SymSys 100, Fall 2015: Homework 3
Due Wednesday November 11, 2015 at 5PM
to the dropbox on the CourseWork website:

Submission rules (designed to streamline grading):

A. Please ensure that your homework has your name at the top of the file.

B. Please submit your homework in .pdf format (rather than e.g. Microsoft Word’s .docx format). If you don’t know how to do this, it’s worth learning: look on Google for a conversion method specific to your word processing program.

C. The filename of your submission should be in the following format:

[LASTNAME]-[FIRSTNAME]-HW3.pdf

replacing “[LASTNAME]” with your last name and “[FIRSTNAME]” with your first name.

Remember, failure to follow these instructions will result in a deduction of 10% per requirement.

Please answer the following questions. Pay attention to the word limits.

1. (50 words or less; 9 points) Suppose there are only 5 candidates left in the presidential race: Republicans Donald Trump and Marco Rubio, and Democrats Bernie Sanders, Lincoln Chafee, and Hillary Clinton. The probability of each of them winning the presidency is as follows:

- P(“Donald Trump wins”) = .2
- P(“Marco Rubio wins”) = .2
- P(“Hillary Clinton wins”) = .2
- P(“Bernie Sanders wins”) = .3
- P(“Lincoln Chafee wins”) = .1

Calculate the following. You may, but need not, show intermediate calculations.

a. What is the probability that Trump loses?

b. What is the probability that a Democratic candidate wins?
c. Suppose we know that a Republican will win. What is the probability that Rubio wins?

2. (200 words or less; 18 points) A number of U.S. states have debated laws requiring doctors to recommend HIV testing to all patients during routine medical screenings. The CDC estimates that 417 of every 100,000 U.S. adults between the ages of 18 and 65 are living with HIV, with rates varying considerably by locale: some states have rates as low as 40 per 100,000, and the highest (D.C.) is estimated to have a rate of 3,500 per 100,000.

Suppose that the test administered detects 99.5% of true positives (995 per 1000) and has a false positive rate of 15 per 1000.

a. Imagine that every adult in the U.S. were tested. Use Bayes’ rule to estimate the probability that a randomly chosen individual with a positive test result actually has HIV — i.e., \( P(\text{HIV} | \text{positive test}) \).

b. If universal testing were implemented only in D.C., what would the probability be that a randomly chosen D.C.-resident with a positive test result actually has HIV?

c. Which policy would you recommend implementing?

You may, but need not, show intermediate calculations.

3. (200 words or less; 13 points) Read the final project description posted on the course website (or click here for a direct link).

a. Groups of up to 3 individuals are permitted for this project. If you choose to work in a group, identify the individuals in your group.

b. Identify the individual or lab whose research you will focus on in your project, and give a brief (2-3 sentence) explanation of how their research is relevant to the course content or some other aspect of Symbolic Systems research, broadly construed.

4. (500 words or less; 30 points) On Tuesday 11/3 we discussed inductive arguments where people’s judgments of argument strength appear to be related to the conditional probability of the conclusion given the premises. We noted that people’s willingness to make such inferences is often sensitive to a background model of the process by which properties are transmitted. For example, people are inclined to generalize from “Mice carry the toxoplasmosis gondii parasite” to “Owls carry the toxoplasmosis gondii parasite”, much more than they are to generalize from “Mice have sesamoid bones” to “Owls have sesamoid bones”. As we discussed, this is presumably related to the fact that owls eat mice, and people know that parasites can be directly transmitted from prey to predator.

Shafto & Coley (2003) found that people were more willing to generalize disease properties from prey to predator than vice versa. For example, people judged the inference from a premise like “Seals carry leptospirosis” to a conclusion like “Great white sharks carry leptospirosis” to be stronger than the converse inference.

a. Why would there be an asymmetry here? Situate your answer with respect to two or more of the subcomponents of Bayes’ rule (posterior, likelihood, prior).
b. Shafto & Coley (2003) also found that — where arguments involving marine animals was concerned — the difference in strength between prey-to-predator vs. predator-to-prey arguments was greater for commercial fishermen than for college undergraduates. In light of your answer to (a), why might this be?

5. (500 words or less; 30 points) On Thursday 10/29 and Tuesday 11/3 we discussed the inductive problem posed by the general lack of negative evidence in concept acquisition — for example, the fact that children are rarely told that something is “not a dog”.

   a. Suppose that children learned categories in a purely deductive fashion. Why would a lack of negative evidence pose a severe problem?
   b. How can Bayesian reasoning about the process by which labels are generated help reduce uncertainty, and what kind of learning bias does this reasoning lead to?
   c. At which level of the Marr hierarchy does the explanation you gave in (b) lie?

References