Playing Atari With 6 Neurons

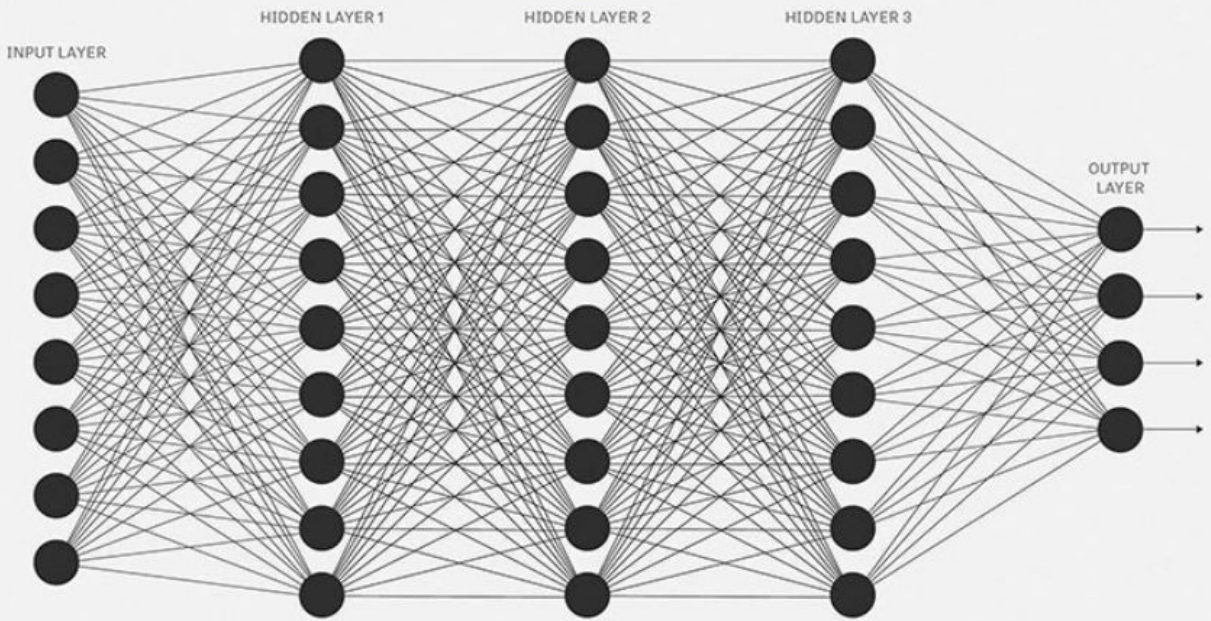
by Giuseppe Cuccu, Julian Togelius, Philippe Cudre-Mauroux

Presented by Marc Soda

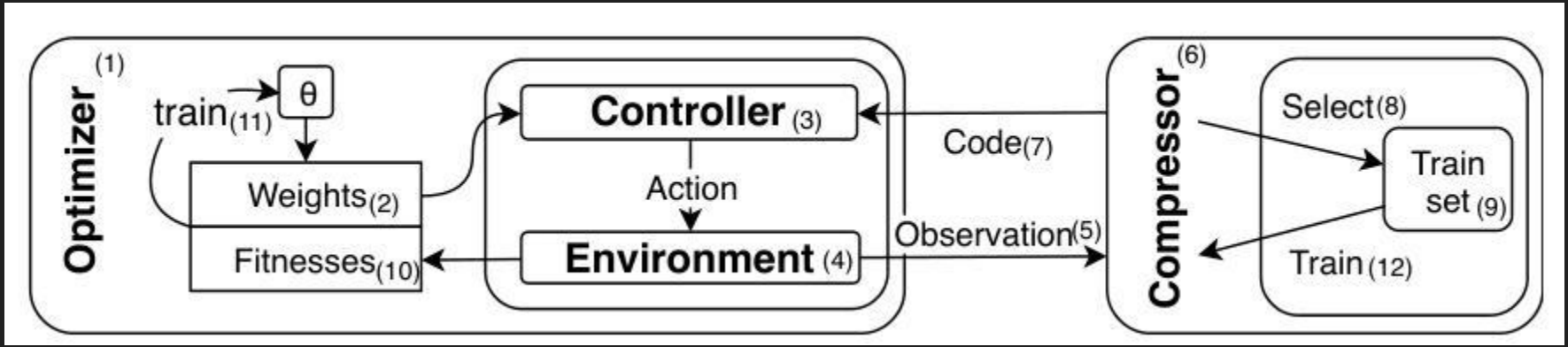
Research

- Novel approach to reinforcement learning
- What is reinforcement learning?
- Goals

