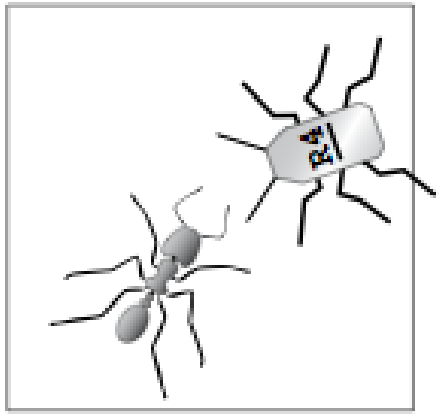
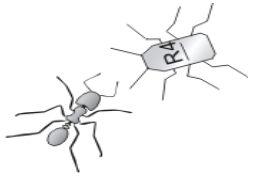
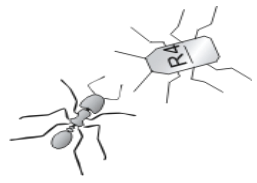
Evolution and Learning





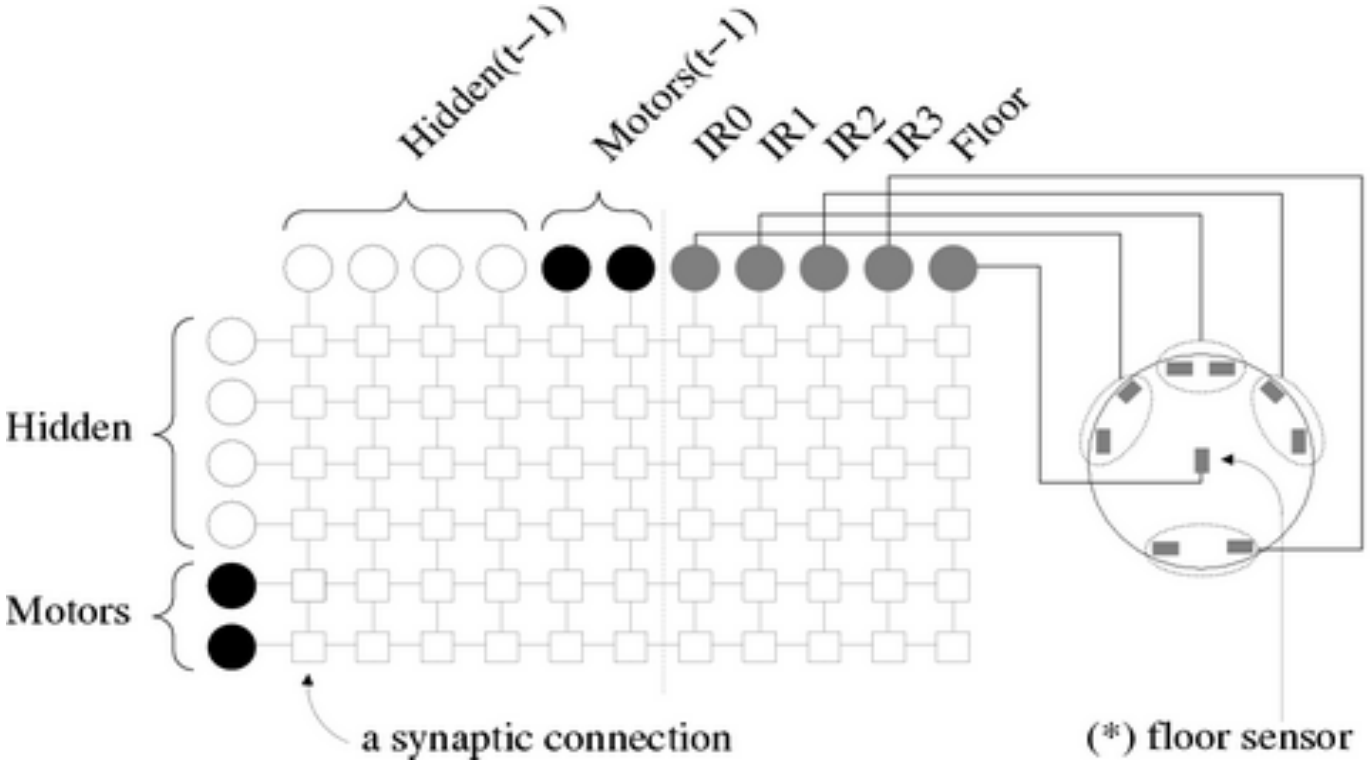
What you will learn in this class

- Advantages and costs of learning in evolution
- How learning can help and guide evolution
- How evolution can help learning
- Darwinian vs Lamarckian evolution
- The Baldwin effect
- Evolution of learning rules
- Evolution of reward-based learning with neuro-modulation



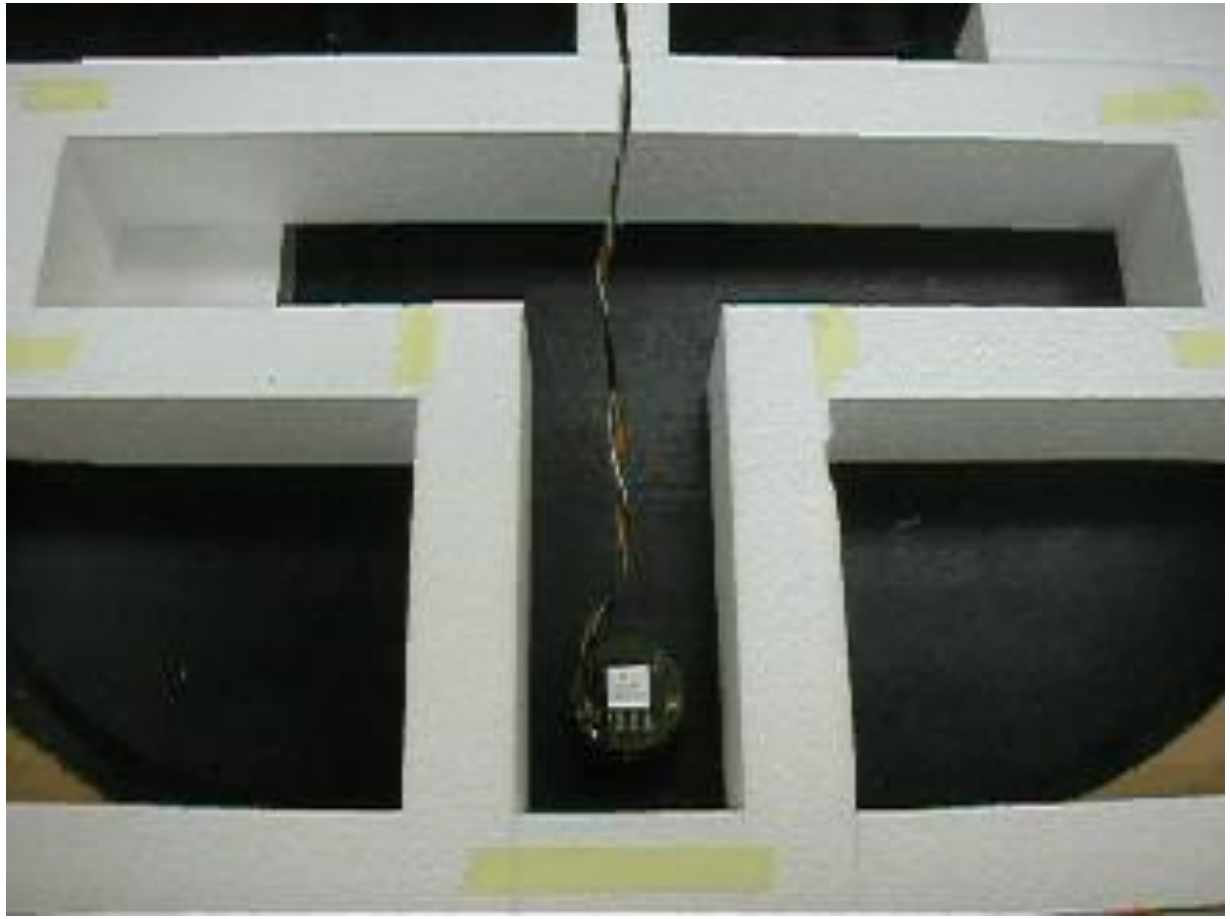
Apparent learning (no change of connection weights)

The T-maze (Tolman, 1925)

