TDT4171 Assignment 2

Quang-Tam Huynh

February 2017

Part A

The set of unobserved variables is variables that are in the domain, but that we can't observe. In the umbrella domain the unobserved variable is the weather as we cannot observe it, but must infer from the known variables.
Unobserved: R_t

The set of observed variables is variables that are in the domain, and that we can observe. In the umbrella domain the observed variable is the umbrella. We use the observable variables to calculate the probabilities of the unobservable variables.
Observed: U_t

$$P(X_t|X_{t-1}) = \begin{pmatrix} 7 & 3 \\ 3 & 7 \end{pmatrix}$$

$$P(E_t|X_t) = \begin{pmatrix} 0.9 & 0 \\ 0 & 0.2 \end{pmatrix}$$

Markov assumption — that the current state depends on only a finite fixed number of previous states. Yes the assumption is reasonable in this domain as it is not reasonable to take into account how the weather were 10 years ago to calculate the weather today.

Part B

```
Day 1: [ 0.81818182  0.18181818]
Day 2: [ 0.88335704  0.11664296]
Day 3: [ 0.19066794  0.80933206]
Day 4: [ 0.730794    0.269206  ]
Day 5: [ 0.86733889  0.13266111]
```

The probability of rain on the 5^{th} day given the sequence of observation is 86.733%

Part C

```
Backward message Day 6:  [ 1.  1.]
Backward message Day 5:  [ 0.69  0.41]
Backward message Day 4:  [ 0.4593  0.2437]
Backward message Day 3:  [ 0.303981  0.158129]
Backward message Day 2:  [ 0.20099577  0.10421293]
Backward message Day 1:  [ 0.13288011  0.06885867]
```

$$P(X_1|e_{1:5}) = \langle 0.867, 0.133 \rangle$$

The probability for rain on the first day given the sequence of observations is 86.734%

Viterbi

The Viterbi algorithm is an algorithm for finding the most likely sequence given the evidence. The algorithm calculates the probability of an event on each day and makes a guess at what if there is rain a day or not based on the calculated probability. the algorithm bases the probability for that there is rain on day t on the previous day t-1 and takes the max on the current day.
If the evidence is {False, True, False, True, False} we get the probabilites:

```
xT:  [0.11111111111111112 0.23999999999999999 0.016799999999999999 0.018815999999999999 0.00131712]
xF:  [0.88888888888888895 0.12444444444444445 0.069688888888888897 0.009756444444444446 0.00546360]
```

xT is the probability for rain and xF is the probability that it doesn't rain. From these numbers we can see that the most likely sequence is:
{noRain, Rain, noRain, Rain, noRain} Which matches the evidence!

My implementation of the Viterbi algorithm first uses the forward algorithm to calculate the initial values. Then it uses a for-loop to update the probabilities day by day. There is an if-else statement to change the numbers depending of the evidence is true or false.

Task D

The source code is attached in a zip and delivered to Itslearning together with the report. The source code is written in Python 2.7 using numpy and should be sufficiently commented.