Parallel Transfer Learning in Multi-Agent Systems: What, when and how to transfer?

Adam Taylor, Ivana Dusparic, Maxime Guériau and Siobhan Clarke

School of Computer Science and Statistics,
Trinity College Dublin, Ireland

tayloral@tcd.ie, *ivana.dusparic@scss.tcd.ie*, *maxime.gueriau@scss.tcd.ie*, *siobhan.clarke@scss.tcd.ie*

@ivanadusparic @maximegueriau @SClarkeTCD

2019 International Joint Conference on Neural Networks
Budapest, Hungary
July 17th 2019

Fondúireacht Eolaíochta Éireann
Dá bhfuil romhainn
Science Foundation Ireland
For what's next

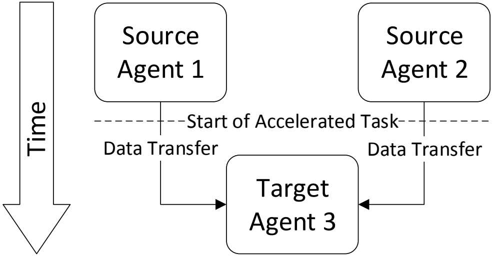
IRISH RESEARCH COUNCIL
An Chomhairle um Thaighde in Éirinn

Trinity College Dublin
Coláiste na Tríonóide, Baile Átha Cliath
The University of Dublin

Reinforcement Learning (RL) in Multi-Agent Systems (MAS)

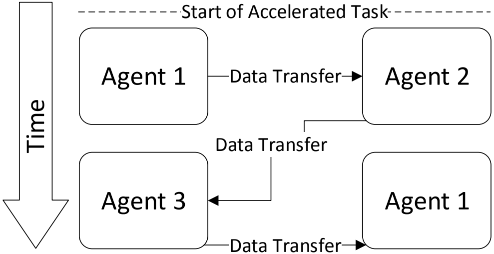
- Learning takes time:
  - ★ Performance of RL-based systems depends on the speed of learning
  - ✘ More agents interaction implies slower learning
- Real-world problems are dynamic and non-stationary:
  - ✘ Once learning is complete, changes can invalidate the learnt policy, so re-learning is required
  - ✓ Additional knowledge can be used in MAS [1]

Transfer Learning (TL)



- ✓ TL [1] was applied in MAS [2] to transfer learnt policies offline and for multi-agent coordination [3]
- ✗ Agents require pre-training before transferring fully learnt policies, applied to similar tasks

Parallel Transfer Learning (PTL)



✓ PTL enables online knowledge transfers [4] and can be used for multi-objective problems [5]

★ PTL can result in higher risk of negative transfers (*e.g.* learning wrong behaviours or un-learning correct ones)

Challenges for PTL

- ★ How to minimise bad transfers?
- ★ How to reuse knowledge for different tasks?
- ★ What information should be transferred?
- ★ At which frequency?
- ★ How to integrate knowledge in the locally learnt behaviours without over-writing correct or converged tasks?

Context and objectives

- Reinforcement Learning in Multi-Agent Systems
- Transfer Learning
- Parallel Transfer Learning
- Challenges

Parallel Transfer Learning

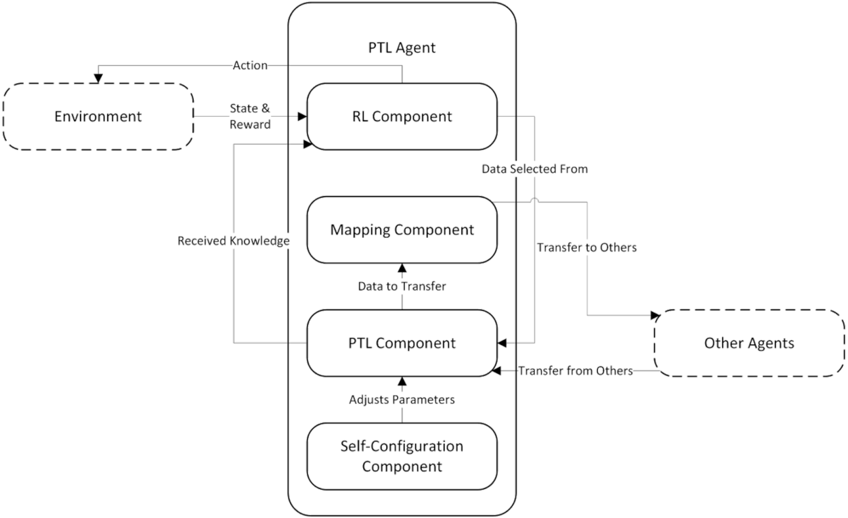
- Components
- Data selection methods
- Merge methods

Evaluation

- Scenarios
- Impact of PTL parameters
- Effects of learning time

Conclusions and future work

PTL components



Data selection methods (SM)

What and when to transfer?