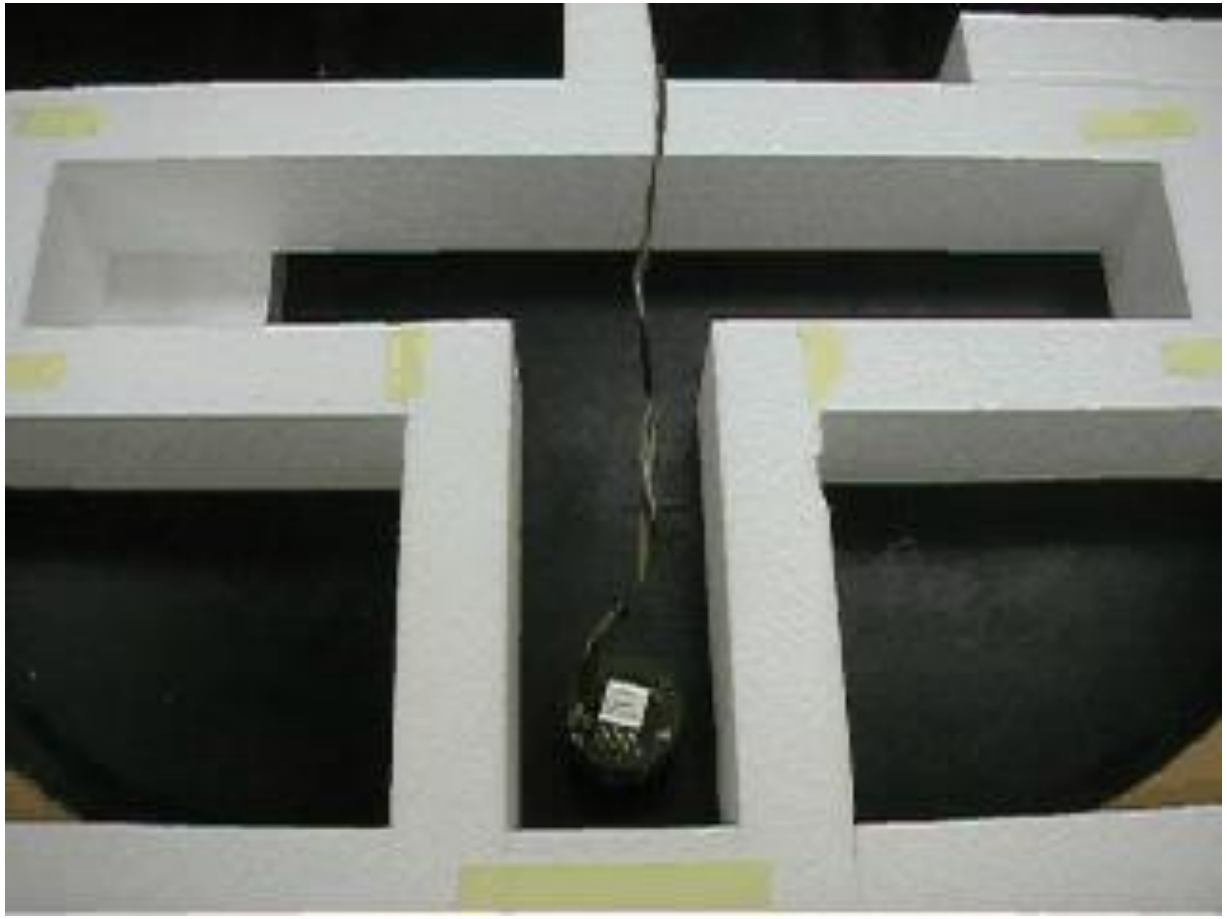


Blynel and Floreano, 2003

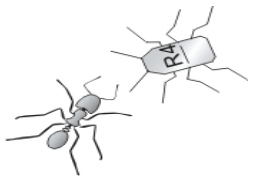
One-shot memorization without weight change



Trial 1



Trial 2

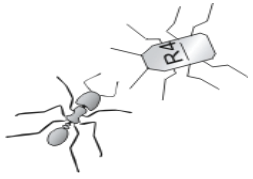


Evolutionary advantages and costs of learning

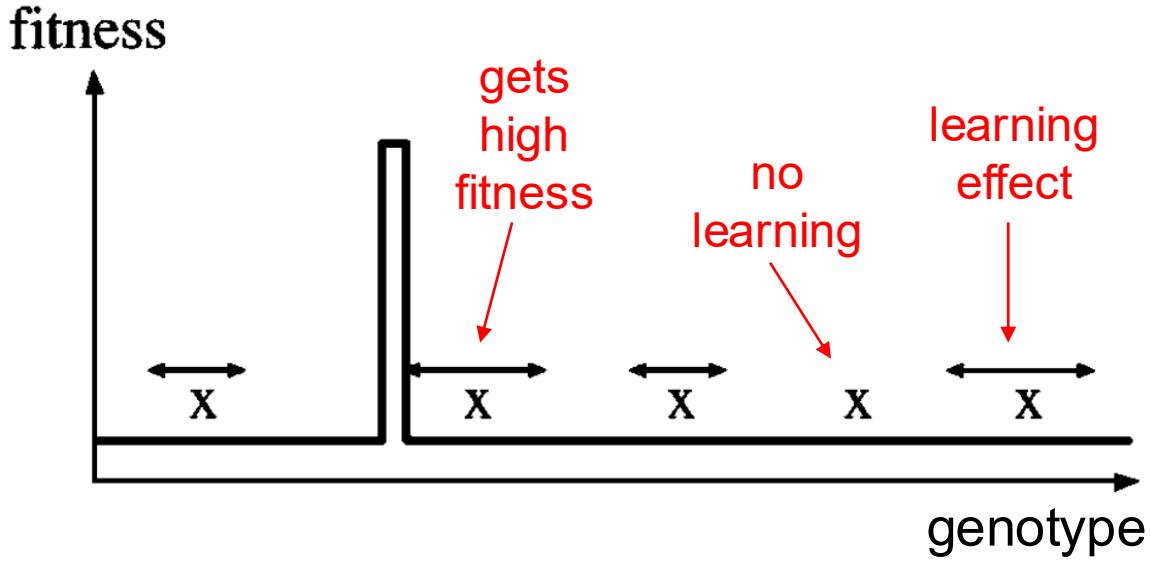
Evolution and learning are both adaptive mechanisms, but have important differences:

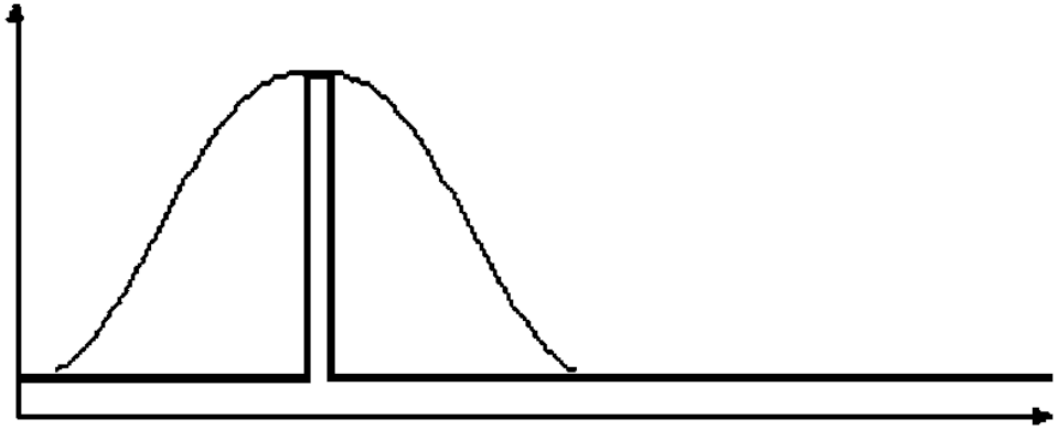
- They take place at different time scales
- They use different processes
- Evolution operates on the genotype, learning operates on the phenotype

ADVANTAGES of learning	COSTS of learning
It can capture environmental change that occurs faster than generation time	It implies a delay in the ability of improving fitness
It can help and guide evolution	It can learn things that are wrong or delay fitness improvement
It can enable shorter genotypes	It requires tutoring, energy, may imply physical damage



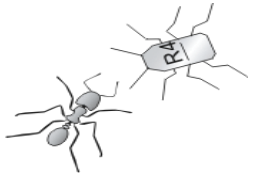
How learning can help and guide evolution *(Hinton and Nowlan, 1987)*





Learning process explores the surroundings in the fitness landscape (individuals who learn may obtain high fitness even if their genes are in low fitness region)

As a consequence, the fitness landscape becomes smoother and displays a "gradient" towards peaks of high fitness, resulting in faster and better evolution



How evolution can help learning

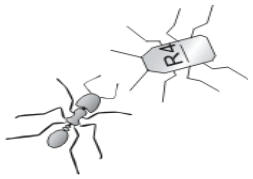
Find initial network weights for better and faster learning

Find good set of learning hyperparameters (initialization range, learning rate, momentum, etc.)

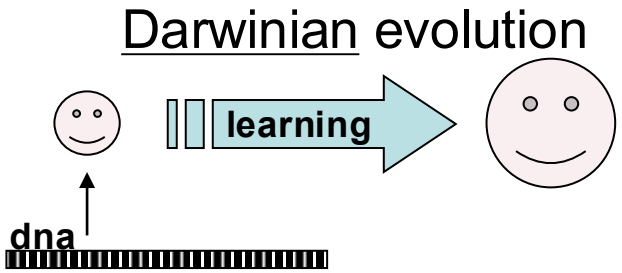
Find suitable learning algorithms

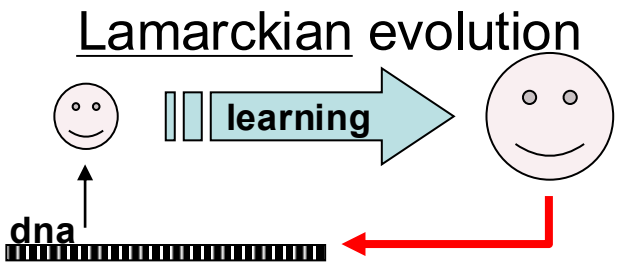
Find network morphology for better and faster learning

Simplify learning problem by co-evolving suitable sensors and bodies



Darwinian vs. Lamarckian evolution



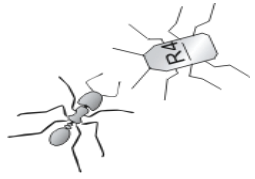


Phenotypic changes cannot be transmitted to the DNA.
Learned abilities cannot be inherited by offsping

Phenotypic changes can be transmitted to the DNA.
Learned abilities can be inherited by offspring.
(No biological evidence for Lamarckian evolution)

In static environments, Lamarckian evolution can produce better and faster results [Lund, 1999].

In dynamic environments, Lamarckian evolution can get stuck in local minima [Sasaki & Tokoro, 1997, 1999].



