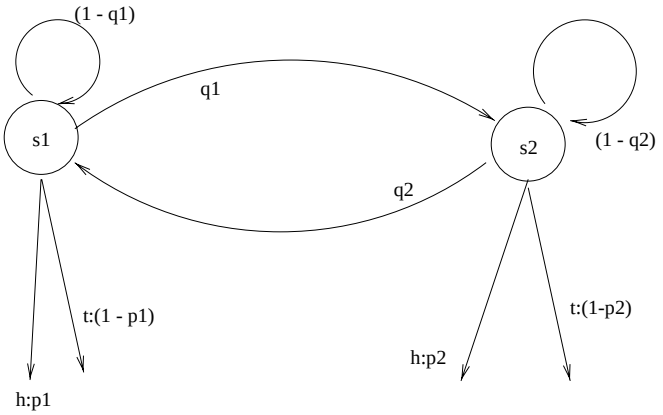
CS 530 — Principles of AI
Written Exercises
Out: November 6, 2003
Due: November 18, 2003

Problem 1. The goal of this problem is to familiarize yourself with HMMs, and to get a sense of some of the strengths and weaknesses of HMMs as models of the world.

We model the following situation. We are observing an agent. The agent has two coins, possibly unfair; the agent always starts with a specific coin c_1. At each step, the agent flips the coin, and then selects a new coin for the next step—perhaps the same one as last time with some probability, perhaps the other one. You get to observe the flip, which is either heads *h* or tails *t*. You do not know which coin was used, so you never know what state the agent is in.

Overall then we have an HMM which can be visualized as follows:



There are two states, S_1 and S_2, corresponding to the coin in use; there are two possible observations *h* and *t*. The system starts out *in state* S_1 or S_2 with equal probability, so the initial state probability distribution π is $[0.5, 0.5]$. There are four additional distinct probabilistic parameters: p_1 is the probability of seeing a head in state S_1; p_2 is the probability of seeing a head in state S_2; q_1 is the probability of moving from state S_1 to state S_2 and q_2 is the probability of moving from state S_2 to state S_1.

1a. Suppose the coin in S_1 *always* comes up heads and that in state S_2 always comes up tails, but that each state changes to the other (or stays the same) with probability .5. Draw a specialized version of the HMM above to describe this situation, filling in specific numbers for the probabilities and omitting transitions with zero probability.

1b. For a given sequence of heads and tails $O_{1,T}$ of length T, how many paths through the HMM of problem 1a generate $O_{1,T}$ with nonzero probability?

1c. For a given sequence of heads and tails $O_{1,T}$ of length T, what is the probability of observing $O_{1,T}$ according to the HMM model of problem 1a? Ie., give the numerical value of $P(O_{1,T})$.

1d. Now suppose the coins in states S_1 and S_2 are both fair: they come up heads or tails with probability .5. Assume we switch from one state to the other with a common probability *q*. Again, draw a specialized version of the introductory HMM to describe this situation, filling in specific numbers for the probabilities and omitting transitions with zero probability.

1e. Suppose the probability of seeing the first *k* symbols and being in state S_1 is *a* and the probability of seeing the first *k* symbols and being in state S_2 is *b*. We can capture this by using the

following notation:

$$P(O_{1,k}, q_{k+1} = S_1) = a$$
$$P(O_{1,k}, q_{k+1} = S_2) = b$$

Use this notation to calculate

$$P(O_{1,k+1}, q_{k+2} = S_1)$$
$$P(O_{1,k+1}, q_{k+2} = S_2)$$

as a function of a and b, assuming the HMM model of problem 1d.

1f. Use your answer to the previous problem to write $P(O_{1,k+1})$—the probability of seeing the first $k+1$ observations in the HMM model of problem 1d—as a function of $P(O_{1,k})$, the probability of seeing the first k observations.

1g. Therefore, what probability does this second model assign to the sequence $O_{1,T}$?

