Optimal Gaze Control Through Reinforcement Learning

Adrian Ionita, Jeremy Wyatt

Background

Biological systems have been known to achieve near optimal decision making when faced with noisy sensory and motor systems. Human ocular fixations have been found to aid current task activity. As acuity exists only in a central location of the eye, the fovea, the remainder surface provides mostly imprecise or uncertain information.

Using a robotics model, we investigate how optimal gaze behaviour can be learned from positive and negative feedback to increase performance on a relevant task.

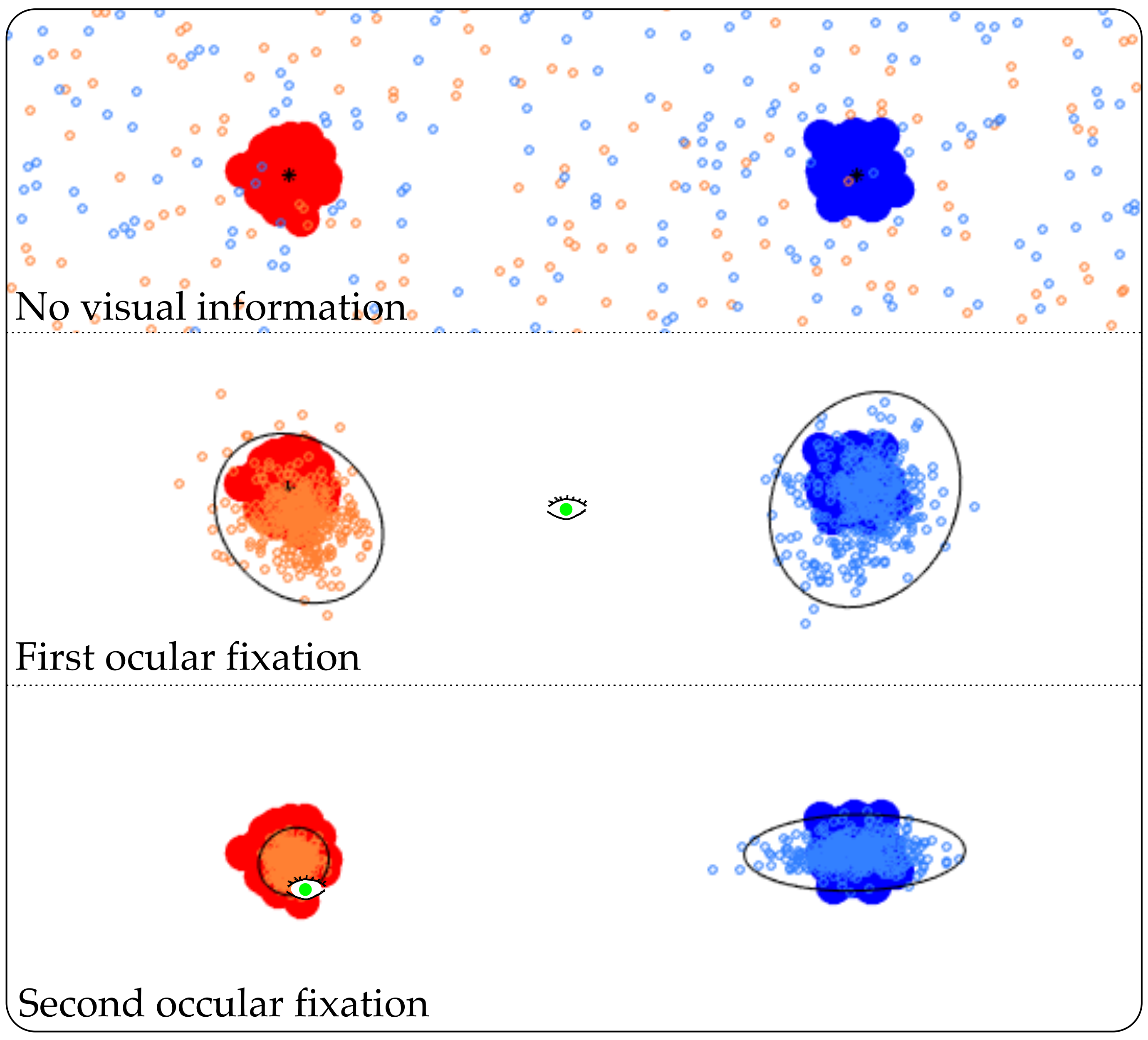
Context

Optimal behaviour is assumed to direct gaze in the direction which maximizes current task reward. We test this hypothesis in an experiment similar to a visual perception task.