The Baldwin effect

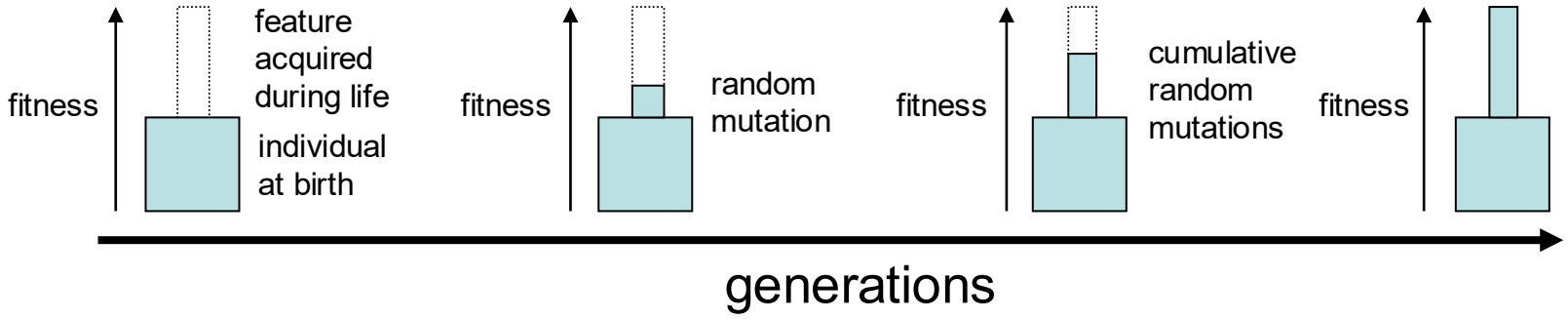
The Baldwin effect [Baldwin, 1896; Morgan, 1896; Waddington, 1942] describes a phenomenon whereby learned features can *indirectly* transfer to the DNA. It has been reported also in evolution of artificial systems [Mayley, 1997; Ackley and Littman, 1991]. Here is how it works:

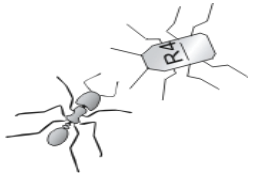
1- Learning is good for survival and thus is selected and maintained by evolution

2- But learning has evolutionary costs

3- Therefore, individuals with mutations that are primitive sketch of abilities that would normally be learned, have a selective advantage with respect to those that must learn them.

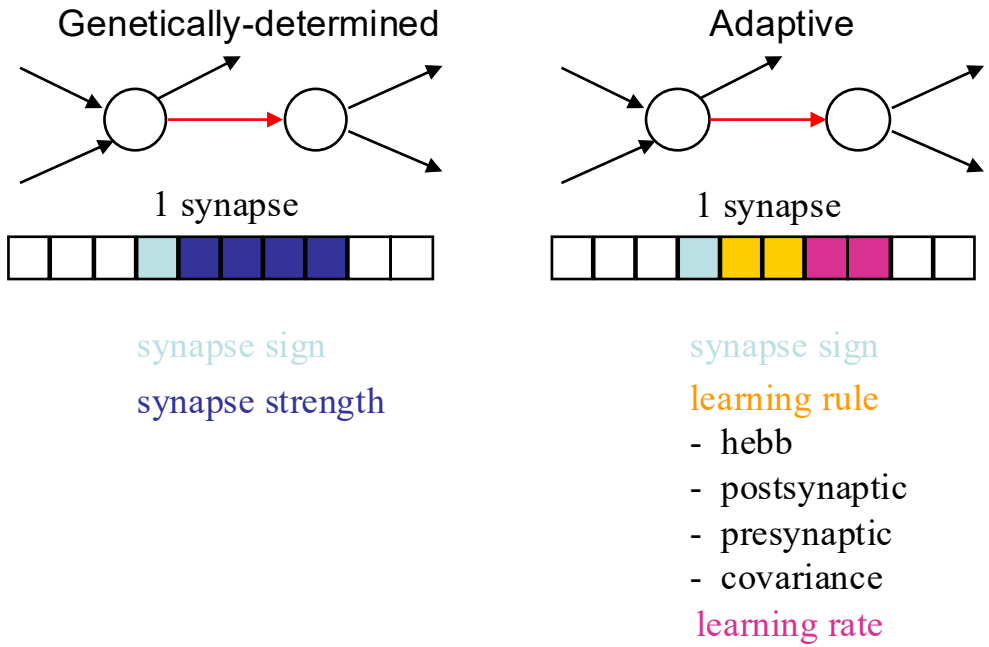
4- Gradually, individuals are born with fully-fledged abilities that had to be learned in early generations

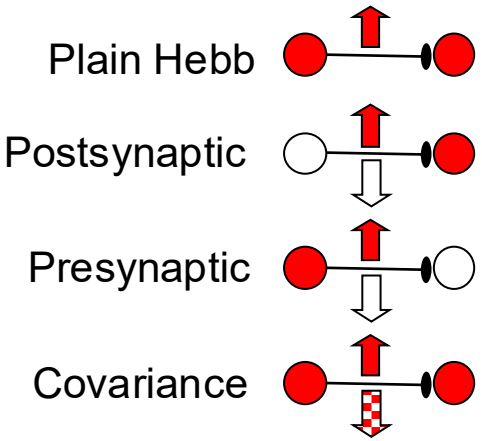




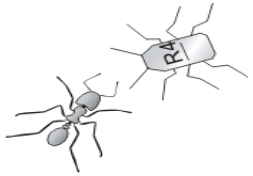
Evolution of Learning Algorithms

Genotype encodes variations of Hebbian learning rules for each connection or each neuron
Connection weights of newborn individuals are always initialized to random values (no Baldwin effect)
Neural network learns during life time using learning rules described in its genotype

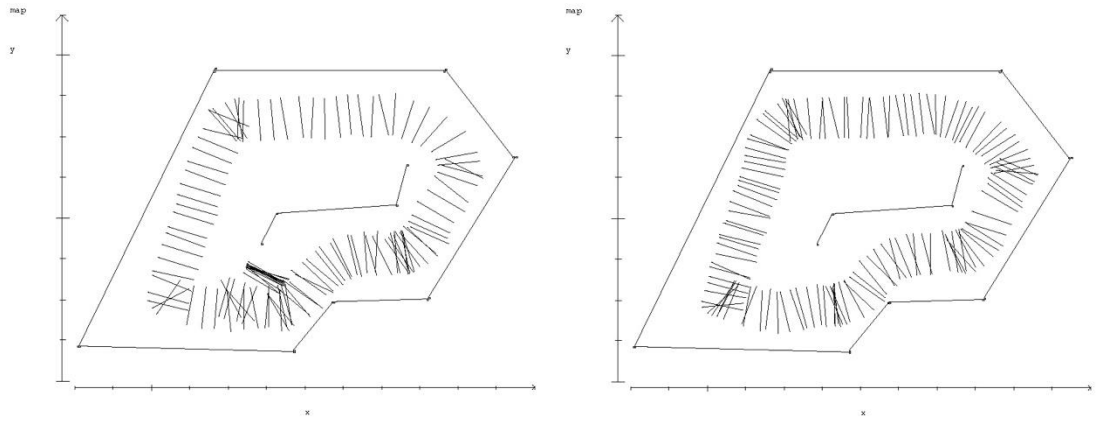




- A neural network can use different learning rules in different parts
- There is no need of teacher or reinforcement learning, no gradient descent and local minima
- Individuals are selected not only to solve a specific problem, but also for their ability to learn in real time

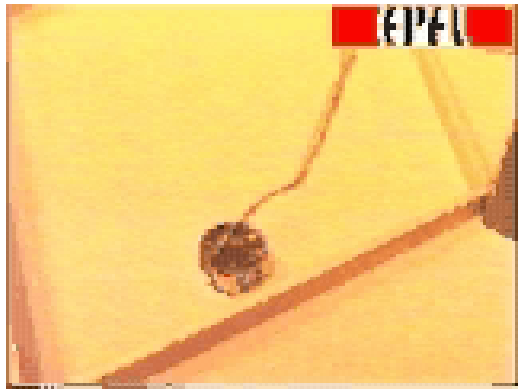


Online Adaptation

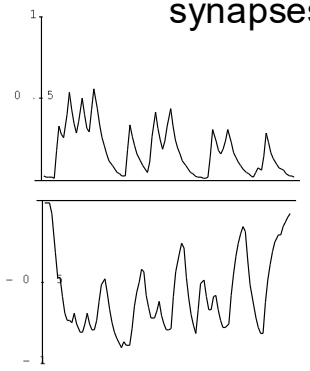


In addition, they perform well in different environments by developing suitable strategies. Contrary to conventional models, several synapses continue to change, but the overal pattern of change is dynamically stable.

