CPSC-340 Machine Learning and Data Mining 2009

Homework # 3

Due Wednesday, Nov 4 1pm.

NAME:____

Signature:______________________________________

STD. NUM: ____________________________________

General guidelines for homeworks:

You are encouraged to discuss the problems with others in the class, but all write-ups are to be done on your own.

Homework grades will be based not only on getting the "correct answer," but also on good writing style and clear presentation of your solution. It is your responsibility to make sure that the graders can easily follow your line of reasoning.

Try every problem. Even if you can't solve the problem, you will receive partial credit for explaining why you got stuck on a promising line of attack. More importantly, you will get valuable feedback that will help you learn the material.

Please acknowledge the people with whom you discussed the problems and what sources you used to help you solve the problem (e.g. books from the library). This won't affect your grade but is important as academic honesty.

When dealing with python exercises, please attach a printout with all your code and show your results clearly.

1. **Hidden Markov Models for Assistive Technologies**

Assistive technologies are automated devices to help the elderly and disabled, and are an extremely active area of robotics (and one where UBC researchers and alums are involved). An example is the Misubishi Wakamaru domestic robot, which assists the elderly.

HMMs are an important tool for these robots, as they need to infer and predict a human being's internal states and motivations from their actions. For this exercise, you will be implementing an HMM for a robot pet, which is designed to play with or entertain its owner when it thinks the owner is sad, and remain unobtrusive when the owner is happy. Of course, the robot cannot directly access the human's internal state, but it can observe his or her actions.

For the sake of this exercise, humans have two hidden states: happy (`H`) and sad (`S`), which the robot will need to deduce from five actions: cooking (`CO`), crying (`CR`), sleeping (`SL`), socializing (`SO`), and watching TV (`W`).

From a series of experiments, the robot designers have determined that as people move from one task to another, sad people remain sad 80% of the time and happy people remain happy 90% of the time (the remainder of the time they switch from sad to happy or happy to sad). Initially, 60% of people are happy. That is, $P(x_0 = H) = 0.6$.

The designers have also determined that for happy and sad people, the activity breakdown is as follows:

	S	H
CO	10%	30%
CR	20%	0%
SL	40%	30%
SO	0%	30%
W	30%	10%

Let $\theta = p(x_t|x_{t-1})$ and $\Phi = p(y_t|x_t)$ denote the transition and observation probability tables respectively. Implement a Python function that allows you to do inference and prediction for a given Φ, θ, $p(x_0)$ and y for T time steps. You must allow for prediction then T is greater than the number of observations y.

Your function should look something like this:

```
def HMM(px0, theta, phi, y, T):

    ??? (setup) ???

    for t in xrange(T):
        if t < len(y):
            ???
        else:
            ???

    return px
```

(i) Filtering:

For the sequence of observations $y_{1:5} = \{SO, SO, CO, W, SL\}$, what are the predictions of the user's internal state at each time step? In other words, what is the filtering distribution $p(x_t|y_{1:t})$ for $t = 1, 2, 3, 4$ and 5?

(ii) Prediction:

Given the above sequence of observations, what is the state prediction $p(x_8|y_{1:5})$?

(iii) Smoothing: Show that the smoothing distribution $p(x_t|y_{1:T})$, where $t \leq T$, can be written in terms of the filtering distribution $p(x_t|y_{1:t})$, state predictive distribution $p(x_t|y_{1:t})$ and the transition model $\theta = p(x_t|x_{t-1})$ as follows:

$$p(x_t|y_{1:T}) = p(x_t|y_{1:t}) \sum_{x_{t+1}} \frac{p(x_{t+1}|y_{1:T})p(x_{t+1}|x_t)}{p(x_{t+1}|y_{1:t})}$$

Hint: First marginalize x_{t+1} from $p(x_t, x_{t+1}|y_{1:T})$, apply conditional probability rules and then use Bayes rule.

This completes the well known **Forward-Backward algorithm** for HMMs. In the forward pass we compute all the filtering distributions. In the backward pass, starting with $p(x_T|y_{1:T})$, we compute all the smoothing distributions. The smoothing distributions are also known as the beliefs. The Forward-Backward algorithm is a special instance of the junction tree algorithm; a very popular one! It has played a central role in speech recognition.