# CSS 485 Midterm — Winter 2021

Prof. Michael Stiber

Wednesday, February 24, 2021

Answer the questions in the spaces provided on the question sheets. If you run out of room for an answer, continue on the back of the page (and indicate so on the front). Please *write neatly* and *show all of your work*; credit cannot be given for unreadable or unsupported answers.

Name: _____

**Problem 1**.

ANNs are called *Artificial* Neural Networks because, though they take their architectural inspiration from biological neural networks, they are *not* meant to be biological simulations. Briefly discuss two *non-trivial* differences between biological and artificial neural networks.

10

*Answer*:

**Problem 2**.

Show that any multi-layer perceptron with a `purelin` output function can be rewritten as a single-layer perceptron.

10

*Answer*:

**Problem 3**.

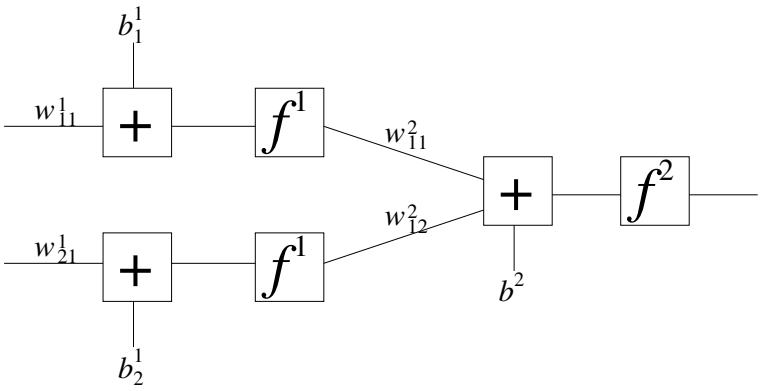
When we apply the Hebb rule to a linear associator, we call the resulting neural network an *associative memory*. Why?

5

*Answer*: .....

**Problem 4**.

For the network shown below, the initial weights and biases are chosen to be: $w_{11}^1 = 0.5$, $w_{21}^1 = 1$, $b_1^1 = 0$, $b_2^1 = 1$, $w_{11}^2 = 1$, $w_{12}^2 = 0.25$, $b^2 = -1$, $f^1(n) = n^3$, $f^2(n) = 1/n^2$ (note that the neurons' output functions are different for the two layers, that the superscripts in these functions are powers of $n$ — not layer numbers, and that neither is a sigmoid). The input/target pair $\left\{p = \begin{bmatrix} 1 & -1 \end{bmatrix}^\mathsf{T}, t = 1\right\}$ is presented. Perform one iteration of backpropagation with learning rate $\alpha = 0.5$. Remember that the "sensitivity" for an output element is computed as $s_i^M = \partial \hat{F}/\partial n_i^M = -2\dot{f}^M\left(n_i^M\right)(t_i - a_i)$ and for a hidden unit as $s_j^m = \partial \hat{F}/\partial n_j^m = \dot{f}^m\left(n_j^m\right)\sum_l s_l^{m+1} w_{jl}^{m+1}$. *Please show your work.*




20

*Answer*:

**Problem 5**.

Answer the following questions about *overfitting*:

a. Briefly describe overfitting and why it is a problem. 5

*Answer*:

b. Draw a simple example showing well-fitted and overfitted data and label the various parts. 5

*Answer*:

c. Briefly describe two ways of preventing overfitting in neural networks and why they work. 10

*Answer*:

**Problem 6**.

In this course, we have discussed vector spaces in general, as well as vector spaces that have specific meanings in terms of artificial neural networks. Two of the latter are the *input vector space* and the *parameter vector space* (where the parameters are the network weights and biases). This question is about the input vector space (i.e., the space within which each input pattern is a vector). Describe this vector space, making sure you relate the dimensionality of the space and identity of each axis (input vector components or basis vectors) to things like the physical quantities and sensors used to capture the input pattern data. 10

*Answer*: