Artificial Evolution of Plastic Neural Networks: a few Key Concepts

Jean-Baptiste Mouret and Paul Tonelli

Abstract This paper introduces a hierarchy of concepts to classify the goals and the methods of works that mix neuro-evolution and synaptic plasticity. We propose definitions of "behavioral robustness" and oppose it to "reward-based behavioral changes"; we then distinguish the switch between behaviors and the acquisition of new behaviors. Last, we formalize the concept of "synaptic General Learning Abilities" (sGLA) and that of "synaptic Transitive Learning Abilities (sTLA)". For each concept, we review the literature to identify the main experimental setups and the typical studies.

1 Introduction

The abilities of animals to adapt to new environments is one of the most fascinating aspects of Nature and it may be what most clearly separates animals from current machines. Natural adaptive processes are classically divided into three main categories, each of them having been a continuous source of inspiration in artificial intelligence and robotics [8]: evolution, development and learning. While studying each of these processes independently have been widely successful, there is a growing interest in understanding how they benefit from each other.

In particular, a large amount of work has been devoted to understand both the biology of learning (e.g. [19, 29]) and the design of learning algorithms for artificial neural networks (e.g. [9]); concurrently, evolution-inspired algorithms have been successfully employed to automatically design small "nervous systems" for robots [7, 10, 12, 13, 25, 26], sometimes by taking inspiration from development processes [10, 13, 16, 25]. A comparatively few papers proposed to combine the artificial evolution of neural networks with synaptic plasticity to evolve

Jean-Baptiste Mouret and Paul Tonelli
ISIR, Université Pierre et Marie Curie-Paris 6, CNRS UMR 7222, Paris, France
e-mail: {mouret,tonelli}@isir.upmc.fr
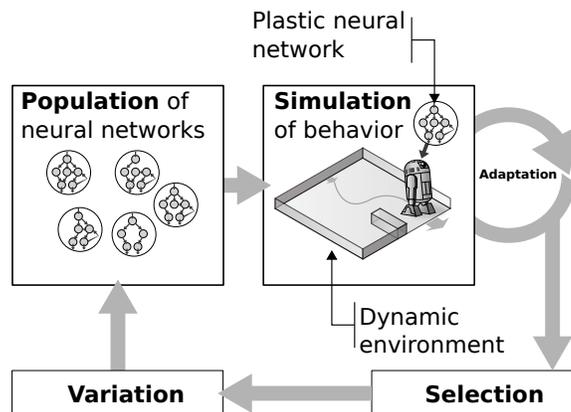
Fig. 1 The artificial evolution of plastic neural networks relies on the classic evolutionary loop used in neuro-evolution. The algorithm starts with a population of genotypes that are thereafter developed into plastic neural networks. The topology of the neural network is sometimes evolved [17, 18, 21–24, 28]. Typical plastic neural networks use a variant of the Hebb's rule to adapt the weight during the "lifetime" of the agent. The fitness of the agent is most of the time evaluated in a dynamic environment that requires the agent to adapt its behavior. The agent is therefore usually directly selected for its adaptive abilities.

artificial agents that can adapt their "artificial nervous system" during their "lifetime" [7, 15–18, 21–23, 28, 30] (Fig.1). However, the analysis of these works shows that they often address different challenges in very different situations, while using the same terminology (e.g. "learning", "robustness" or "generalization").

The goal of the present chapter is to provide a set of definitions to make as clear as possible current and future works that involve the evolution of such plastic artificial neural networks (ANNs) to control agents (simulated or real robots). While some definitions and some distinctions are novel, the main contribution of the present paper is to isolate each concept and to present them in a coherent framework. For each definition, we will provide examples of typical setups and current results. Figure 2 displays the hierarchy of the concepts that will be introduced; it can serve as a guide to the paper.

