3 Decision Problems 3.1 Utility Theory and 3.2 Decision Networks

$$EU(a|o) = \sum_{s'} (P(s'|a,o)U(s'))$$
$$EU^*(o) = \max_a \{EU(a|o)\}$$
$$VOI(O'|o) = \left(\sum_{o'} (P(o'|o)EU^*(o,o'))\right) - EU^*(o)$$

3.3 Games

$$U([a_1 : p_1; \cdots ; a_n : pn]) = \sum_{i=1}^{n} (p_i U(a_i))$$

Dominant Strategy - strategy that is best response to all possible opposing strategies s_{-i}.
Dominant Strategy Equilibrium - All agents have a dominant strategy.
Nash Equilibrium - No agent can benefit by switching strategies when all other agents keep their strategy.
Logit level k strategy
level 0 choose uniformly
level k assume other agent acting at level k-1 and choose based on logit distribution

$$P(a_i) = e^{\lambda U_i(a_i, s_{-i})}$$

4 Sequential Problems

4.1 Formulation MDP

$$U = \sum_{t=0}^{n-1} (r_t)$$
$$U = \sum_{t=0}^{\infty} (\gamma^t r_t)$$
$$U = \lim_{n\to\infty} \left(\frac{1}{n} \sum_{t=0}^{n} (r_t)\right)$$

4.2 Dynamic Programming
Policy Evaluation

$$U_t^\pi(s) = R(s, \pi(s)) + \gamma \sum_{s'} \left(T(s'|s, \pi(s))U_{t-1}^\pi(s')\right)$$

Policy Iteration

$$\pi_{k+1}(s) = \operatorname{argmax}_a \left\{ R(s,a) + \gamma \sum_{s'} (T(s'|s,a)U^{\pi_k}(s')) \right\}$$

Value Iteration

$$U_{k+1}(s) = \max_a \left\{ R(s,a) + \gamma \sum_{s'} (T(s'|s,a)U_k(s')) \right\}$$
$$\pi(s) = \operatorname{argmax}_a \left\{ R(s,a) + \gamma \sum_{s'} (T(s'|s,a)U^*(s')) \right\}$$
$$U(s) = \max_a \left\{ R(s,a) + \gamma \sum_{s'} (T(s'|s,a)U(s')) \right\}$$