TTIC 31230: Fundamentals of Deep Learning
Problem set 1
Due Friday, October 9 (11:59pm)

This problem set involves understanding and modifying the education framework (EdF). Everyone should start by installing Anaconda with Python 3.5 (https://www.anaconda.com/products/individual). Then open a terminal and go to the problem set directory. Then enter "jupyter notebook". This should open a window in your browser from which you can open PS1.ipynb (for Interactive PYthon NoteBook). You can also open the source code edf.py of the framework. It is about four pages of Python.

Problem 1.

a. Show that the backpropagation algorithm is correct in the sense that the derivative of the loss function with respect to the value of a component equals the sum over the users of that component of the derivative of the loss with respect to that input of the user. In particular, show that for the loss $\ell(g(x), h(x))$, where ℓ, g and h are arbitrary differentiable functions, we have that $\partial\ell/\partial x$ equals the sum over the users of x of the gradient of the loss with respect to the user's input.

b. Show that backpropagation fails for computing second derivatives – for $\ell(g(x), h(x))$, we do *not* have that $\partial^2\ell/\partial x^2$ is the sum of second derivatives of users of x.

Problem 2. Examine the implementation of Softmax in edf.py. Give a formal derivation of the correctness of the implementation of its backward method.

Problem 3. You are to expand the framework EdF by adding an implementation of $\text{ReLU}(x) = \max(0, x)$ and of $\tanh(x) = (e^x - e^{-x})/(e^x + e^{-x})$. Fill the missing code for the forward/backward methods of ReLU and tanh in the jupyter notebook. Then, run the following cells to train networks using ReLU and tanh as activation functions.

You should turn in your complete jupyter notebook and a write up of the solutions to problems 1 and 2.