1h. The standard HMM training procedure, which we covered in class, converges to a local optimum. It finds an HMM model near the random initial starting point which makes the likelihood of the training data as high as possible. Use the examples of this problem to comment on the following question. Suppose that in the actual situation, the agent is in fact flipping fair coins. Can you expect the standard HMM training procedure to learn this?

1i. The limits of HMM training come up in practical AI experiements. Jeff Siskind (UIUC) has used HMMs to analyze visual inputs of human action. The observations in each HMM are quantitative features extracted from video data. The states describe the progress of an action over time. Picking something up, for example, might involve states of motion of a hand, followed by states where the hand is stationary and in contact with an object, followed by states where the hand and the object are moving together. In this approach you get a set of HMMs for all the different available actions, and recognize what's happening in video data by choosing the model with the highest posterior probability given the observations.

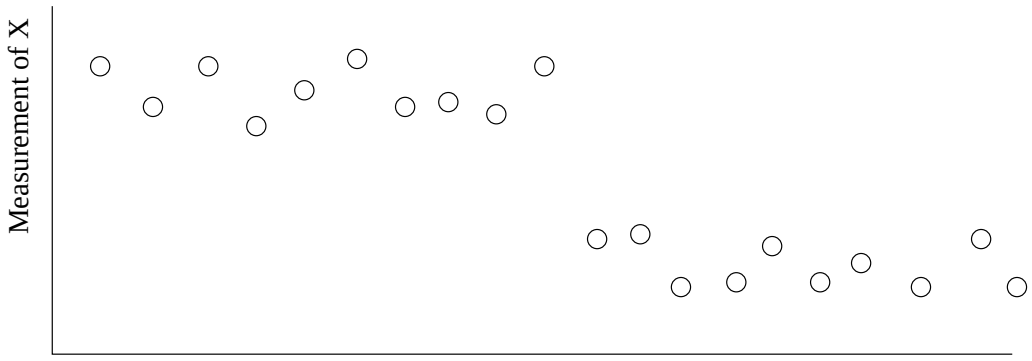
It sounds good, but Siskind had difficulties. The standard training algorithm for HMMs was not doing what he would have liked. It was learning "high frequency" information when learning a general HMM network from his data. In other words, it was assigning meaning to states spuriously to capture rapid changes of observations over time.

Use the examples of this problem to comment on Siskind's difficulty.

Problem 2. This problem offers a case study in the contrasts between using different models to keep track of a changing environment. Our agent is interested in estimating the value of a variable X at each point in time. You might think of this variable as the income or success that the agent is achieving in each cycle of deliberation and action. X takes on one of two values over time: there is a "normal mode" in which the value of X is high and a "failure mode" in which the value of X is low.

Our agent acts in this environment in trials of a fixed length of time. In each trial, the environment starts out normal, but there may be a failure at any time; once there is a failure, the failure persists until the end of the trial. At each stage, the agent makes an observation that gives noisy information about the value of X.

In all, then, within a trial, the agent might see a pattern of observations such as that shown below.



2a. Suppose our agent is designed to keep track of the value of X by just updating a running estimate. This corresponds to a coarse, general model in which the agent assumes that the actual value of X is more similar to the values seen in more recent observations. One algorithm for computing this estimate (related to so-called *reinforcement learning* algorithms) works as follows. Initially the system's estimate of X, $\hat{X}$ is the first data point w_1. Thereafter, on making an observation X_i, the system updates $\hat{X}$ by

$$\hat{X} \leftarrow w_i/3 + 2\hat{X}/3$$

Thus, we write w_i for the reward observed on step i, and write $w_{i,j}$ for the sequence of rewards observed on steps i through j. Use a graph such as that above to describe how the agent's estimate changes over time according to this algorithm.

