Departamento de Informática
Universidad Carlos III de Madrid

Grado en Ingeniería Informática

Artificial Intelligence

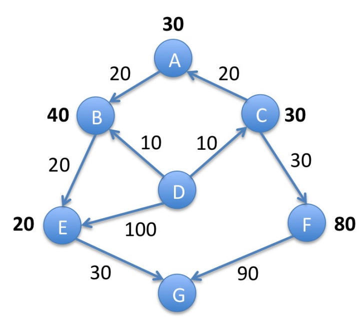
Mayo 2013

General instructions

- Time assigned to the exam is **2 hours**
- Teachers will not answer any question about the exam
- You cannot leave the classroom during the exam, unless you have finished it
- Exams cannot be answered using a pencil

1. **(1 point)** What will be the main problem when using breadth-first search to solve any high-complexity task? And what are its main advantages over depth-first search?

2. **(1 point)** In the following graph, we provide information about the heuristic value of each node, $h(\cdot)$, and the cost of each operator (arc), $k(n_i, n_j)$. Describe the first four cycles of the A* algorithm (with check for repeated states) when solving the problem of going from node D to node G. Show separately each of the resulting trees and clearly show the values of the relevant functions at each node.



3. **(1 point)** What algorithm(s) would you use if you had not a heuristic value (or function) and would like to compute the optimal solution in terms of cost?

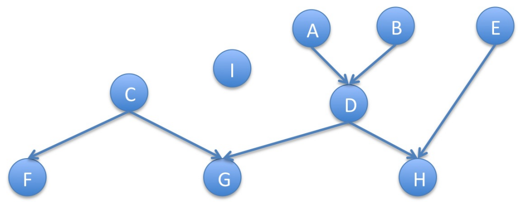
4. **(1 point)** What is the main property that we use to reduce complexity in computing joint probabilities when using bayesian networks?

5. **(1 point)** A factory of smartphones produces five defective phones for every 1000. They generate smartphones equipped with three different operating systems. Over 100 smartphones they build, 40 use iOS, 40 use Android, and the remaining use Windows Phone. They have done studies on failures when they are shown to customers in a demo. If the phone is defective, then 50% of them fail if they use iOS, 40% fail if they use Android, and 70% fail if they use Windows Phone. If the phone is not defective, then those failures are reduced to 0.5% for iOS, 0.3% for Android and 0.9% for Windows Phone. Finally, if customers encounter a failure in a demo, 30% of the customers will buy the phone anyway. Define the corresponding bayesian network with the variables, their relationships and conditional probability tables.

6. **(1 point)** What does the value function $V(s)$ express for Markov Decision Processes? What method would you use to compute it if you know the state transition and reward functions?

7. **(1 point)** Suppose you want to control the air conditioning system at home with a fuzzy system. You can measure the humidity of the room (low, high) and the temperature (low, medium, high). And you can control whether to increase or decrease the air conditioning temperature. Define graphically the needed membership functions. You can choose the shape and limits of each function, but they should clearly appear in the graphics. Define two fuzzy rules to control such a system.

8. **(1 point)** Given the following bayesian network, formally define the joint probability of $P(A, B, C, D, E, F, G, H, I)$.



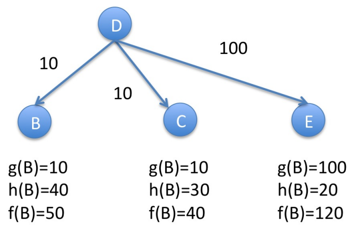
9. **(1 point)** If you would like to use a genetic algorithm for solving an optimization problem: what kind of representation would you use for individuals? what three operators would you most probably use? what do you have to define for assessing the value of each individual?

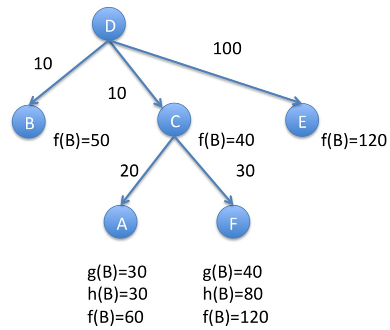
10. **(1 point)** When you work for a car manufacturer company, you are in charge of the development of an automatic recogniser of traffic signs. One of the subtasks to build such system consists on: given an image of the road, your software must return true or false depending on whether there is a stop sign or not in the image. Very easily, images can be analyzed to obtain three vectors (histograms) of 256 values each for the three main colors (red, green and blue). You decide to use a bayesian-based machine learning technique for this task, Naïve Bayes. Define formally the training instances (examples). What is the main assumption made by Naïve Bayes?