- **Most visited states**: provides a frequent estimate of the important states
- **Converged states**: provides the highest confidence information
- **Best/Worst states**: share states that are important based on their value
- **Visit threshold**: enables a greater degree of confidence
- **Greatest change**: selects the state in which the most significant learning has occurred since the last transfer

Merge methods (MM)

How and when to incorporate received knowledge?

- We use a decaying linear combination of received and local knowledge:

$$scaledRecieved = Recieved\ Q(S_t, A_t) \times (numCV - currVC)$$

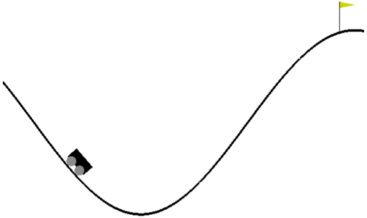
$$scaledLocal = Q(S_t, A_t) \times currVC$$

$$Q(S_t, A_t) = \frac{scaledRecieved + scaledLocal}{numCV}$$

- where *currVC* is the current visits count for that state and *numCV* is the number of confidence visits required

Evaluation: scenarios

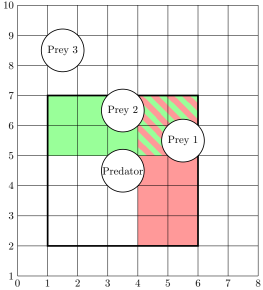
Mountain Car



Parameter	**Value/Range**
State: Position	$[-1.2, 0.6]$ **(goal at 0.6)**
State: Velocity	$[-0.07, 0.07]$
Actions	Left -1, Neutral 0 or Right 1
Reward	10 if at the goal, -1 otherwise
Q-learning	$\alpha = 0.1$ and $\gamma = 0.01$

Cooperative predator-prey pursuit

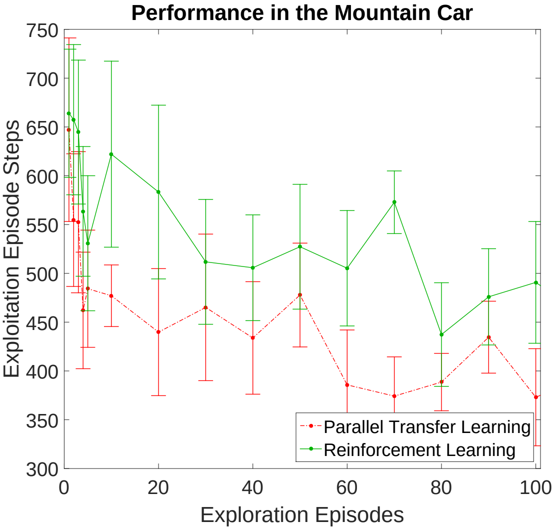
- 2^4 states (4 directions and binary detection)
- 5 actions
- Q-learning: $\alpha = 0.95$ and $\gamma = 0.2$



Impact of PTL parameters (SM and MM) in Mountain Car

Selection Method (SM)	**Confidence (MM)**	**Source**	**Target**	**Difference**
Best States	***1**	569.45	382.29	187.16
	4	616.59	511.93	104.66
	7	476.96	451.52	25.44
	10	529.69	526.05	3.64
Converged	1	524.04	525.52	−1.48
	***4**	660.17	397.79	262.38
	7	401.83	514.88	−113.05
	10	527.03	491.44	35.59
Greatest Change	1	416.9	514.75	−97.85
	***4**	475.32	483.81	−8.49
	7	518.3	539.8	−21.5
	10	500.17	492.49	7.68
Visit threshold	1	496.63	502.91	−6.28
	4	418.08	438.75	−20.67
	7	534.75	511.33	23.42
	***10**	508.7	444.08	64.62
Most Visits	1	438.3	514.18	−75.88
	4	364.85	441.4	−76.55
	***7**	472.09	523.09	−51
	10	403.16	559.53	−156.37

