WORKSHEET FOR EXERCISES FROM CHAPTER 13

EXERCISE 13.1

1. **What was the sum of squared error at the end of training?**

The total SSE at the end of training was approximately 0.02.

2. **How many sweeps of training were required before the program stopped training the network?**

The network converged after 49 epochs of training.

3. **Given your answers to questions 1 and 2, what are the implications of this particular network for the claim that perceptrons are unable to represent solutions to linearly nonseparable problems?**

Clearly, it is not true that perceptrons in general are incapable of learning linearly nonseparable problems, because this perceptron learned XOR. In order to say whether a perceptron will have trouble (or not) with a problem, one must also know the properties of the activation function. A perceptron with a Gaussian activation function can learn XOR.

4. **Remembering that this network uses the Gaussian equation described in Chapter 9, and that the bias or threshold of the output unit is equal to the value of μ in the equation, examine the two connection weights that feed into the output unit, as well as the bias of the unit. How does this perceptron compute this particular logical operation?**

In this network, when the incoming signal is opposite in value to μ the total net input is 0, and the unit will turn on. The value of μ is –0.86, and the two connection weights are both equal to 0.86. So, the only situation that will turn the unit on is when one, and only one, of the input units is on, producing a net input of –0.86 + 0.87 = 0.01, which is close enough to 0. If both input units are on, the net input is 0.88, and if both are off, the net input is –0.86 (i.e., μ); both these values are far enough away from 0 to turn the output unit off.

EXERCISE 13.2

1. **What was the sum of squared error at the end of training?**

The total SSE at the end of training was approximately 0.02.

2. **How many epochs of training were required before the program stopped training the network?**

The network converged after 35 sweeps.

3. **Given your answers to questions 1 and 2, what are the implications of this particular network for the claim that perceptrons are unable to represent solutions to linearly nonseparable problems?**

Again, a perceptron with a Gaussian activation function was able to solve a linearly nonseparable problem. This indicates that one can only make claims about what perceptrons can or cannot do by considering the properties of their activation functions.

4. **Remembering that this network uses the Gaussian equation described in Chapter 9, and that the bias or threshold of the output unit is equal to the value of μ in the equation, examine the two connection weights that feed into the output unit, as well as the bias of the unit. How does this perceptron compute this particular predicate?**

The bias is very small, equal to 0.02. The weights from the three horizontal line units are essentially zero. The weights from "kitty corner" vertical units are equal in magnitude, but opposite in sign. The only time that the unit will turn on is when two of these "kitty corner" units are on at the same time, cancelling each other's signal out, and causing a total net input of about zero. This only happens for connected figures.

5. **Many would argue – legitimately – that this exercise is just a parlour trick. To get a sense of what I mean by this, answer the following question: What is the relationship between how this network solves the connectedness problem and how the previous perceptron solved the XOR problem? To answer this question, you should have already generated answers to question 4 in both this exercise and in the previous one.**

The logical structures of this problem and of XOR are the same. Both networks work by using signals to cancel out net input. In XOR, one signal cancels μ, but two signals or no signals do not. In the second problem, signals from paired units (units that must both be on in order to make a connected figure) cancel each other out to turn a unit on; however, units that are turned on for unconnected figures do not cancel out, therefore turning off the output unit.