# Co-evolution of Initial Configuration and Control in Evolutionary Robotics

Kasper Stoy

IT University of Copenhagen, Denmark
ksty@itu.dk

## Abstract

In evolutionary robotics, we evaluate individuals by placing them in an initial configuration in the environment, and then measure their fitness over a period of time. The choice of initial configuration has a direct impact on the fitness of an individual and thereby also the overarching evolutionary process. In this paper, we propose the concept of dynamic initial configurations, which is an initial configuration that is neither random nor fixed, but develops dynamically in response to the evolutionary process. As an example we have implemented a competitive co-evolutionary algorithm where initial configurations and controllers are evolved together to solve an obstacle avoidance task of a mobile robot. We show that, while a evolutionary approach taken from literature consistently fails, the co-evolutionary approach succeeds in 22 out of 25 runs. This example demonstrates the benefit of dynamic initial configurations, but more work is needed to establish if the concept generalizes to more complex tasks, environments and morphologies.

## Introduction

In evolutionary robotics (Nolfi and Floreano, 2000; Doncieux et al., 2015), many factors influence the performance of an evolutionary robotics setup among others the sensory-motor system of the robots, the genotype-to-phenotype mapping, the fitness function and the environment (Fehérvári et al., 2013). In this paper we focus on one of these factors, namely, the choice of initial configuration of the robot in the environment at the onset of fitness evaluation.

A controller should perform well independent of its initial configuration. It is therefore common to gain coverage by sampling from all allowable initial configurations (e.g. all valid positions and orientations in the environment). While this can work, it hinders the evolutionary process because the fitness evaluations become noisy and randomly chosen initial configurations may not be optimal for differentiating the fitness of individuals. Another used alternative is a fixed initial configurations as this clearly reduce noise, but unfortunately this often leads to over-fitting to the initial configuration. On the basis of these two approaches, we hypothesize that dynamic initial configurations can address some of these challenges. A dynamic initial configuration is one that can adapt dynamically over generations to ideally provide good opportunities for differentiating between individuals, counter over-fitting by ensuring coverage, and reduce noise in fitness evaluation.

In this paper we contribute a first attempt at implementing dynamic initial configurations using co-evolution between controller and initial configuration. We use the canonical task of evolution of obstacle avoidance of a mobile robot as evaluation of our approach. We use a challenging variation of this task where obstacles are rare due to the environment being a large, empty arena. In our experiments, we find that a standard evolutionary approach fails to evolve obstacle avoidance. In contrast, the co-evolutionary approach evolves controllers successfully in 22 out of 25 cases, because the co-evolutionary process continuously selects initial configurations that differentiate controllers and prevents over-fitting.

Co-evolution of the initial configuration has synergies with incremental evolution (Gomez and Miikkulainen, 1997) of task (Rossi and Eiben, 2014), morphology or environment (Bongard, 2011; Wang et al., 2019). These approaches require encoding of the task or environment to be used by the evolutionary process and thus are complex, but arguably can also handle more complex, open-ended problems. In contrast, we only encode the configuration which is simple and practical and well suited for problems where the task-environment is given.

## Experimental Setup

We have developed a kinematic simulation of the differential drive Khepera IV robot (Soares et al., 2016) using simple ray-casting to model the eight infrared sensors. The environment is a walled, square arena with a side length of $8m$. The evolutionary robotic system follows that of (Mondada and Floreano, 1996) unless otherwise mentioned. The controller consists of two perceptrons with eight inputs, two recurrent connections between the outputs and one bias unit. The corresponding chromosome consists of 22 genes encoding the weights of the neural network. For the initial configuration population we use a chromosome of 3 genes encoding position and orientation. The fitness function for obstacle avoidance is adapted to explicitly reward forward movement in order to prevent the near optimal solution where the robot moves back and forth repeatedly. The fitness function for the population of initial configurations is one minus the fitness of the controller. Hence, the system is a competitive co-evolutionary system.

For the evolutionary process we use a single point crossover with a probability of 0.6, a mutation with probability 0.02 which replaces a gene with a random number which for the weights is normalized to be between 8 and -8. We use an elitism of 8 and the rest of a generation is selected using tournament selection with a tournament size of 2. We use 256 individuals and run the process for 500 generations where each individual is evaluated for 5 minutes. For the initial configuration part, we use 128 individuals and an elitism of 4. The initial configuration is limited to be inside the arena and at least $20cm$ from the wall, which is just outside of sensor range which is $15cm$ from the center of the robot. For the co-evolutionary experiment we evolve the controller and the initial configuration alternately one generation at a time. For fitness evaluation the most fit individual from the other population is used.

We performed two evolutionary experiments with this setup. In the first experiment, which works as a base case, we use random initial configurations and in the second experiment we co-evolve the initial configurations.

