WORKSHEET FOR EXERCISES FROM CHAPTER 21

RECORD YOUR DATA FROM THE FIRST STUDY IN TABLE 21-1

Run Number	Number Of Hidden Units	Converged?	Total Sweeps
1	5		
2	5		
3	5		
4	5		
5	5		
6	4		
7	4		
8	4		
9	4		
10	4		
11	3		
12	3		
13	3		
14	3		
15	3		
16	2		
17	2		
18	2		
19	2		
20	2		

Table 21-1. Record for the results of the first simulation.

EXERCISE 21.1

1. **On the basis of the results that you have recorded in Table 21-1, what would you conclude about the number of hidden value units that are required to solve this problem? What evidence would you use to defend this conclusion?**

2. **Examine the "sweeps to converge" column for the simulations that converged in Table 21-1. Can you make any conclusions about what the effect of reducing the number of hidden units is on the amount of time that is required to learn the problem? If you believe that there is an effect, then why do you think this effect emerges?**

3. **Imagine that you wanted to determine how many hidden units are required to solve this problem without conducting an experiment like the one that you have just performed. Speculate about how you might predict the minimum number of hidden units that are required.**

CREATE YOUR VERSION OF TABLE 21-2 BY RECORDING DATA FOR THE SECOND SIMULATION

Run Number	Total Sweeps	Converged?	SSE At End	Misses At End
1				
2				
3				
4				
5				
6				
7				
8				
9				
10				

Table 21-2. Record of your data for the second exercise.

EXERCISE 21.2

1. **From your simulation records, what would you estimate the probability of having a network converge to a solution when these settings are used?**

2. **Can you explain why in the majority of runs the network fails to solve this problem, but in some cases a solution is found? What is the difference between runs that might affect whether the network will converge or not?**

3. **How many hidden value units would you now say are required to solve the 5-bit parity problem?**

4. **Assuming that you were able to train a network to solve the problem, examine its connection weights and biases, as well as the responses of the two hidden units to each pattern that was presented. Can you explain how this network is solving the parity problem? If you cannot come up with an explanation – and it might be difficult! – can you describe the source of the problem?**