# COMP 557: Assignment #6

Due on Monday, November 14, 2016

**Misiura Mikita & Lee Call**

# Problem 1. Hidden Markov Models

**In this problem we will denote 'had enough sleep' event as E, 'Has red eyes' as R and 'Sleeps in class' as S.**

**Part 1**

For our model we are already given $\pi$:

$$\pi = [0.7, 0.3], \tag{1}$$

and A:

| | t | f |
|---|---|---|
| t | 0.8 | 0.2 |
| f | 0.3 | 0.7 |

We only need to calculate B by uniting two observations into one observational variable O. We define O as (R, S). Thus, $O \in \{tt, tf, ft, ff\}$. To compute $P(O|E)$ we used the following formulas:

$$P(O = tt|E = t) = \frac{P(R = t|E = t) * P(S = t|E = t)}{P(R = t|E = t) * P(S = t|E = t) + P(R = t|E = f) * P(S = t|E = f)} \tag{2}$$

$$P(O = tf|E = t) = \frac{P(R = t|E = t) * P(S = f|E = t)}{P(R = t|E = t) * P(S = f|E = t) + P(R = t|E = f) * P(S = f|E = f)} \tag{3}$$

$$P(O = ft|E = t) = \frac{P(R = f|E = t) * P(S = t|E = t)}{P(R = f|E = t) * P(S = t|E = t) + P(R = f|E = f) * P(S = t|E = f)} \tag{4}$$

$$P(O = ff|E = t) = \frac{P(R = f|E = t) * P(S = f|E = t)}{P(R = f|E = t) * P(S = f|E = t) + P(R = f|E = f) * P(S = f|E = f)} \tag{5}$$

The results are:

| $O$ | $P(E = t\|O)$ |
|---|---|
| tt | 0.087 |
| tf | 0.269 |
| ft | 0.471 |
| ff | 0.774 |

**Part 2**

- $X_0 = [0.7, 0.3]$ as given.
- $X_1$. First, we use transition matrix A to determine probabilities given no O:

$$X_1 = [0.7 \cdot 0.8 + 0.3 \cdot 0.3, 0.3 \cdot 0.7 + 0.7 \cdot 0.2] = [0.650, 0.350] \tag{6}$$

Second, we use evidence and matrix B to refine probabilities given O(1)=ff:

$$X_1 = [\alpha \cdot 0.650 \cdot 0.774, \alpha \cdot 0.350 \cdot 0.223] = [0.864, 0.136] \tag{7}$$

- $X_2$. Same steps:

$$X_2 = [0.864 \cdot 0.8 + 0.136 \cdot 0.3, 0.136 \cdot 0.7 + 0.864 \cdot 0.2] = [0.732, 0.268] \tag{8}$$

O(2)=tf:

$$X_2 = [\alpha \cdot 0.732 \cdot 0.269, \alpha \cdot 0.268 \cdot 0.731] = [0.501, 0.499] \tag{9}$$

- $X_3$. Same steps:

$$X_3 = [0.501 \cdot 0.8 + 0.499 \cdot 0.3, 0.499 \cdot 0.7 + 0.501 \cdot 0.2] = [0.550, 0.450] \tag{10}$$

O(3)=tt:

$$X_3 = [\alpha \cdot 0.550 \cdot 0.087, \alpha \cdot 0.450 \cdot 0.913] = [0.104, 0.896] \tag{11}$$

**Part 3**

Backward computation:

- $\beta_3 = [1, 1]$
- $\beta_2 = \alpha[0.8 \cdot 0.087 \cdot 1 + 0.2 \cdot 0.913 \cdot 1, 0.3 \cdot 0.087 \cdot 1 + 0.7 \cdot 0.913 \cdot 1] = [0.275, 0.725]$
- $\beta_1 = \alpha[0.8 \cdot 0.269 \cdot 0.275 + 0.2 \cdot 0.731 \cdot 0.725, 0.3 \cdot 0.269 \cdot 0.275 + 0.7 \cdot 0.731 \cdot 0.725] = [0.296, 0.704]$

Worward computation:

- $X_1 = \alpha[0.864 \cdot 0.296, 0.136 \cdot 0.704] = [0.728, 0.272]$
- $X_2 = \alpha[0.501 \cdot 0.275, 0.499 \cdot 0.725] = [0.276, 0.724]$
- $X_3 = \alpha[0.104 \cdot 1, 0.896 \cdot 1] = [0.104, 0.896]$

**Part 4**

Smoothed probabilities are shifted towards our belief that student did not have enough sleep. This happens, first, because our latest state $X_3$ says with 90% probability that student did not have enough sleep, and second, it significantly changes $X_2$ because O(2) is now supported by O(3).