2 Synaptic Plasticity

In neuroscience, plasticity (or neuroplasticity) is the ability of the brain and nervous systems to change structurally and functionally as a result of their interaction with the environment. Plasticity is typically observed during phases of development and learning. Trappenberg [29] defines two kinds of plasticity: structural plasticity and synaptic (or functional) plasticity.
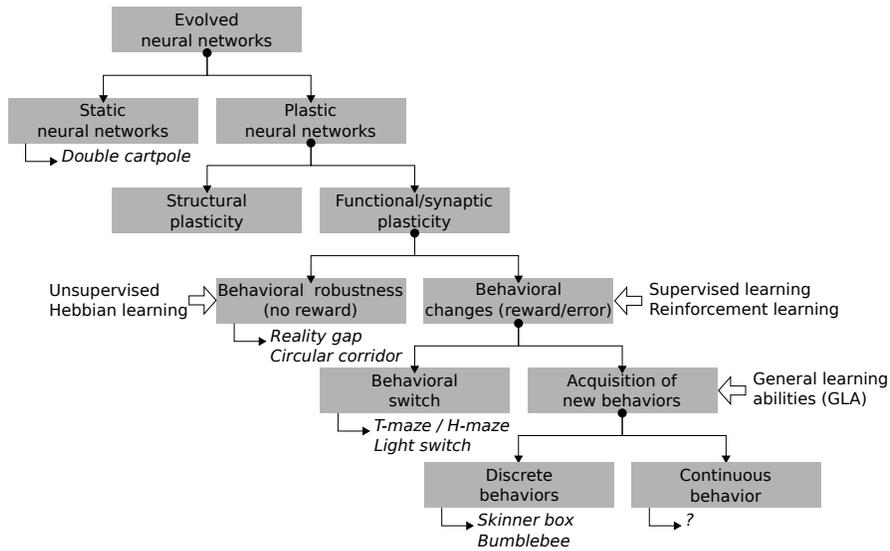
Fig. 2 Hierarchy of concepts described the present paper. See text for a definition of each gray box.

Definition 1 (Structural plasticity). Structural plasticity is the mechanism describing generation of new connections and thereby redefining the topology of the network.

Definition 2 (Synaptic plasticity). Synaptic plasticity is the mechanism of changing strength values of existing connections. It is sometimes termed "functionnal plasticity" [29].

Nolfi et al. [16] investigated *structural plasticity* in a system in which the genotype contained developmental instructions for the construction of a neural network. Genes specified (1) the position of each neuron and (2) instructions that described how axons and branching segments grew. These instructions were executed when a neuron was sufficiently stimulated by its surrounding neurons and by the agent's environment. The authors observed different phenotypes when the same genotype was used in two different environments and concluded that their approach increased the adaptive capabilities of their organisms. Several other authors evolved neural networks while letting them grow axons depending on their location (e.g. [10, 25]) but the enviroment was not taken into account.

Most studies on the evolution of plastic neural networks instead focused on *synaptic plasticity* [2, 6, 14, 30], maybe because of the prominence of learning algorithms that only adapt weights in the Machine Learning literature. Most of the works that do not rely on machine learning algorithms (e.g. the backpropagation algorithm) [3, 15] use variants of the "Hebb's rule" [2, 6, 17, 28, 30], which posits that the simultaneous activation of two neurons strengthens the synapse that link them.

Definition 3 (Hebb's rule). Let us denote by i and j two neurons[1], a_i and a_j their respective activation level, w_{ij} the synaptic weight of the connection from i to j and η a learning rate that describes how fast the change occurs. According to Hebb's rule, w_{ij} should be modified as follows:

$$w_{ij}(t+1) = w_{ij}(t) + \Delta w_{ij} \tag{1}$$

$$\Delta w_{ij} = \eta \cdot a_i \cdot a_j \tag{2}$$

Hebb's rule is often extended to include more complex combinations of pre- and post-synaptic activities [2, 6, 17, 29, 30].

Definition 4 (Extended Hebbian rule).

$$\Delta w_{ij} = f(a_i, a_j, w_{ij}) \tag{3}$$

Many different $f()$ have been investigated; one of the simplest extended Hebbian rule consists in linearly combining pre- and post-synaptic activities [14, 17, 21, 30]:

$$\Delta w_{ij} = A \cdot a_i \cdot a_j + B \cdot a_i + C \cdot a_j + D \tag{4}$$

where A, B, C and D are four real numbers. Most authors select a set of values for A, B, C and D, but several different combinations can also be used in the same neural network. For instance, Urzelai and Floreano proposed to evolve the plasticity rules that govern each synapse of fully connnected neural networks [30].