## Results

The standard evolutionary process based on random initial configurations quickly achieves a high average fitness of around 0.73 at generation 15 and then does not improve further consistently across 25 runs. By inspecting the trajectories of the best individual from the last generation of each evolutionary run, we find that the 14 individuals that encounter a wall collide with it (we ignore the remaining 11 individuals which do not come near the wall during evaluation). Hence, the controllers fail to learn the obstacle avoidance task. Instead the controllers have just learned to move forward. The reasons for this are a) an individual is unlikely to encounter a wall, b) if it encounters the wall it will have a low fitness and thus low chance of survival. In fact, we find that only below an arena side length of 3m the probabilities of encountering an obstacle is high enough to support the evolution of obstacle avoidance. We also tried to place the robots close to the wall, but randomize the orientation. In which case the individuals facing the arena always get higher fitness and therefore are selected over robots facing the wall.

In the second experiment we co-evolve the initial configurations with the controllers. The average fitness of 25 runs reaches 0.6 at generation 75 and stays constant. The low fitness is a consequence of the co-evolutionary process that places robots in difficult initial configurations as can be seen

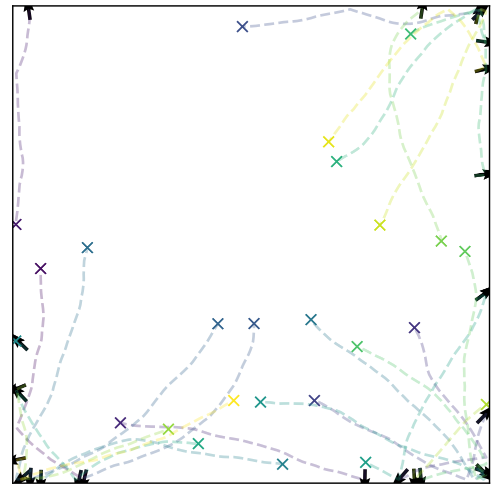

Figure 1: The robot arena (8m x 8m) and the trajectories of the best individual of each evolutionary run (25). The arrows and crosses represent initial and final configurations, respectively. The initial configurations are evolved to be close to and facing the wall.

in Figure 1. We can also see that 22 out of 25 runs (88%) lead to individuals able to perform obstacle avoidance and steer away from the wall (counting the crosses not placed at the wall). Hence, the co-evolutionary process works by finding challenging configurations in the environment, which most effectively differentiate individuals from the onset of fitness evaluation while ensuring coverage. This allows obstacle avoidance controllers to be evolved even in a large, empty environment. In order to validate the hypothesis that it is beneficial for the robot to be placed in a challenging configuration we tried a fixed initial configurations close to and facing a corner. For this specific initial configuration, the robots would learn the task in 80% of the runs supporting the hypothesis. However, in contrast to the co-evolved controllers, these controllers are over-fitted to the specific initial configuration and do not generalize to other initial configurations.

## Conclusion and Perspective

In conclusion, we find that for this particular task, where fitness differentiating features are rare, dynamic initial configurations are able to solve the task by finding these features with ease while preventing over-fitting. While our results are promising, this work is only preliminary as it still remains to be seen how the controllers handle the reality gap (Jakobi et al., 1995). In general, the work also raises questions about the choice of initial configurations for tasks with a natural fixed starting configuration (e.g. maze solving, walking, etc.). However, it remains to be seen if dynamic initial configurations can generalize to these domains.

## References


Bongard, J. C. (2011). Morphological and environmental scaffolding synergize when evolving robot controllers: Artificial life/robotics/evolvable hardware. In *Proceedings of the 13th Annual Conference on Genetic and Evolutionary Computation*, page 179–186, New York, NY, USA. Association for Computing Machinery.

Doncieux, S., Bredeche, N., Mouret, J.-B., and Eiben, A. E. G. (2015). Evolutionary robotics: What, why, and where to. *Frontiers in Robotics and AI*, 2:4.

Fehérvári, I., Trianni, V., and Elmenreich, W. (2013). On the effects of the robot configuration on evolving coordinated motion behaviors. In *2013 IEEE Congress on Evolutionary Computation*, pages 1209–1216.

Gomez, F. and Miikkulainen, R. (1997). Incremental evolution of complex general behavior. *Adaptive Behavior*, 5(3-4):317–342.

Jakobi, N., Husbands, P., and Harvey, I. (1995). Noise and the reality gap: The use of simulation in evolutionary robotics. *Lecture Notes in Computer Science (Lecture Notes in Artificial Intelligence)*, 929:704–720.

Mondada, F. and Floreano, D. (1996). Evolution and mobile autonomous robotics. *Towards Evolvable Hardware. Lecture Notes in Computer Science*, pages 221–249.

Nolfi, S. and Floreano, D. (2000). *Evolutionary Robotics: The Biology, Intelligence, and Technology of Self-Organizing Machines*. MIT Press.

Rossi, C. and Eiben, A. (2014). Simultaneous versus incremental learning of multiple skills by modular robots. *Evol. Intel.*, 7:119–131.

Soares, J. M., Navarro, I., and Martinoli, A. (2016). The Khepera IV mobile robot: Performance evaluation, sensory data and software toolbox. In Reis, L. P., Moreira, A. P., Lima, P. U., Montano, L., and Muñoz-Martinez, V., editors, *Robot 2015: Second Iberian Robotics Conference*, pages 767–781, Cham. Springer International Publishing.

Wang, R., Lehman, J., Clune, J., and Stanley, K. O. (2019). Paired open-ended trailblazer (poet): Endlessly generating increasingly complex and diverse learning environments and their solutions.