6CCS3AIN, 2019, Tutorial 05 (Version 1.0)

1. Consider the umbrella world example from the lecture. The lecture slides explain the computation of the filtered probability of rain on days 1 and 2 given the observation on an umbrella on days 1 and 2.

   Use the results for the filtered probability for day 2 (from the slides) to predict the probability of rain on days 3 to 10, and comment on the results that you obtain.

   Hint: The model is easy to program into a spreadsheet.

2. Now use the predicted probability for rain on day 3 (given umbrellas on days 1 and 2) which you calculated in the previous question to extend the filtering computation to day 3. Calculate the filtered probability of rain on day 3:

   (a) If an umbrella is observed on day 3.

   (b) If an umbrella is not observed on day 3

   Comment on your answers.

3. Following on from the previous question, predict the probability of rain on day 4 given that an umbrella was observed on days 1–3.

   Then predict the probability of rain on days 5–10, and comment on the results that you obtain.

   Hint: Again, a spreadsheet makes this easy to calculate.

4. What are the filtered probabilities for days 1, 2 and 3 if an umbrella is not observed on day 2 (that is umbrellas are observed on day 1 and day 3, but not on day 2)?

5. The slides show how to carry out a smoothing calculation on the umbrella example for the case in which an umbrella is observed on both day 1 and day 2.

   Extend that example to the case in which an umbrella is observed on the third day as well. For this you will need to compute the smoothed probabilities for days 1 and 2.

6. How does the smoothed probability for days 1 and 2 change if an umbrella is not observed on day 2 (that is umbrellas are observed on day 1 and day 3)?

7. This is the optional computational part.

   This part once again makes use of `pomegranate`. If you don't have this installed, see the installation notes in Tutorial 02. YOu will also need to download the file `umbrella.py` from KEATS. This is a `pomegranate` version of the umbrella example.

   (a) Run the code you just downloaded:

   `python umbrella.py`

   This runs the example with the evidence of an umbrella seen on Day 1 and no information about umbrellas on Day 2. The example only contains the network for Days 1 and 2. You should get:

```
Evidence is:  Umbrella on Day 1.

Rain0:  (('y', 0.6272727272727271), ('n', 0.37272727272727296))
Rain1:  (('y', 0.8181818181818179), ('n', 0.18181818181818207))
Rain2:  (('y', 0.6272727272727271), ('n', 0.3727272727272729))
Umbrella2:  (('y', 0.6390909090909089), ('n', 0.36090909090909107))
```

   Now, `pomegranate` does not use the ideas of prediction and smoothing that we talked about in class, and it doesn't use the message-passing methods either. Rather it treats the model it is given as a standard Bayesian network, and applies its usual (maximum likelihood estimation) inference to get the answers. Then `umbrella.py`, like `wetGrass.py` before it, reports the values of every variable that is not evidence.

   Since we have evidence for `Umbrella1`, that is we saw an umbrella on Day 1, and there is nothing reported for Umbrella on Day 2, what the output of the model shows is:

   - The filtered probability of rain on Day 1, `Rain1`. Just as in the Lecture.
   - The predicted probability of rain on Day 2, `Rain2`. Again this is just as in the Lecture.

- The smoothed probability of rain on Day 0, `Rain0`. We didn't calculate this in the lecture, but we could have. We see exactly the same pattern as we did when computing the smoothed probability of Day 1. Because we have strong evidence of rain on Day $n + 1$, and because rain tends to lead to rain, rain on Day n becomes more likely than before we had evidence. (In this case the probabilty of rain on Day 0 increases frrom 0.5 to 0.627 because we have evidence of rain on Day 1.)
- The predicted probability of umbrella on Day 2, `Umbrella2`. Again we didn't calculate this in the lecture, but it is easy to calculate from the probability of rain on Day 2 and the sensor model.

(b) Modify the evidence in `umbrella.py` so that you can compute the probabilities of rain in days 1 and 2 given all possible combinations of umbrella or not on days 1 and 2.

Remember that to show "no umbrella on Day 2", you need to set `Umbrella2` to `'n'`. Setting the value to `None` says you have no evidence about whether `Umbrella2` is true or false.

(c) Now extend the Bayesian network in `umbrella.py` so that you can calculate the results for Day 3, and use this to check your answers for questions 2, 4, 5 and 6.