CPSC422 Formula Sheet

1 Value of Information and Control

- Value of information of r.v. X for decision D = E(utility of knowing X) $-$ E(utility of not knowing X)
- Value of control of r.v. X = E(utility of X being a decision variable) $-$ E(utility of X being a r.v.)

2 Markov Models

- Markov assumption: $P(S_{t+1}|S_0...S_t) = P(S_{t+1}|S_t)$
- Stationary assumption: $P(S_{t+1}|S_t)$ is constant
- Expected value/total reward of MDPs: $\sum_t^T P(S_t) * R(S_t)$

3 Value Iteration and POMDPs

- Expected value of following policy π in state s: $V^\pi(s) = Q^\pi(s, \pi(s))$
- Expected value of performing a in s, then following π: $Q^\pi(s,a) = R(s) + \gamma \sum_{s'}(P(s'|s,a) * V^\pi(s'))$
- Value iteration algorithm: $V^{k+1}(s) = R(s) + \gamma max_a \sum_{s'}(P(s'|s,a) * V^k(s'))$
- Optimal policy: $\pi^*(s) = argmax_a \sum_{s'} P(s'|s,a) * V^{\pi^*}(s')$
- Belief state update: $b'(s') = \alpha P(e|s') \sum_s P(s'|a,s) b(s)$

4 Reinforcement Learning

- Q-value true relation: $Q(s,a) = R(s) + \gamma \sum_{s'} P(s'|s,a) V^{\pi^*}(s')$, where $V^{\pi^*}(s) = max_a Q(s,a)$
- Estimate by TD: $A_k = A_{k-1} + \alpha_k(v_k - A_{k-1})$ where $A_k = \frac{v_1 + ... + v_k}{k}$ and $(v_k - A_{k-1})$ is TD-error
- Q-value approx. by TD: $Q^i[s,a] = Q^{i-1}[s,a] + \alpha_k(r + \gamma max_{a'} Q^{i-1}[s',a'] - Q^{i-1}[s,a])$
- ϵ-greedy exploration: choose random action with $P(\epsilon)$, best action with $P(1-\epsilon)$
- τ-controlled softmax exploration: choose action a with $P(\frac{e^{Q[s,a]/\tau}}{\sum_a e^{Q[s,a]/\tau}})$
- SARSA: $Q^i[s,a] = Q^{i-1}[s,a] + \alpha_k(r + \gamma Q^{i-1}[s',a'] - Q^{i-1}[s,a])$

5 Belief and Markov Networks

- Prior (forward) sampling: $P(X_1 = x_1, ..., X_k = x_k) \approx \frac{\text{number of samples w/ } X_1 = x_1, ..., X_k = x_k}{\text{number of samples}}$
- Rejection sampling: $P(y|E=e) \approx \frac{\text{number of } y \text{ w/o samples where } E != e}{\text{number of } y}$
- $P(x_i'|mb(X_i) = \alpha P(x_i'|par(X_i)) \prod_{Z_j \in child(X_i)} P(z_j|par(Z_j))$
- Number of samples to ensure max sampling error ϵ at frequency δ: $n > \frac{-ln\frac{\delta}{2}}{2\epsilon^2}$
- Inference in CRFs: $P(Y|X) = \frac{1}{Z(X)} \prod_{i=1}^m \Phi_i(D_i)$, where $Z(X) = \sum_Y \prod_{i=1}^m \Phi_i(D_i)$

6 Temporal Models

- Filtering: $P(X_t|e_{0:t}) = \alpha P(e_t|X_t)P(X_t|e_{0:t-1})$ where $P(X_t|e_{0:t-1}) = \sum_{x_{t-1}} P(X_t|x_{t-1})P(x_{t-1}|e_{0:t-1})$
- Prediction: $P(X_{t+k+1}|e_{0:t}) = \sum_{x_{t+k}} P(X_{t+k+1}|x_{t+k})P(x_{t+k}|e_{0:t})$
- Smoothing: $P(X_k|e_{0:t}) where 1 \leq k \leq t = P(X_k|e_{0:k})P(e_{k+1:t}|X_k)$
- Most likely sequence: $argmax_{x_{1:T}} P(X_{1:T}|e_{1:T})$
- P(most likely seq.): $\max_{x_1,...,x_{t-1}} P(x_1, ..., x_t|e_{1:t}) = P(e_t|x_t) \max_{x_{t-1}} P(x_t|x_{t-1}) \max_{x_1,...,x_{t-2}} P(x_1, ..., x_{t-2}, e_{1:t-1})$

7 Ontologies

- Lesk similarity: number of common words between two glosses
- Extended Lesk similarity: $\sum_{i \in P} p_i^2$ for each common phrase of length p between two glosses
- Path-length similarity: $sim_{path}(c_1, c_2) = \frac{1}{pathlen(c1,c2)}$
- Probability of a concept: $P(c) = \frac{count(c)}{N}$, where N is the number of all things
- Information content of a concept: $IC(c) = -logP(c)$
- Least common subsumer of 2 concepts: $LCS(c1, c2)$
- Resnik similarity: $sim_{resnik}(c_1, c_2) = -logP(LCS(c_1, c_2))$
- Jiang-Conrath distance: $dist_{JC}(c_1, c_2) = 2logP(LCS(c_1, c_2)) - logP(c_1) - logP(c_2)$
- Point-wise mutual information: $assoc_{PMI}(w, w_i) = log_2 \frac{P(w,w_i)}{P(w)P(w_i)}$
- Cosine similarity: $sim_{cos}(\vec{v}, \vec{w}) = \frac{\vec{v} \cdot \vec{w}}{||\vec{v}|| \times ||\vec{w}||}$
- Jaccard similarity: $sim_{Jaccard}(\vec{v}, \vec{w}) = \frac{\sum_{i=1}^{N} min(v_i,w_i)}{\sum_{i=1}^{N} max(v_i,w_i)}$

8 Probabilistic Context Free Grammar

- $P(\text{tree}) = \prod_{\text{node} \in \text{all nodes}} P(\text{expansion for node})$
- $P(\text{sentence}) = \sum_{\text{tree} \in \text{all possible trees}} P(\text{tree})$
- Probability of grammar rule: $P(A \to \alpha|A) = \frac{count(A \to \alpha)}{count(A)}$
- $\hat{\text{tree}}(\text{sentence}) = argmax_{\text{tree} \in \text{parse-trees(sentence)}} P(\text{tree})$

9 Markov Logics

- $P(x) = \frac{1}{Z} exp(\sum_i w_i f_i(x_i))$
- $\Phi_i(pw) = e^{w_i f_i(pw)}$, where $f_i(pw) = 1$ when formula=T in pw, 0 otherwise,
- and w_i =weight of f_i
- $P(pw) = \frac{1}{Z} \prod_c \Phi_c(pw_c) = \frac{1}{Z} exp(\sum_i w_i n_i(pw))$
- $n_i(pw)$ =number of true groundings of formula i in pw
- Most likely state of world: $\arg\max_{pw} P(pw) = \arg\max_{pw} \frac{1}{Z} exp(\sum_i w_i n_i(pw)) = \arg\max_{pw} \sum_i w_i n_i(pw)$