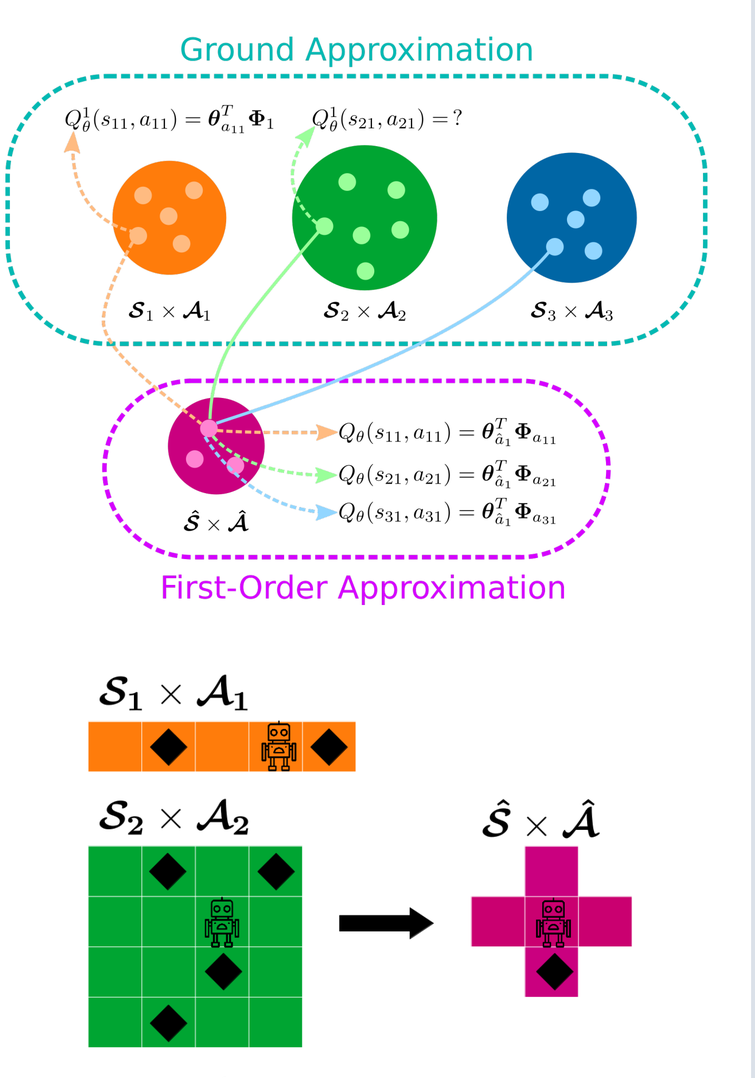
First-Order Function Approximation for Transfer Learning in Relational MDPs

Jun Hao Alvin Ng, Ronald P. A. Petrick
Edinburgh Centre for Robotics
Department of Computer Science, Heriot-Watt University
School of Informatics, University of Edinburgh
Alvin.Ng@hw.ac.uk, R.Petrick@hw.ac.uk

Overview



Mixed Approximation

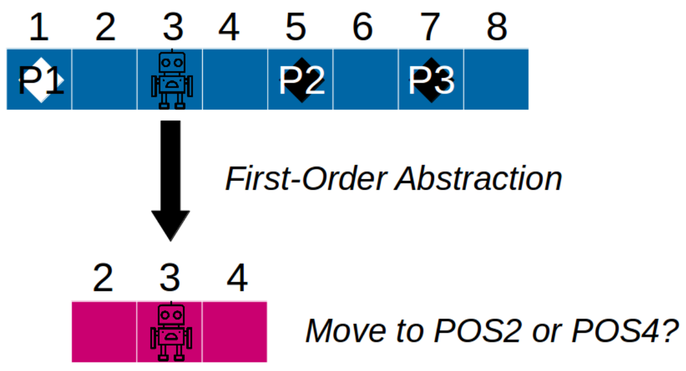
1. Combines first-order and ground approximation
2. Trained "independently" and concurrently
3. Achieves generalisation and finer granularity
4. Ground approximation serves as implicit tiebreaker

