Evolving fast-learning neural controllers for a mobile robot

Dario Floreano
Cognitive Technology Laboratory
AREA Science Park
Padriciano, 34012 Trieste, Italy

Francesco Mondada
Microcomputing Laboratory
Swiss Federal Institute of Technology
1015 Lausanne, Switzerland

Several recent experiments have shown that evolutionary algorithms can be employed for developing interesting neural controllers in real situated agents, that is mobile robots that interact with an environment. However, an important drawback of artificial evolution is slow adaptation. In unpredictable real-world situations, it would be preferable to have faster adaptation mechanisms, while still retaining all the advantages offered by evolutionary development. Rather than evolving fixed-weights neural networks, here we suggest to evolve *learning neural controllers*, that is neural networks that can learn during the "life" of the agent without external supervision according to evolved genetic specifications. In our preliminary explorations, we have used the miniature mobile robot Khepera reproducing the same simple experimental setup described in [Floreano e Mondada, *IEEE Trans. on SMC-Part B*, 23, 1996]. The robot is placed in a maze where the selection criterion is a three-term function maximised by a combination of low sensor activity (obstacle avoidance), same direction of wheel rotation (straight trajectory), and average wheel rotation speed. The genotype codes a set of individual parameters for each synaptic link of a fully connected recurrent network (11 nodes: 8 clamped to the infrared sensors, 2 linked to the motors, and 1 functioning as an "internal" node): whether the link is excitatory or inhibitory, whether its effect on the postsynaptic unit is driving or modulatory, its learning rate between 0.0 and 1.0), and its learning modality (one among four Hebbian rules: "pure Hebb", presynaptic, postsynaptic, and covariance). The synaptic rules are designed so that they cannot invert the synaptic sign and have an intrinsic decay factor that limits excessive synaptic strenghtening. When a genetic string is decoded into the corresponding neural controller and downloaded into the robot, its synaptic strenghts are initialised to small random values in the range between 0.0 to 0.1. The robot is let free to act for 24 s while its own synaptic values change according to the genetic instructions and its fitness is automatically calculated; however, the final synaptic strenghts are not coded back into its own genotype. After 50 generations (approx. one day), we detected a number of neural controllers that displayed a smooth navigation consisting of straight trajectories and obstacle avoidance. Since their synaptic values were always randomly initialised, these controllers "learned" the appropriate behaviour in the first 2 or 3 seconds of the agent life. Successive modifications consisted in minor behavioural changes that reflected the geometry of the environment. If the test environment had long corridors, the agent would gain speed, otherwise it would move more carefully with slight changes of directions to sample the presence of obstacles. The analysis of one particular controller showed a number of interesting properties. After the initial seconds, synapses displayed a dynamic stability, that is they continuously and rapidly changed while the overall behaviour remained stable. Those few synapses that learned according to pure Hebb rule were responsible for the basic behaviour of forward motion, whereas presynaptic and postsynaptic learning controlled changes in directions and quick halts in the presence of novel obstacles. Only 18% of the synapses where inhibitory, but their role was crucial in determining behavioural changes. Finally, the controller was very robust to external world changes: one could interact at any moment with the robot, pick it up, cover its sensors with the hands for minutes, but the robot would still quickly restore its navigation ability when put back in the environment.