A synapse can also be strengthened or weakened as a result of the firing of a third, modulatory inter-neuron (e.g. dopaminergic neurons) [1, 14, 22]. To reflect this phenomenon, two kinds of neurons can be distinguished: modulatory neurons and modulated neurons. Inputs of each neuron are divided into modulatory inputs and signal inputs; the sum of the modulatory inputs of j governs the modulation of the all non-modulatory connections to j:

Definition 5 (Modulated Hebbian rule). Let us denote by $I_j^{(m)}$ the set of modulatory inputs of neuron j and by $I_s^{(j)}$ the set of non-modulatory inputs. Each incoming connection of neuron j is modified as follows:

$$m_j = \tanh\left(\sum_{k \in I_j^{(m)}} w_{kj} a_k\right) \tag{5}$$

$$\forall i \in I_s^{(j)}, \Delta w_{ij} = m_j \cdot f(a_i, a_j, w_{ij}) \tag{6}$$

In addition to its biological realism, this weight adaptation rule makes easier to use rewards signals (for instance, plasticity could be enabled only when a reward signal

---

[1] We focus our discussion on classic neurons (as used in classic machine learning) and population-based models of neurons (e.g. leaky integrators) because they are the neuron models that are used by most of the community. Spiking neuron models can make use of other plasticity mechanisms (e.g. STDP) that will not be described here.

is on). It also leads to networks in which only a part of the synapses are changed during the day-to-day life of the agent. These two features make such networks match more closely some of the current actor-critic models of reinforcement learning used in computational neuroscience [19].

Modulated Hebbian plasticity has been used several times when evolving plastic neural networks [11, 14, 18, 21, 22]. In these simulations, experiments in reward-based scenarios where modulatory neurons were enabled achieved better learning in comparison to those where modulatory neurons were disabled [21].

3 Robustness and Reward-based Scenarios

A major goal when evolving neuro-controllers is to evolve neural networks that keep performing the same optimal (or pseudo-optimal) behavior when their morphology or their environment change. For instance, a robot can be damaged, gears can wear out over time or the light conditions can change: in all these situations, it is desirable for an evolved controller to compensate these changes by adapting itself; we will call this ability *behavioral robustness*.

Definition 6 (Behavioral robustness). An agent displays behavioral robustness when it keeps the same qualitative behavior, notwithstanding environmental and morphological changes. Behavioral robustness does not usually involve a reward/ punishment system.

In a typical work that combines synaptic plasticity, evolution and behavioral robustness, Urzelai and Floreano [30] evolved neuro-controllers with plastic synapses to solve a light-switching task *in which there was no reward*; they then investigated whether these controllers were able to cope with four types of environmental changes: new sensory appearances, transfer from simulations to physical robots, transfer across different robotic platforms and re-arrangement of environmental layout. The plastic ANNs were able to overcome these four kinds of change, contrary to a classic ANN with fixed weights.

However, as highlighted by Urzelai and Floreano, "these behaviors were not learned in the classic meaning of the term because they were not necessarily retained forever". Actually, synaptic weights were continuously changing such that the robot performed several sub-behaviors in sequence; the evolutionary algorithm therefore opportunistically used plasticity to enhance the dynamic power of the ANN. These high-frequency changes of synaptic weights appear different from what we observe in natural system (in particular in the basal ganglia), in which synaptic weights tend to hold the same value for a long period, once stabilized [5, 27].

Besides robustness, an even more desirable property for an evolved agent is the ability to change its behavior according to external stimuli and, in particular, according to rewards and punishments. For instance, one can imagine a robot in a T-maze that must go to the end of the maze where a reward has been put [17, 18, 21]. The robot should first randomly try different trajectory. Then, once the reward would

have been found a few times, the robot should go directly to the reward. Nonetheless, if the reward is moved somewhere else, the robot should change its behavior to match the new position of the reward. Once the robot will have found the optimal behavior (the behavior that maximizes the reward), the synaptic weights of its controller should not change anymore. This ability to adapt in a reward-based scenario can be more formally defined as follows:

Definition 7 (Behavioral change). A plastic agent is capable of behavioral changes in a reward-based scenario if and only if:

- a change of reward makes it adopt a qualitatively new behavior;
- the synaptic weights do not significantly change once an optimal behavior has been reached.

