Optimal Gaze Control Through Reinforcement Learning

Adrian Ionita, Jeremy Wyatt

Background

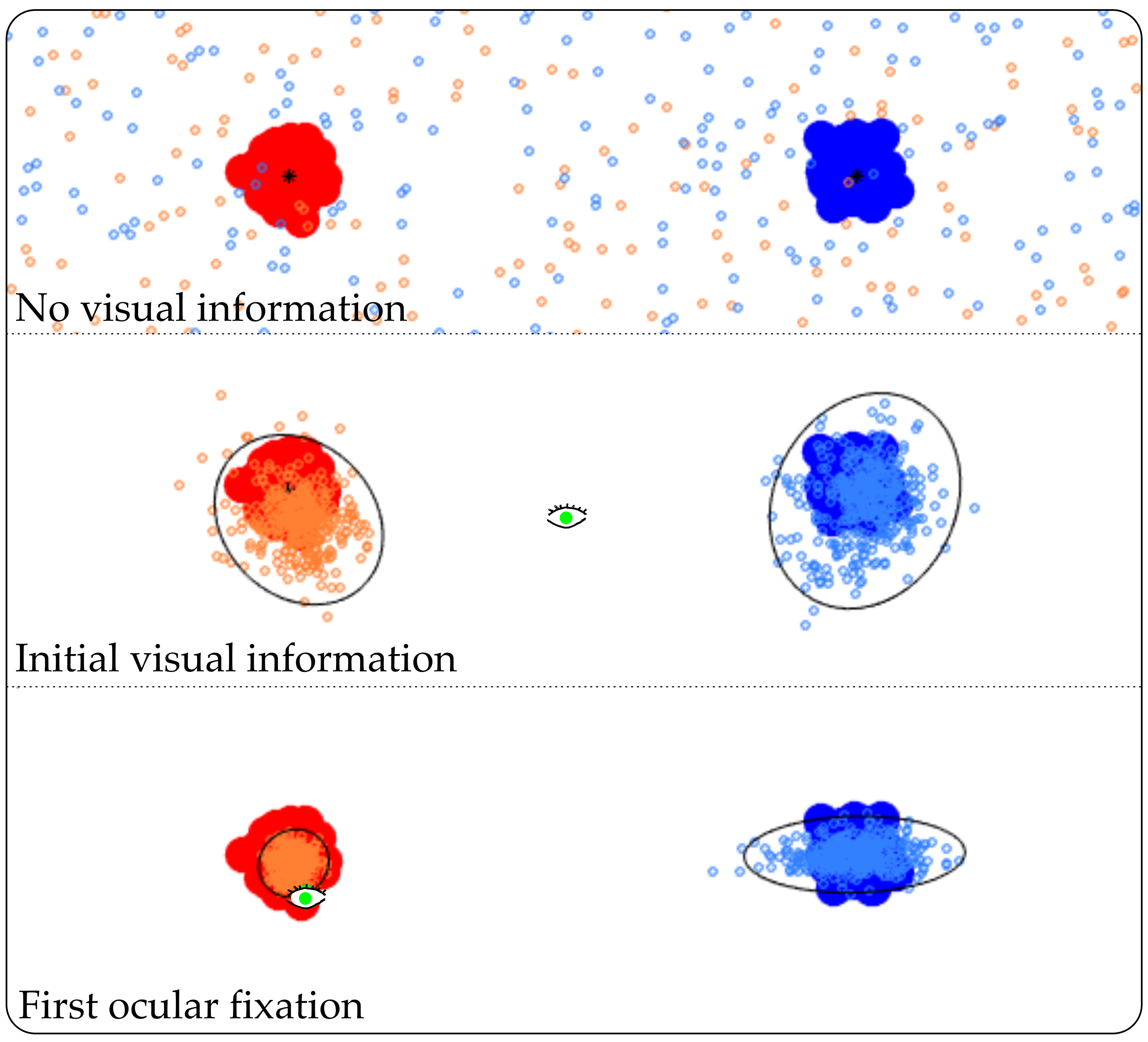
· Biological systems achieve near optimal decisions whilst exposed to noisy sensory and motor systems
· Human ocular fixations aid current task activity
· Acuity exists only in a central location of the eye, the fovea
· We investigate how optimal gaze behaviour can be learned from positive and negative feedback to increase performance on a relevant task