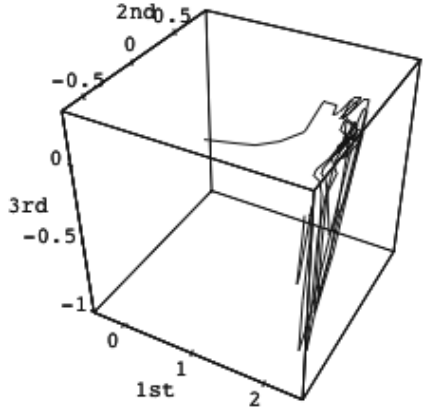
Test in new environment

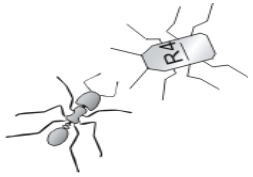


Continuously changing synapses



Dynamic stability

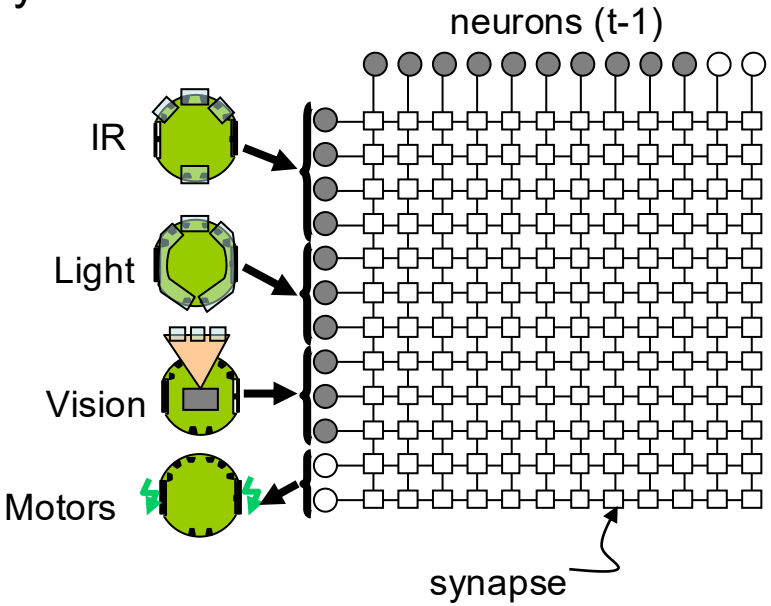




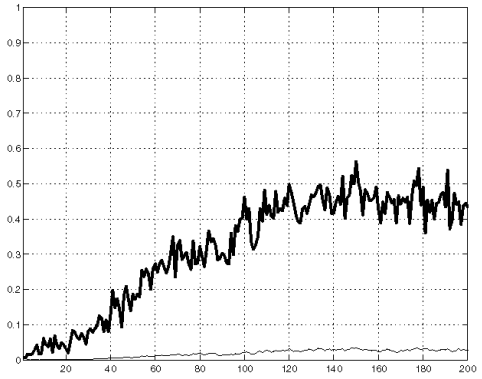
A Sequential Task

A Khepera robot is evolved to switch on a light and go under the light, but this sequence of actions is not directly rewarded by the fitness function.

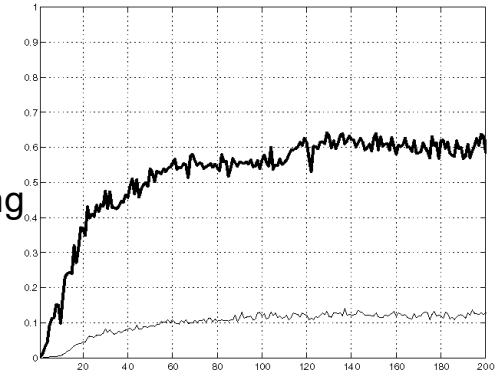
Fitness= time_gray_light / total_time



evolution of weights

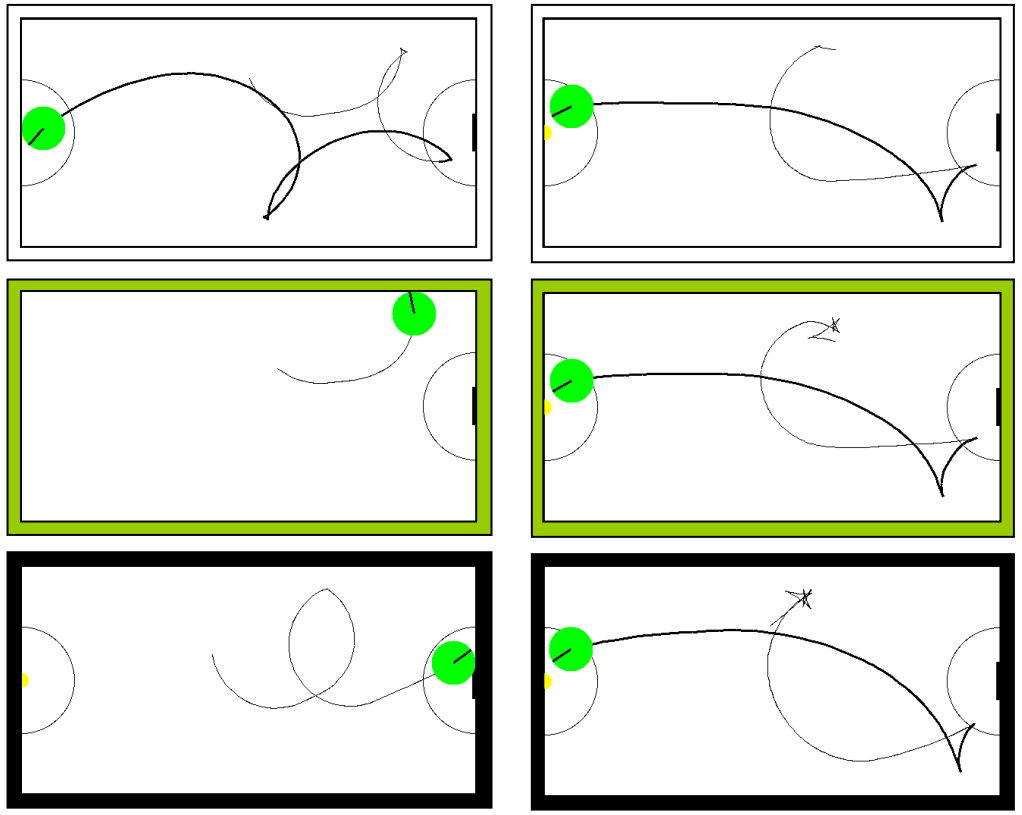


evolution of learning rules



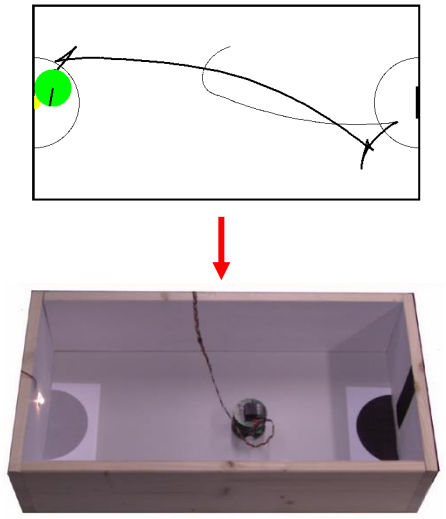
Robustness to Color Change

Evolved adaptive individuals can cope with new colours of the walls whereas genetically-determined individuals fail.



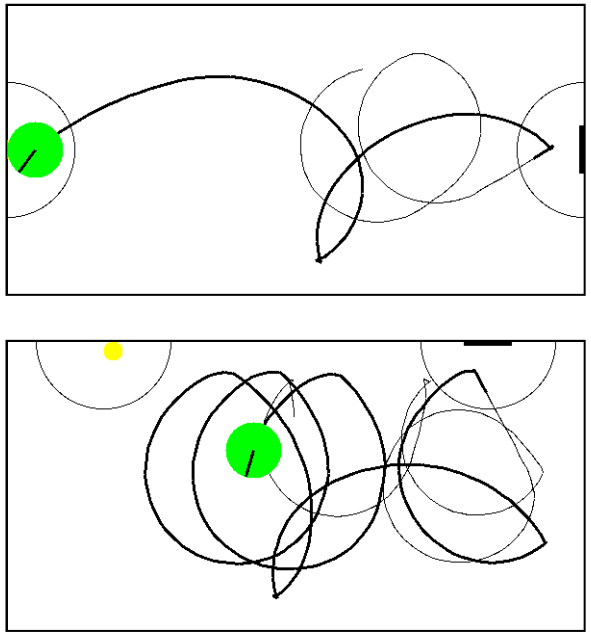
Genetically-determined Adaptive

Similarly, evolved adaptive individuals transfer smoothly from simulated to real world.

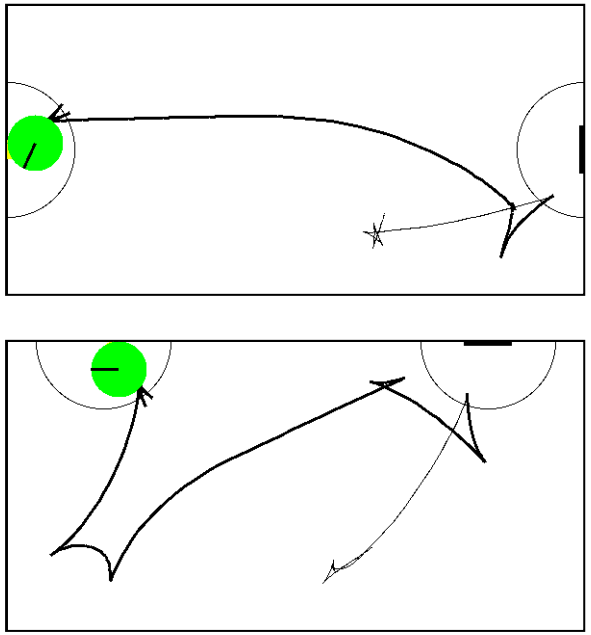


Robustness to Layout Modification

Evolved adaptive individuals can cope with new positions of the two landmarks whereas genetically-determined inviduals cannot.

