Ruben Glatt

Improving Deep Reinforcement Learning with Knowledge Transfer

Dipl.-Ing. Mechatronics & Master Mechanical Engineering
2nd year D.Sc. Candidate (expected to graduate 03/2019)
@ Escola Politécnica da Universidade de São Paulo
Advised by Anna Helena Reali Costa

Research interests

- Intelligent agents utilizing scalable architectures for lifelong learning
- Increasingly on AI influence on society and ethics

Future goals

- Create positive impact in the world
- Live forever (preferably in New York)
- More short-term:
  - **Internship late 2017/early 2018 (3-6 month)**

Find out about me: http://www.cowhi.org
Talk to me: ruben.glatt@usp.br

Ruben Glatt

IMPROVING DEEP REINFORCEMENT LEARNING WITH KNOWLEDGE TRANSFER

Escola Politécnica da Universidade de São Paulo
Laboratório de Técnicas Inteligentes
Departamento de Engenharia de Computação

AAAI-17
Doctoral Consortium

San Francisco
04/02/2017

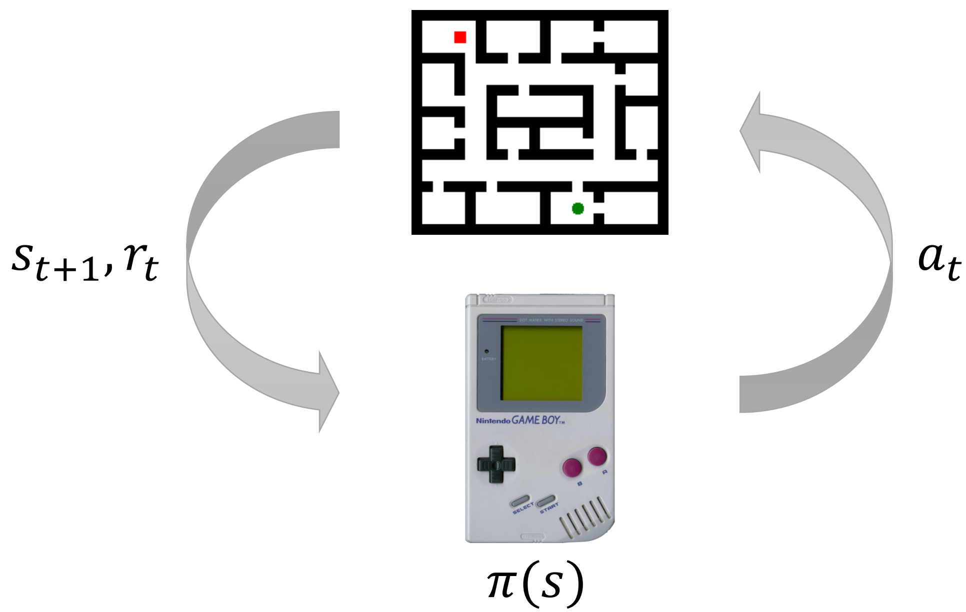
Agenda

- Content
  - Context
    - RL, DRL, Observations, TL
  - Proposal
    - Motivation, Key concept, Intended contribution, Related work
  - Progress
    - Preliminary results, Next steps, Going further
  - Wrap up
    - Summary, Acknowledgements

Context

Reinforcement Learning

- Solve sequential decision-making problems
- Learn policy to maximize total future reward
  - Balance exploration and exploitation during training
  - Learn from feedback of environment

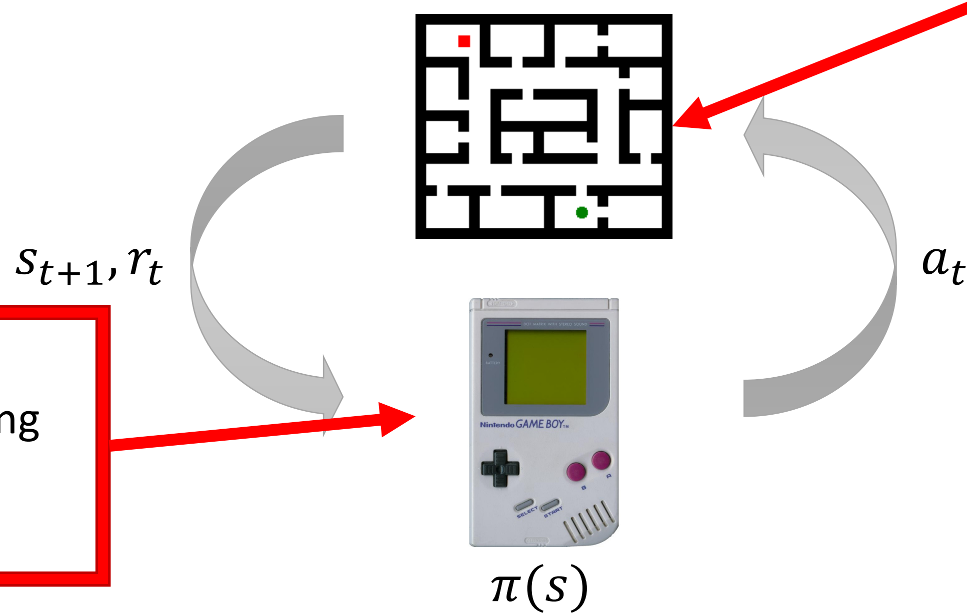


[Sutton & Barto, 1998]

Reinforcement Learning

- Solve sequential decision-making problems
- Learn policy to maximize total future reward
  - Balance exploration and exploitation during training
  - Learn from feedback of environment

Low dimensional problem space (coordinates, few parameters)

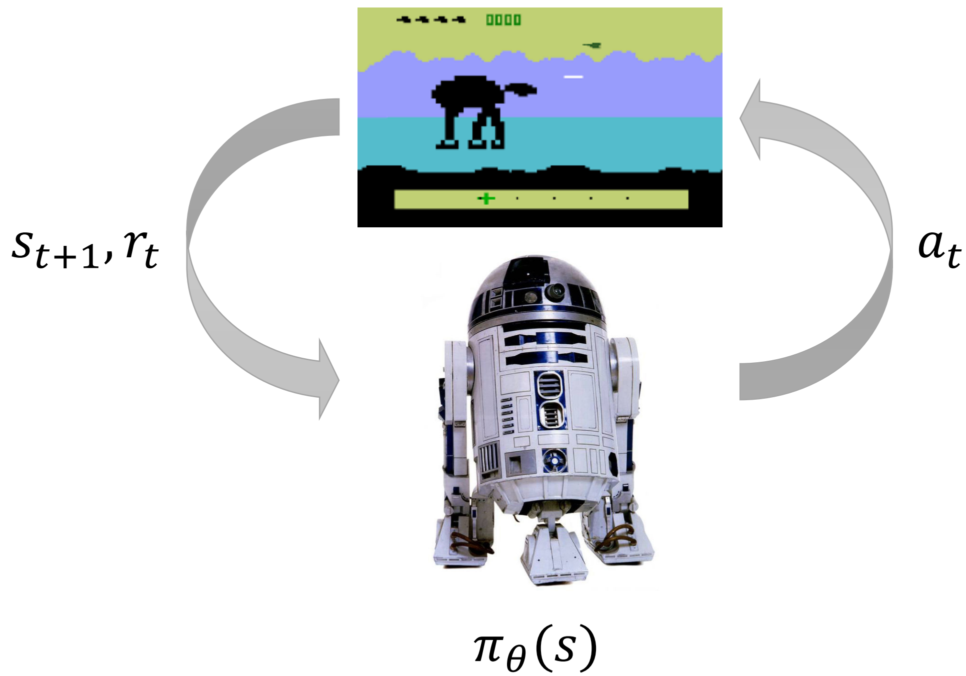


Little processing power

[Sutton & Barto, 1998]

Deep Reinforcement Learning

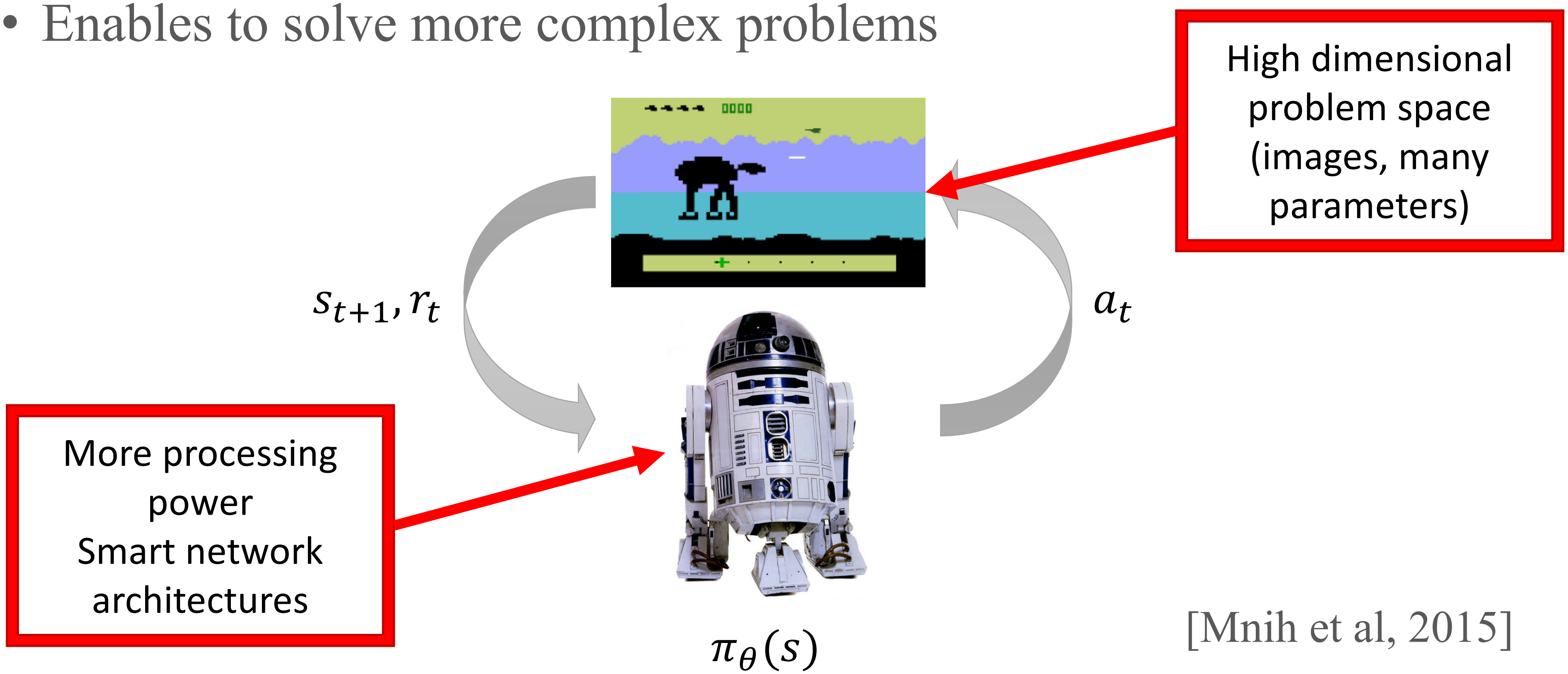
- Same as classic Reinforcement Learning, but …
  - Uses Deep Neural Networks as function approximator
  - State abstraction and value/policy approximation within single network
  - Enables to solve more complex problems



