Computational Intelligence and Machine Learning Group

UNIVERSITY OF PISA
PHD IN COMPUTER SCIENCE
MAURIANA PESARESI SEMINARS

CONTINUAL REINFORCEMENT LEARNING: TOWARDS MULTI-TASK AND GENERIC AGENTS

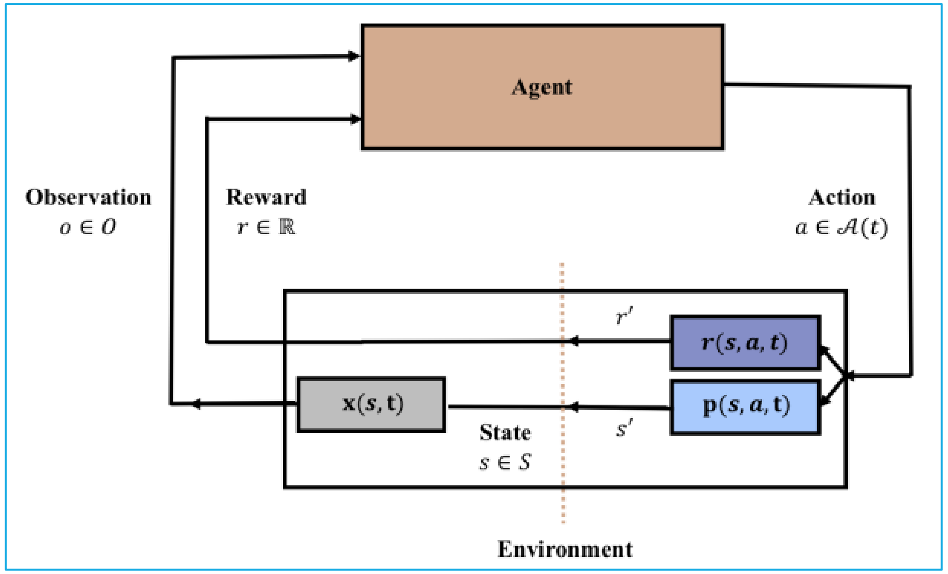
ELIA PICCOLI

CONTENTS OVERVIEW

REINFORCEMENT LEARNING

Reinforcement Learning (RL)[1] formulates the learning process as a sequence of interactions with the environment.

$$\langle S, A, R, P, \gamma \rangle$$



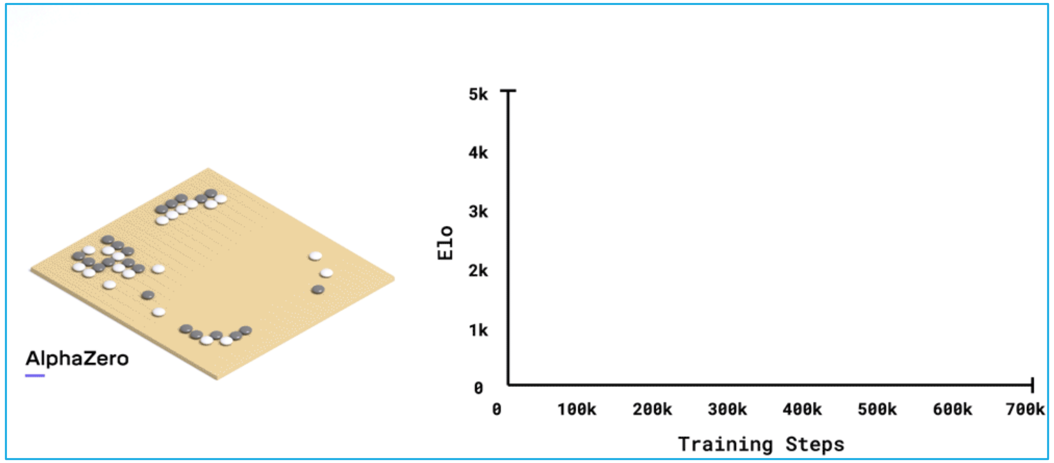
The agent's goal is to maximize the sum of *discounted reward*:

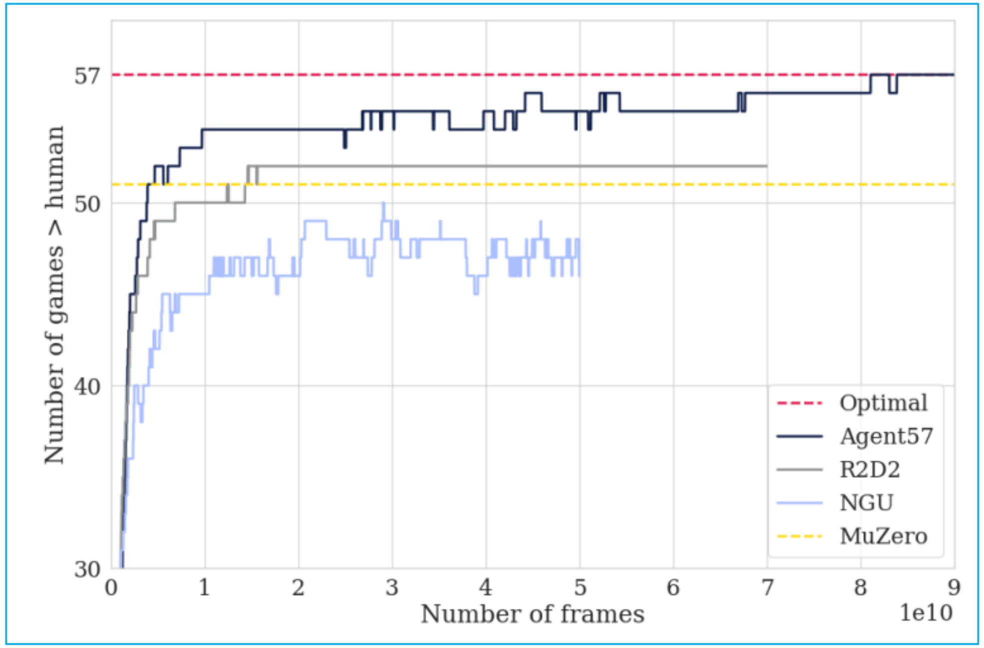
$$G_t = r_{t+1} + \gamma r_{t+2} + \gamma^2 r_{t+3} + \ldots = \sum_{k=0}^{\infty} \gamma^k r_{t+k+1}$$

[1] Richard S Sutton and Andrew G Barto. *"Reinforcement learning: An introduction"*. MIT press, 2018.

REINFORCEMENT LEARNING – PROBLEM

Reinforcement Learning has achieved astonishing results reaching <u>super-human performances</u> and even beating professional players in different games or scenarios, e.g., *AlphaZero*[2] or *Agent57*[3].