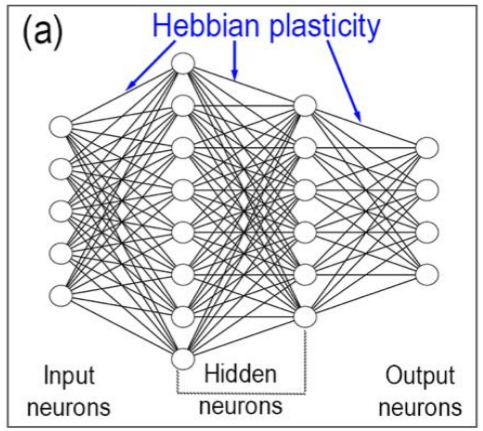


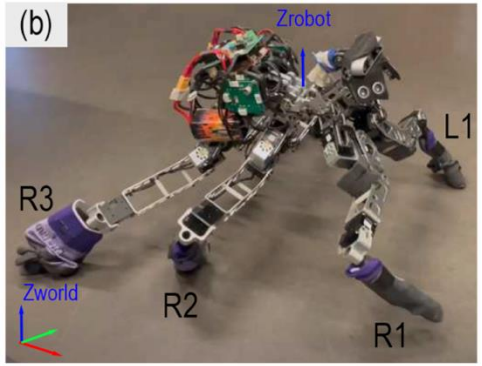
Genetically-determined

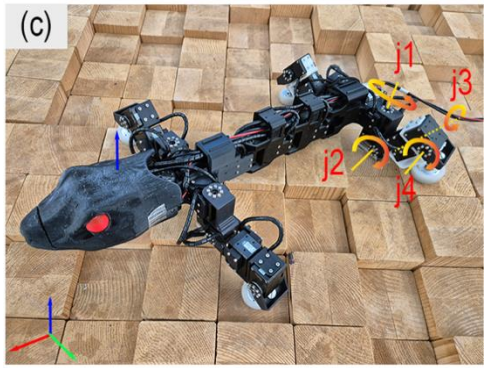


Adaptive

Polynomial Representation of Hebbian Rules



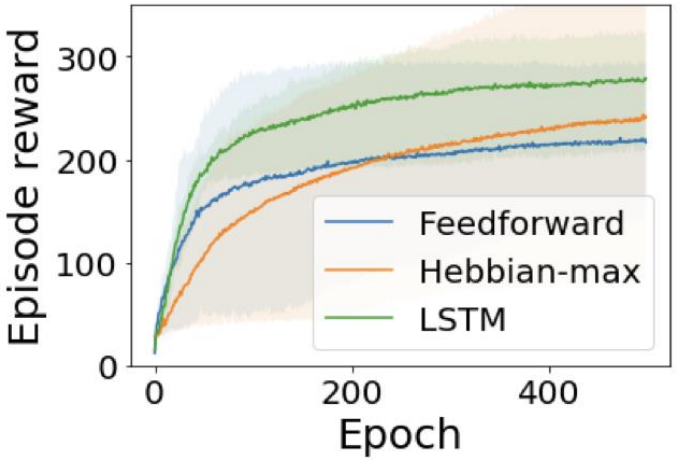




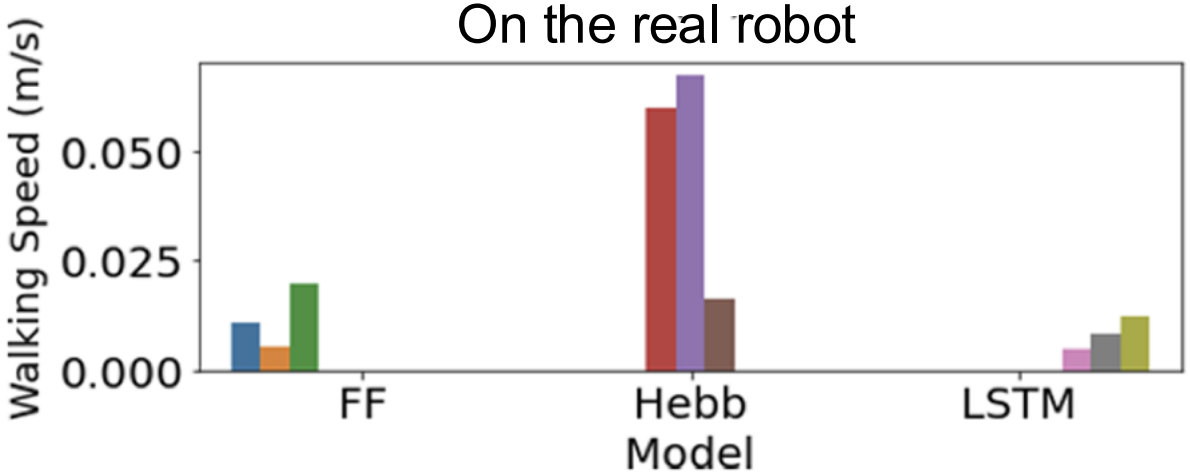
$$\Delta w_{ij}^{k} = \eta \cdot (Ao_i o_j + Bo_i + Co_j + D)$$

These parameters are genetically encoded and evolved for every weight

Evolution in simulation



On the real robot



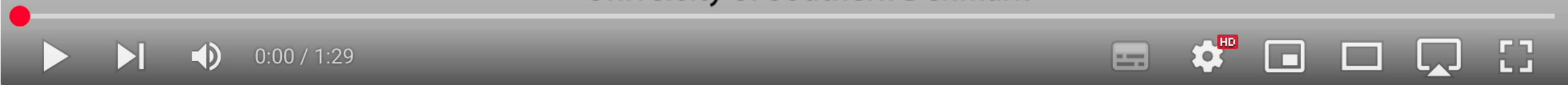
Leung et al (2025) *Bio-Inspired Plastic Neural Networks for Zero-Shot Out-of-Distribution Generalization in Complex Animal-Inspired Robots*, https://arxiv.org/pdf/2503.12406

https://haomachai.com/2025/03/02/bio-inspired-plastic-neural-networks-for-zero-shot-out-of-distribution-generalization-in-complex-animal-inspired-robots/

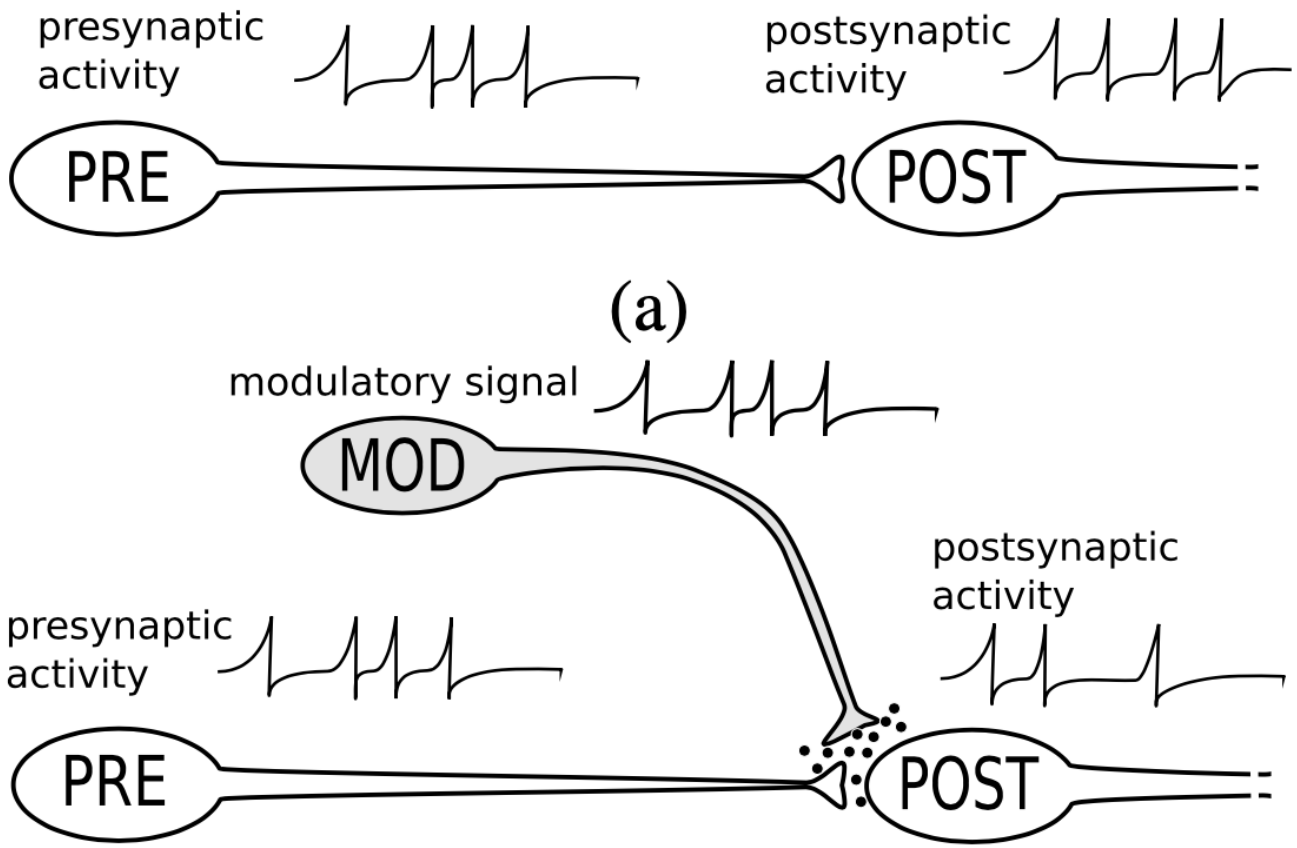
Bio-Inspired Plastic Neural Networks for Zero-Shot Out-of-Distribution Generalization in Complex Animal-Inspired Robots

Binggwong Leung[1], Worasuchad Haomachai[1], Joachim Winter Pedersen[2]
Sebastian Risi[2,*], and Poramate Manoonpong[1,3,*]