[Mnih et al, 2015]

Deep Reinforcement Learning

- Same as classic Reinforcement Learning, but …
  - Uses Deep Neural Networks as function approximator
  - State abstraction and value/policy approximation within single network
  - Enables to solve more complex problems



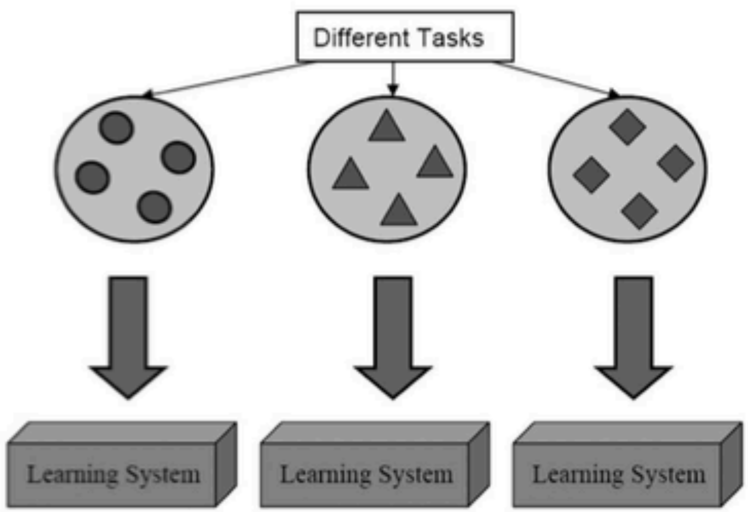
[Mnih et al, 2015]

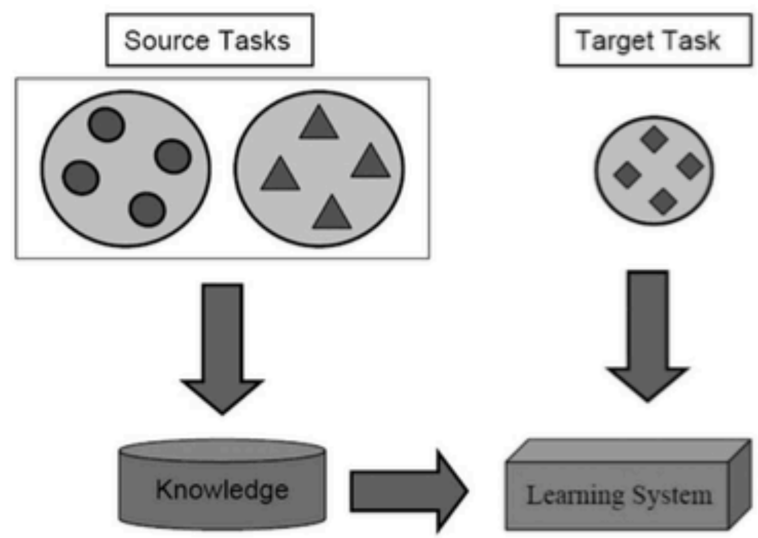
Initial observations

- Amazing results for single task learning with DRL
  - Atari [Mnih et al, 2015], Robotics [Levine et al, 2016], 3-D environments [Mirowski et al, 2016], …
- BUT…
  - Every new task has to be learned from scratch
  - Sample inefficient
  - Still relatively slow
  - No usage of prior knowledge
  - Few publications on multi-task learning with DRL

