Intro to AI, CSCI 3202 -- A brief walk through the second half of the semester

Note that this is **not** meant to be an exhaustive listing of material to study for the final exam.

- **Bayesian reasoning**
  - Independence, conditional independence
  - Prior distribution, likelihood function, posterior distribution
  - Prior sampling
  - Approximate Bayesian computation ("rejection" sampling)
  - Know how this works. How would you create samples and estimate probabilities from those samples? If you were given a data frame of samples, could you estimate various marginal and/or conditional probabilities?
- **Bayesian networks**
  - How would you construct a reasonable network to model a set of variables?
  - What are the primary components that make up a Bayesian network? (nodes, connections (parent/child relationships), and CPTs)
  - How do variables relate to one another? Given their Markov blanket?
  - 3 canonical forms
  - How do you calculate probabilities, given a Bayesian network and (1) a query variable, and (2) evidence?
  - What is "explaining away"?
- **Markov models**
  - What is the Markov property? (of order 1, or 2, or …)
  - How do Markov chains relate to Bayesian networks?
  - What is a random walk? How can we use it to sample from some process?
  - What is the transition probability matrix? How can we use it to calculate the probability of making a transition from state s to state s' in 2 steps, or 3, or…?
  - What is the stationary distribution? How does it relate to the transition probability matrix? (for a discrete state space)
  - Using the mini-forward algorithm (AKA, Law of Total Probability on steroids) to get $P(X_t)$ in terms of $P(X_t \mid X_{t-1})$ and $P(X_{t-1})$. And $P(X_{t-1})$ in terms of $P(X_{t-1} \mid X_{t-2})$ and $P(X_{t-2})$. And so on…
  - What is detailed balance? How does it relate to stationarity?
- **Hidden Markov models**
  - Know what an HMM is, why it is a special kind of a Markov model
  - Filtering
  - Smoothing
  - Forward-backward algorithm
  - Viterbi algorithm (just know what it is, and if you were given a bunch of relevant (log)-likelihoods and prior probabilities, could you decide what the most likely path is? I.e., the shortest path analogy)

- Could you do a few iterations of the algorithm to update state estimates (filtering and/or smoothing) based on transition probabilities for the Markov process X_n and set of observations (evidence) $\{E_n\}$? (Like we did in class.)

- **Markov decision processes**
  - Value iteration, policy iteration (policy evaluation, policy improvement)
  - Be comfortable to do an iteration (or a few) of value iteration and/or policy iteration by hand for a simple MDP
  - Given a set of utilities (for example, output from value iteration) and transition model (probabilities), can you compute the optimal policy for each state?
  - What does the linear system of equations from the Bellman equations look like, that policy iteration needs to solve to obtain an exact solution?
- **Reinforcement learning**
  - Know what the difference between passive and active RL are. What do we know, what do we not know, and what do we estimate?
  - Model-free vs model-based -- what do these mean?
  - Could you identify one type of learning agent from another, if I gave you some pseudocode?
  - Passive RL
    - Direct utility estimation
      - What are the estimates of $U^{\pi}(s)$, given a set of samples?
    - Active dynamic programming
      - What are the estimates of $U^{\pi}(s)$, given a set of samples at some particular point within the sampling?
      - What is our estimate of P(s' | s, a) at some particular point within the sampling?
    - Temporal difference learning
      - What are the estimates of $U^{\pi}(s)$, given a set of samples at some particular point within the sampling?
      - What's up with that learning rate α? Why would we want it to decrease as we continue training?
  - Active RL
    - Exploration versus exploitation: What is an exploration function, and why would we use it?
    - What does "GLIE" mean and how does it relate to exploration/exploitation?
    - Q-learning
      - What are our estimates of Q(s, a)? How do they relate to utility of a state s, U(s)?
      - Given this set of training episodes, what are our estimates of Q(s,a)?
    - (If time) epsilon-greedy agents, how they balance exploration/exploitation