Notable setups in which authors evolved plastic neuro-controllers for behavioral changes are the T-maze [17, 18, 21], the bumblebee foraging task [14], the "dangerous foraging task" [24] and the Skinner box [28].

4 Learning Abilities in Discrete Environment

The main challenge when evolving plastic agents for behavioral change is to make them able to learn *new behaviors* in *unknown situations* and, in particular, in situations that have never been encountered during the evolutionary process. Put differently, selecting agents for their abilitities to switch between alternatives is not sufficient; the evolved agent must also be placed in completely new situations to assess its ability to find an optimal behavior in a situation for which it has never been selected.
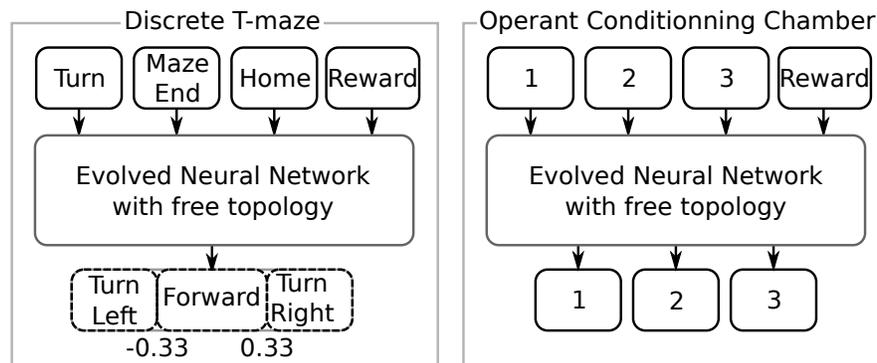
Fig. 3 Learning the best-rewarding behavior in a discrete T-maze is equivalent to a Skinner box (Operant Conditioning Chamber, left): in both cases, the challenge is to associate the right stimulus to the right action.

We previously introduced a theoretical framework to characterize and analyze the learning abilities of evolved plastic neural networks [28]; we will rely on this framework in the remainder of this paper. For the sake of simplicity, we focus on a discrete world, with discrete stimuli and discrete actions. The canonical setup, inspired by experiments in operant conditioning, is the Skinner Box [20]: an agent is placed in a cage with n stimuli (lights), m actions (levers), positive rewards (food) and punishments (electric shocks). The goal of the agent is to learn the right associations between each stimulus and each action. This task encompasses most discrete reward-based scenarios (Fig. 3). For instance, the discrete T-maze experiment [17, 18, 21–23] can be described as a special case of a Skinner box.

More formally, an evolved neural network $N(I, \lambda)$ must adapt several synaptic weights $\lambda \in \mathbb{R}^z$ such that each input pattern $I \in [0, 1]^n$ is associated to the best rewarded output vector $K \in [0, 1]^m$. The adaptation is performed by a learning function such that $\lambda = g(\lambda_r, I, R_{I,K})$, where λ_r is a random vector in $\mathbb{R}^z$ and $R_{I,K}$ the reward function. These notations lead to the following definitions:

Definition 8 (Association set). An association set $A = \big\{ (I_1, K_1), \cdots, (I_n, K_n) \big\}$ is a list of associations that covers all the possible input patterns. The set of all association sets is denoted $\mathbb{A}$.

Definition 9 (Fitness associations set). The fitness associations set $\mathbb{F}_{\mathbb{A}} = \{A_1 \cdots A_k\}$ is the set of the association sets that are used during the fitness evaluation.

For a given topology, some association sets may not be learnable by only changing synaptic weights. This case occurs in particular when the topology of neural networks are evolved: if there is no selective pressure to maintain a connection, it can easily disappear; but this connection may be required to learn a similar but different association set. Some association sets may also be not learnable because they require specific topologies. For instance, the XOR function requires a hidden layer of neurons to be computed.