Transfer Learning

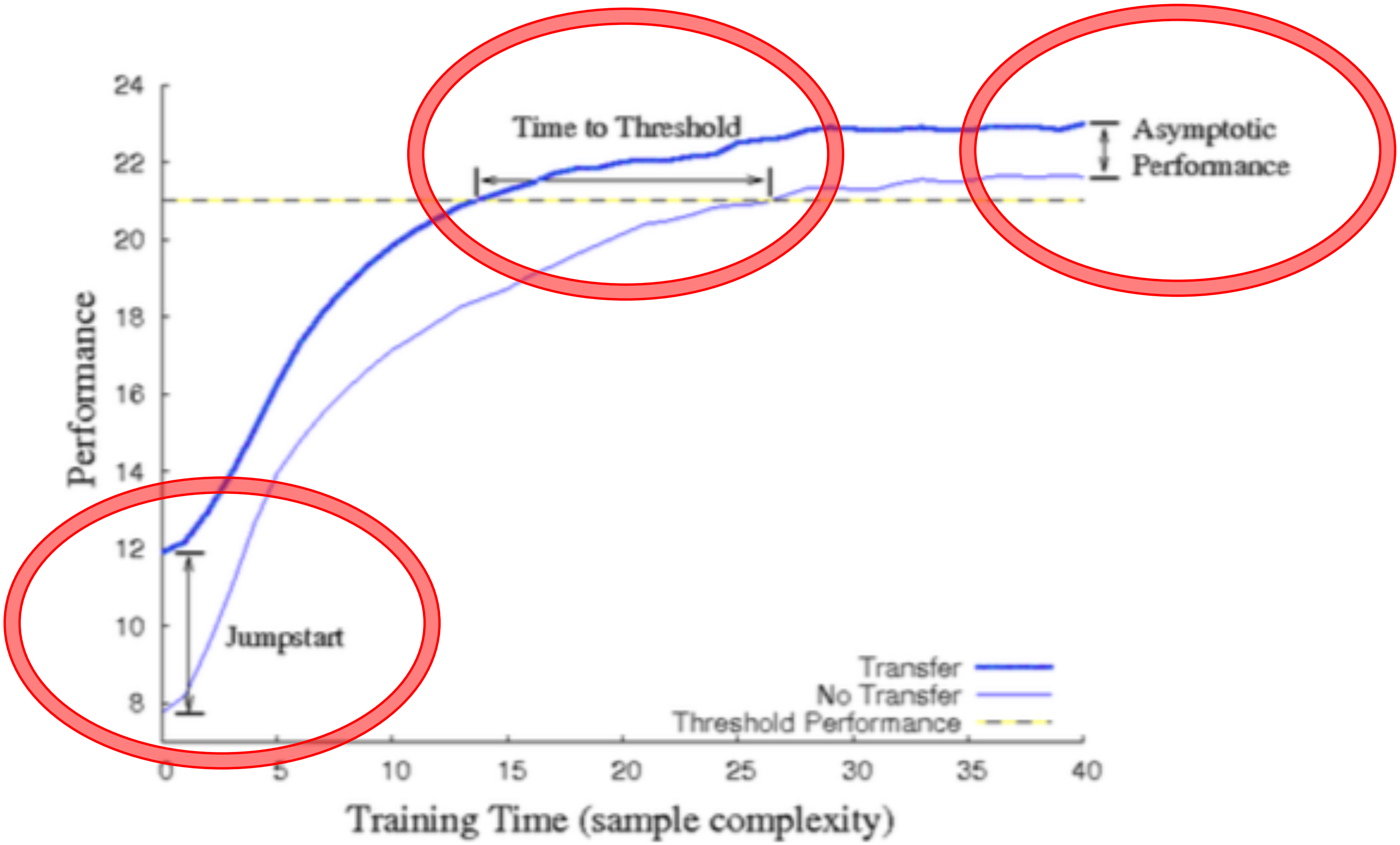
- Reuse existing knowledge while learning new task





[Pan & Yang, 2010]

Transfer Learning for RL



[Taylor & Stone, 2009]

Proposal

Motivation

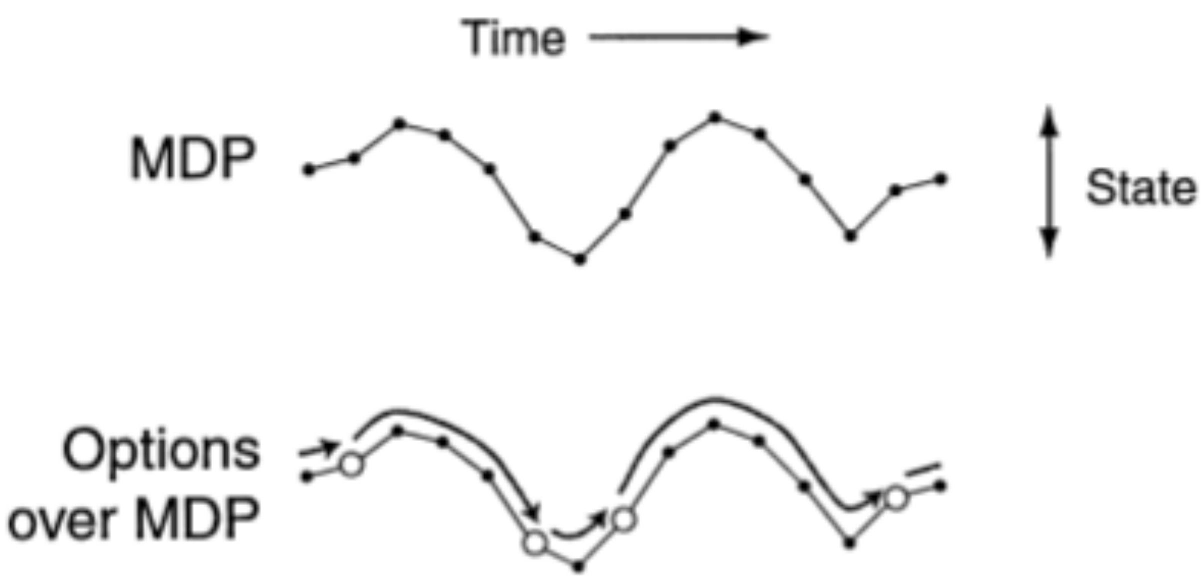
- Abstract goal
  - Improve and speed up DRL for multi-task with knowledge transfer

