

Advanced topics in Deep Reinforcement Learning  
Week 7 slides

# When Do Transformers Shine in RL?

## Decoupling Memory from Credit Assignment

Thorben Klamt - 02.05.2024

Advanced Topics in Reinforcement Learning, Theresa Eimer, Prof. Dr. rer. nat. Marius Lindauer,  
Gottfried Wilhelm Leibniz University Hannover

# When Do Transformers Shine in RL? Decoupling Memory from Credit Assignment (2024)

- Two major long-term dependency tasks in RL
  - learning effective representations of (past and present) observations
  - determining how actions influence future returns
- Strongly enhance the memory capability of RL algorithms
- Transformer-based agents are more sample-efficient in long-term memory high-dimensional tasks

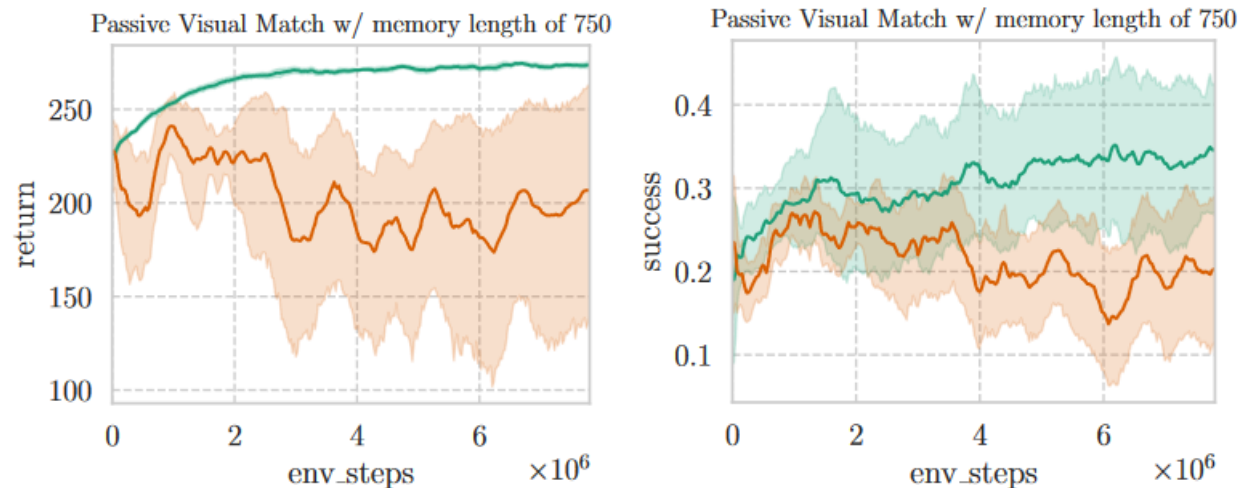


Fig.1: In Passive Visual Match with long memory length, Transformer-based agents show a slight advantage over LSTM-based agents in the memory capability (as measured by success rate), while being much higher in total rewards (as measured by return)

# When Do Transformers Shine in RL? Decoupling Memory from Credit Assignment (2024)

- Transformer-based agents are more sample-efficient in long-term memory high-dimensional tasks
- But without (consistent) improvements in temporal credit assignment capabilities
- Increasing the number of layers (and heads) in Transformers aids in temporal, but not long-term credit assignment

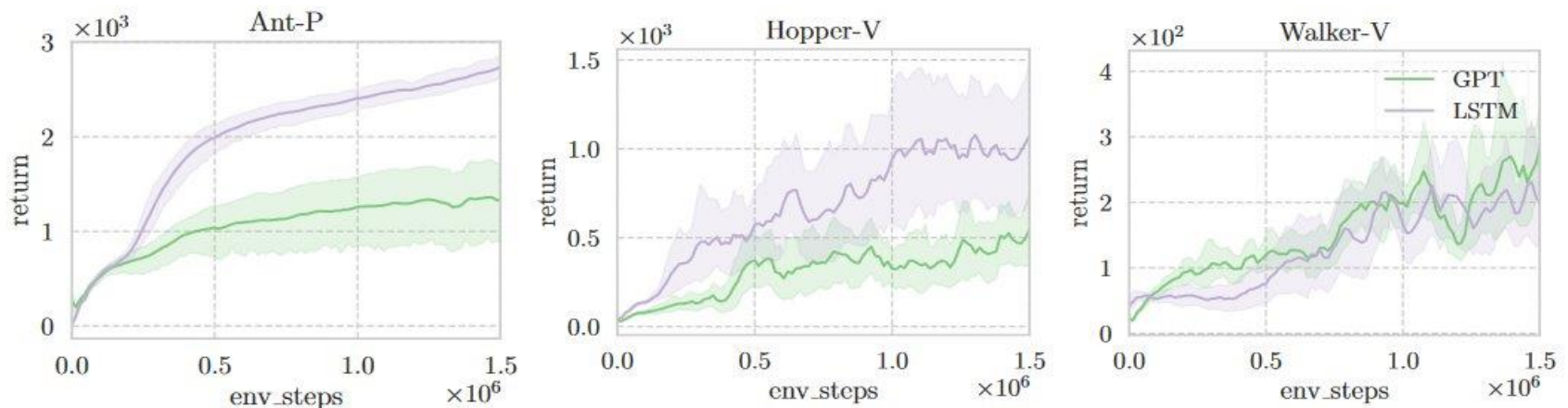


Fig.2: Transformer-based RL is sample-inefficient compared to LSTM-based RL in most PyBullet occlusion tasks. The tag -P refers to occluded position, while -V refers to occluded velocity. Velocities or positions can be approximately deduced or retrieved from one another, requiring short-term memory.