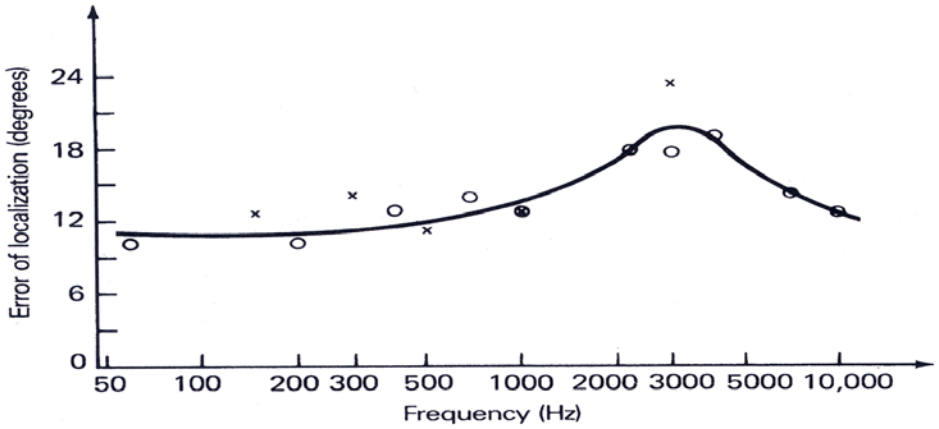
CS786 Quiz 3

100 points

Q1. See the graph below



(a) Can you explain why the error in sound localization rises and then falls? (6 points)

This is because there are two separate mechanisms used for sound localization. (2 points) Low frequency sounds are localized using inter-aural time difference (ITD), viz. difference in the time it takes for a particular sound to reach either ear (need not mention the name, knowing the concept is what counts) (1 point). High frequency sounds are localized using inter-aural intensity difference (IID), viz. difference in the intensity of sound reaching the two ears (1 point). The region centered around 3000 Hz in the sound spectrum is where the ITD mechanism has started to fail, and the IID mechanism has not yet started working well (2 points).

(b) Why does it rise and fall at the particular frequency it does? (8 points)

Head size (3 points for saying just this much).

Lower frequency sounds have considerably higher wavelengths, and high frequency sounds have much shorter wavelengths. (1 point) Small wavelength sounds are blocked by the human head (1 point), since they bump into it and are absorbed by the flesh. Large wavelength sounds can wrap around the head (1 point), since their larger wave front allows them to cover greater distances perpendicular to their direction of travel. The human head has a circumference of about 0.3 meters, and a diameter of about 0.1 meters (1 point). A frequency of 3000 Hz corresponds to a wavelength of about 0.1 meters. (1 point) Sound waves with wavelengths larger than 0.1 meters wrap around peoples' heads. Thus, the ITD mechanism takes over for such sound waves (frequencies larger than 3 kHz). For lower frequencies, the IID mechanism stays active (1 point). For a differently sized head, the change-over point would be different.

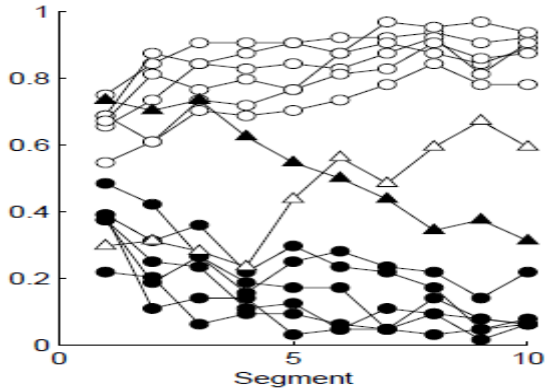
(c) How do people differentiate between sound sources that are right in front of them and right behind them? (2 points)

The pinna (outer covering) of the ear is asymmetric (1 point) and distorts sounds coming from the front differently than sounds coming from the back of the head (1 point).

(d) What are visemes, and how do they help in learning to speak? (4 points)

Visemes are the set of unique mouth shapes made during the construction of speech sounds (1 point). Multiple phonemes (basic speech sounds) may correspond to the same viseme (1 point), but the context of surrounding visemes may permit an expert observer, for example, a professional lip-reader, to understand what someone is saying by just looking at their lips. It has been conjectured that the viseme-phoneme mapping helps pre-verbal children learn how to speak (1 point) - by mimicking the mouth shapes they see associated with the corresponding sounds. (1 point)

Q2. Results from the Smith and Minda (1986) experiment proved challenging for classical categorization theories to model. In this experiment, participants' saw two types of stimuli associated with corresponding labels, say 'white' and 'black'. The graph below plots the probability that participants called stimuli of any kind 'white' over multiple presentations of these stimuli across time. Circles represent typical stimuli of each variety; triangles represent atypical stimuli in each variety.



