CS 3630 Project 2

Name: Matthew Yang
GT email: MattYang@gatech.edu
GT username: myang349
GTID: 903447357

1. What is the difference between CDF and PMF?

PMF stands for “probability mass function.” The PMF tells you the probability of your random variable taking on a specific discrete value. PMF obeys the properties pi >= 0, and sum(all pi’s) = 1, where pi is the probability of your random variable taking on a specific ith discrete value.

On the other hand, the CDF, or cumulative distribution function, gives the “probability associated with the SUBSET of outcomes with index less than or equal to a given index i.”

The key difference is that PMF gives the probability for one discrete value, while CDF gives the summed probability for a subset of discrete values. CDF(xi) is the sum of PMF(xj) for all j <= i. Note: CDF(xn) = 1.

2. How did you implement maximum-probable explanation?

Step 1. Convert the Bayes Net into a factor graph. For me, my process is to draw unknown variables as nodes, then turn each term in the joint probability distribution into a factor and draw them in the factor graph, connecting factors to the nodes that are related.

Step 2. Each look-up table is essentially the elimination of one variable from the factor graph. I wrote out the product factor for each table. Product 1 is P(s1) * Likelihood(s1; o1) * Transition(s2 | s1, a1), Product 2 is v1[s2] * Likelihood(…) * Transition(…), and Product 3 is v2[s3] * Likelihood(…)

In a generalized version, to create a specific look-up table, the outer loop iterates over every key (if not the last table). Then, for each key, iterate over the other variables, say j, in the product factor. Then find the j that maximizes product(j), and record g[key] = j and v[key] = product(j).

After repeating this process for every table, you will have a chain of look-up tables. Then, you start with the MPE for state S_n. Then let S_{n-1} = Table[S_n], then S_{n-2} = Table[S_{n-1}], etc. Repeat until you have every state, and that is the MPE. I do believe my code is pretty well commented in the extra credit MPE implementation, if that helps explain my thought process.

3. What is the purpose of sampling and how is it used?

There are different types of sampling, but basically the idea is that sampling can simulate the behavior of a variable (or probability based behavior)

Inverse transform sampling allows us to simulate a single variable.

Ancestral sampling can simulate from a Bayes Net.

In this project, we used sampling in the methods sample_from_pmf(), sample_from_sensor_model(), sample_from_transition_model(), and sample_from_dbn() to quantify/illustrate how our sensor model, transition model, and dbn behave.

Over a large number of samples, we can get a pretty good idea of the behavior of our models.

4. What is a factor graph?

A factor graph is a probabilistic graphical model. Specifically, in this class we have learned how to convert Bayes nets into a data structure (factor graph) that gives us evidence that is “all implicit in the factors.”

The advantage is that, given a factor graph defining a density, we can easily evaluate it for any given value. As we’ve seen previously, the Bayes filter is very inefficient if we are interested in something such as the MPE for all states. That is, the entries required to in tabulating to find the MPE is exponential in the number of states.

On the other hand, the properties of factor graphs offer us an efficient way to perform such computations.

5. What did you learn in this project?

I learned a lot about Python, as I was a totally noob prior to this class. (Examples include tuples vs. lists, basic for loops, list manipulation, etc.)

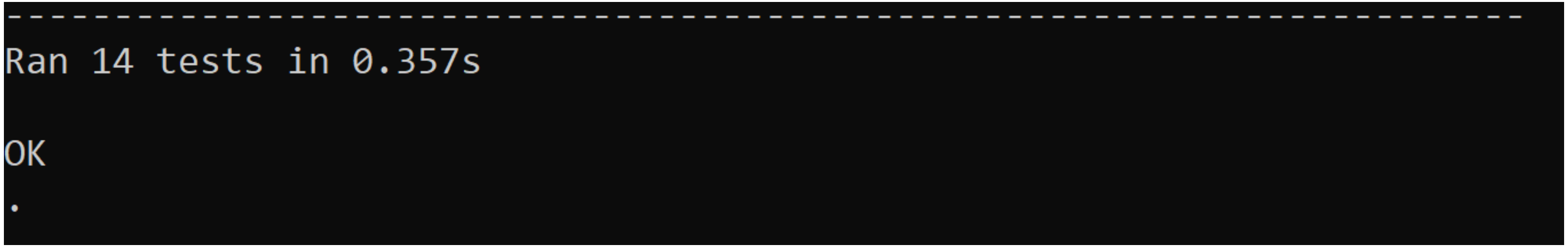
I also gained a solid understanding of the MPE algorithm and turning Bayes' nets into factor graphs, something that would’ve been very helpful on the last quiz ☹

I also learned about the power of the NumPy library, and how to use a few of its functions (np.zeros(), np.random.choice(), etc.)