Context

· We assume optimal behaviour directs gaze in the direction which maximises reward
· The hypothesis is tested in an experiment similar to a visual perception task
· Two objects of different reward values are presented
· Reward is only given when the agent correctly grasps at the object's location
· The agent is allowed to grasp either object but is only allowed one visual fixation

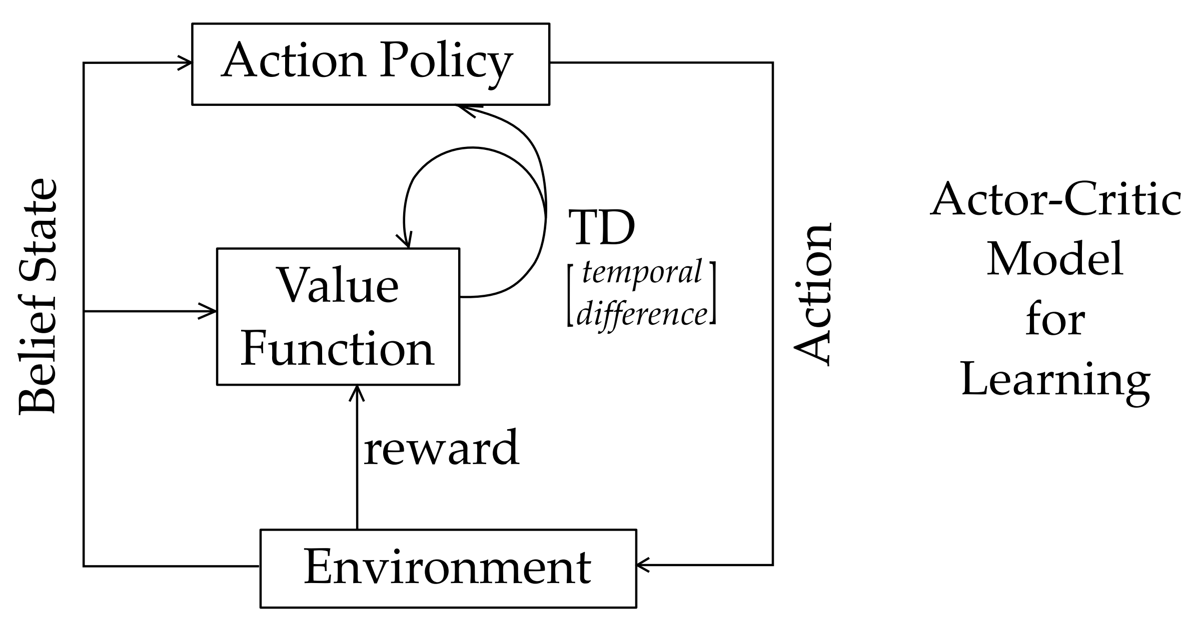
Representing Uncertainty

• A particle filter covers possible object locations (belief state)
• The spread of particles represents the level of uncertainty
• The real object locations (**red** and **blue**) are unkown to the agent