The agent is presented with two objects of different reward values. Reward is given when grasping at an object's true location. Consequently, a large penalty is given if the target is missed. The agent can attempt to grasp both objects, but it is allowed a single visual fixation.

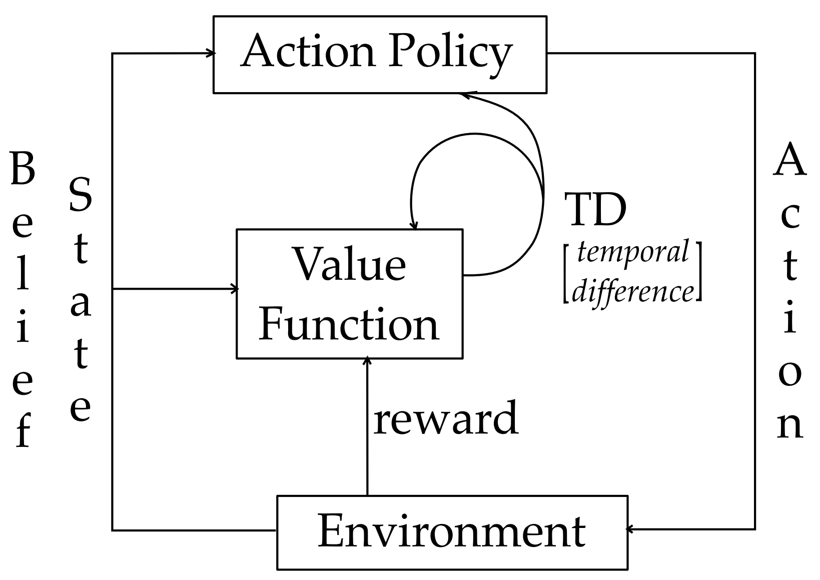
Representing Uncertainty

• A particle filter covers possible object locations (belief state)



Reinforcement Learning

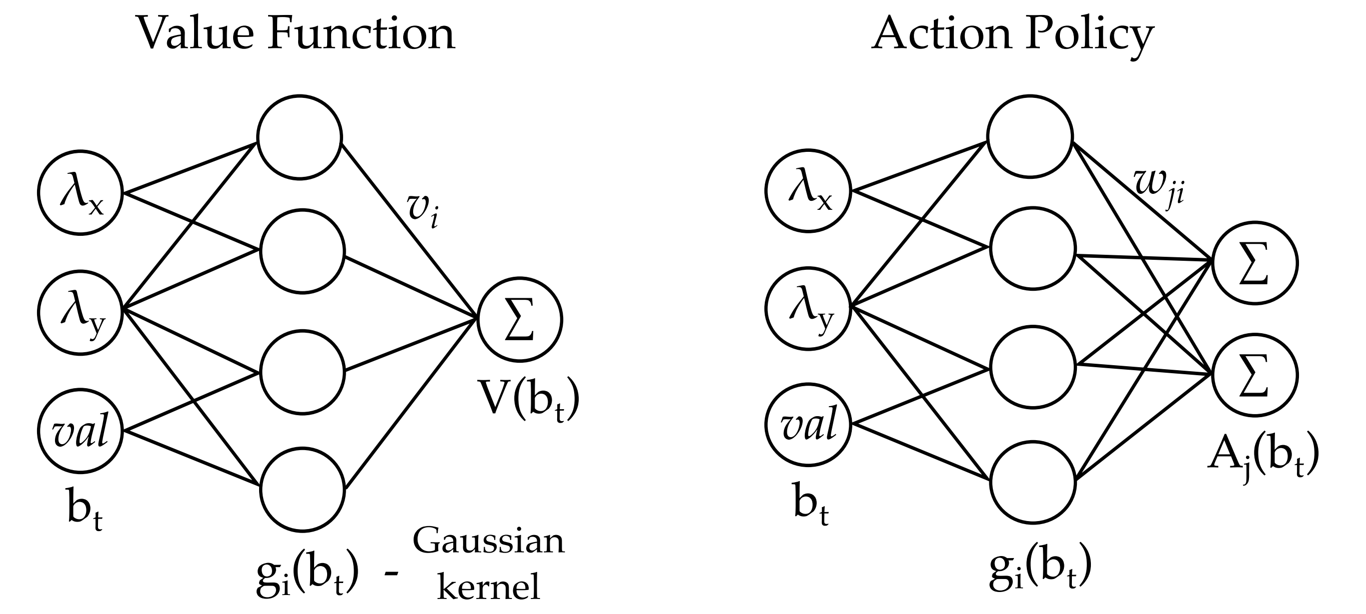
• Agent learns optimal behaviour using an actor-critic model