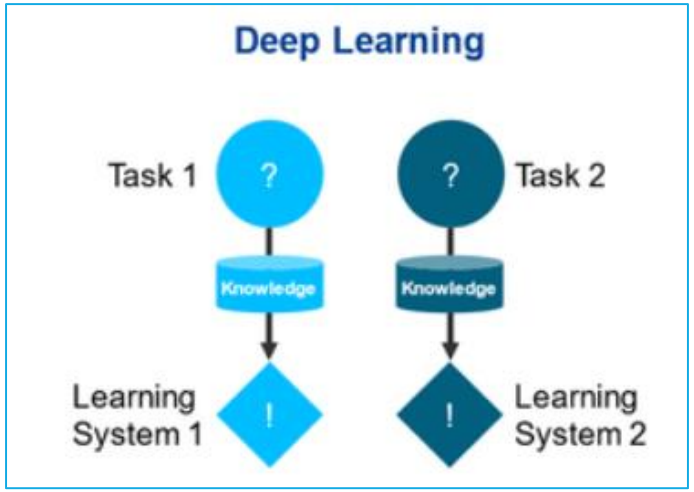


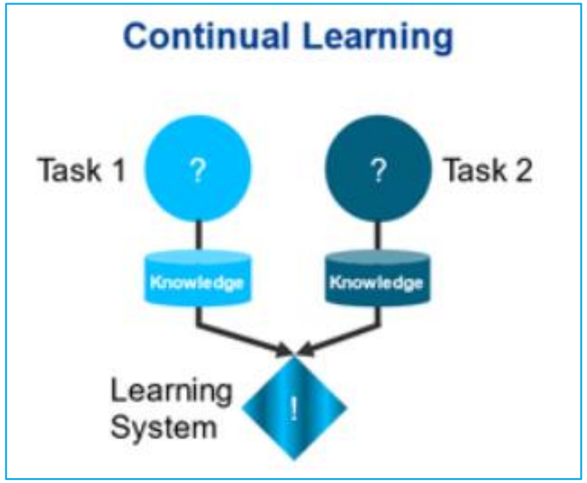
[2] Silver, David, et al. *"A general reinforcement learning algorithm that masters chess, shogi, and Go through self-play"*, *Science* (2018)
[3] Badia, A. P., et. al. *"Agent57: Outperforming the atari human benchmark"*. In International conference on machine learning (2020)
GIF from https://www.deepmind.com/blog/alphazero-shedding-new-light-on-chess-shogi-and-go

CONTINUAL LEARNING

Continual Learning (CL) seeks to create models that can *incrementally learn* from a sequence of tasks. Key aspects are the *sequential* and *non-stationary* nature of the learning process

The challenge: Learn without *catastrophic forgetting*.





Maschler, Benjamin, et. al. *"Deep transfer learning for industrial automation: a review and discussion of new techniques for data-driven machine learning."* (2021)