Definition 10 (Learnable set). Given a suitable reward function $R_{I,K}$, an association set $A \in \mathbb{A}$ is said to be learnable by the neural network N, if and only if $\forall \lambda_r \in \mathbb{R}^z$ and $\forall (I, K) \in A$, $\exists \lambda = g(\lambda_r, I, R_{I,K})$ such that $N(I, \lambda) = K$. The set of all learnable sets for N is denoted $\mathbb{L}_N$.

Definition 11 (sGLA). A plastic ANN is said to possess synaptic General Learning Abilities (sGLA) if and only if $\forall A \in \mathbb{A}, A \in \mathbb{L}_N$.

Although it does not use Hebbian learning, the multi-layer perceptron with the back-propagation algorithm is an example of a neural network with synaptic General Learning Abilities. At the opposite, a neural network in which each input is connected to only one output can learn only one association set.

To evolve a plastic ANN with sGLA, the simplest method is to check the learnability of each association set during the fitness evaluation, that is the fitness associations set is equal to the set of all the association sets; this approach has often been followed by the authors who evolved agents to solve the T-maze task [17, 18, 21–23].

We propose to call such approaches the *evolution of behavioral switches* to distinguish it from the evolution of more general learning abilities.

Definition 12 (Evolution of behavioral switches). A study deals with the evolution of behavioral switches when the fitness associations set $\mathbb{F}_\mathbb{A}$ is equals to the set of all association sets that are possible for the studied task (called $\mathbb{A}$): $\mathbb{F}_\mathbb{A} = \mathbb{A}$

However, a plastic ANN that can cope with unknown situations must have sGLA while only a subset of the possible association sets (i.e. a subset of problems from the same problem class) has been used during the evolutionary process.

Definition 13 (Evolution of sGLA for unknown situations). A study deals with the evolution of synaptic General Learning Abilities for unknown situations when the goal of the authors is (1) to use a fitness associations set that is smaller than the set of all association sets ($card(\mathbb{F}_\mathbb{A}) < card(\mathbb{A})$) and (2) to show that all associations sets are learnable with evolved solutions ($\forall A \in \mathbb{A}, A \in \mathbb{L}_N$):
$card(\mathbb{F}_\mathbb{A}) < card(\mathbb{A})$ and $\forall A \in \mathbb{A}, A \in \mathbb{L}_N$.

At first sight, Nature relies on the long lifetime of animals (compared to the "lifetime" of artificial agents) and on the large size of the populations to obtain a stochastic evaluation of virtually every possible scenarios. This probably explains why most authors tried to obtain agents with sGLA by using a large, often randomized subset of the association sets in the fitness association set. In supervised learning, Chalmers [3] assessed how well an evolved plastic ANN can cope with situations never encountered during the evolution. In his experiments, he evolved the learning rule for a small single-layer ANN (5 inputs, 1 output) and his analysis showed that at least 10 sets of input/output patterns (among 30 possible sets) were required to evolve an algorithm that correctly learns on 10 unknown sets. In reinforcement learning, Niv et al. [14] evolved plastic ANNs to solve a bumblebee-inspired foraging task in which simulated bees must select flowers by recognizing their color. To promote general learning abilities, they randomly assigned rewards to colors at each generation and they showed that the resulting ANNs successfully learned unknown color/reward associations. In the "dangerous foraging task", Stanley et al. [24] similarly randomized the parameters of the fitness function to avoid overspecialized behaviors.

However, the encoding and the development process may also play a key role in allowing the adaptation to situations which have never been encountered before [28]. Intuitively, a very regular network may repeat the same adaptation structure many times whereas it was only required once by the fitness; it could therefore "propagate" the adaptation structure. Since most developmental encoding are designed to generate very regular structures [4, 13, 28], using such encodings could substantially reduce the number of evaluations required to obtain general learning abilities. In the ideal case, we should be able to show that the developmental process implies that if a few association sets have been successfully learned, then all the other sets have a high probability of being learnable. Such networks will be said to possess "synaptic Transitive Learning Abilities".

