

Workshop 11: Smoothing

(Version 1.1)

1 Overview

This workshop relates to Lecture 09. In that lecture, we discussed Dynamic Bayesian Networks (DBNs). Here you will see how DBNs work using a mixture of Excel, and the Python Pomegranate package. In particular, you will carry out a mixture of filtering, prediction and smoothing tasks on the umbrella network that we studied in the lecture.

2 Excel for filtering and prediction (redux)

The spreadsheet `umbrella-smoothing.xls` that can be found on Blackboard models the umbrella example over the first 3 days.

The first tab/sheet “Filtering D3” provides the solution to two of the problems from last week.

First, it gives an answer to the problem from Section 6 on predicting rain. The upper cells of the sheet provide a prediction forward until Day 10 (when the predicted probability of Rain has converged to 0.5). BTW, why do we know that this means it had converged?

Second, the tab/sheet also provides the computation of the filtered probability of rain for Day 3. This is the answer to the problem from Section 7.

(Note that the tab/sheet “Filtering D2” provides predictions from Day 3 onwards — it filters to Day 2 and then predicts — so provides another prediction forward to Day 10.)

3 Excel for smoothing (Day 1)

The tab/sheet “Smoothing D1” gives the smoothing calculation for Day 1.

The forward message is computed just as for filtering (it is the same message after all). The backward message (at the bottom of the sheet) is computed exactly as on Slide 26 and Slide 28 (the layout is similar to that on Slide 28 which hopefully makes it easy to see the correspondence).

The smoothed probability is then just the product of the forward and backward messages, normalised. This is the calculation on Slide 27 (and 28).

Look at what happens when the probabilities of umbrella/not umbrella on Days 1 and 2 vary.

4 Excel for smoothing (Day 2)

The tab/sheet “Smoothing D2” gives the smoothing calculation for Day 2.

Compared to the calculation for Day 1, this involves predicting forward another day, entering evidence, and computing the backward message.

Again, look at what happens when the probabilities of umbrella/not umbrella on Days 1 and 2 vary.

5 Using pomegranate for Smoothing

As in Workshop 8 we will use a Python package called `pomegranate`, which provides support for probabilistic reasoning. If you have not installed it before, you will need to do so with:

```
pip install pomegranate==v0.14.9
```

Then you can run the version of the umbrella model in `umbrella.py`. `pomegranate` can only solve Bayesian networks (not Dynamic Bayesian Networks), so we have to unroll the whole example to the depth that we want.

The file `umbrella.py`, which you can download from Blackboard, has the network unrolled to a depth of 2 days. To do smoothing, we need to add:

```
Rain3 = ConditionalProbabilityTable([
    ['y', 'y', 0.7],
    ['y', 'n', 0.3],
    ['n', 'y', 0.3],
    ['n', 'n', 0.7]], [Rain2])
```

to tell the model about rain on Day 3. We also need to add the elements:

```
# Node
s6 = Node(Rain3, name="Rain3")
# State
model.add_states(s6)
# Edge
model.add_edge(s4, s6)
```

before we `model.bake()` and start to use the model.

Note that we only call `model.bake()` once the last elements are entered.

Now that we have the model entered, we can ask it questions. We tell the model that we saw umbrellas on Days 1 and 2:

```
# Umbrellas on Day 1 and 2:
scenario = [[None, None, 'y', None, 'y', None]]
3# Run the model
results = model.predict_proba(scenario)
# Ask for the probability of rain on Day 1:
print(results[0][1].items())
# And the probability of rain on Day 2:
print(results[0][3].items())
```

This should give exactly the results from the lecture (Slide 28), and from the Excel model.

Note that we didn't tell `pomegranate` to do smoothing. As we saw last time with Day 0, it (in effect) always runs the backwards propagation and gives us smoothed probabilities for all days before the latest piece of evidence.

I said "in effect" because `pomegranate` doesn't do the computation the way we studied. It just computes the probability of every hidden variable given the evidence. To see this, try looking up the probability of rain on Day 3.

6 More smoothing using Excel

Now go back to the Excel spreadsheet and calculate the smoothed probability for Day 3.

7 Day 3 using pomegranate

Extend the pomegranate model to compute the smoothed probability of rain Day 3, given both umbrella being true and false on Day 4 (that is two separate calculations).

Check your result against the values from the Excel model.

8 Version list

- Version 1.0, 12th December 2022.
- Version 1.1, 16th November 2023.
 - Fix pomegranate version
 - Fix typos