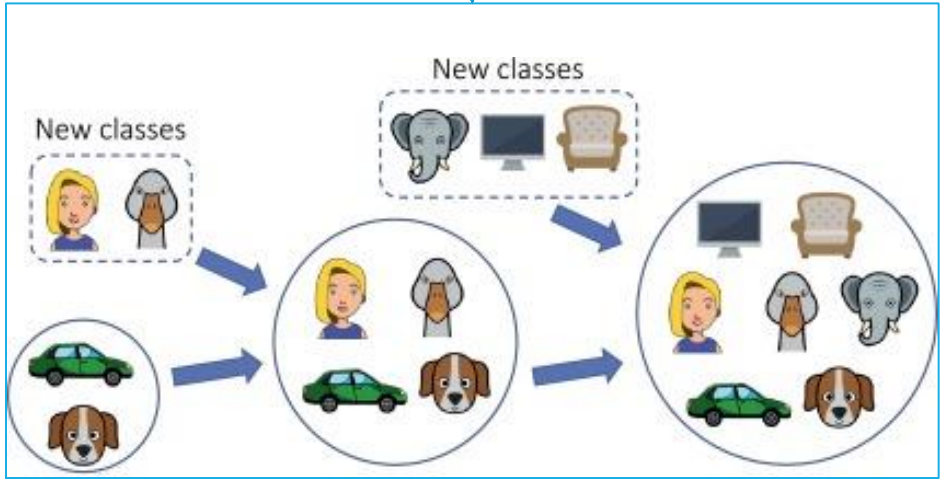
CONTINUAL LEARNING - SCENARIOS

Continual Learning Scenarios

Single Incremental Task

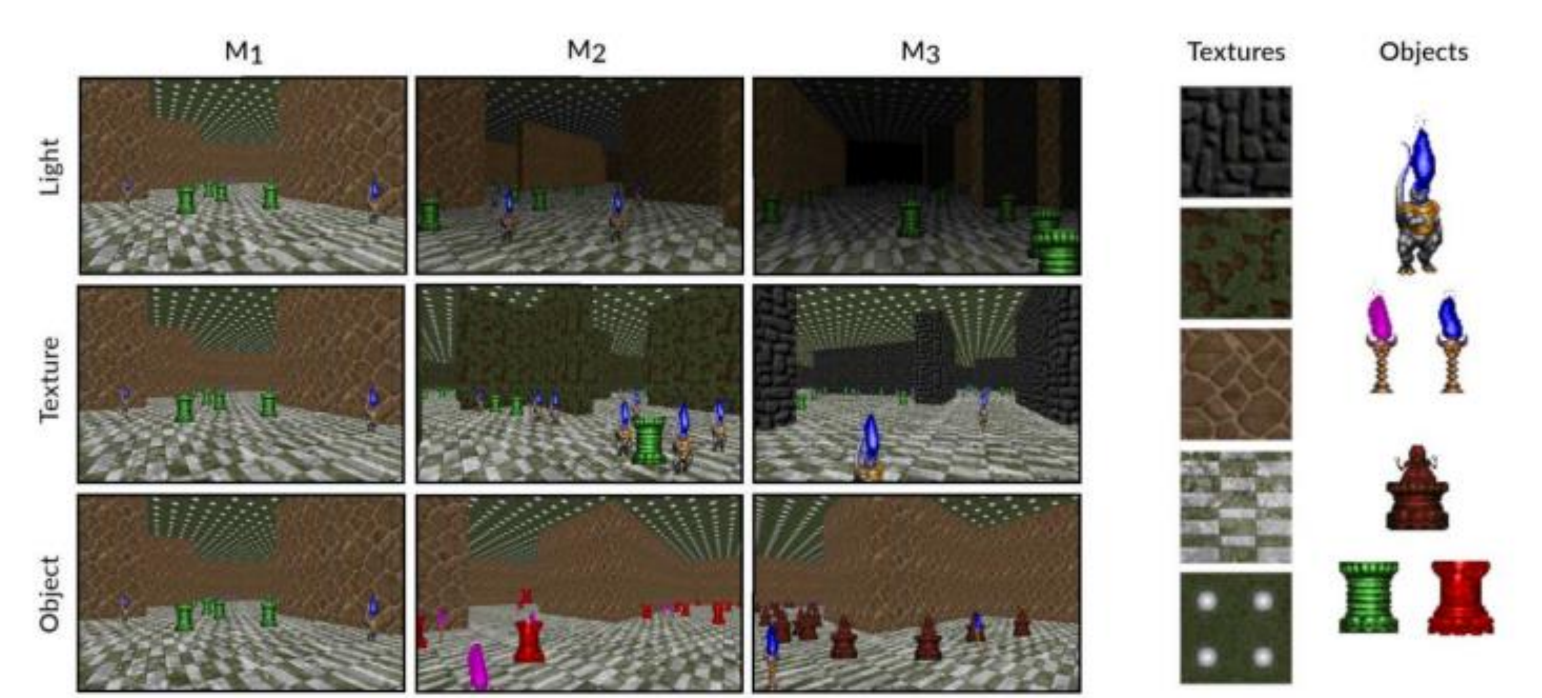
Multi-Task

Multi Incremental Task



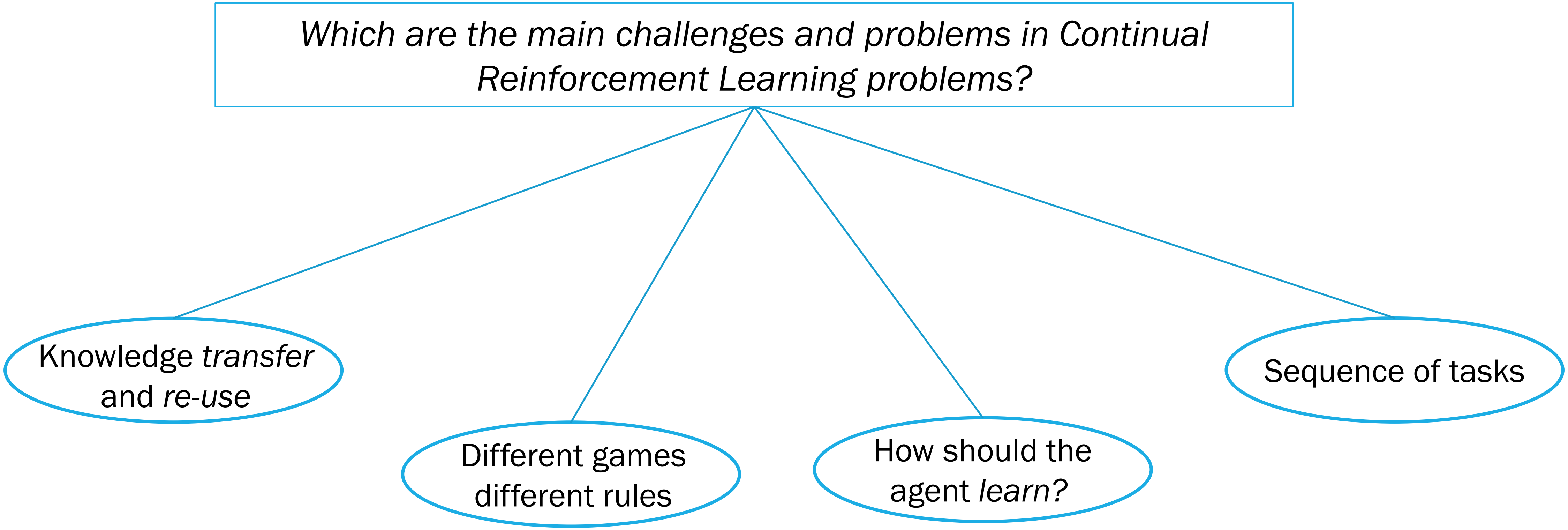
CONTINUAL REINFORCEMENT LEARNING

Continual Reinforcement Learning (CRL) leverages *different scenarios* to train agents that can learn to solve *multiple task simultaneously*.



Lomonaco, Vincenzo, et al. *"Continual reinforcement learning in 3d non-stationary environments."* (2020)

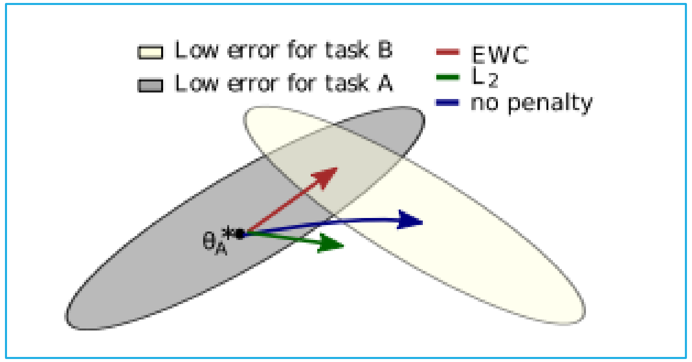
CONTINUAL REINFORCEMENT LEARNING - PROBLEMS