Definition 14 (sTLA). Let us denote by $\mathbb{T}_N$ a subset of the learnable association set $\mathbb{A}$. A plastic ANN is said to possess synaptic Transitive Learning Abilities (sTLA) if and only if $\exists\, \mathbb{T}_N \subset \mathbb{A}$ such that the following implication is true:

$$\mathbb{T}_N \subset \mathbb{L}_N \Rightarrow \mathbb{L}_N = \mathbb{A}$$

$p = card(\mathbb{T}_N)$ will be called the "sTLA-level".

Definition 15 (Optimal-sTLA). A plastic ANN is said to possess Optimal synaptic Transitive Learning Abilities (optimal-sTLA) if and only if it possesses sTLA and $card(\mathbb{T}_N) = 1$.

The sTLA-level of certain families of topologies (i.e. topologies generated by a specific genetic encoding) can possibly be computed theoretically. It can also be easily evaluated by a succession of evolutionary experiments: (1) select p association sets; (2) evolve ANNs that successfully learns the p association sets; (3) check the sGLA of optimal ANNs; (4) if optimal ANNs do not possess sGLA, then increase p and start again.

Using this method, Tonelli and Mouret [28] showed that a very regular map-based encoding proposed in [13] have a TLA-level or 1 or 2. Preliminary experiments suggest that other generative encodings such as HyperNEAT [4, 25] could also possess a low TLA-level. Overall, the concept of sTLA highlights how evolution, learning and development are interwoven.

All these definitions are well suited for discrete worlds and discrete behaviors. Future work should try propose equivalent definitions for continuous worlds.

5 Concluding Remarks

With the rise of computing power, it is now easy to simulate artificial agents for enough time for them to learn and to evolve; this allows the study of well-defined scientific questions with modern experimental and statistical techniques. Nevertheless, future work in this direction will have to precisely define what they work on: do they aim at behavioral robustness or at behavioral change? how do they evaluate the general learning abilities of the evolved agents? do the evolved neural network manage to learn in unknown situations? what is the role of the encoding in the final result? The definitions proposed in the present paper will hopefully help to design a methodology to answer such questions.

The present work also highlights open questions and avenues for future researches:

- Should future work focus more on structural plasticity? This approach to plasticity may be more complex but it may also allow agents to learn new skills without forgetting the old ones (because the previous structure is not deleted).
- How to evaluate learning abilities in continuous world and with continuous behaviors?

- What are the links between encodings, plasticity and learnability? [28] provides first answers but only for simple and discrete scenarios.

6 Acknowledgement

References

1. Bailey, C.H., Giustetto, M., Huang, Y.Y., Hawkins, R.D., Kandel, E.R.: Is heterosynaptic modulation essential for stabilizing Hebbian plasticity and memory? Nature reviews. Neuroscience **1**(1), 11–20 (2000)
2. Blynel, J., Floreano, D.: Levels of dynamics and adaptive behavior in evolutionary neural controllers. In: From Animals to Animat 7: Proceedings of the 7th conference on Simulation of Adaptive Behavior (SAB'02), pp. 272–281 (2002)
3. Chalmers, D.: The evolution of learning: An experiment in genetic connectionism. Connectionist Models Summer School, (1990)
4. Clune, J., Stanley, K.O., Pennock, R.T., Ofria, C.: On the Performance of Indirect Encoding Across the Continuum of Regularity. IEEE Transaction On Evolutionary Computation p. TODO (2011)
5. Daw, N.D., Niv, Y., Dayan, P.: Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. Nature neuroscience **8**(12), 1704–11 (2005)
6. Floreano, D.: Evolution of plastic neurocontrollers for situated agents. In: From Animals to Animat 4: Proceedings of the 4th conference on Simulation of Adaptive Behavior (SAB'96) (1996)
7. Floreano, D., Dürr, P., Mattiussi, C.: Neuroevolution: from architectures to learning. Evolutionary Intelligence **1**(1), 47–62 (2008)
8. Floreano, D., Mattiussi, C.: Bio-inspired artificial intelligence: theories, methods, and technologies. The MIT Press (2008)
9. Haykin, S.: Neural networks: a comprehensive foundation. Prentice hall (1999)
10. Kodjabachian, J., Meyer, J.A.: Evolution and development of neural controllers for locomotion, gradient-following, and obstacle-avoidance in artificial insects. IEEE transactions on neural networks **9**(5), 796–812 (1998)
11. Kondo, T.: Evolutionary design and behavior analysis of neuromodulatory neural networks for mobile robots control. Applied Soft Computing **7**(1), 189–202 (2007)
12. Mouret, J.B., Doncieux, S.: Encouraging Behavioral Diversity in Evolutionary Robotics: an Empirical Study. Evolutionary Computation **20**(1), 91–133 (2012)
13. Mouret, J.B., Doncieux, S., Girard, B.: Importing the computational neuroscience toolbox into neuro-evolution—application to basal ganglia. GECCO'10: Proceedings of the 12th annual conference on Genetic and Evolutionary Computation pp. 587–594 (2010)
14. Niv, Y., Joel, D., Meilijson, I., Ruppin, E.: Evolution of Reinforcement Learning in Uncertain Environments: A Simple Explanation for Complex Foraging Behaviors. Adaptive Behavior **10**(1), 5–24 (2002)
15. Nolfi, S.: How Learning and Evolution Interact: The Case of a Learning Task which Differs from the Evolutionary Task. Adaptive Behavior p. TODO (1999)
16. Nolfi, S., Miglino, O., Parisi, D.: Phenotypic plasticity in evolving neural networks. In: From Perception to Action Conference, Proceedings, pp. 146–157. IEEE (1994)

