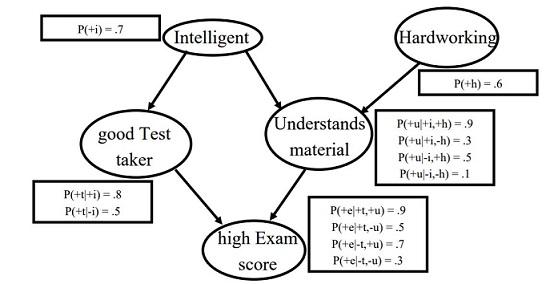
**Assignment 4**

**Question 1: Bayesian Networks (BN)**



Consider the above network:

a) Compute the probability that a student who did well on the test actually understood the material, that is, compute P (+u| + e).

|  |  |  |  |
| --- | --- | --- | --- |
| +t | +u | +e | -e |
| T | T | 0.9 | 0.1 |
| T | F | 0.5 | 0.5 |
| F | T | 0.7 | 0.3 |
| F | F | 0.3 | 0.7 |

P(+u | +e) = P(+u, +e) / P(+e)

= (P(+u | +e) \* P(+e)) / (P(+e | +t) \* P(+t)) + (P(+e | +u) \* P(+u))

= (0.9 \* 0.7) / ((0.5 \* 0.56) + (0.7 \* 0.51))

= 0.989

b) Are T and U independent? Explain your answer.

* Yes because both nodes T and U are conditionally independent. All the paths between the nodes A and B are separated by node E. They don’t connect to each other directly.

c) Compute the probability that a student has high exam score, that is, P (+e).

P(+t) = P(+t, +i)

= P(+t | +i) \* P(+i)

= 0.8 \* 0.7

= 0.56

P(+u) = P(+u, +i, +h)

= (P(+u | +i) / P(+i)) + (P(+u | +h) / P(+h))

= (0.3 \* 0.7) + (0.5 \* 0.6)

= 0.51

P(+e) = P(+e, +t, +u)

= (P(+e | +t) \* P(+t)) + (P(+e | +u) \* P(+u))

= (P(+e | +t) \* P(+t)) + (P(+e | +u) \* P(+u))

= (0.5 \* 0.56) + (0.7 \* 0.51)

= 0.637

d) Compute the probability that a student is intelligent given the student is a good test taker.

P(+i | +t) = (P(+t | +i) \* P(+i)) / P(+t)

= (0.8 \* 0.7) / (0.8 + 0.5)

= 0.431

e) Compute the probability that a student is a hard working given the student is intelligent and understands material.

P(+h | +i, +u) = (P(+u, | +h, +i) \* P(+i | +h) \* P(+h)) / (P(+u | +i) \* P(+i))

= (P(+u, | +h, +i) \* (P(+i) + P(+h)) \* P(+h)) / ((P(+u, +i)/P(+i)) \* P(+i))

= (P(+u, | +h, +i) \* (P(+i) + P(+h)) \* P(+h)) / P(+u, +i)

= (0.9 \* (0.7 + 0.6) \* 0.6) / (0.9 + 0.3)

= 0.585

f) Compute the probability that a student is a good test taker given the student is intelligent.

P(+t | +i) = (P(+i | +t) \* P(+t)) / P(+i)

= ((P(+t | +i) \* P(+i))/P(+t)) \* P(+t)) / P(+i)

= P(+t | +i)

= 0.8

**Question 2 Fuzzy Inference System**

1. Consider the fuzzy rule, R : if u is A and v is B then w is C, where A=(0, 2, 4), B=(3, 4, 5) and C=(3, 4, 5)

(Note: (x, y, z) denotes the triangle vertices for each fuzzy variable)

Find inference result C' when input is a =3, b=4 by using Larsen method.

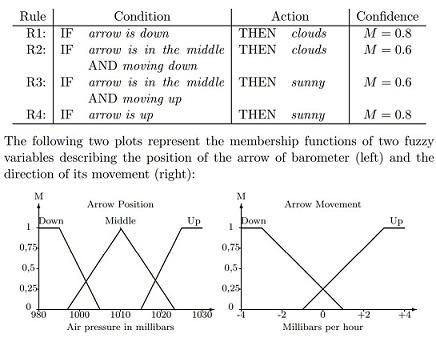
R = (A and B) → C

A’ = 3 / (2 + 4) = 0.5

B’ = 4 / (3 + 5) = 0.5

C’ = 4 / (3 + 5) = 0.5

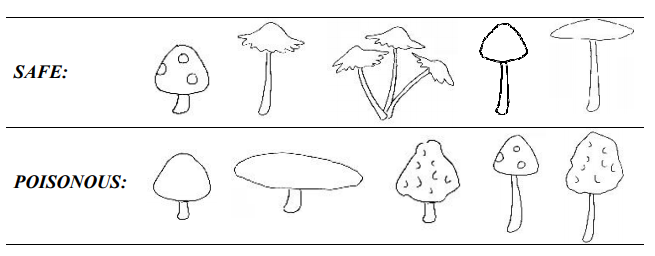
C’ = 0.5

****

b) Consider the above fuzzy expert system for weather forecast with confidence defining the membership degree. The air pressure is measured in millibars, and the speed of its change in millibars per hour. Answer the following questions:

1. How much is the arrow **Down**, **Up** or in the **Middle** if it indicates that the pressure is 1020 millibars? Use membership functions on the graphs.
   1. Down = 0
   2. Middle = 0.25
   3. Up = 0.5
2. How much is the arrow moving **Down** or **Up** if the pressure changes −2 millibars ever hour? Use membership functions on the graphs.
   1. Down = 0.75
   2. Up = 0
3. Using the membership values found above and confidences of the rules in the table calculate the degree of confidence in that the sky is clear or cloudy.
   1. input for air pressure is 1020 → max membership function value is Up for 0.5
   2. input for pressure change is -2 → max membership function value is Down for 0.75
   3. aggregation value = ((0.5 \* 0.8) + (0.75 \* 0.8)) / 2 = (0.4 + 0.6) / 2 = 0.5

**Question 3: Data Mining (DM)**

Consider the task of learning to identify mushrooms that are SAFE or POISONOUS to eat based on a set of physical features. Four Boolean and discrete valued features that you could use are: STEM = {short, long}, BELL = {rounded, flat}, TEXTURE = {plain, spots, bumpy, ruffles}, and NUMBER = {single, multiple}. Consider using these features on the following training data:

Use information gain to choose between TEXTURE and NUMBER to use in a decision stump (a decision tree with only one internal node at the root). Draw a diagram of the learned stump, and break ties at classification nodes by labeling them as POISONOUS (just to be on the safe side). Show all your work for partial credit.



S = [6, 4]

= Entropy = [6/10(+), 4/10(-)]

= - 3/5(log2)(3/5) - 2/5(log2)(2/5)

Info of all training samples = 0.97

****

Info[5, 4] Info[0, 1]

= Entropy[5/9(+), 4/9(-)] = 0 bits

= - 5/9 (log2)(5/9) - 4/9(log2)(4/9)

= 0.99 bits

Average weighted info of subtree = (0.99 \* 9/10) + (0 \* 1/10) = 0.89

Gain (S, Outlook) = 0.97 - 0.89 = 0.08

****

Info[3, 1] [2, 0] = 0 bits

Entropy[3/4(+), 1/4(-)]

= - 3/4 (log2)(3/4) - 1/4(log2)(1/4)

= 0.81 bits

Info[2, 2] Info[0, 0] = 0 bits

= Entropy[2/4(+), 2/4(-)]

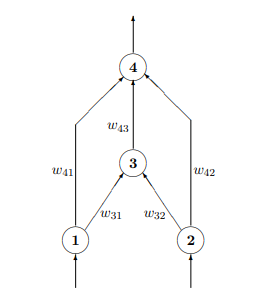
= - 1/2 (log2)(1/2) - 1/2(log2)(1/2)

= 1 bits

Average weighted info of subtree = (0.81 \* 4/10) + (1 \* 4/10) + (0) + (0 \* 2 / 10) = 0.72

Gain (S, Outlook) = 0.97 - 0.72 = 0.25

**Question 4: Neural Networks (NN)**



Consider the above multi-layer perceptron network, where all the units have binary inputs (0 or 1) and binary outputs (0 or 1). The weights for this network are w31=1, w32=1, w41=-1, W42=-1 and w43 =3. The threshold of both input units (1 and 2) is 0.5, so the output of these units is exactly the same as the input. The threshold of the hidden unit (3) is 1.5 and the threshold of the output unit (4) is −0.5. Which of the following Boolean functions can be computed by this network?

1. AND , b. OR, c. XOR, d. All of the above answers, e. None of the above answers

Show the steps of your solution.

-------------------------------------------------------------------------------------------------------------------------------------------

Weights are

w31=1, w32=1, w41=-1, w42 =-1, and w43 =3

input (1 and 2 = 0.5)

activation function for 3 = (input 1 \* weight of 31) + (input 2 \* weight of 32)

= (1 \* 0.5) + (1 \* 0.5)

= 1

activation function for 4 = (input 1 \* weight of 41)

= (-1 \* 0.5)

= -0.5

activation function for 4 = (input 2 \* weight of 42)

= (-1 \* 0.5)

= -0.5

1 is enough to pass the excitation point to use the hidden neuron of 3 since input 1 and 2 together with the identical weights have an activation function that is lower than the the threshold value of 3 which is 1.5.

However, both 1 and 2 can use their respective weights of 41 or 42 to reach the output neuron 4. So through this both paths input can reach the output allowing the XOR function to be calculated.