- Main questions
  - How to guide learning to balance between past and current knowledge?
  - How can we reuse existing knowledge?
  - How to extract and conserve knowledge?

- Challenges
  - Knowledge selection (what knowledge should be transferred)
  - Knowledge source selection (what seems to be similar)
  - Knowledge management (what to keep/discard)

Key concept: Options

- Options (also: skills, macro-actions)
  - Temporally extracted actions for sub-task solutions
  - Defined as initiation set, option policy, and termination condition
  - Flexible knowledge presentation for better generalization



[Sutton et al, 1999]

Intended contribution

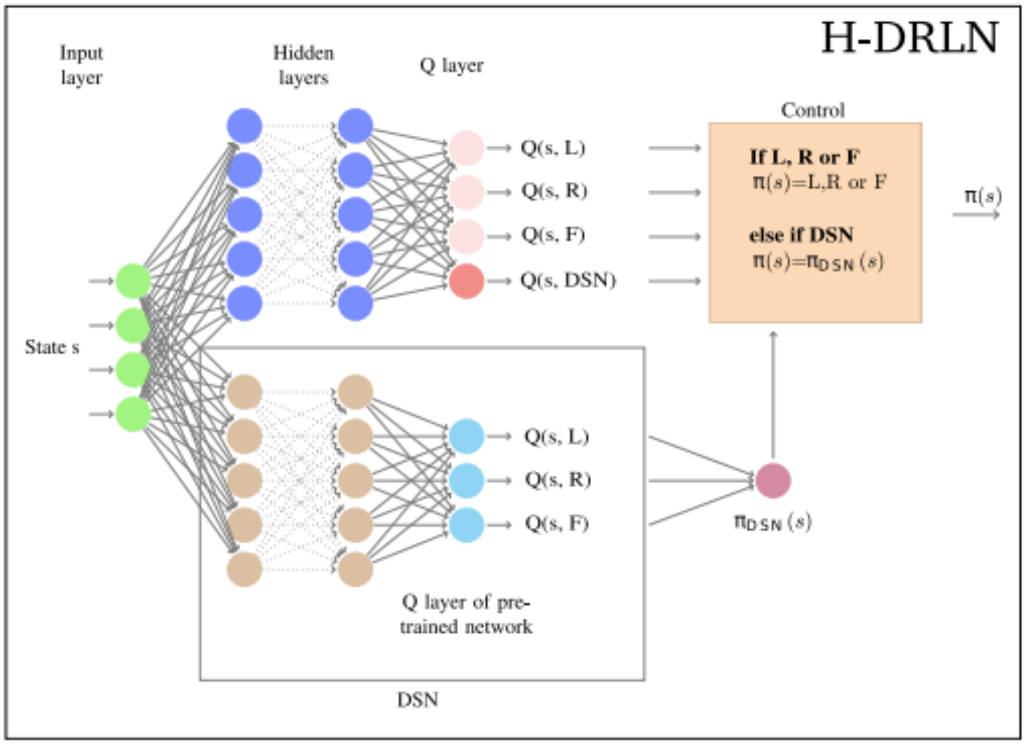
- Intended outcome
  - Framework for option extraction and reuse for learning multiple tasks using DRL approaches

- Research goals
  - Learn to autonomously extract stochastic options
  - Build self-extending option library
  - Design an importance metric to choose right option for targeted knowledge transfer when learning new task

- Why DRL?
  - Integrate learning in single network structure

Related work

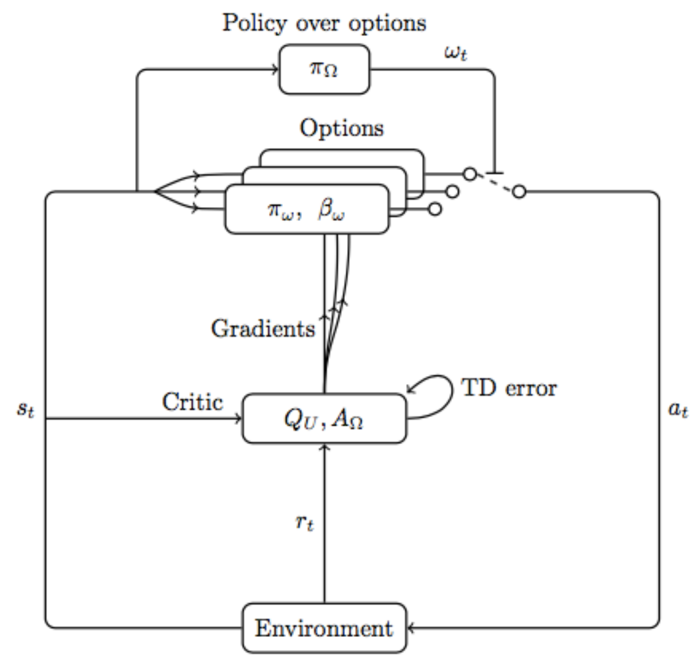
- Deep Skill Networks
  - Learns to use primitive actions or options for hierarchical tasks
  - Deep Skill Module: Library of a-priori trained options