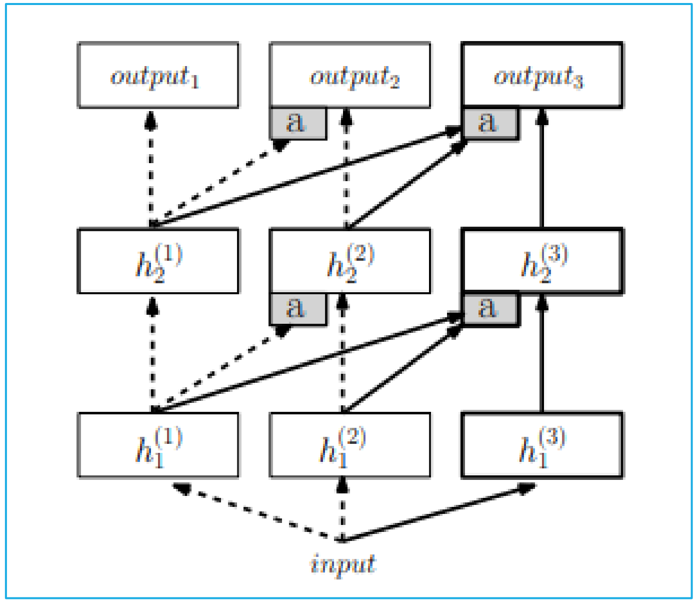


CONTINUAL REINFORCEMENT LEARNING – SOLUTIONS 1

Two different approaches that can be used to tackle these problems are: *Elastic Weight Consolidation (EWC)*[6] and *Progressive Neural Networks (PNN)*[7].

$$\mathcal{L}(\theta) = \mathcal{L}_B(\theta) + \sum_i \frac{\lambda}{2} F_i(\theta_i - \theta^*_{A,i})^2$$

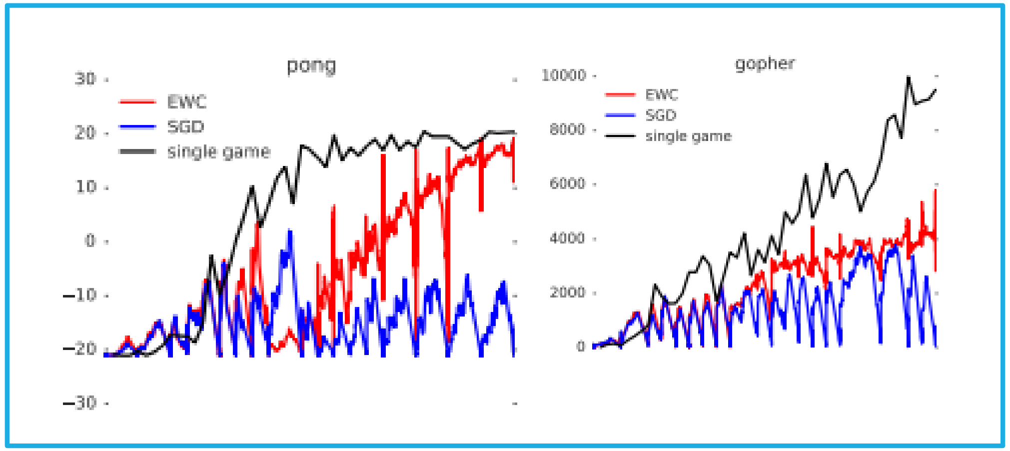


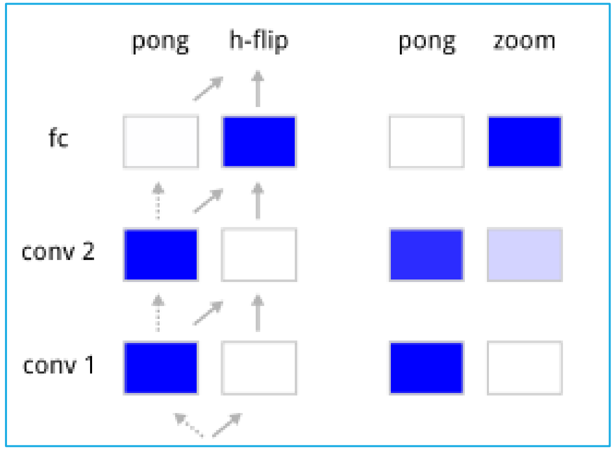


[6] Kirkpatrick, James, et al. *"Overcoming catastrophic forgetting in neural networks."* (2017)
[7] Rusu, Andrei A., et al. *"Progressive neural networks."* (2016)

CONTINUAL REINFORCEMENT LEARNING – SOLUTIONS 2

Two different approaches that can be used to tackle these problems are: *Elastic Weight Consolidation (EWC)*[6] and *Progressive Neural Networks (PNN)*[7].





