Problem Set 04

For problems 1 and 2, record yourself saying the word mystery [mɪsˈtɪri]/[M IH S T ER IY].

1. (20 points) Estimate the formants of the center part of [t] in your rendition of mystery using wavesurfer\(^1\) format tracker (right click, Create pane, formant plot). Examine the spectrum and tracks to verify that wavesurfer did a good job of estimating the formants. (Turn in a screen shot showing the formants.)

The chart on the last page of this assignment is from Peterson and Barney (1952) and shows the first formant vs the second formant for English vowels. Place a copy of the chart in the materials that you hand in and mark clearly where the formants for your [t] fall.

2. (20 points) For the consonant [m,t] from the word mystery, describe the place and manner of articulation for each one.

3. (20 points) Consider the time series plot of /a/ (as in cot) below. Showing your work, calculate your F0 in Hz (cycles/s). You will obtain a more accurate estimate if you use several cycles of your vocal folds opening and closing (be sure to divide by the number of cycles).

\(^1\) Wavesurfer is available at [http://www.speech.kth.se/wavesurfer/](http://www.speech.kth.se/wavesurfer/).
4. (20 points) Consider the computation graph below where $x$ is an input vector, $y$ is the desired output, $o$ is the output, and $L$ is the squared error loss function with a scalar modifier. Given input $x=[3,4]'$ and $y=7$

\[ L = \frac{1}{2}(o - y)^2 \]
\[ o = \sigma(w_1c + w_2d) \]
\[ c = x_1^2 + 2x_2 + 5 \]
\[ d = x_2 + 1 \]

Compute

a. Partial derivatives $\frac{\partial L}{\partial w_1}$ and $\frac{\partial L}{\partial w_2}$.

Note that our concrete example of the chain rule in class did not show how to propagate across an activation function. You simply take the derivative of the node output with respect to the input. Several slides have been added to the deep nets slides to provide an example of this (slides 53-57). While you are free to derive the derivative of the sigmoid function, it is given in your book (eq. 3.34).

b. Normally, we would average gradients over our minibatch. We will simplify and assume that the batch size is 1. Given input $x$ and output $y$, above and a learning rate of 0.01, what will the new weights $w$ be? You do not need to adapt the weights of nodes $c$ and $d$ although in a standard backpropagation we would update these as well.

REFERENCES

Fig. 8. Frequency of second formant *versus* frequency of first formant for ten vowels by 76 speakers.