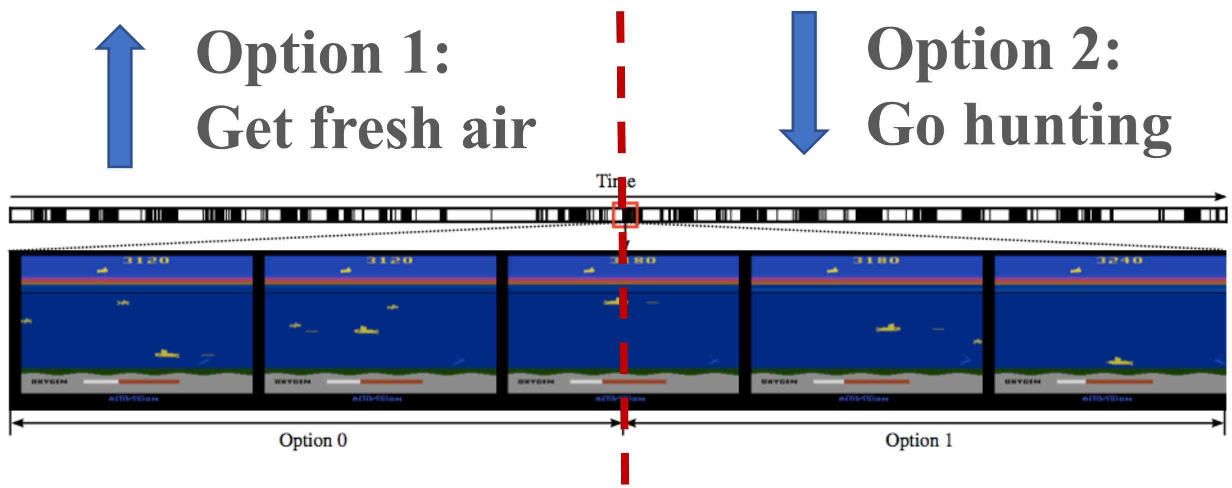


[Tessler et al, 2016]

Related work

- The Option-Critic framework
  - Learns options policies and termination function “on the go”
  - No need to provide any additional rewards or sub goals



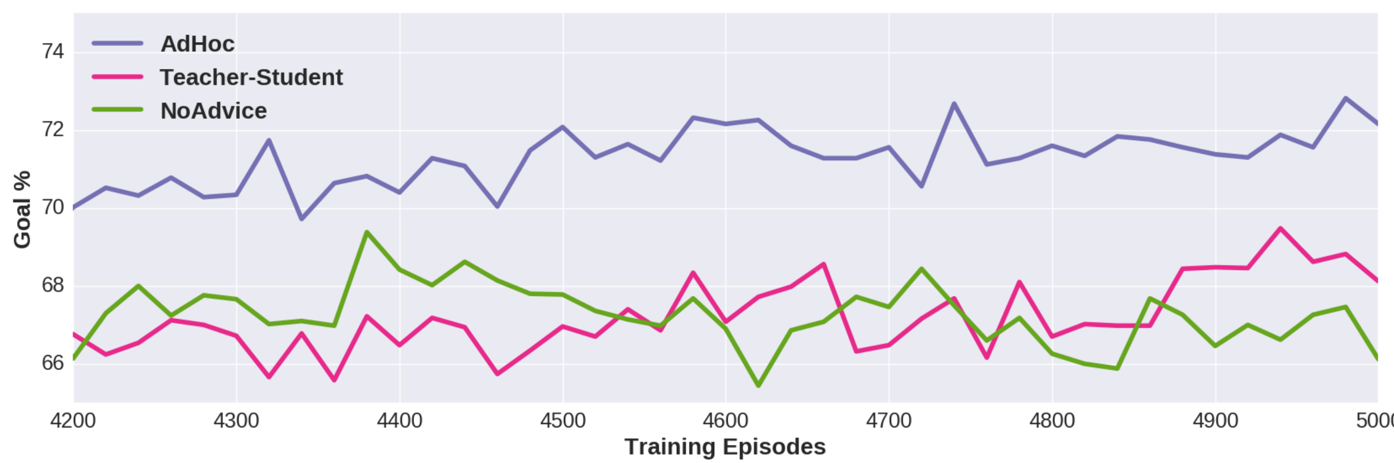


[Bacon et al, 2017]

Progress

Preliminary work

- Multiagent RL based on teacher-student framework
  - Knowledge transfer in form of advice through AdHoc communication
  - Advisor and advisee can learn simultaneously



[Silva et al, 2017]