Policy with Mixed Approximation

$$Q^{sum}_\theta(s, a) = Q^{gnd}_\theta(s, a) + Q^{fo}_\theta(s, a)$$

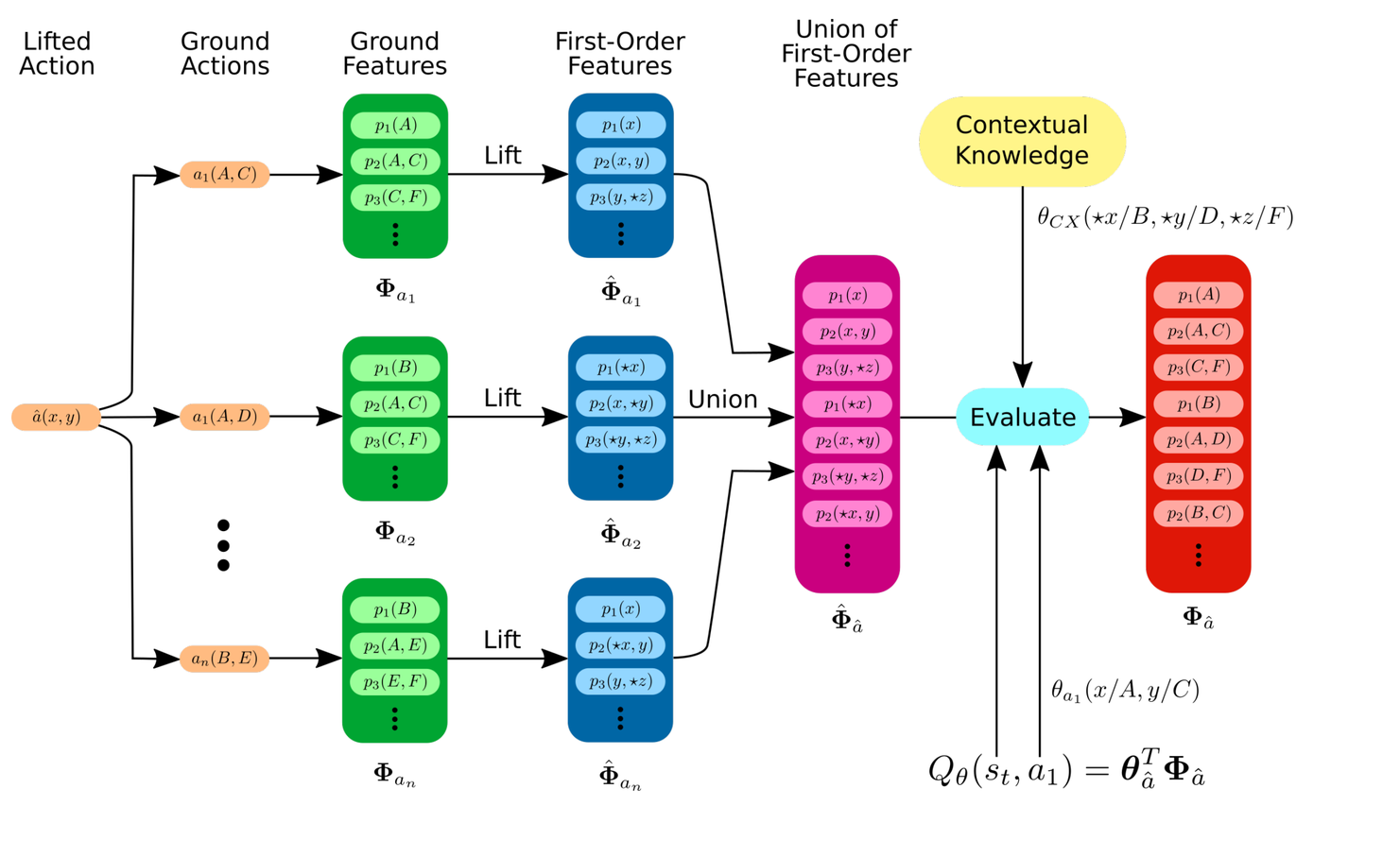
$$Q^{switch}_\theta(s, a) = \begin{cases} Q^{fo}_\theta(s, a), & \text{if } episode \leq E \\ Q^{gnd}_\theta(s, a), & \text{otherwise} \end{cases}$$

Plateaus



$\phi_f = \texttt{robot\_at}(pos) \wedge \texttt{ADJACENT}(pos, \star pos) \wedge \texttt{POI\_AT}(\star poi, \star pos)$

First-Order Approximation



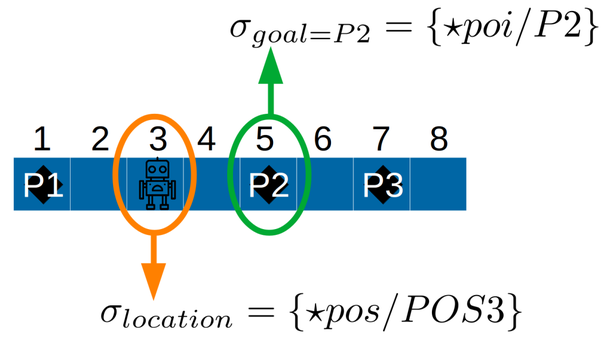
Contextual Grounding of Free Variables

Goal Context

- Use objects in fluents representing (unachieved) goals
- Fluents are determined from terminal state or reward function

