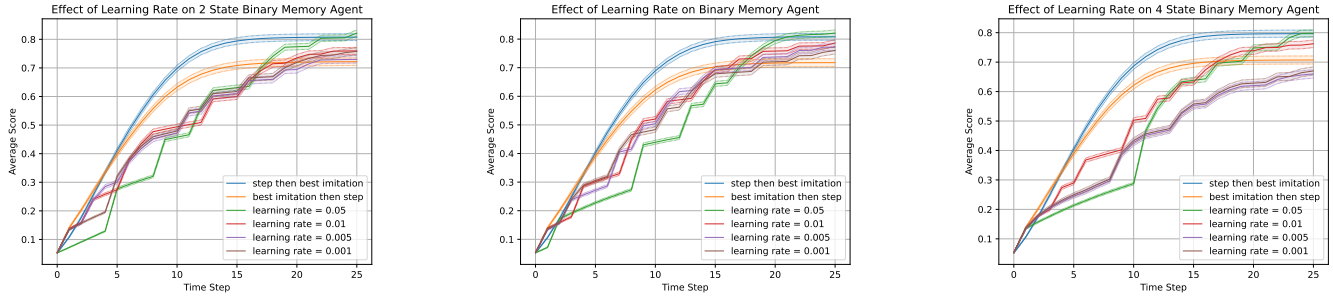
Notes

Constants

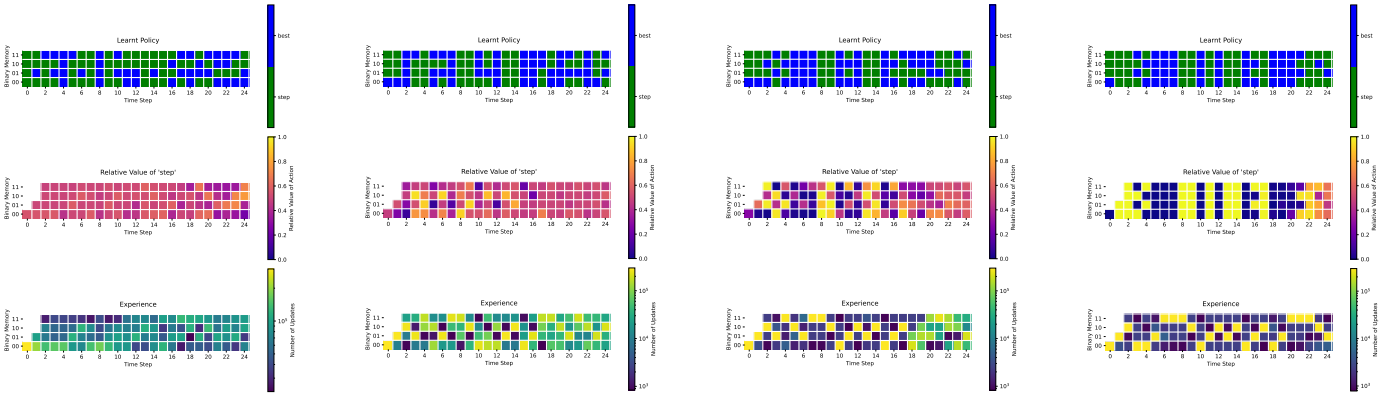
- deadline of 25
- regular graph
  - 30 nodes
  - 5 degrees per node
- trained over 10,000 episodes
- tested over 1,000 episodes