[6] Kirkpatrick, James, et al. *"Overcoming catastrophic forgetting in neural networks."* (2017)
[7] Rusu, Andrei A., et al. *"Progressive neural networks."* (2016)

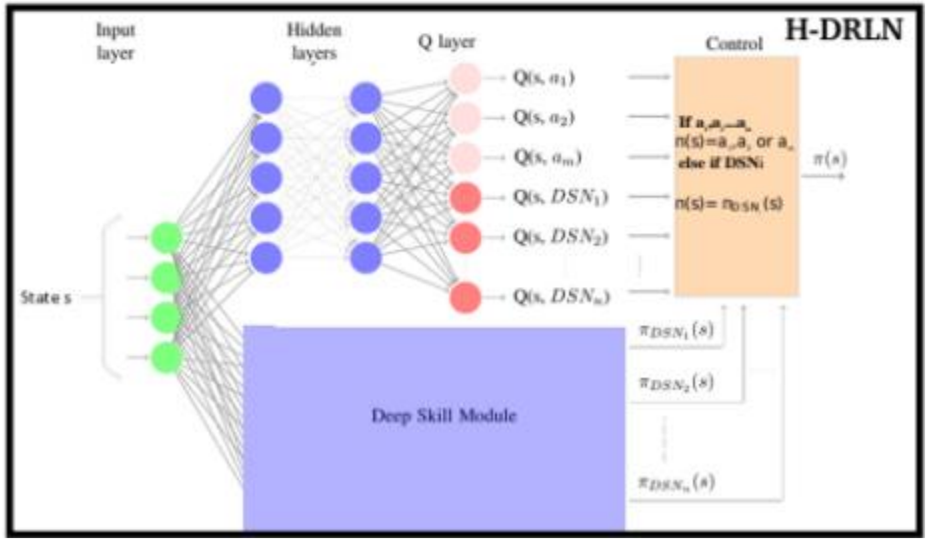
CONTINUAL REINFORCEMENT LEARNING – SKILLS

Another solution is to consider the *compositionality* of tasks: in order to solve a task, it can be divided into several *sub-problems* that has to be solved. The agent learns to combine partial solutions to reach the desired behavior.

Semi-Markov Decision Processes (SMDPs)[8]

$$< S, \Sigma, R, P, \gamma >$$

each skill defined by a tuple (I, π, β)

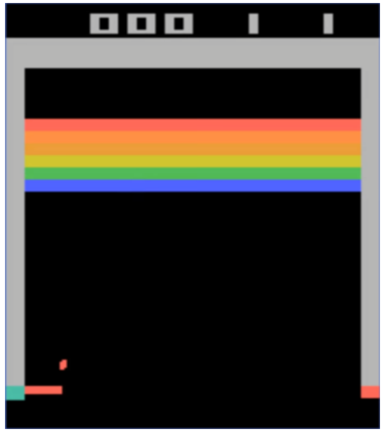
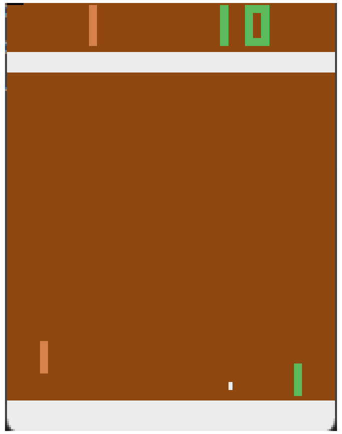
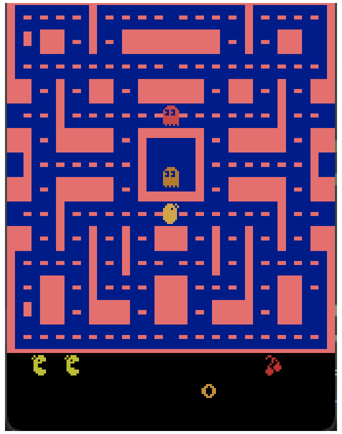
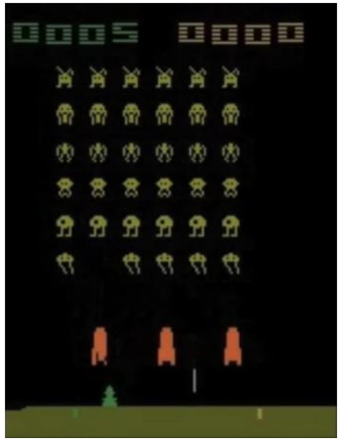