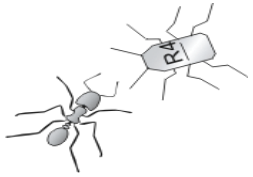
[1]Vidyasirimedhi Institute of Science and Technology
[2]IT University of Copenhagen
[3]University of Southern Denmark



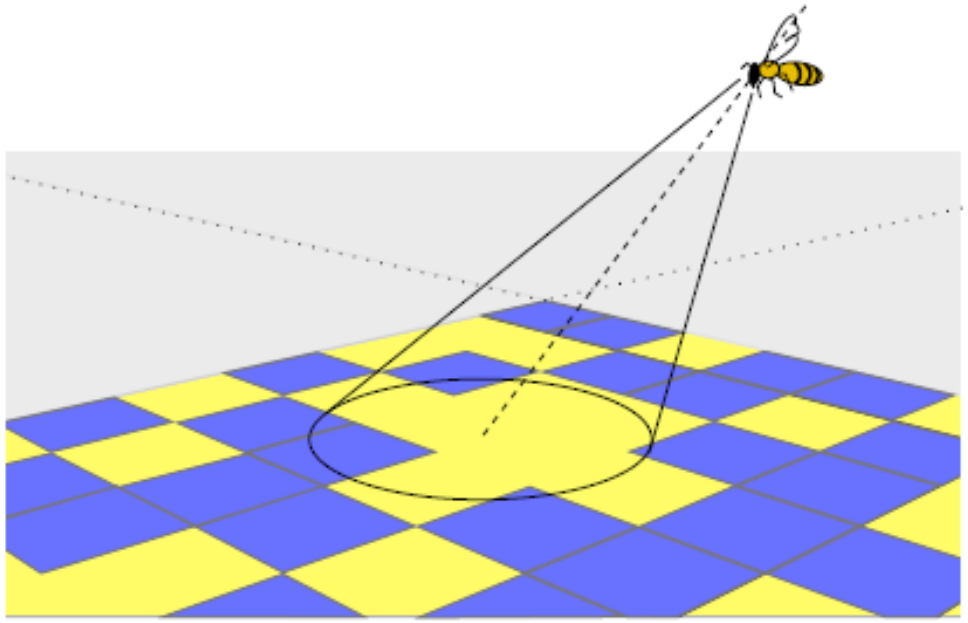
Neuromodulation of synaptic plasticity



Bailey, M. Giustetto, Y.-Y. Huang, R. D. Hawkins, and E. R. Kandel. (2000) Is heterosynaptic modulation essential for stabilizing Hebbian plasticity and memory? *Nature Reviews Neuroscience*, 1(1):11–20



Reward-based behavioural choice



Scenario	Nectar of the high rewarding flower	Nectar of the low rewarding flower
1	$0.8\mu l$	$0.3\mu l$
2	$0.7\mu l$	$1.0\mu l$ with P=0.2 $0.0\mu l$ with P=0.8
3	$1.6\mu l$ with P=0.75 $0.0\mu l$ with P=0.25	$0.8\mu l$ with P=0.75 $0.0\mu l$ with P=0.25
4	$0.8\mu l$ with P=0.75 $0.0\mu l$ with P=0.25	$0.8\mu l$ with P=0.25 $0.0\mu l$ with P=0.75

G% Percentage of GREY colour under the cone-view

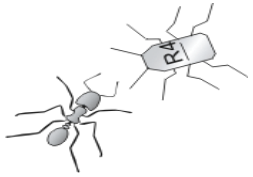
B% Percentage of BLUE colour under the cone-view