Location Context

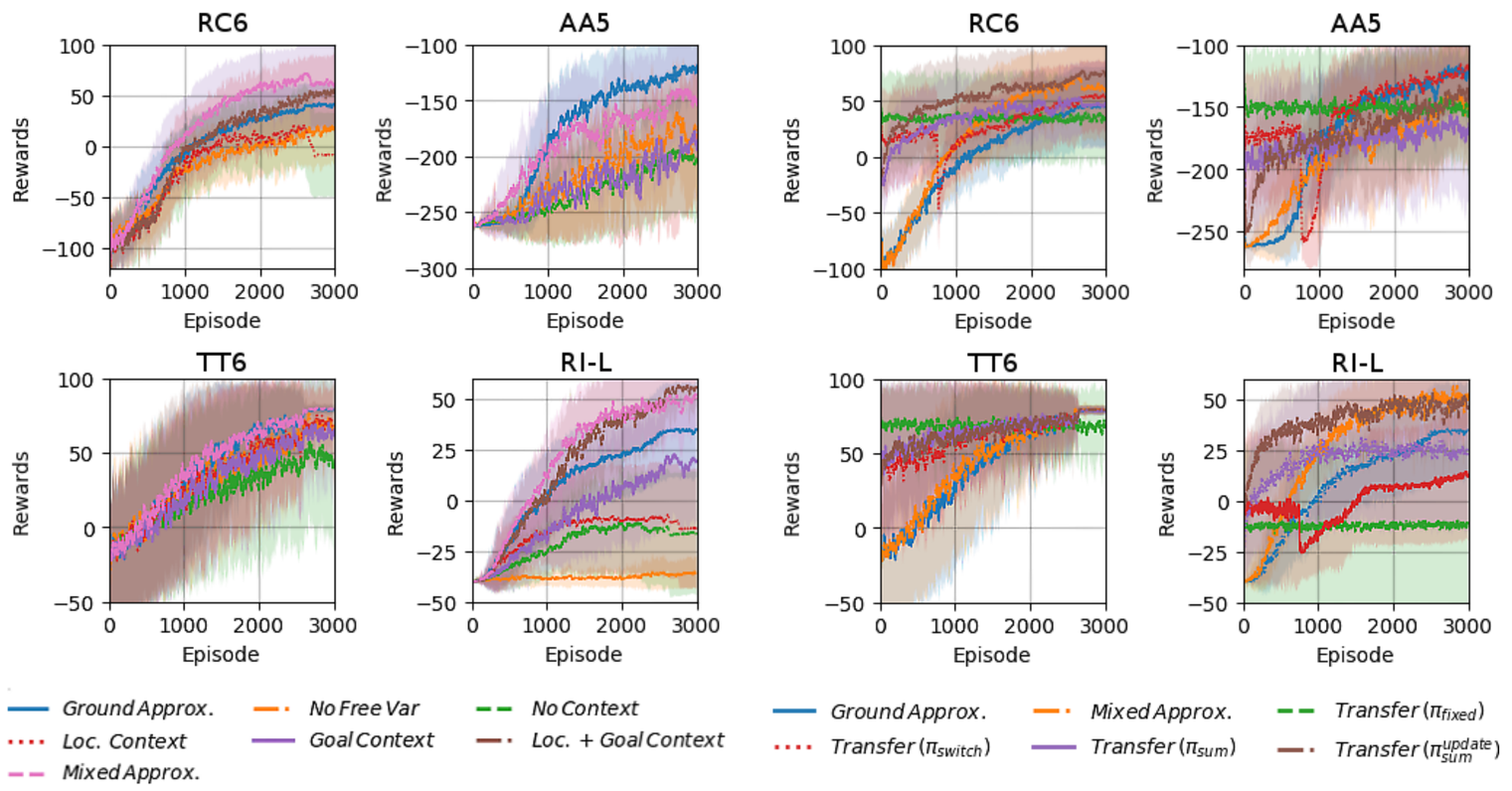
- Use location of agent
- Assume agent can only interact with objects in its proximity



Experimental Results

Ablation Study

Transfer Learning



RC6: Recon 6, **AA5**: Academic Advising 5, **TT6**: Triangle Tireworld 6, **RI-L**: Robot-Inspection (large-scale)