[8] Richard S Sutton, et. al. *"Between MDPs and semi-MDPs: A framework for temporal abstraction in reinforcement learning"* (1999)
Tessler, Chen, et al. *"A deep hierarchical approach to lifelong learning in Minecraft."* (2017)

SKILLS – NEW SKILL DEFINITION

Which are the main characteristics that a skill should have?

- Task Agnostic
- Unsupervised
- Easy access to data
- Compute useful information

FINAL OVERVIEW

Key ideas highlighted today:

- Introduced, even if very briefly, the main learning approaches: *Reinforcement Learning*, *Continual Learning* and *Continual Reinforcement Learning.*
- Highlighted the *different challenges* that emerge while tackling these problems.
- Analyzed several *state-of-the-art* solutions to Continual Reinforcement Learning tasks.

IDEAS FOR DISCUSSION/FUTURE WORKS

- How can we scale up how systems? CRL is *computational resources hungry!*

- How can we make the learning process more *human-like*?

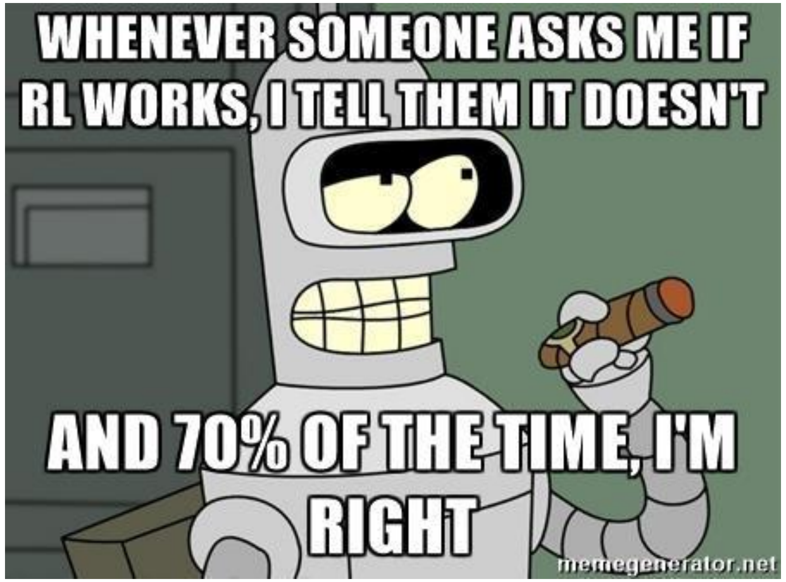
- How can we define *knowledge and abilities* to ease agents' learning process? Can we *transfer* and *re-use* knowledge?

- Should we change how we *see and approach* the problem of learning multiple tasks simultaneously?

Q&A AND DISCUSSION

THANK YOU FOR YOUR ATTENTION!

Q&A



PERSONAL INFO:

ELIA PICCOLI

HTTPS://ELIAPICCOLI.GITHUB.IO/

ELIA.PICCOLI@PHD.UNIPI.IT