17. Risi, S., Stanley, K.O.: Indirectly Encoding Neural Plasticity as a Pattern of Local Rules. From Animals to Animat 11: Proceedings of the 11th conference on Simulation of Adaptive Behavior (SAB'10) pp. 533–543 (2010)
18. Risi, S., Vanderbleek, S.D., Hughes, C.E., Stanley, K.O.: How novelty search escapes the deceptive trap of learning to learn. In: GECCO'09: Proceedings of the 11th annual conference on Genetic and Evolutionary Computation (2009)
19. Schultz, W., Dayan, P., Montague, P.R.: A Neural Substrate of Prediction and Reward. Science **275**(5306), 1593–1599 (1997)
20. Skinner, B.F.: Operant behavior. American Psychologist **18**(8), 503 (1963)
21. Soltoggio, A., Bullinaria, J., Mattiussi, C., Dürr, P., Floreano, D.: Evolutionary advantages of neuromodulated plasticity in dynamic, reward-based scenarios. Artificial Life **11**, 569 (2008)
22. Soltoggio, A., Dürr, P., Mattiussi, C., Floreano, D.: Evolving neuromodulatory topologies for reinforcement learning-like problems. In: Proceedings of IEEE Congress on Evolutionary Computation (CEC) (2007)
23. Soltoggio, A., Jones, B.: Novelty of behaviour as a basis for the neuro-evolution of operant reward learning. In: GECCO'09: Proceedings of the 11th annual conference on Genetic and Evolutionary Computation (2009)
24. Stanley, K.O., Bryant, B.D., Miikkulainen, R.: Evolving adaptive neural networks with and without adaptive synapses. In: Proceedings of IEEE Congress on Evolutionary Computation (CEC), pp. 2257–2564 (2003)
25. Stanley, K.O., D'Ambrosio, D., Gauci, J.: A Hypercube-Based Indirect Encoding for Evolving Large-Scale Neural Networks. Artificial Life **15**(2), 185–212 (2009)
26. Stanley, K.O., Miikkulainen, R.: Evolving neural networks through augmenting topologies. Evolutionary Computation **10**(2), 99–127 (2002)
27. Sutton, R., Barto, A.G.: Reinforcement learning: An introduction. The MIT press (1998)
28. Tonelli, P., Mouret, J.B.: On the Relationships between Synaptic Plasticity and Generative Systems. In: GECCO'11: Proceedings of the 13th annual conference on Genetic and Evolutionary Computation, pp. 259–266 (2011)
29. Trappenberg, T.P.: Fundamentals of Computational Neuroscience (2. ed.). Oxford University Press (2009)
30. Urzelai, J., Floreano, D.: Evolution of adaptive synapses: Robots with fast adaptive behavior in new environments. Evolutionary Computation **9**(4), 495–524 (2001)