisting in collecting some balls in an arena and putting it into a basket located at a particular place (fig. 1). The robot is a simple two-wheeled robot that has the following sensors: two wall distance sensors; two bumpers; two object detection sensors; two basket detection sensors; one carry ball sensor. The effectors are the following: left wheel motor; right wheel motor; catch ball motor. Once a ball is thrown, even not in front of the basket, it is removed from the arena. Figure 1 shows the arena with the details of the sensor configuration of the robot.

The chosen states are depicted on figure 2. We have also built a simple goal objective $o_g(x)$, that takes the number of collected balls $n_c(i)$ during evaluation i, summed over the n_{eval} evaluations of individual x. Last, we have defined an objective $o_r(x)$ with a random value for comparison purpose: this objective takes a random value between 0 and 10. We end up with five different experiment setups which only differ on the objectives to maximize. The *Control* experiment maximizes $o_g(x)$ only and all other experiments add one or two objectives: *random* experiment adds $o_r(x)$, *diversity* experiments adds $o_{bd}(x)$, *transition* adds $o_{st}(x)$ and *transition+diversity* adds $o_{bd}(x)$ and $o_{st}(x)$.

Individuals are tested in three different setups characterized by the robot starting points which are represented on figure 1. Robot controllers are neural networks whose structure and parameters are generated via the EA with a direct encoding. The EA we used is NSGA-II [2]. 20 runs with 2000 generations were launched for each configuration. The mean, max and min numbers of collected balls over the twenty launched runs at some particular generations are the following (best values are in bold font, standard deviation are also provided):

name	stat	500	1000	1500	2000
control	mean	0.35	0.6	0.6	0.75
	max	2	2	2	3
	min	0	0	0	0
	stdev	1.5	1.6	1.6	1.7
random	mean	0.35	0.5	0.6	0.7
	max	2	2	2	3
	min	0	0	0	0
	stdev	1.4	1.5	1.7	1.8
diversity	mean	2.7	3.95	**4.55**	**4.8**
	max	**6**	**9**	**9**	**9**
	min	**1**	**1**	**1**	**1**
	stdev	2.6	3.2	4.5	5.6
transition	mean	**3.5**	4.2	**4.55**	4.75
	max	**6**	8	8	8
	min	**1**	**1**	**1**	**1**
	stdev	2.9	3.2	4.5	4.5
transition+diversity	mean	3.05	**4.25**	**4.55**	**4.8**
	max	**6**	8	8	8
	min	**1**	**1**	**1**	**1**
	stdev	2.0	4.0	5.0	5.2

Control and *random* experiments (group 1) generate controllers that can collect no more than one ball per evaluation, whereas the other experiments (group 2) have all generated individuals able to do it. The statistical difference between these two groups of experiments is significant (T-Test) and the difference between members of a same group is not significant.

The two proposed additional objectives have been able to drive the evolutionary process to solve the task in a single evolutionary run, i.e. without any behavioral decomposition and incremental approach. Using a direct encoding, neural network controllers have been evolved that are able to catch a ball, navigate towards a hidden basket, reach the basket and release the ball inside it before going back exploring and finding new balls to catch and doing this cycle again until no balls are available or the evaluation time has elapsed. Both objectives are independent from any controller encoding and are compatible in particular with indirect encodings. They require to define states on the basis of sensorial data. The first objective requires to know what the desired state transitions are, i.e. what state the robot should reach from a particular state.The second objective requires less a priori knowledge and aims at encouraging behavioral diversity. The two objectives gave similar results. Being simpler and based on less a priori knowledge, the behavioral diversity objective appears as the most promising approach.

REFERENCES

[1] E. D. de Jong, R. A. Watson, and J. B. Pollack. Reducing bloat and promoting diversity using multi-objective methods. *Proceedings of the Genetic and Evolutionary Computation Conference*, pages 11–18, 2001.

[2] K. Deb, S. Agrawal, A. Pratab, and T. Meyarivan. A Fast Elitist Non-Dominated Sorting Genetic Algorithm for Multi-Objective Optimization: NSGA-II. In Marc Schoenauer et al., editor, *Proceedings of the PPSN VI Conference*, pages 849–858, Paris, France, 2000. Springer. LNCS No. 1917.

[3] Julia Handl, Simon Lovell, and Joshua Knowles. Multiobjectivization by decomposition of scalar cost functions. In *Proceedings of the PPSN X Conference*, pages 31–40. 2008.

[4] J. Kodjabachian and J.-A. Meyer. Evolution and development of neural networks controlling locomotion, gradient-following, and obstacle-avoidance in artificial insects. *IEEE Transactions on Neural Networks*, 9:796–812, 1997.

[5] J.B. Mouret and S. Doncieux. Incremental evolution of animats' behaviors as a multi-objective optimization. In *From Animals to Animats 10*, volume 5040 of *Lecture Notes in Computer Science*, pages 210–219. Springer, 2008.