(a) Can you interpret this graph in terms of the behavior of the experiment's participants? (10 points)

The graph shows that (a) humans increasingly get better at discriminating the two stimuli categories over multiple trials (3 points), and (ii) they start out predicting the atypical stimuli in the opposite category (2 points), but then switch over to the correct category eventually (2 points), (c) with this switchover happening much slower than the overall trend in performance improvement (3 points).

(b) Can you explain why its conclusions cannot be explained by either prototype or exemplar models of categorization? (10 points)

Prototype models would predict no improvement in the ability to predict the atypical stimulus' category correctly, since categorization would be determined based on distance from the prototypical stimulus, which would be close to the typical stimuli psychophysically (3 points). In prototype models, the atypical stimuli are predicted in the opposite category (2 points).

Exemplar models predict no performance improvement, since the set of exemplars for the categories don't change over the course of testing without feedback (3 points). Atypical stimuli would be continually weakly misclassified if people were behaving like exemplar models (2 points).

Its ok to give full marks if they just draw the prediction graph for prototype models, i.e. slide 4 from Lec 22. Same for exemplar model predictions, i.e. slide 5 from Lec 22.

(b) Anderson's Rational model of categorization could fit this data, by allowing multiple clusters to be fit per stimulus category. Describe how this model works. Be sure to specify what lets it decide how many clusters to use to fit the data? (10 points)

Just need a description of the Dirichlet process mixture model using the Chinese Restaurant Process construction for the Dirichlet prior. (see slides 23-26 in Lec 22) +2 points for connecting the mapping of items to clusters as a sample from a multinomial distribution, +2 points for pointing out that the Dirichlet distribution is the conjugate of the multinomial distribution, +1 point for pointing out that conjugacy is helpful in the getting the sequential Bayesian math of the DPMM to work, +2 points for describing the CRP prior accurately, +2 point for explaining the role of the concentration parameter, i.e. high value of the parameter makes spinning off new clusters more likely, small values make new clusters less likely, +1 point for pointing out the exchangeability assumption, viz. that the order in which observations are seen don't matter.

(c) Draw a plate notation diagram of the hierarchical Dirichlet process model of categorization, dereferencing all notation (10 points).

This is from the slides. See slide 29 in Lec 22. 4 points for getting all the plates correct (with notation), 3 points for labelling all the variables correctly (-0.5 points for each error), 3 points for getting all the arrows right (-0.5 points for each error).

Q3. Assume that you have built a search engine that crawls 500k web pages, on average each one being 1000 words long. I want you to use a simple query likelihood model of document relevance to figure out which document would be most responsive to my query "computational cognitive science":

Document 1, which is 2000 words long, and contains two instances of 'computational', one of 'cognitive' and none of 'science'.

Document 2, which is 1000 words long, and contains one instance each of all the words.

(a) Calculate the query likelihood for both documents, assuming Dirichlet smoothing with parameters $\mu = 1000$, $\alpha = 0.1$ and collection-wide frequencies {1350, 1200, 10800} for the three terms computational, cognitive and science respectively. (15 points)

(It's ok to leave the answer as an expression)

Entirely from the slides. This is just a formula memory check. The formula is the log P(Q|D) calculation on slide 15 in Lec 24. f_(q,D) are the counts of the individual query terms, c_q are the collection wide frequencies of the three query words, mu is given and |D| is either document's length. I found document 1 more likely in my calculation, but you can do your own. 5 marks for getting the formula right, 5 more for sticking the actual numbers in correctly, and 5 more for trying to assess which quantity is higher.

(b) How would you change this model, such that documents that contain terms like 'cognition', 'compute' etc. may also be returned? (15 points)

I would replace the original unigram probability calculation $P(q_i|D) = \frac{f_{q_i,D}}{|D|}$ with something like $P(q_i|D) = \frac{\sum_x sim(x,q_i) f_{x,D}}{|D|}$ where x are all words in memory. In practice, we'd only consider words with sim(x,q_i) greater than some threshold, and this sim value would be obtained from some semantic association data.

Full marks if someone spells out a semantic association based solution like this one. 8 marks if someone says semantic association should be used, but can't spell out how mathematically.

Q4.

(a) What empirical variable did Cooper and Shepherd's mental rotation experiment measure? (5 points)

The reaction time it took for people to say whether two shapes were the same, with a rotational difference. (2 points for saying only reaction time; 3 points for the remaining description)

(b) What is the significance of the results of this experiment? (5 points)

The experiment provides evidence for the existence of a non-propositional representation of visual items (3 points) that can be cognitively manipulated (2 points).