Lastly, I learned a LOT about probability, as I am taking a prob-stat class concurrently with this one and don’t have much prior experience.

6. Screenshot and paste the output of running your unit tests here

<Insert Screenshot here.>

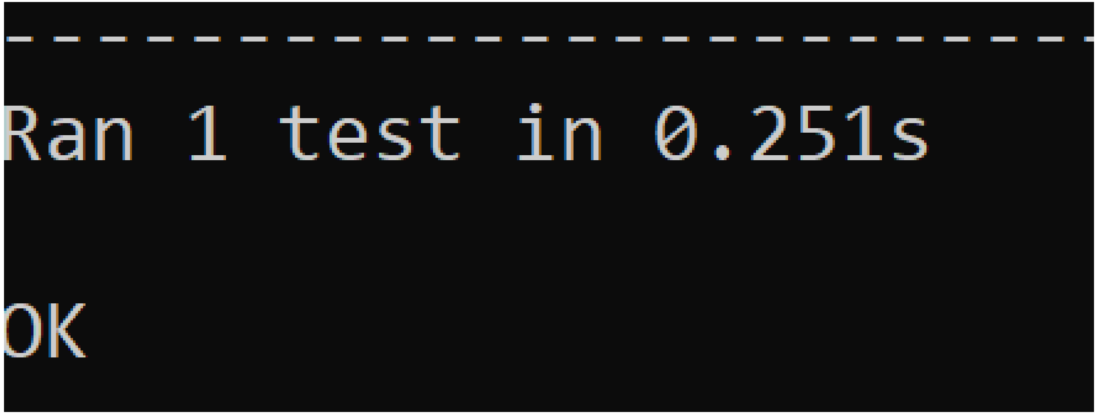


```
----------------------------------------------------------------------
Ran 14 tests in 0.357s

OK
.
```

7. Extra-Credit - MPE: Screenshot and paste the output of running your extra-credit unit tests here

<Insert Screenshot here.>

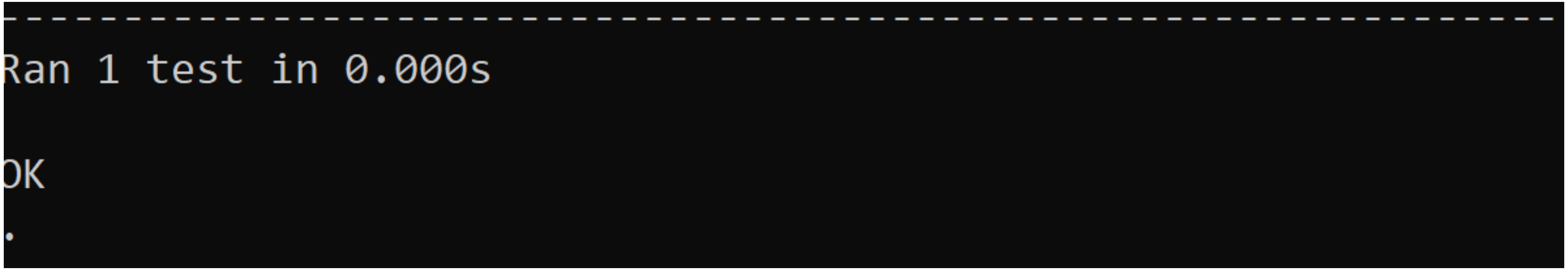


```
------------------------
Ran 1 test in 0.251s

OK
```

8. Extra-Credit - Portal: Screenshot and paste the output of running your extra-credit unit tests here

<Insert Screenshot here.>



```
-------------------------------------------------------------------
Ran 1 test in 0.000s

OK
.
```