Online Evolution of Adaptive Robot Behaviour

Fernando Silva, Instituto de Telecomunicações, Lisboa, Portugal & LabMAg, Faculdade de Ciências, Universidade de Lisboa, Lisboa, Portugal

Paulo Urbano, LabMAg, Faculdade de Ciências, Universidade de Lisboa, Lisboa, Portugal

Anders Lyhne Christensen, Instituto de Telecomunicações, Lisboa, Portugal & Instituto Universitário de Lisboa (ISCTE-IUL), Lisboa, Portugal

ABSTRACT

The authors propose and evaluate a novel approach to the online synthesis of neural controllers for autonomous robots. The authors combine online evolution of weights and network topology with neuromodulated learning. The authors demonstrate our method through a series of simulation-based experiments in which an e-puck-like robot must perform a dynamic concurrent foraging task. In this task, scattered food items periodically change their nutritive value or become poisonous. The authors demonstrate that the online evolutionary process, both with and without neuromodulation, is capable of generating controllers well adapted to the periodic task changes. The authors show that when neuromodulated learning is combined with evolution, neural controllers are synthesised faster than by evolution alone. An analysis of the evolved solutions reveals that neuromodulation allows for a more effective expression of a given topology’s potential due to the active modification of internal dynamics. Neuromodulated networks learn abstractions of the task and different modes of operation that are triggered by external stimulus.

Keywords: Adaptation, Artificial Neural Network, Neuroevolution, Neuromodulation, odNEAT, Online Evolution, Online Learning, Synaptic Plasticity

INTRODUCTION

The development of control systems for autonomous robotics has progressed significantly since the manual programming in low-level languages of assembly robotics (Lozano-Perez, 1983). In particular, evolutionary computation techniques have been widely studied with the purpose of automating the design of robotic systems (Floreano & Keller, 2010). In evolutionary robotics (ER), robot controllers are typically based on artificial neural networks (ANNs) due to their capacity to tolerate noise in sensors. The parameters of the ANN, namely the connection weights, the neuron bias terms, and occasionally the topology, are optimised by an evolutionary algorithm (EA), a process termed neuroevolution (Yao, 1999; Floreano et al., 2008).

Evolutionary synthesis of neurocontrollers is usually performed offline in simulation, which presents a number of limitations. When a suitable neurocontroller is found, it is deployed on real robots. Since no evolution or adaptation
takes place online, the controllers are fixed solutions that remain static throughout the robot’s lifetime. If environmental conditions or task parameters become distinct from those encountered during offline evolution, the evolved controllers may be incapable of solving the task as they have no means to adapt.

Online evolution is a process of continuous adaptation that allows robots to modify their behaviour in order to respond to changes in the task or in environmental conditions. An EA is executed on the robots themselves while they perform their tasks. This way, robots may be capable of long-term self-adaptation in a completely autonomous manner. In recent years, different approaches to online evolution have been proposed, see for instance (Watson et al., 2002; Bianco & Nolfi, 2004; Bredeche et al., 2012). Notwithstanding, in such contributions, online neuroevolution has been limited to evolving weights in fixed-topology ANNs. In a recent study (Silva et al., 2012), we proposed a novel approach called odNEAT. odNEAT is an efficient online, distributed, and decentralised EA for online evolution in groups of robots. odNEAT optimises both weights and network topology as part of a continuous evolutionary process.

Evolution is a form of adaptation that acts at genotypic level. Controllers produced are static in the sense that they do not change their parameters while they are controlling the robot. Whereas evolution produces phylogenetic adaptation, online learning operates on a shorter time-scale. Learning acts at phenotypic level and gives each individual controller the capability to self-adjust during task execution. Several studies indicate that learning can accelerate the evolution of good solutions, a phenomenon known as the Baldwin effect (Hinton & Nowlan, 1987).

Agents controlled by ANNs can learn from experience by dynamically changing their internal synaptic strengths. This mechanism is inspired by how organisms in nature adapt to cope with dynamic and unstructured environments as a result of synaptic plasticity (Niv et al., 2002). In this article, we synthesise behavioural control for autonomous robots based on online evolution and online learning. We combine odNEAT, capable of efficient evolution of weights and network topology, with neuromodulation (Soltoggio et al., 2008). In biological organisms, neuromodulation is a form of synaptic modification involving modulatory neurons that diffuse chemicals at target synapses. Neuromodulation has been suggested as essential for stabilising classical Hebbian plasticity and memory (Bailey et al., 2000). The combination of online evolution and neuromodulated learning allows the evolutionary process to explore two distinct types of plasticity: (i) structural plasticity, the generation of new connections and neurons, which in turn redefines the network topology, and (ii) synaptic plasticity that changes the strength of existing connections in a given topology.

We investigate if synaptic plasticity through neuromodulation facilitates online adaptation to novel and dynamic environmental conditions. To that end, we setup a simulated experiment in which an e-puck-like robot (Mondada et al., 2009) must perform a dynamic concurrent foraging task, an extension of the dangerous foraging task described by Stanley et al. (2003). The robot must locate and consume scattered food items. When a food item is consumed, a new item of the same type is randomly placed in the environment. At regular time intervals, food items change their nutritive value or become poisonous. Besides learning to forage, the robot must also be able to change its foraging policy when necessary in order to survive. It should be noted that in our domain, the concurrent foraging, there are simultaneously two types of food items in the environment. The task continuously changes over time as nutritious items become poisonous and vice-versa. The concurrent foraging task therefore differs significantly from the dangerous foraging task, which is based on repeated trials of experiments that contain only a single type of food, either nutritious or poisonous, and in which the robot does not have to explore the environment to survive.

The article is organised as follows. In the section titled Background, we discuss previous
Related Content

Structural Alignment of RNAs with Pseudoknots
[www.igi-global.com/chapter/structural-alignment-rnas-pseudoknots/52332?camid=4v1a](www.igi-global.com/chapter/structural-alignment-rnas-pseudoknots/52332?camid=4v1a)

Feature Selection Based on Clonal Selection Algorithm: Evaluation and Application
[www.igi-global.com/chapter/feature-selection-based-clonal-selection/19645?camid=4v1a](www.igi-global.com/chapter/feature-selection-based-clonal-selection/19645?camid=4v1a)
Linear Time Solution to Prime Factorization by Tissue P Systems with Cell Division
*International Journal of Natural Computing Research* (pp. 49-60).
[www.igi-global.com/article/linear-time-solution-prime-factorization/58066?camid=4v1a](www.igi-global.com/article/linear-time-solution-prime-factorization/58066?camid=4v1a)

Lost Student Tracking in an Incomplete and Imprecise Information Environment Using Soft Computing Paradigm
*International Journal of Artificial Life Research* (pp. 32-48).