Actor-Critic Model for Learning

Neural Network Architecture

• Radial basis function networks are used to learn the value function and action policy from reward
• Value function returns expected reward for a belief state (b_t)
• Policy returns action probability for the same belief state (b_t)

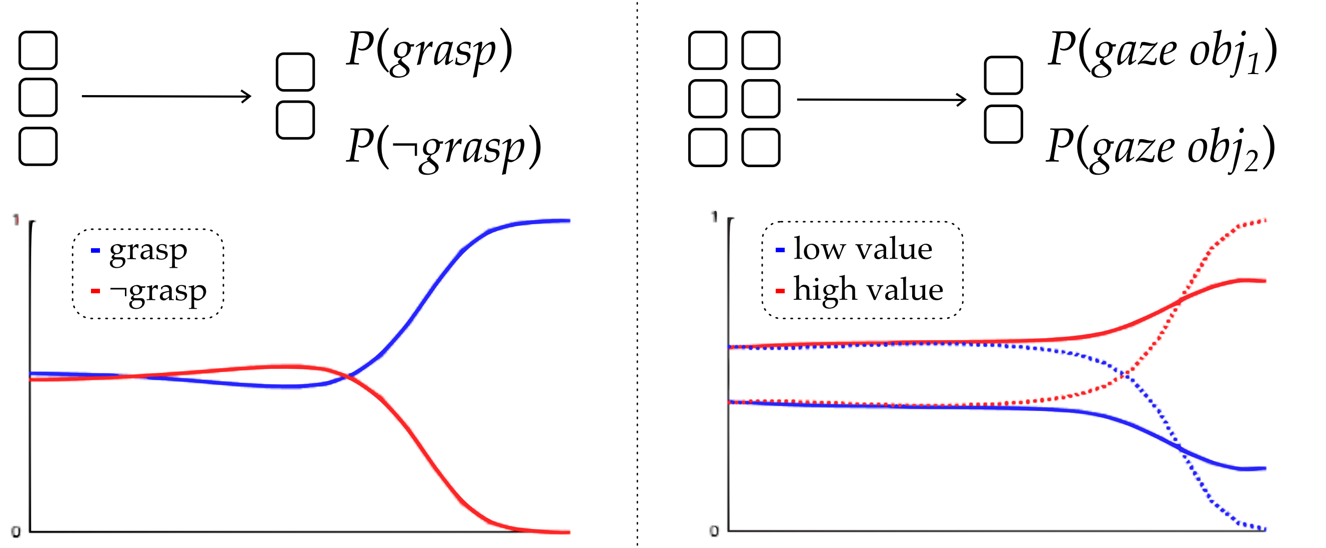


$\Delta v_i = \eta_1 \mathrm{TD}\ g_i(b_t)$ $\Delta w_{ji} = \eta_2 \mathrm{TD}\ g_i(b_t)$