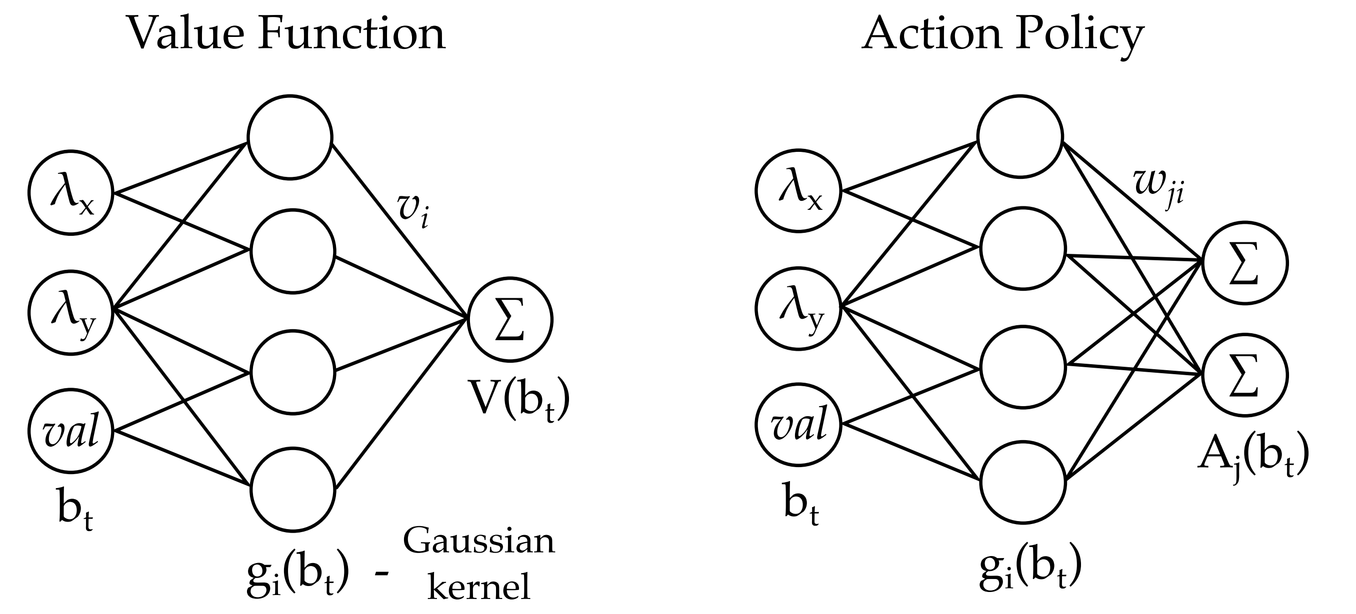
Reinforcement Learning

• Agent learns optimal behaviour using an actor-critic model



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)

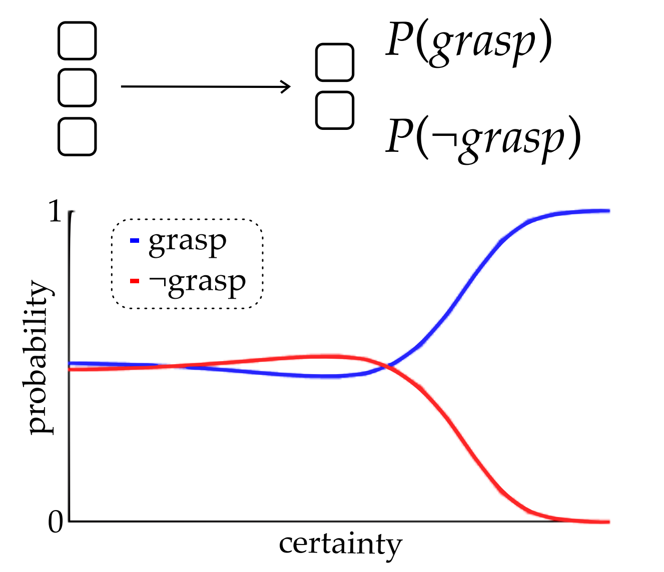


$\Delta v_i = \eta_1 \mathrm{TD}\, g_i(b_t)$ $\Delta w_{ji} = \eta_2 \mathrm{TD}\, g_i(b_t)$

b_t = eigenvalues of particle spread and object value

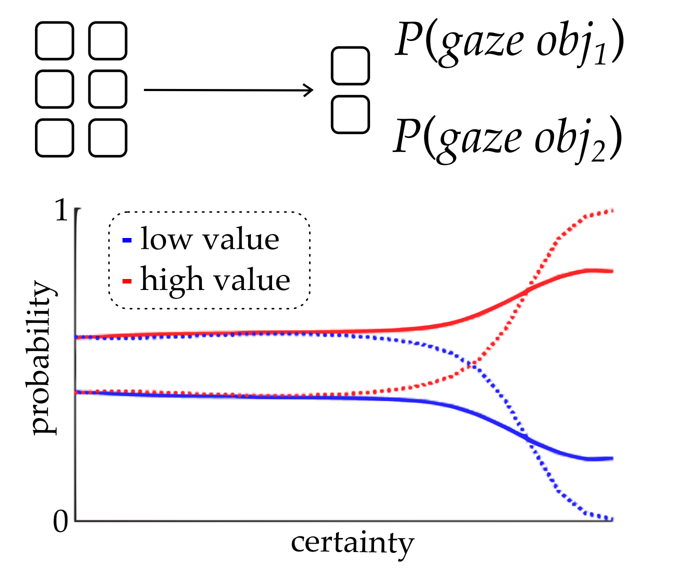
•Two pairs of networks are used, for grasp and gaze respectively