Effects of learning time in Mountain Car



- SM: Best states; MM confidence: 1
- ✓ PTL performs better than RL from 10^{th} learning episode
- ✓ PTL converges around episode 60 vs. 80 for RL
- ✓ Final performance is improved by 27% after 150 exploration episodes

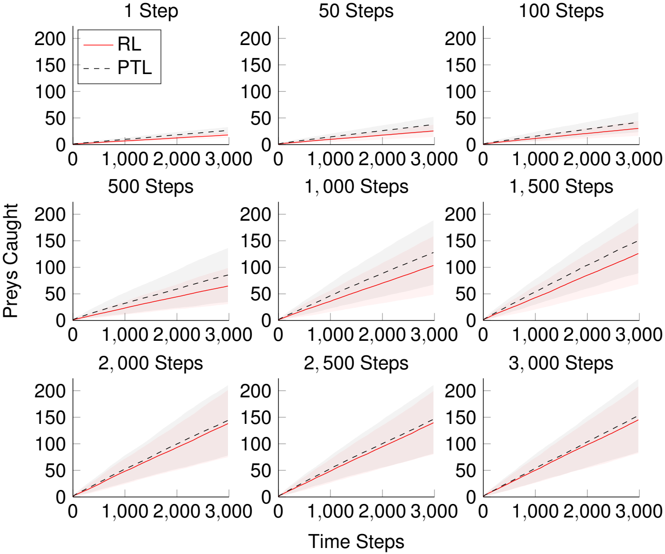
Effects of learning time in Cooperative predator-prey pursuit

- SM: Greatest Change;
  MM set as a probability: 0.6

✘ PTL and RL perform poorly in the early training (as knowledge is missing)

✓ PTL catches more preys between episodes 500 and 1500 (around 20% more)

✓ RL and PTL converge to similar level of performance



Conclusions and future work

- ✓ PTL enables online knowledge sharing between RL agents in MAS
- ✓ It increases the use of early acquired knowledge
- ✓ PTL speeds up the overall learning process
- ✗ Like RL and TL, PTL is sensitive to parameter selection, which could be done autonomously: *i.e.*, by learning (or sharing) transfer parameters

Thank you for your attention

Maxime GUÉRIAU
PhD., Research Fellow
School of Computer Science and Statistics
Trinity College Dublin
Ireland

maxime.gueriau@scss.tcd.ie
`maxime.gueriau.fr`
@maximegueriau

Fondúireacht Eolaíochta Éireann
Dá bhfuil romhainn
Science Foundation Ireland
For what's next

IRISH RESEARCH COUNCIL
An Chomhairle um Thaighde in Éirinn

Trinity College Dublin
Coláiste na Tríonóide, Baile Átha Cliath
The University of Dublin

References I

[1] Matthew E Taylor and Peter Stone. Transfer Learning for Reinforcement Learning Domains: A Survey. The Journal of Machine Learning Research, 10:1633–1685, 2009.

[2] Georgios Boutsioukis, Ioannis Partalas, and Ioannis Vlahavas. Transfer learning in multi-agent reinforcement learning domains. In 9th European Workshop on Reinforcement Learning, 2011.

[3] Peter Vrancx, Yann-Michaël De Hauwere, and Ann Nowé. Transfer learning for multi-agent coordination. In ICAART 2011 - Proceedings of the 3rd International Conference on Agents and Artificial Intelligence, Volume 2 - Agents, Rome, Italy, January 28-30, 2011, pages 263–272, 2011.

[4] A. Taylor, I. Dusparic, E. Galván-López, S. Clarke, and V. Cahill. Transfer learning in multi-agent systems through parallel transfer. In Theoretically Grounded Transfer Learning Workshop at the International Conference on Machine Learning (ICML), July 2013.

[5] A. Taylor, I. Dusparic, E. Galván-López, S. Clarke, and V. Cahill. Accelerating learning in multi-objective systems through transfer learning. In 2014 International Joint Conference on Neural Networks (IJCNN), pages 2298–2305, July 2014. doi: 10.1109/IJCNN.2014.6889438.