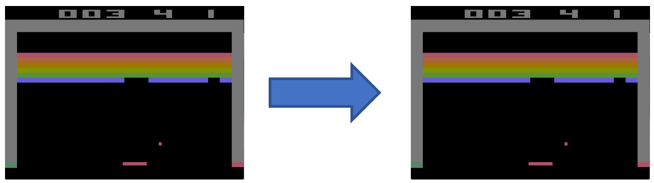
Preliminary work

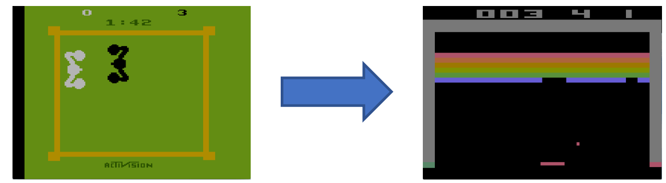
- "Baby steps" in TL for DRL

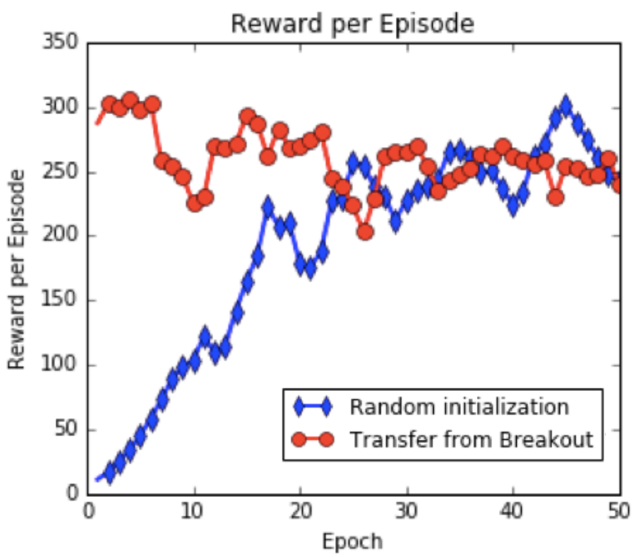
- Knowledge transfer through simple parameter initialization from previously learned tasks
  - Confirms results from classic domain
    - Transfer works when source task well chosen
    - Highlight importance of being able to start learning of new task with good set of parameters
  - But…
    - No autonomous source task selection

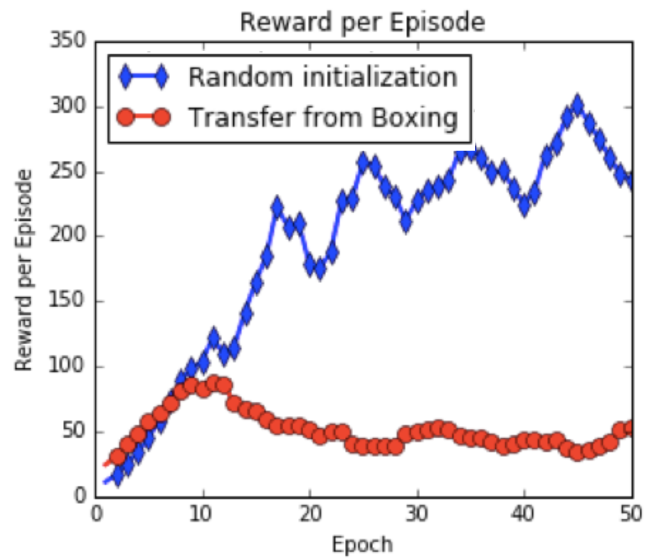
[Glatt et al, 2016]