Deep neural network



System Overview



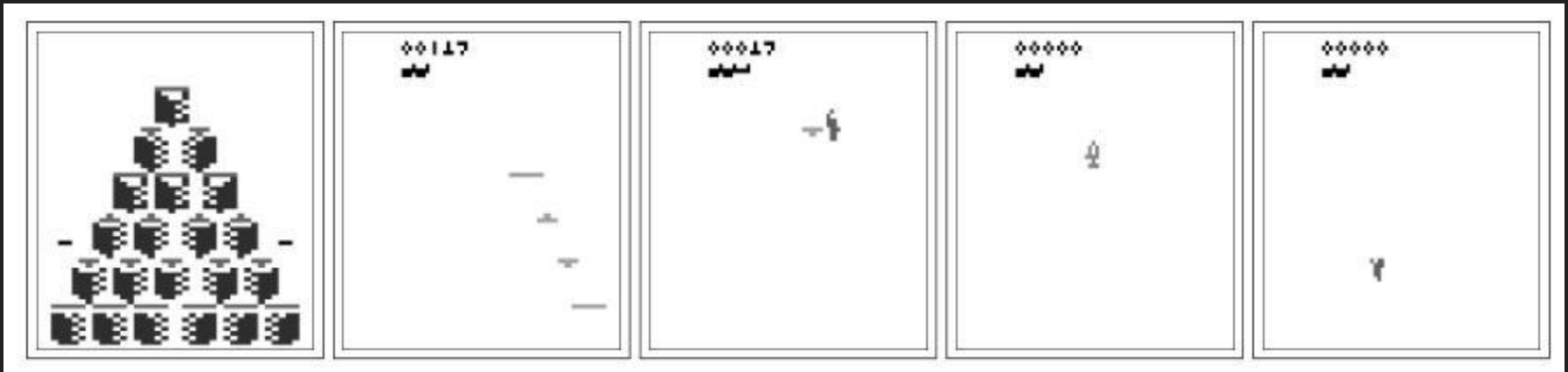
- Best explained with an example

Optimizer

- Initializes weights of new controllers (populations) based on fitness of previous generation controllers
- Uses XNES algorithm
  - neuroevolution

Compressor - Image Processor

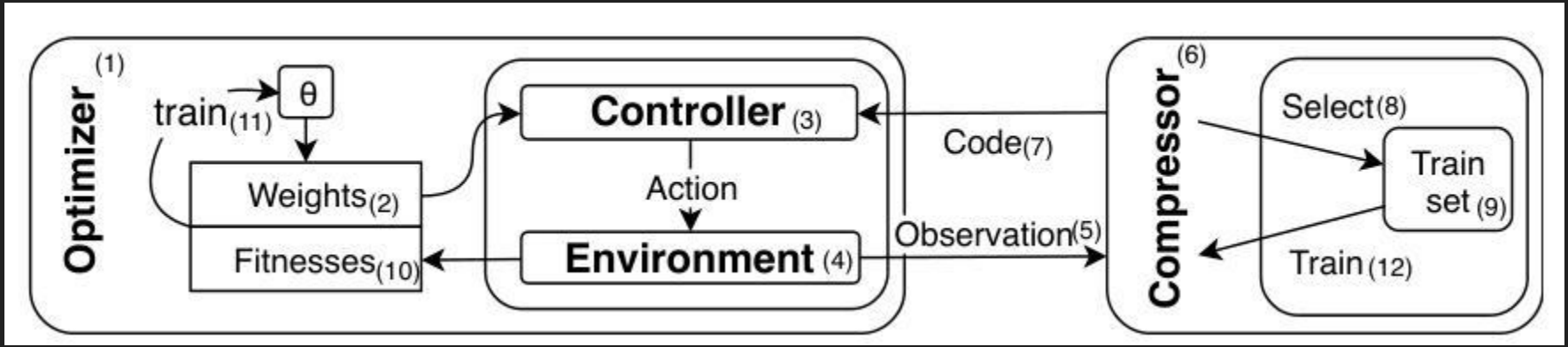
- Takes pixels of every frame and compresses them into “code”
- This code is fed into the controller
- Trained between generations
- IDVQ, DRSC
- Qbert:



Controller - Policy Approximator

- Input: compressor “code”
- Output: next action of the agent
- Difference: much much smaller, less complex
  - Because of decoupling of image processing and policy approximation

System Overview



Results - Scores

Game	HyperNeat	OpenAI ES	GA (1B)	NSRA-ES	**IDVQ+DRSC+XNES**	# of neurons
DemonAttack	3590	1166.5	-	-	325	6
FishingDerby	-49	-49	-	-	-10	18
Frostbite	2260	370	4536	3785	300	18
Kangaroo	800	11200	3790	-	1200	18
NameThisGame	6742	4503	-	-	920	6
Phoenix	1762	4041	-	-	4600	8
Qbert	695	147.5	-	1350	1250	6
Seaquest	716	1390	798	960	320	18
SpaceInvaders	1251	678.5	-	-	830	6
TimePilot	7340	4970	-	-	4600	10

Results - Network Characteristics

	HyperNeat	OpenAI ES	GA (1B)	NSRA-ES	**IDVQ+DRSC+XNES**
# neurons	~3034	~650	~650	~650	**~18**
# hidden layers	2	3	3	3	**0**
# connections	~906k	~436k	~436k	~436k	**~3k**

Conclusion