2b. Now we consider an alternative design for our agent: it has a model of failure, and reasons by finding the best estimates for this model given the evidence it has. We design the model to interpret a set of data like those graphed above. We assume that the maximum length of the sequence is given in advance as a parameter T. The events we can hypothesize are that there was a failure after some specific step t between step 1 and step $T-1$. Write the event that there was a failure at time t as $F = t$; we can think of F as a discrete random variable taking on any of $T-1$ possible values. (Note that we stop at $T-1$ because we would only be able to detect whether there was a failure *after* observation T if we obtained another observation at time $T+1$! Note also that the first try is always a successful case: we can only hypothesize a failure that starts afterwards!)

We assume that the model assigns a prior probability to each of the following hypotheses:

$$P(F=1),\ldots,P(F=T-1)$$

Regard this as a prior distribution $P(F)$ for variable F.

We also formalize a probabilistic model of our observations in terms of two parameters, X_s and X_f, which are statistically independent of one another and of F; X_s gives the rate of reward in situations of success; X_f gives the rate of reward in failure mode. Specifically, we assume that for any i and j, w_i is conditionally independent of w_j given F, X_s and X_f. If there is no failure before step i, the reward on step i is normally distributed with mean X_s (for success) and variance $\sigma^2 = 1$. On the other hand, if there is a failure after any step up to step $i-1$, the reward on step i is normally distributed with mean X_f (for failure) and (again) variance $\sigma^2 = 1$.

We can summarize this mathematically as follows:

$$F \geq i \Rightarrow w_i \sim N(X_s, 1)$$
$$F < i \Rightarrow w_i \sim N(X_f, 1)$$

Specifically, this gives us the second ingredient of the model:

$$p(w_i|F,X_s,X_f) = \begin{cases} ce^{-\frac{(w_i-X_f)^2}{2}} \text{ if } F=t \text{ and } t<i \\ ce^{-\frac{(w_i-X_s)^2}{2}} \text{ otherwise} \end{cases}$$

We assume priors for X_s and X_f that are uniformly distributed on some interval, so $p(X_s) = p(X_f) = h$.

Our agent will reason incrementally about failure. After n steps, the agent must consider specific alternative hypotheses about a failure that may have occurred: $F=1$, $F=2$, $\ldots F=n-1$. The agent will also lump together all the possibilities of later failure as $F \geq n$. This is because if any of these hypotheses is true, the agent will not yet have any evidence of the failure yet.

The rest of problem 2b walks you through the derivation of the maximum a posteriori joint estimate for F, X_s and X_f given a sequence of observations $w_{1,n}$. Before proceeding, note that the formalization developed so far just spells out some details that are implicit in the problem statement, and that you probably used to interpret the graphs above. Your intuitions provide a good sense of what kind of answers to expect here.

b1. Use Bayes's theorem to find an expression for the quantity $p(F,X_s,X_f|w_{1,n})$ in terms of $p(w_{1,n}|F,X_s,X_f)$, and other densities drawn from the model.

b2. Rewrite your answer to b1, so that you consider the individual observations separately, and expand the conditional definition of $p(w_i|F,X_s,X_f)$. (Use two product operators.)

b3. Rewrite your answer to b2, dropping all constant factors in anticipation of maximization, and take the logarithm.

b4. As a function of F, say what value of X_s maximizes the quantity described in your answer to b3? This is the best estimate of X_s.

b5. Find the best estimate for X_f the analogous way—noting that for $F \geq n$, you actually have no evidence about X_f.

b6. Since F is a discrete set of hypotheses, you have to maximize by enumerating the alternatives. What estimation procedure results?

2c. Suppose the failure occurs right after observation $n-1$, so $F = n-1$. The question we want to address is how different the nth observation has to be from previous observations in order to recognize the failure immediately, using the model. In other words, when do we prefer the hypothesis $F = n-1$ to the collection of hypotheses $F \geq n$ (using our best estimates for X_s and X_f) based on the nth data point? Provide an answer, using the rough assumption that we have enough data that the average of the first $n-1$ observations is about the same as the average of the first n observations.

2d. Use a graph such as that sketched initally to describe how the agent's estimate changes over time according to the algorithm derived in problem 2b; assume the criterion in problem 2c is met at the moment of failure when the value of X changes. Contrast this with the coarse model of problem 2a.