Y% Percentage of YELLOW colour under the cone-view

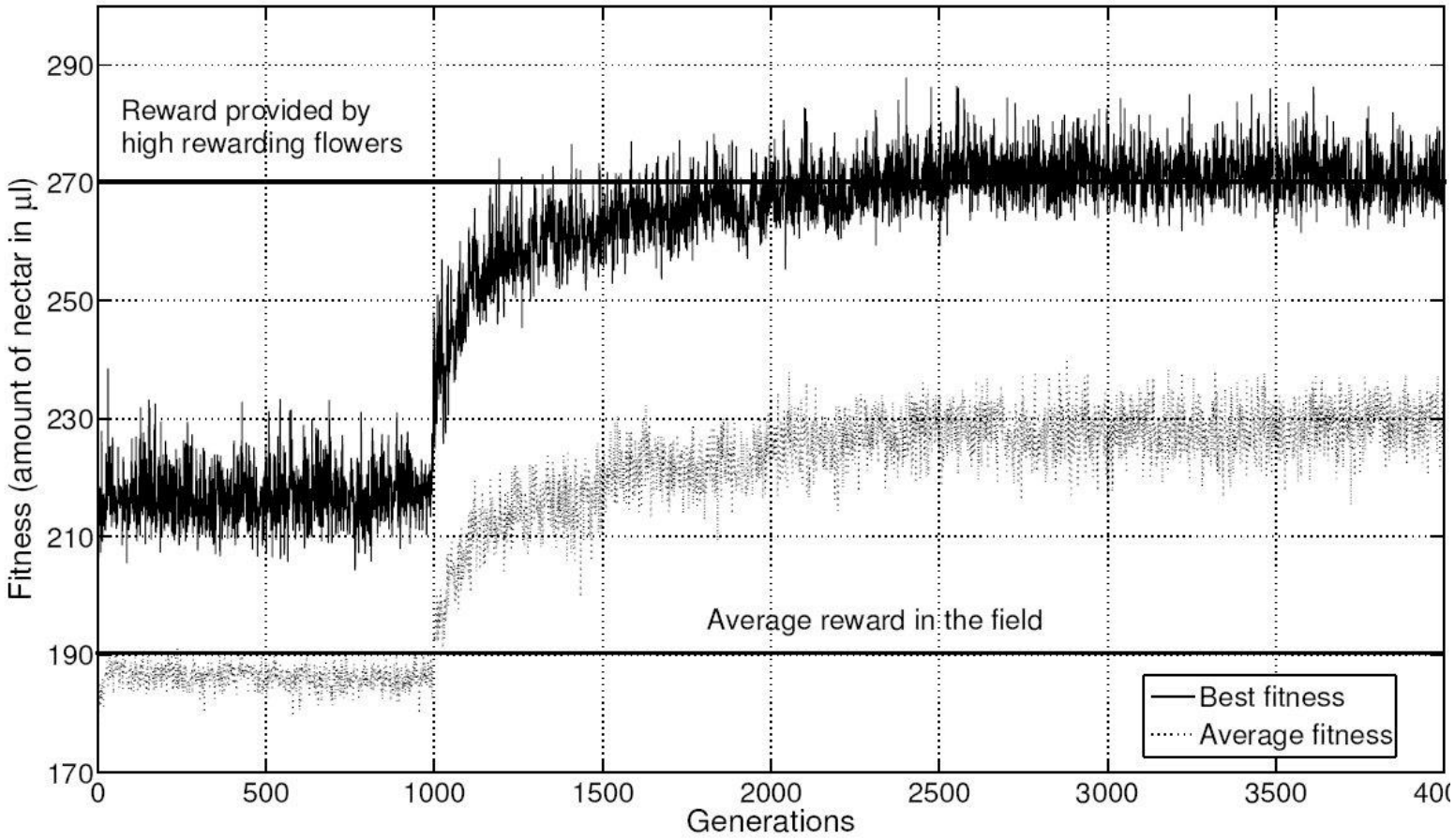
R Reward received upon landing

L Landing signal

$$fitness = \square\ rewards$$

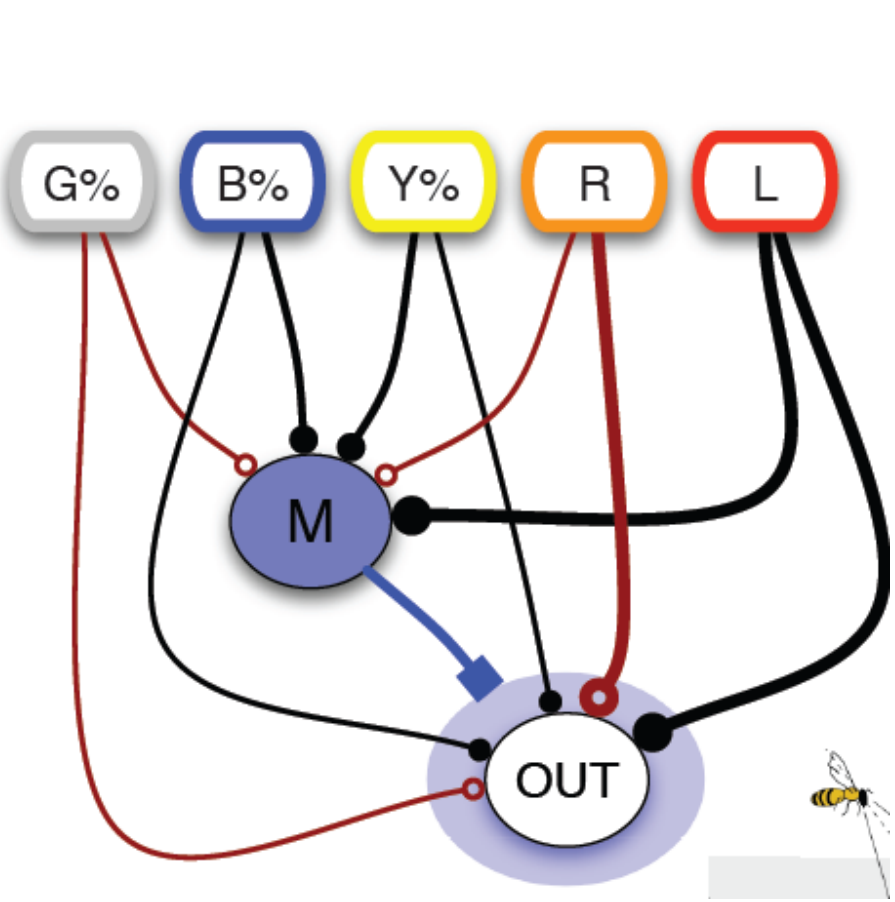


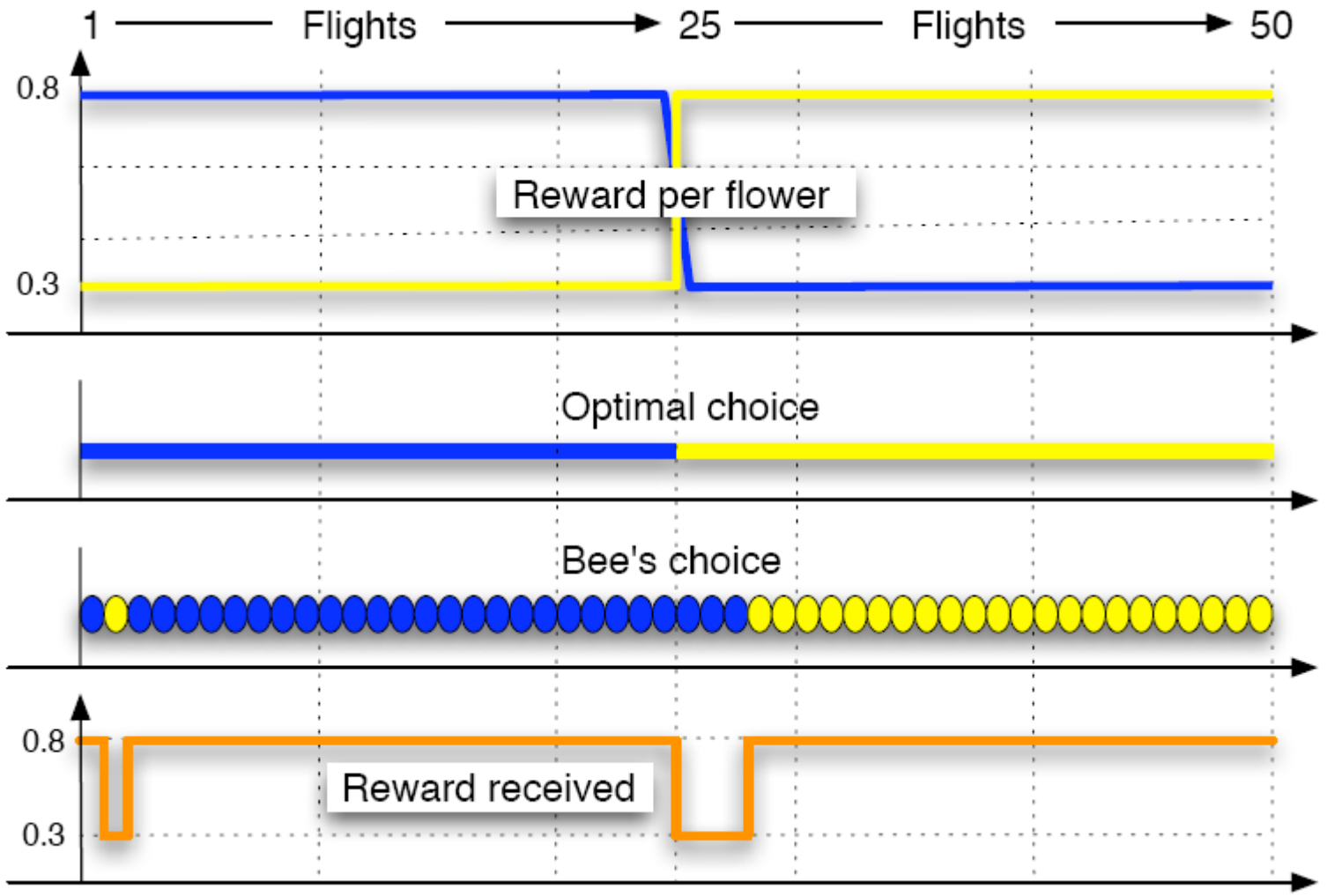
Evolutionary discovery of modulated plasticity



A. Soltoggio, P. Durr, C. Mattiussi and D. Floreano (2007) Evolving neuromodulatory topologies for reinforcement learning-like problems, *IEEE Congress on Evolutionary Computation*, 2471-2478

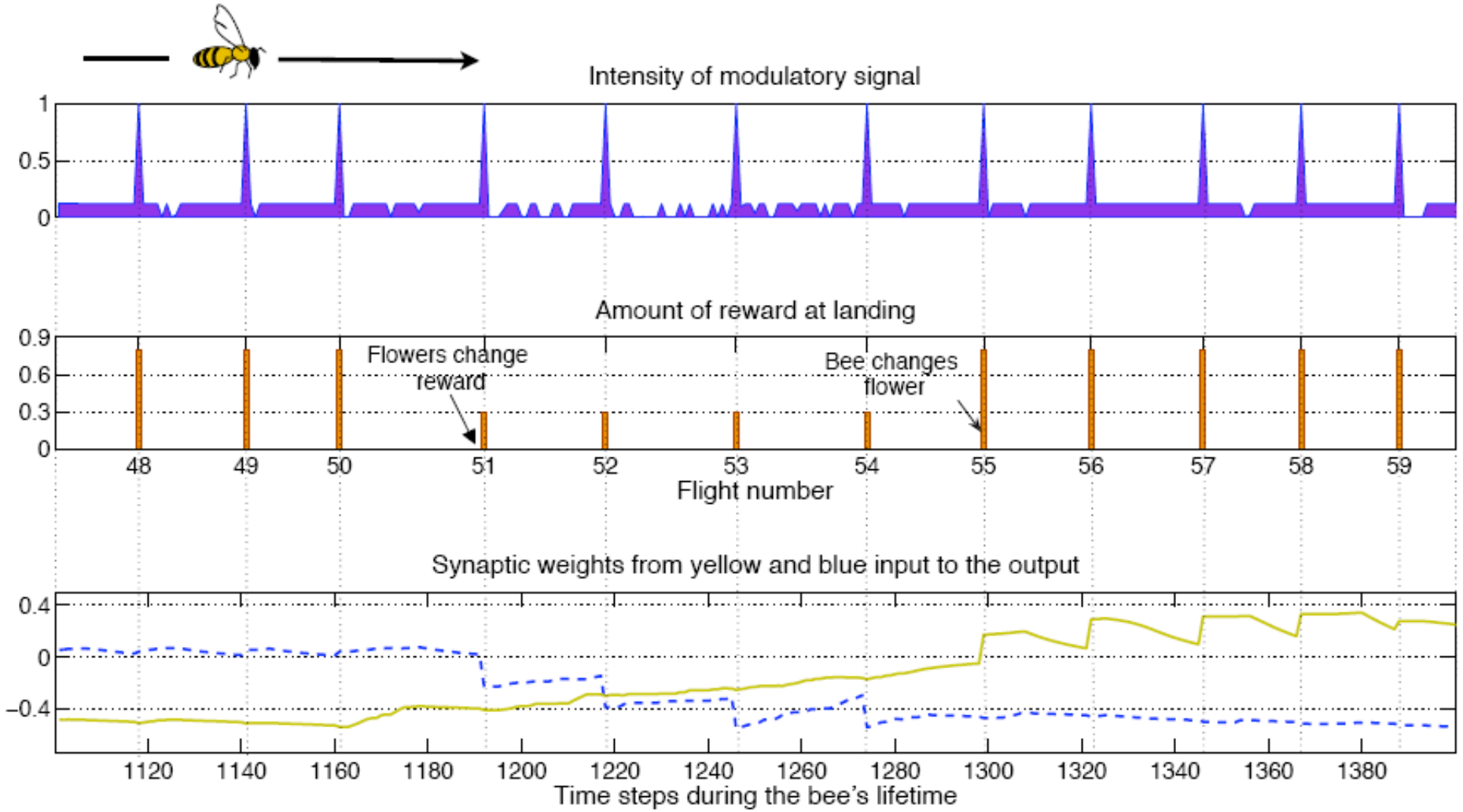
Best evolved individual





A. Soltoggio, P. Durr, C. Mattiussi and D. Floreano (2007) Evolving neuromodulatory topologies for reinforcement learning-like problems, *IEEE Congress on Evolutionary Computation*, 2471-2478