b_t = eigenvalues of particle spread and object value

•Two pairs of networks are used, for grasp and gaze respectively

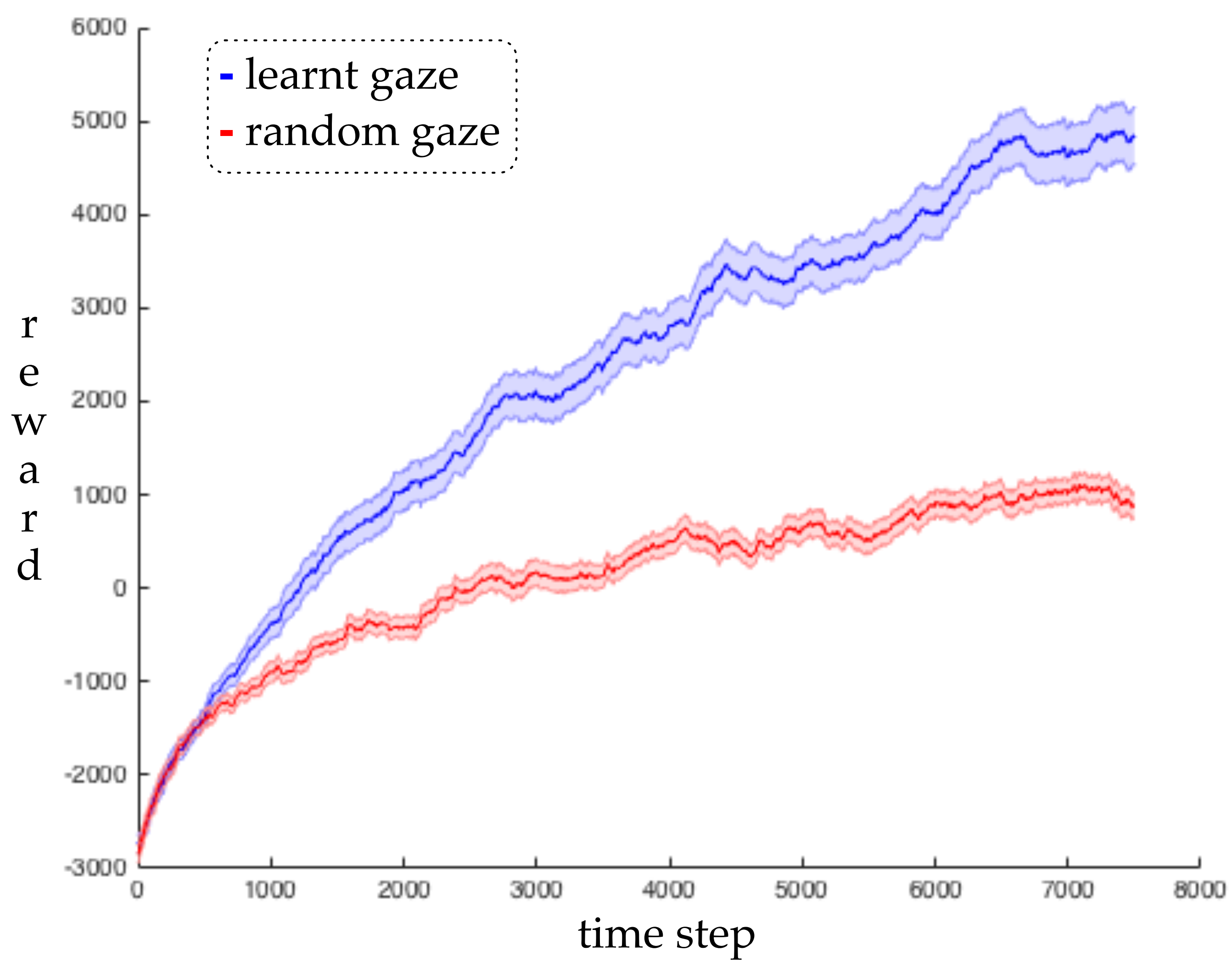
Grasp Policy **Gaze Policy**



reward = from environment
TD = reward + $V(b_{t+1})$ - $V(b_t)$

reward=$\sum[P(grasp_{t+1})-P(grasp_t)]$
TD = reward - $V(b_t)$

Results



References

[1] Nunez-Varela, J. and Wyatt, J. L. (2013) 'Models of gaze control for manipulation tasks', ACM Transactions on Applied Perception, 10(4), pp. 1-22. doi: 10.1145/2536764.2536767.

[2] Rao, R. P. N. (2010). Decision making under uncertainty: A neural model based on partially observable Markov decision processes. Frontiers in Computational Neuroscience, 4, . doi:10.3389/fncom.2010.00146