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

Engaging Content
Engaging People

3rd International Tutorial & Workshop on Interactive Adaptive Learning, Würzburg, 2019

Combating Stagnation in Reinforcement Learning Through 'Guided Learning' With 'Taught-Response Memory'

Keith Tunstead, Joeran Beel

Introduction

One of the primary problems with training any kind of modern AI in a Reinforcement Learning environment is stagnation. Stagnation occurs when the agent ceases to make progress in solving the current task prior to either the goal or the agents maximum effectiveness being reached. The reduction of stagnation is an important topic for reducing training times and increasing overall performance in cases where training times are limited.

Current AI implementations place a large focus on the idea of fully automated learning however, as we push ever closer to more general purpose AI, direct human intervention will likely become increasingly important.

Fully automated learning is very analogous to leaving a child to his/her own devices and expecting them to reason about the world on their own. There is only a finite amount of progress that that child would be able to make, primarily on a trial and error basis.

Guided Learning

1) Allows an agent to 'ask for help' when it encounters stagnation.

2) Guidance is then optionally given by a supervisor, either an expert agent or human.

3) The guidance is then encoded as a separate neural network, a Taught Response Memory (TRM), that is independent of the underlying Reinforcement Learning algorithm.

4) When the agent encounters a 'similar' situation as that of when it received guidance it can then 'recall' the TRM.

5) Each TRM is plastic and can vary over time. This allows a TRM to tend towards a more optimal solution for a single stimulus or towards its applicability, more generally, to other stimuli (TRM adaptation).

- Using neural networks to implement TRMs presents the problem of multiple actions conflicting with each other. Single Action TRMs (SATRMs) are a natural solution to this problem.

- Depending on the encoding of the stimulus, TRMs can be applied not only to different levels of a game but to other games entirely, provided such games are similar (i.e. other platform games).

Future Work

1) Building Guided Learning using more state of the art Reinforcement Learning algorithms.

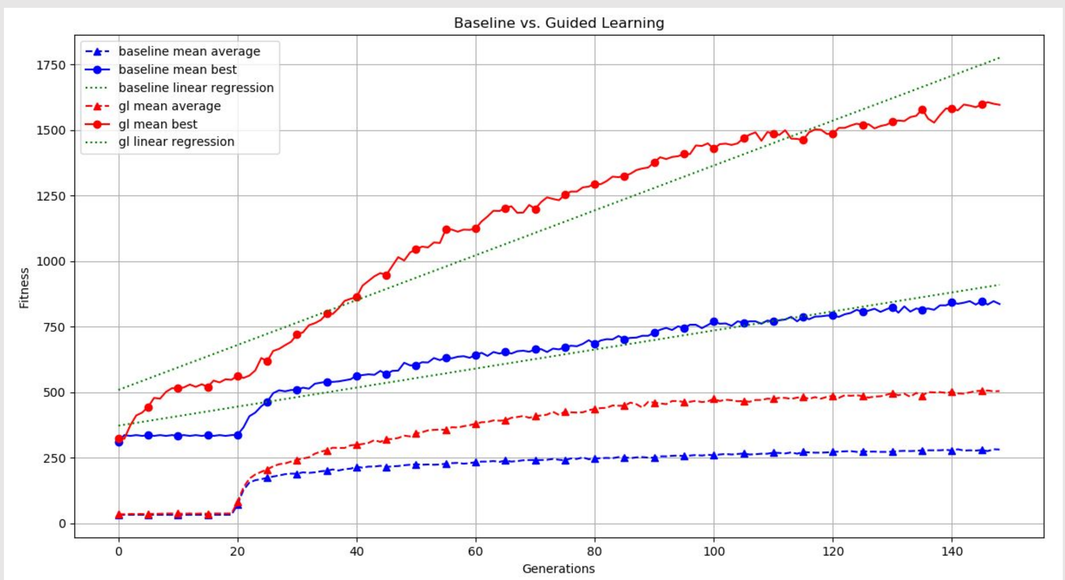
2) Using a more generalized encoding of the stimulus to allow TRMs to be re-used more readily.

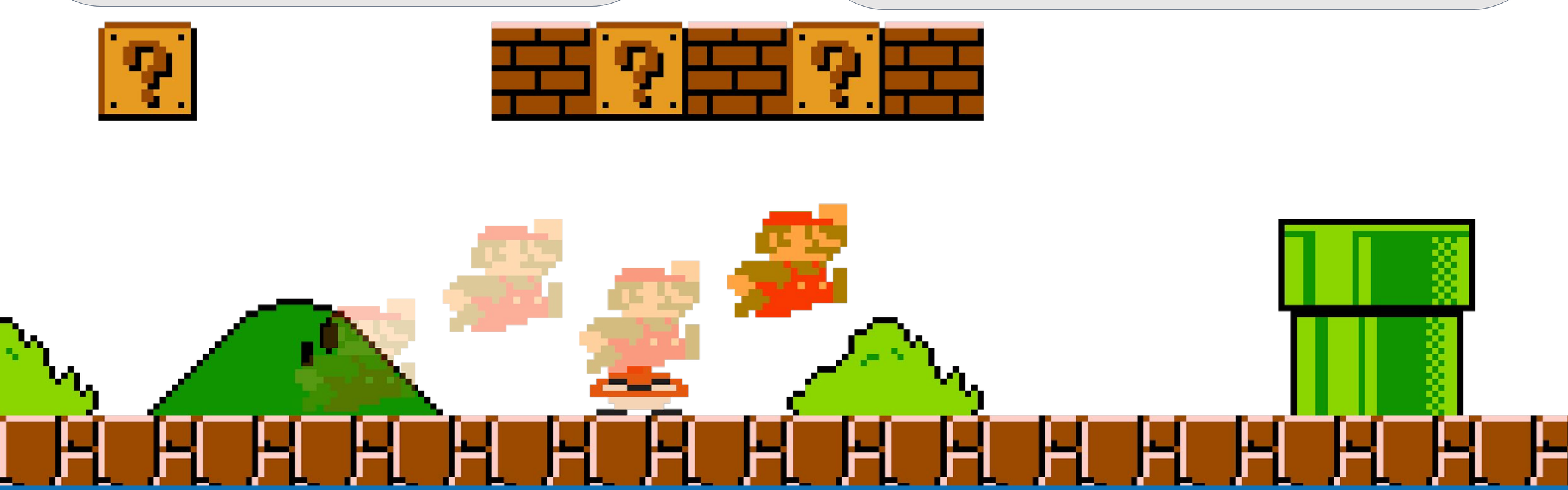
3) Implementing TRM adaptation.

4) Taking advantage of poorly performing TRMs as a method of showing the agent what *not* to do.

5) Run-time optimization, currently ~2x baseline.

Results





https://scss.tcd.ie/joeran.beel/ tunstek@tcd.ie