Choosing the Right Number of Trials for a Minimal Simulation

Adrian Trenaman
National University of Ireland, Maynooth, Co. Kildare, Ireland
Phone: +353-1-7083354, Email: trenaman@cs.may.ie

Abstract

This paper reports on experience gathered in the evolution of robot-controllers using minimal simulations. As a result of the random noise employed in the minimal simulation, the fitness landscape proved deceptive and large numbers of trials were necessary to provide an accurate depiction of fitness.

Using a minimal simulation neural networks were evolved to control a Khepera [2] robot in a T-Maze, where the robot chooses to turn left or right depending on its previous experience [1]. Fitness was measured as the length travelled in the T-maze, with extra points awarded for turning the correct way at the junction. This was averaged over a number of trials where the corridor-width and starting angle of the Khepera was varied to ensure robustness of the solutions. Initially the number of trials was set to 10, and the best-fitness curve for a typical run is shown in Figure 1. The curve is highly erratic and it was found that high-scoring individuals did so by chance, employing simple reactive wall-following strategies rather than more-sophisticated memory-based behaviour.

To determine the effect of varying the number of trials with robustness of the fitness evaluator, the minimum, maximum, mean and standard deviation of fitness was measured over different numbers of trials, as shown in Table . Clearly more trials are required to reduce the standard deviation and so increase the robustness of solutions.

Using 70 trials to evaluate each indivdiual, much better results were observed, with best, average and worst fitness curves shown in Figure 2. These results suggest that when randomness is employed in a fitness-function, the number of trials must be carefully examined to ensure robustness of solutions.

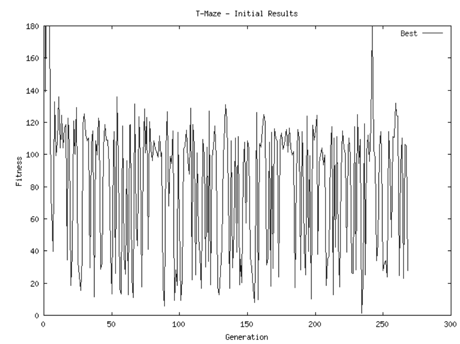

Figure 1: Best fitness curve for initial runs on the T-maze problem when only 12 evaluations were used.

Trials	min	mean	max	stdev
10	1.97068	66.8218	114.544	19.8848
20	32.9452	65.1106	97.5312	14.3681
30	38.7639	62.1924	89.0272	10.8576
40	42.9916	64.7262	90.1784	10.3004
50	42.2339	64.5591	88.7796	9.99436
60	46.2047	65.5714	86.32	8.26861
70	51.7541	63.9567	85.0338	7.12731
80	45.6503	65.8529	78.4363	7.11616
90	48.3097	65.4392	80.2008	6.46012
100	50.8544	65.3903	83.7711	6.65713

Table 1: Robustness of fitness evaluator under different numbers of trials.

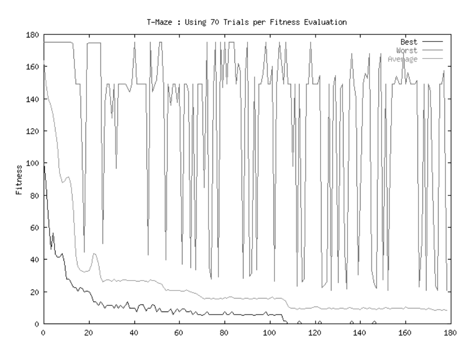

Figure 2: Fitness curves for runs on the T-maze problem when only 70 evaluations were used.

References

[1] Nick Jacobi. Evolutionary robotics and the radical-envelope-of-noise hypothesis. *Adaptive Behaviour*, 6:131–174, 1997.

[2] Mondada et al. K-team. More information available at `www.k-team.com`.