Answer is c. XOR

**Question 5: Genetic Algorithms (GA)**

1. Show, in pseudo code, a simple genetic algorithm with a brief description of each of the main elements.

// basic genetic algorithm

// generate initial population of chromosomes

population = generateMatrixOfChromosomes(numberOfChromosomes)

generation = 0

maxGeneration = 10

fitnessChromosomeMap = {}

// while we haven’t found the ideal chromosome and still within the range of generations keep looking

while generation < maxGeneration

generation += 1

chromosomeRow = 0

// while we have more chromosomes

while chromosomeRow < population.length

// determine the fitness value of each chromosome

fitnessChromosomeMap[chromosomeRow] = calcFitness(population[chromosomeRow])

// filter the list down by eliminating the weakest fit chromosomes

population = filterListForWeakestFitChromosomes(fitnessChromosomeMap)

// apply crossover and mutation logic to the newly filtered chromosomes

population = applyCrossoverAndMutation(population)

fitnessChromosomeMap = {}

return sortByHighestFitnessValue(population).first();

1. Mention the two requirements a problem should satisfy in order to be suitable for solving it by a GA.

* Fitness function must be defined
* Problem solutions can be broken down into steps which will then be encoded as chromosomes.

1. Suppose a genetic algorithm uses chromosomes of the form x = *abcdefgh* with a fixed length of eight genes. Each gene can be any digit between 0 and 9. Let the fitness of individual x be calculated as:

*f(x) = (a + b) − (c + d) + (e + f) − (g + h) ,*

and let the initial population consist of four individuals with the following chromosomes:

*x1 = 6 5 4 1 3 5 3 2*

*x2 = 8 7 1 2 6 6 0 1*

*x3 = 2 3 9 2 1 2 8 5*

*x4 = 4 1 8 5 2 0 9 4*

1. Evaluate the fitness of each individual, showing all your workings, and arrange them in order with the fittest first and the least fit last.

Determine Fitness values

* f(x1) = (6 + 5) − (4 + 1) + (3 + 5) − (3 + 2) = 9
* f(x2) = (8 + 7) − (1 + 2) + (6 + 6) − (0 + 1) = 23
* f(x3) = (2 + 3) − (9 + 2) + (1 + 2) − (8 + 5) = −16
* f(x4) = (4 + 1) − (8 + 5) + (2 + 0) − (9 + 4) = −19

The order from fittest to least fit are

x2, x1, x3, x4

1. Perform the following crossover operations:
2. Cross the fittest two individuals using one–point crossover at the middle point.
   1. x2 = 8712 | 6601, x1 = 6541 | 3532
   2. After cross: x2 = 87123532, x1 = 65416601
3. Cross the second and third fittest individuals using a two–point crossover (points b and f).
   1. x1 = 65 | 4135 | 32, x3 = 23 | 9212 | 85
   2. After cross: x1 = 65921232, x3 = 23413585
4. Cross the first and third fittest individuals (ranked 1st and 3rd) using a uniform crossover.
   1. Swap based on a, d, and f of parents x2 and x3
   2. x2 = **8**71**2**6**6**01, x3 = **2**39**2**1**2**85
   3. After cross: x2 = **2**71**2**6**2**01, x3 = **8**39**2**1**6**85

iii. Suppose the new population consists of the six offspring individuals received by the crossover operations in the above question. Evaluate the fitness of the new population, showing all your workings. Has the overall fitness improved?

New crossover chromosomes:

* O1 = 8 7 1 2 3 5 3 2
* O2 = 6 5 4 1 6 6 0 1
* O3 = 6 5 9 2 1 2 3 2
* O4 = 2 3 4 1 3 5 8 5
* O5 = 2 7 1 2 6 2 0 1
* O6 = 8 3 9 2 1 6 8 5

Determine new fitness values

* f(O1) = (8 + 7) − (1 + 2) + (3 + 5) − (3 + 2) = 15
* f(O2) = (6 + 5) − (4 + 1) + (6 + 6) − (0 + 1) = 17
* f(O3) = (6 + 5) − (9 + 2) + (1 + 2) − (3 + 2) = −2
* f(O4) = (2 + 3) − (4 + 1) + (3 + 5) − (8 + 5) = −5
* f(O5) = (2 + 7) − (1 + 2) + (6 + 2) − (0 + 1) = 13
* f(O6) = (8 + 3) − (9 + 2) + (1 + 6) − (8 + 5) = −6

Yes, overall fitness values improved.

iv. By looking at the fitness function and considering that genes can only be digits between 0 and 9 find the chromosome representing the optimal solution (i.e. with the maximum fitness). Find the value of the maximum fitness

1. By looking at the initial population of the algorithm can you say whether it will be able to reach the optimal solution without the mutation operator?
   1. It varies based on the scenario. Mutation helps add genetic diversity to the population by changing values with a certain level of probability. However, there is no guarantee that these modifications to the traits will lead to a survival advantage or still exist in later generations.
2. Use Matlab and Simulink or a similar tool to simulate the above problem. Provide screenshots with explanation for your work.

Using the DEAP python library I generated a list of chromosomes that each 8 randomly generated numbers between 0 and 9. I then applied the fitness function in the above example as a measurement to each of the chromosomes to then be able to order them by the max fitness value. The ordering was done in every generation so that at the end of the 10th generation we had the highest max value. With each generation a crossover and mutation was also done based on initial rate to maximize the ability of the genetic algorithm to find the best combination of chromosome values.