## Problem 2. Understanding Human Emotions

**Part 1**

- Human emotion: $E \in \{sadness, surprise, joy, disgust, anger, fear\}$.
- Pitch contour: $P \in \{angular, glideup, descending, flat, irregular\}$.
- It is a matrix A of a size 6x6 (number of values in the domain of $E$).
- It is a matrix B of a size 6x5 (size of a domain of $E$ times size of a domain of $P$). This is the number if we only specify probabilities to observe each $P$ for each $E$. We can say that this matrix is twice as large if we also want to specify probabilities of *not* observing each $P$ for each $E$ (which are 1 - P really).
- Hard to say really, most of the time humans display almost no emotions while talking... But for this given domain we would say it should be something like $\pi = [0.1, 0.45, 0.45, 0, 0, 0]$. So small chance for sadness, large for surprise and joy, zero for other strongly negative emotions.

**Part 2**

If we got $n$ observations $E = e_1, e_2, .., e_n$, then

$$P(E) = \sum_X P(E|X) \cdot P(X) \tag{12}$$

So the probability to observe sequence of observations is a sum over all possible sequences of hidden states $X = x_1, x_2, .., x_n$.

The first term is:

$$P(E|X) = P(e_1, e_2, .., e_n|x_1, x_2, .., x_n) = \prod_{i=1}^{n} P(e_i|x_1, x_2, .., x_n) = \prod_{i=1}^{n} P(e_i|x_i) \tag{13}$$

Since all $e_i$ depend only on $x_i$. The second term is:

$$P(X) = P(x_1, x_2, .., x_n) = P(x_1|x_0) \prod_{i=2}^{n} P(x_i|x_{i-1}) \tag{14}$$

So the final formula is:

$$P(E) = \sum_X P(x_1|x_0) \prod_{i=1}^{n} P(e_i|x_i) \cdot \prod_{i=2}^{n} P(x_i|x_{i-1}) \tag{15}$$

Again, the sum is over all possible sequences of $x_1, x_2, .., x_n$.

**Part 3**

Using Bayes rule:

$$P(R = x, E) = P(R = x|E) \cdot P(E) = \Theta_x \cdot \phi_x \tag{16}$$

So:

$$P(R = x|E) = \frac{\Theta_x \cdot \phi_x}{P(E)} \tag{17}$$

And the same for $P(R = y|E)$.

## Problem 3. Conditional Random Fields and Named Entity Recognition

**Part 1**

$$P(y) = P(y_t|y_{-t}) \cdot P(y_{-t}) \tag{18}$$

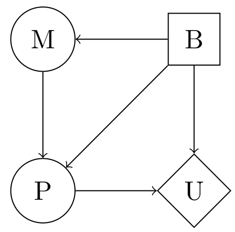
$$P(y_t|y_{-t}) = \frac{P(y)}{P(y_{-t})} \tag{19}$$

$$P(y_t|y_{-t}, x, \theta) = \frac{1}{Z(x;\theta)} \frac{\prod_{t=1}^{T} G_i(y_{i-1}, y_i; x, \theta)}{\prod_{i \neq t} G_i(y_{i-1}, y_i; x, \theta)} = \frac{G_t(y_{t-1}, y_t; x, \theta)}{Z(x;\theta)} \tag{20}$$

However, our grader sometimes shows that our difference is slightly above threshold of 0.05.

## Problem 4. Decision Networks

- Structure of our network:




- Here is the table for our utility values:

| B | P | U |
|---|---|---|
| t | t | 1900 |
| t | f | -100 |
| f | t | 2000 |
| f | f | 0 |

If we choose to buy the book, we set $B = t$ and then:

$$P(p|m) = 0.9 \tag{21}$$

$$P(p|\neg m) = 0.5 \tag{22}$$

$$P(m) = 0.9 \tag{23}$$

And:

$$P(p) = P(p|m) \cdot P(m) + P(p|\neg m) \cdot P(\neg m) = 0.9 \cdot 0.9 + 0.5 \cdot 0.1 = 0.86 \tag{24}$$

Therefore:

$$U(b) = 1900 \cdot 0.86 + (-100) \cdot 0.14 = 1648 \tag{25}$$

If we choose *not* to buy the book, we set $B = f$ and then:

$$P(p|m) = 0.8 \tag{26}$$

$$P(p|\neg m) = 0.3 \tag{27}$$

$$P(m) = 0.7 \tag{28}$$

And:

$$P(p) = P(p|m) \cdot P(m) + P(p|\neg m) \cdot P(\neg m) = 0.8 \cdot 0.7 + 0.3 \cdot 0.3 = 0.65 \tag{29}$$

Therefore:

$$U(\neg b) = 2000 \cdot 0.65 + 0 \cdot 0.35 = 1300 \tag{30}$$

- We got that $U(b) > U(\neg b)$, so poor Sam needs to spend \$100 to buy this book.