Figures



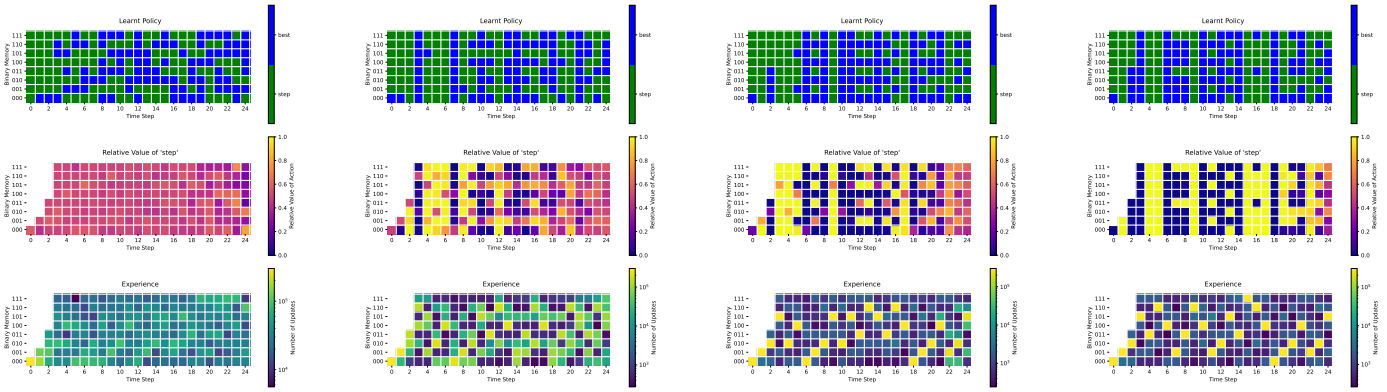
(a) Memory Size of 2 bits (b) Memory Size of 3 bits (c) Memory Size of 4 bits

Figure 1: The Average and 95% confidence of the Average Node Score at each time step.



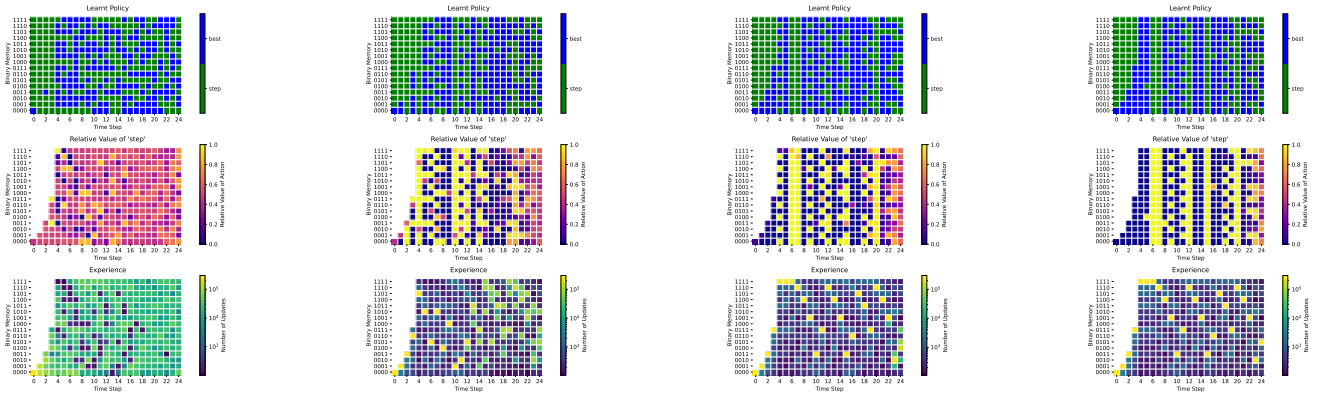
(a) Learning Rate of 0.05 (b) Learning Rate of 0.01 (c) Learning Rate of 0.005 (d) Learning Rate of 0.001

Figure 2: The policy, experiance, and relative value of the action 'step' over best for a binary agents with a memory size of two bits and varying learning rates. The most recent timestep is the least significant bit of the memory. The action 'best' is represented by a '1' and 'step' by a '0'.



(a) Learning Rate of 0.05 (b) Learning Rate of 0.01 (c) Learning Rate of 0.005 (d) Learning Rate of 0.001

Figure 3: The policy, experiance, and relative value of the action 'step' over best for a binary agents with a memory size of three bits and varying learning rates. The most recent timestep is the least significant bit of the memory. The action 'best' is represented by a '1' and 'step' by a '0'.



(a) Learning Rate of 0.05 (b) Learning Rate of 0.01 (c) Learning Rate of 0.005 (d) Learning Rate of 0.001

Figure 4: The policy, experiance, and relative value of the action 'step' over best for a binary agents with a memory size of four bits and varying learning rates. The most recent timestep is the least significant bit of the memory. The action 'best' is represented by a '1' and 'step' by a '0'.