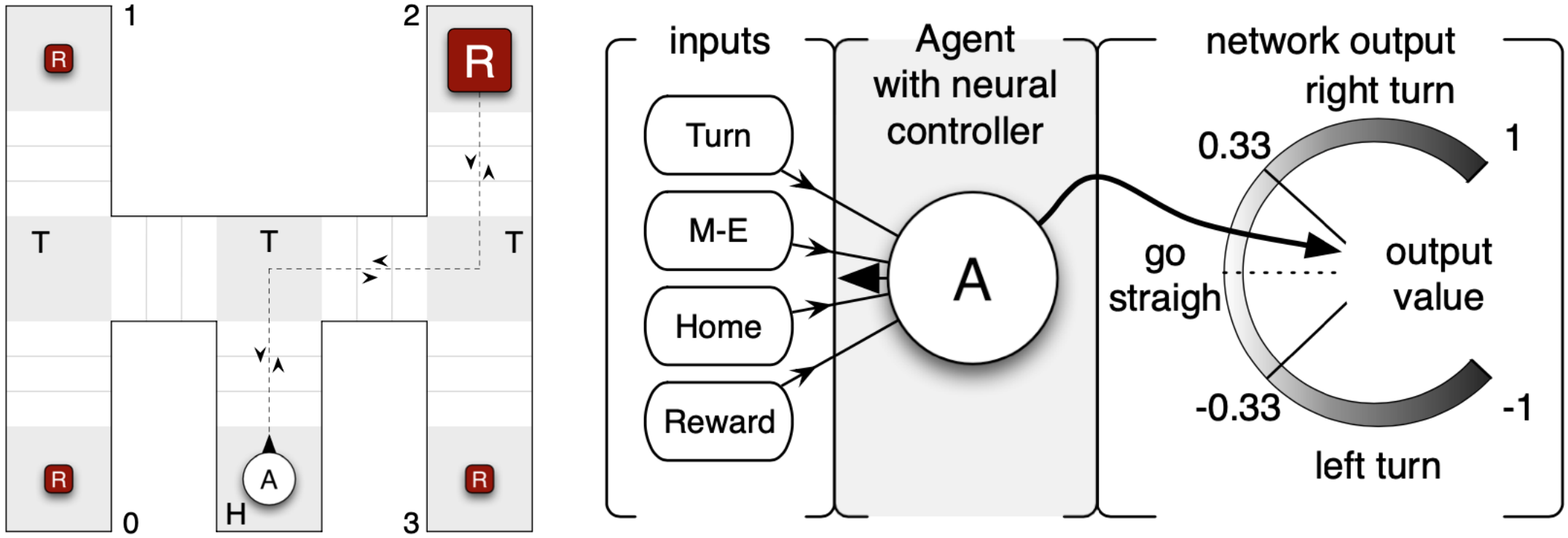
Learning dynamics



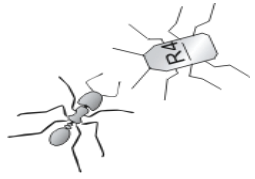
A. Soltoggio, P. Durr, C. Mattiussi and D. Floreano (2007) Evolving neuromodulatory topologies for reinforcement learning-like problems, *IEEE Congress on Evolutionary Computation*, 2471-2478

The double T-maze

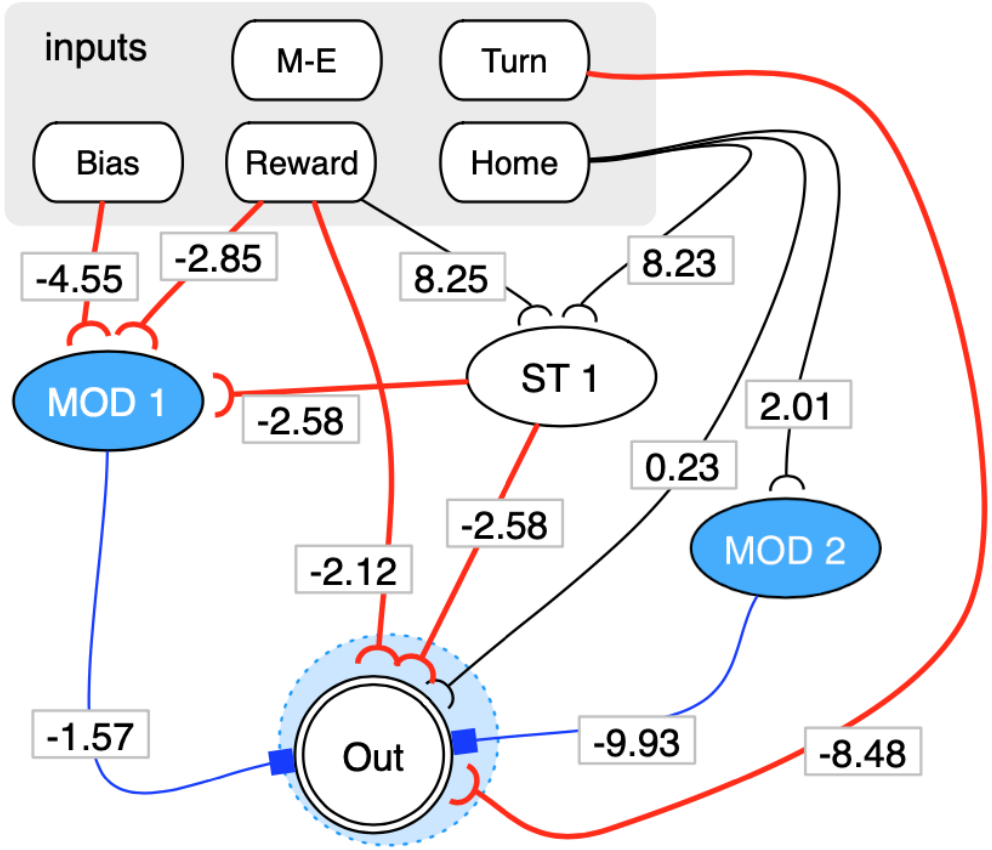
The double T-maze can be memorized too, but requires almost 10 times more generations
In alternative, we can evolve *reinforcement learning rules and architecture*

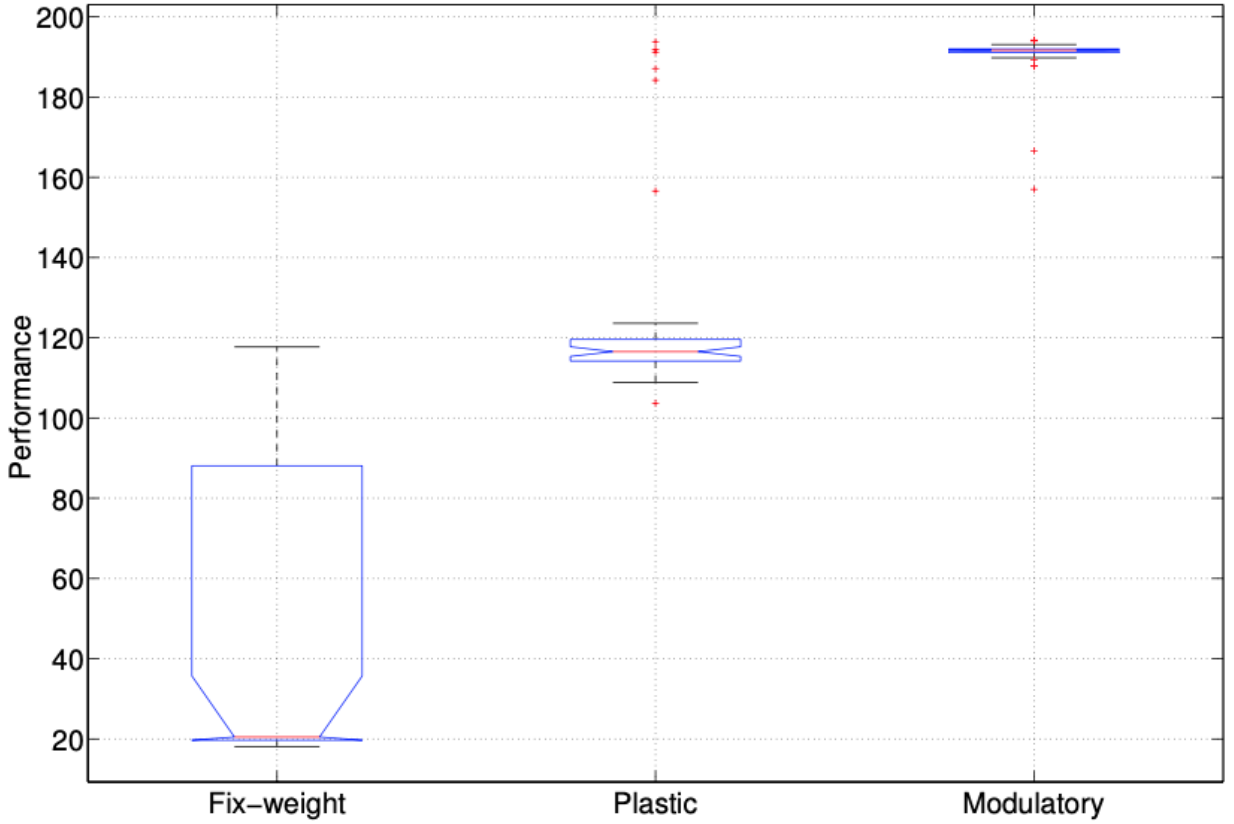


A. Soltoggio; J. A. Bullinaria; C. Mattiussi; P. Dürr; D. Floreano (2008) Evolutionary Advantages of Neuromodulated Plasticity in Dynamic, Reward- based Scenarios. In *Artificial Life XI*, p. 569–576

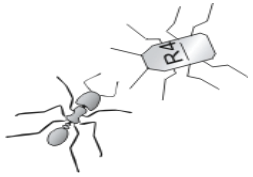


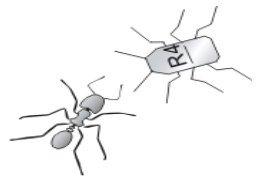
Evolution of reinforcement learning rules & architecture





A. Soltoggio; J. A. Bullinaria; C. Mattiussi; P. Dürr; D. Floreano (2008) Evolutionary Advantages of Neuromodulated Plasticity in Dynamic, Reward- based Scenarios. In *Artificial Life XI*, p. 569–576



Checkpoints

- What is necessary for a neural network to display learning-like behavior without learning?
- List advantages and costs of learning in evolution?
- What is the difference between Lamarckian and Darwinian evolution?
- What is the Baldwin effect?
- Describe ways to combine evolution and learning
- Describe methods for encoding and evolving learning rules
- What are the advantages of evolving learning rules instead of fixed weights?