Second Exam in Learning Systems (CDT407)

Date	2012-06-11
Time	8:10 – 13:30
Allowed material	None
Examiner	Ning Xiong (Phone 151716 /070 8701752)
Grades	0-19 U (did not pass)
	20-27 3 (pass)
	28-34 4 (very good)
	35-40 5 (excellent)

Good Luck!!!

1. Decision Tree Learning (4+4p)

Construct the decision tree for the set of training instances in the left table with attributes x_1, x_2 and classes +, –.
a) Which attribute should be used at the root node and why?
b) Show how the examples D_1, …., D_{10} are sorted down the tree and how they are classified.

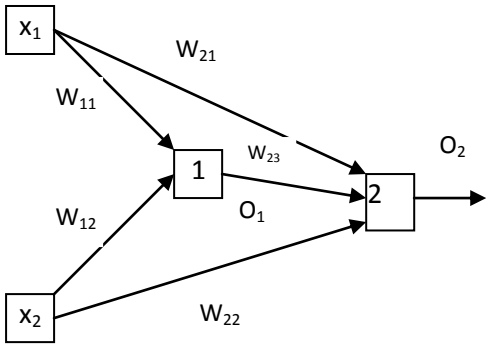
Use the right table to look up the entropy for a subset of training instances. Take the entry that is closest to the decimal number for which you want to calculate the entropy (e.g. if you need to calculate the entropy of 0.33, take the entry for 0.3 in the table which is 0.9). It is sufficient to make approximate calculations rounded to one digit behind the decimal.

Training examples

No.	x_1	x_2	$c(x_1,x_2)$
D_1	0	0	+
D_2	0	0	+
D_3	0	0	+
D_4	0	0	+
D_5	0	1	-
D_6	0	1	-
D_7	0	1	-
D_8	1	0	-
D_9	1	1	+
D_{10}	1	1	+

p	entropy(p)
0.0	0.0
0.1	0.5
0.2	0.7
0.3	0.9
0.4	1.0
0.5	1.0
0.6	1.0
0.7	0.9
0.8	0.7
0.9	0.5
1.0	0.0

2. Artificial Neural Network (5p)



Consider a neural network in the above. The outputs from the two units are given by

$$O_1 = sigmoid(w_{11}x_1 + w_{12}x_2 + c_1)$$

$$O_2 = sigmoid(w_{21}x_1 + w_{23}O_1 + w_{22}x_2 + c_2)$$

Now you are given a training example (x_{10}, x_{20}, t). Suppose that, under this training example, the outputs of the two units are O_{10} and O_{20} respectively, and the current value for weight w_{23} is *s*. The question is how to update the weights and thresholds for the neuron units in light of this training example? Assume the learning rate is defined as γ, your task is to calculate Δw_{ij} and Δc_i in terms of the incremental BP algorithm.

3.Genetic algorithms (2+2+2+2p)

a) Suppose a population has six individuals whose fitness values are illustrated in the table as follows

Individual	Fitness
1	8
2	20
3	13
4	25
5	9
6	25

What are the probabilities of selection for these individuals?

b) Now go to the issue of individual selection based on these selection probabilities. For this purpose a uniform random number from [0, 1] is created. Suppose this created number is equal to 0.7, which individual in the population should be selected according to the roulette wheel scheme and why?

c) What is the role of mutation in genetic algorithms? What would happen without mutation?

d) Now consider mutation on real numbers. It is done in terms of a normal density function. The cumulative probability values derived from this normal density function is given in the following table. Suppose a uniform random number is generated as 0.7736, how should you change the original real number (you need to indicate whether to increase or decrease the original real number and how much)?

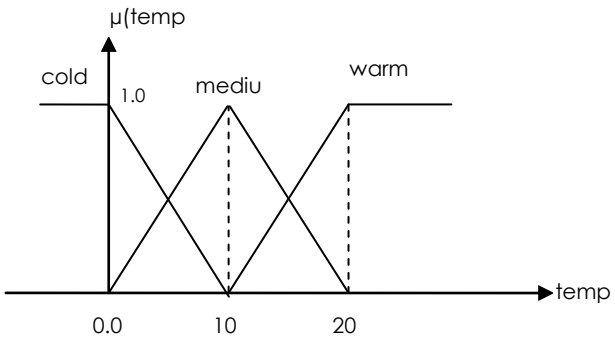
y	P(x<=y)
……..	……..
0.71	0.7611
0.72	0.7642
0.73	0.7673
0.74	0.7704
0.75	0.7734
0.76	0.7764
0.77	0.7794
0.78	0.7823
0.79	0.7852
0.80	0.7881
……	……

4. Fuzzy Systems (2+2+2+2p)

Consider three fuzzy rules to decide whether to turn on heating or not given current temperature:

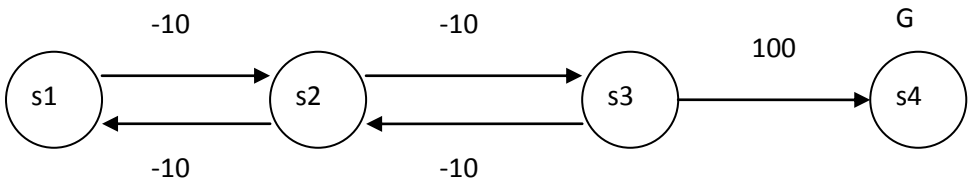
R1: If temp = *cold* Then heating
R2: If temp = *medium* Then heating
R3: If temp = *warm* Then no heating

The membership functions for the linguistic terms *cold*, *medium*, and *warm* are defined as shown in the figure. Suppose the current temperature is 8.0 degree.



a) What are the firing strengths of these fuzzy rules under the current temperature?
b) What are the output fuzzy sets suggested by the fuzzy rules in the current situation?
c) What are the overall output fuzzy set according to the whole fuzzy rule set?
d) What is your final decision and why?

5. Reinforcement learning in deterministic environments (2+2+3p)
Consider the deterministic environment of four states s1, …., s4 as shown in the figure. The state s4 is an absorbing terminal state. In state s1 there is only one way of move till right state s2, while for states s2 and s3 the agent can choose between two alternative actions (right or left move). The agent obtains a reward of +100 when entering into the terminal state s4 and a penalty of -10 for all other situations. Let the discount factor be 0.8.



a) What is the optimal policy of action in the states?
b) What are the best cumulative values $V^*(s)$ that can be achieved from individual states?
c) Suppose that the previous Q table contains full zero entries, how can you update some Q estimatess in the Q table in light of the recorded sequence s1 → s2 → s3 → s4?

6. Reinforcement learning in stochastic environments (2+2p)
Write out the Q-learning rule in a stochastic environment. Please also explain how this Q-learning rule is derived.