Preliminary work





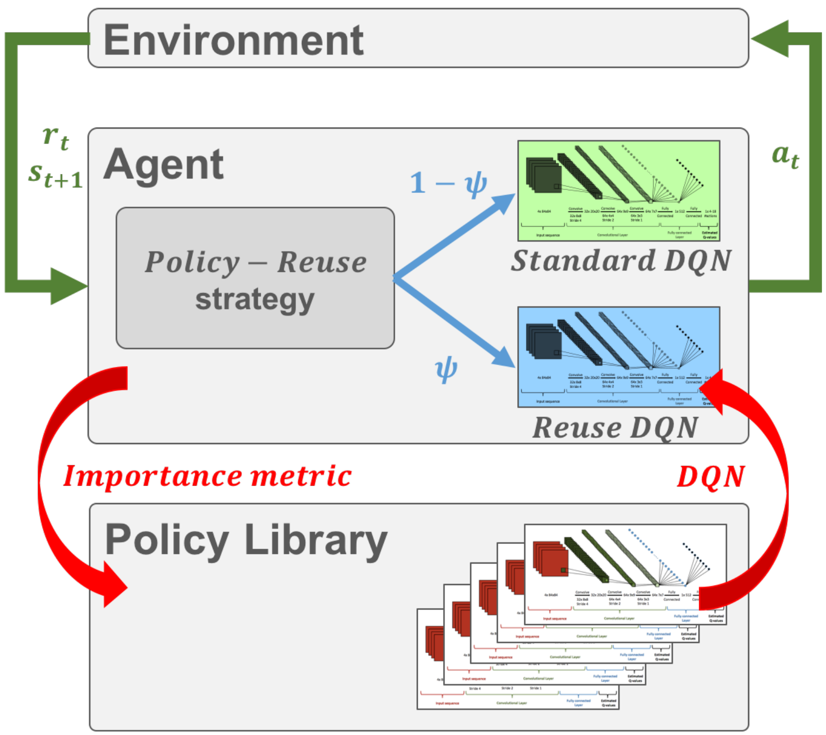


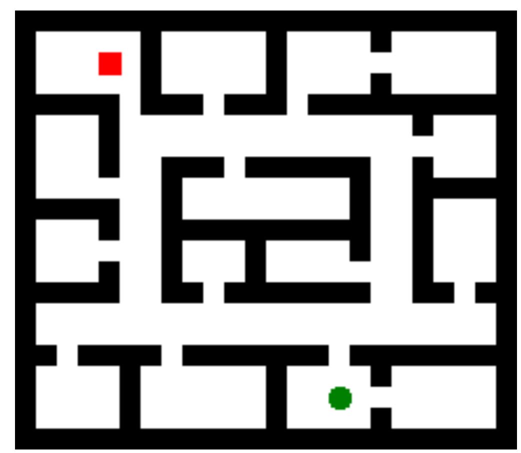


[Glatt et al, 2016]

Next steps

- Investigate Policy Reuse for DRL
- Challenge
  - Network structure for autonomous policy selection





[Fernández & Veloso, 2006]

Going further

- Gain more familiarity with knowledge transfer via options
  - Option-Critic architecture
  - Intra-option policy learning
  - Hierarchical Reinforcement Learning

- Extend Option-Critic architecture to autonomously extend option library
  - Flexible Option-Critic architecture
  - Multi-task learning

Wrap up

Summary

- Facts
  - DRL still slow and sample inefficient
  - TL has great potential to improve DRL
    - Especially options in combination with clever network architectures

- Intended research goal
  - Develop framework that autonomously learns stochastic options and builds an option library for knowledge transfer across tasks

Acknowledgements

- People
  - Anna Helena Reali Costa (Supervisor)
  - Leno Felipe (Lab peer)

- Organizations
  - CAPES (D.Sc. research scholarship)
  - Superintendência de Tecnologia da Informação (HPC resources)
  - Google (Research Award Latin-America 2015 + 2016)
  - Nvidia (GPU Grant Program 2016)

Literature

- References

- **[Bacon et al, 2017]** Bacon, P.-L., Harb, J. and Precup, D. "The option-critic architecture." Accepted at 31st AAAI, 2017.
- **[Fernández & Veloso, 2006]** Fernández, F., and Veloso,M. "Probabilistic policy reuse in a reinforcement learning agent." In *Proc. 5th AAMAS*, pp. 720-727. ACM, 2006.
- **[Glatt et al, 2016]** Glatt, R., Silva, F. L. d., and Costa, A. H. R. "Towards Knowledge Transfer in Deep Reinforcement Learning" In *Proc. 5th BRACIS*, IEEE, 2016.
- **[Levine et al, 2016]** Levine, S., Finn, C., Darrell, T., and Abbeel, P. "End-to-end training of deep visuomotor policies." *Journal of Machine Learning Research* 17, no. 39 (2016): 1-40.
- **[Mirowski et al, 2016]** Mirowski, P., Pascanu, R., Viola, F., Soyer, H., Ballard, A., Banino, A., Denil, M. et al. "Learning to navigate in complex environments." *arXiv preprint arXiv:1611.03673* (2016).
- **[Mnih et al, 2015]** Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A., Veness, J., Bellemare, M. G., Graves, A., et al. "Human-level control through deep reinforcement learning." *Nature* 518, no. 7540 (2015): 529-533.
- **[Pan & Yang, 2010]** Pan, S. J., and Yang, Q. "A survey on transfer learning." *IEEE Transactions on knowledge and data engineering* 22, no. 10 (2010): 1345-1359.
- **[Silva et al, 2017]** Silva, F. L. d., Glatt, R., and Costa, A. H. R. "Simultaneously Learning and Advising in Multiagent Reinforcement Learning" Accepted at *16th AAMAS*, 2017.
- **[Sutton & Barto, 1998]** Sutton, R., and Barto, A. *Introduction to reinforcement learning*. Vol. 135. Cambridge: MIT Press, 1998.
- **[Sutton et al, 1999]** Sutton, R. S, Precup, D., and Singh, S. "Between MDPs and semi-MDPs: A framework for temporal abstraction in reinforcement learning." *Artificial Intelligence*, 112(1), August 1999.
- **[Tayler & Stone, 2009]** Taylor, M., and Stone, P. "Transfer learning for reinforcement learning domains: A survey." *Journal of Machine Learning Research* 10, no. Jul (2009): 1633-1685.
- **[Tessler et al, 2016]** Tessler, C., Givony, S., Zahavy, T., Mankowitz, D., and Mannor, S. "A Deep Hierarchical Approach to Lifelong Learning in Minecraft." *arXiv preprint arXiv:1604.07255* (2016).

Improving Deep Reinforcement Learning with knowledge transfer

GAME OVER

INSERT COINS
TO CONTINUE

http://www.cowhi.org

ruben.glatt@usp.br