Grasp Policy



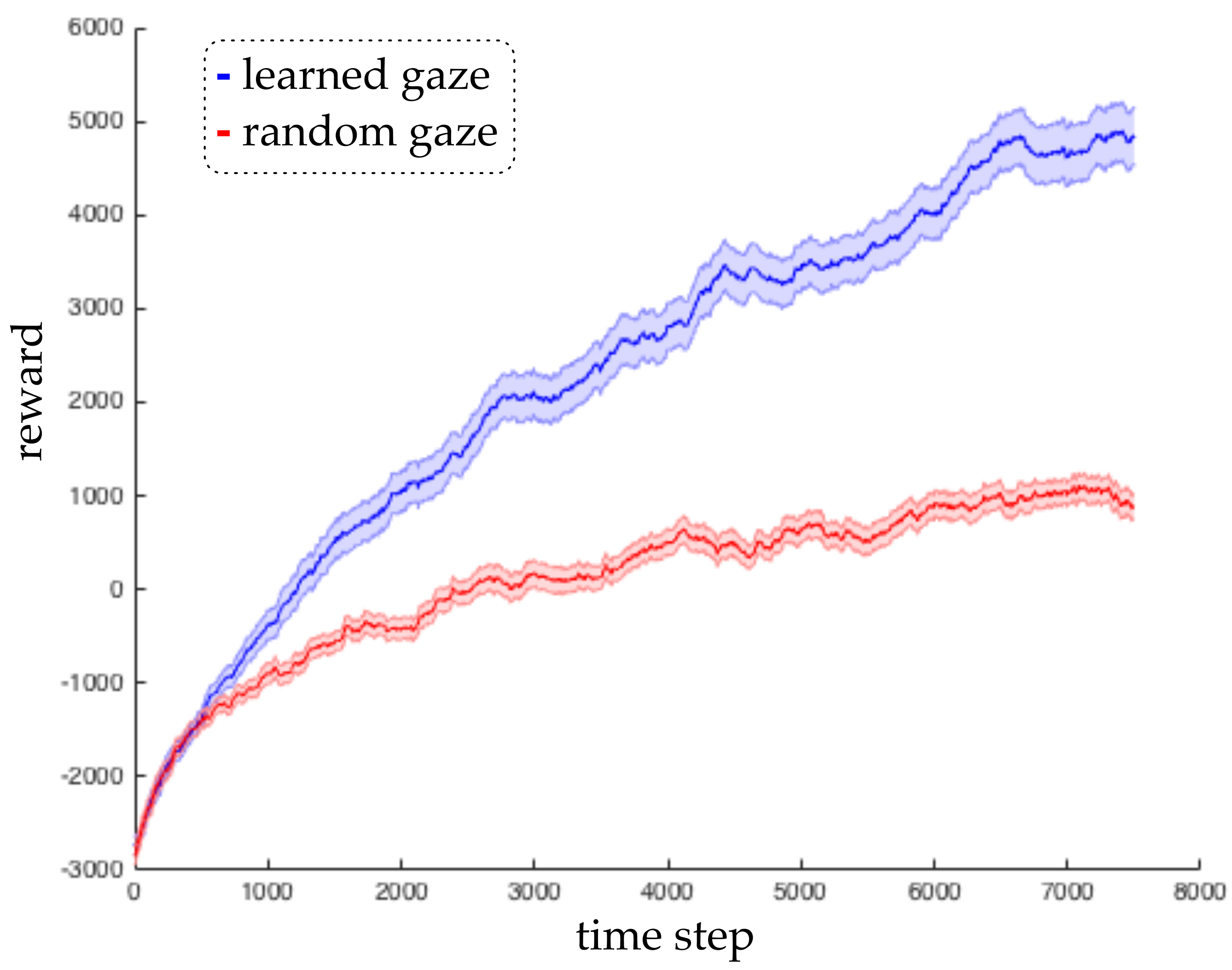
reward = from environment
TD = reward + $V(b_{t+1})$ - $V(b_t)$

Gaze Policy



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