Tutorial 5: Word Recognition

Please work through this tutorial sheet on your own time as much as possible before arriving in tutorial. We encourage you to work together and discuss your methods and solutions.

Word Recognition and Neighborhoods

**Question 1:** In the lectures, we discussed a model of visual word recognition called Bayesian Reader proposed by Norris (2006). In this model, a visual input $I$ is recognized as word $\hat{W}$ by computing:

$$\hat{W} = \arg\max_{W_i} P(W_i|I) = \frac{P(I|W_i)P(W_i)}{P(I)}$$  

(1)

where $W_i$ ranges over all words in the lexicon.

1. Which factors influence word recognition according to equation (1)? Explain the role of each of the terms on the right-hand side of the equation.

2. Lexical decision is the task of deciding for a given input (typically a string of letters) whether it is a word of the language or not. Show how Bayesian Reader can model lexical decision by replacing the MAP computation in equation (1) with a Bayesian summation.

3. Comment on the cognitive plausibility of the Bayesian computation in the previous question. Which additional assumptions may be necessary to increase the cognitive plausibility of this aspect of the model?

4. Is the issue discussed in the previous question a general feature of Bayesian models? Consider the other Bayesian models discussed in the course so far: The Tenenbaum model of concept learning, and the Frank et al. model of word learning.

**Question 2:** Experiments show that the time it takes to recognize a written word depends on the number of orthographic neighbors of the word to be recognized. An orthographic neighbor is defined as a word of the same length that differs from the original string by only one letter.

1. Draw a neighborhood graph for the following words. In such a graph, the words are nodes, and two nodes are connected by an edge if the words represented by the edges are orthographic neighbors.

   cat, bat, fat, cab, bar, bet, but, bus

   What is the maximum number of neighbors a word of three letters can have in principle?

2. Figure 1 plots the number of orthographical neighbors against the reaction time in a lexical decision experiment. Explain the effects that the graph shows. How does Bayesian Reader model them?

3. How would you expect the neighborhood effects shown in Figure 1 to change in non-native speakers? How could you model this with Bayesian Reader?
Figure 1: Reaction time vs. number of neighbors in a lexical decision experiment

References