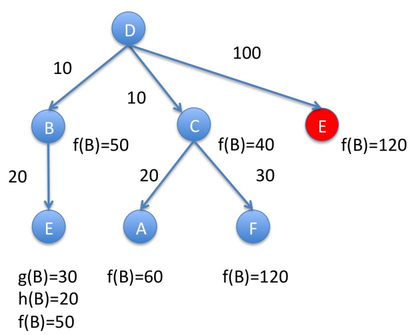
Solutions.
May 2013

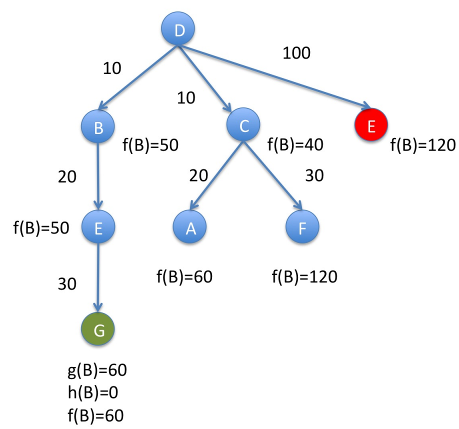
1. Breadth-first search has exponential space requirements, in the worst case. In comparison with depth-first search, it is admissible and complete, given uniform costs and finite number of successors.

2. The first four (and only) steps in running the A* in the example graph are shown next.









3. Either Dijkstra or Branch-and-bound

4. Conditional independence

5. Figure 1 shows the corresponding BBN and its CPTs. Since the question did not specify the last probability, it is unknown.

6. $V(s)$ represents the expected discounted reward over time, or expected cost. If we have the model (state transition and reward functions), we can use Value Iteration, or dynamic programming to solve the corresponding equations.

7. Figure 2 shows examples of possible representations of the fuzzy membership functions for those variables and values.

   Two examples of fuzzy rules in this context would be:

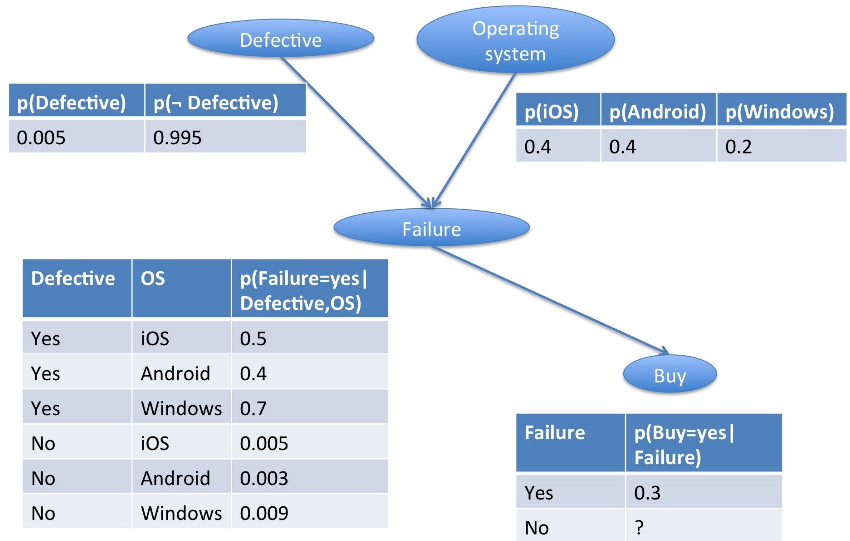


Figure 1: Solution BBN and its CPTs.

- If humidity is low and temperature is high then increase temperature
- If humidity is high and temperature is medium then decrease temperature

8. $p(A, B, C, D, E, F, G, H, I) = p(A)p(B)p(C)p(E)p(I)p(D \mid A, B)p(H \mid D, E)p(G \mid C, D)p(F \mid C)$

9. Genetic algorithms use a string of bits to represent each individual. The most used genetic operators are: mutation, crossover and selection. In order to assess each individual, genetic algorithms use a fitness function.

10. Naïve Bayes, as many other machine learning techniques, take as input a set of instances. Each instance is represented as the values for a set of attributes (characteristics) and a class. In this case, each instance would represent an image taken, there would be 255×3 attributes, and the class would be true or false. The main assumption made by Naïve Bayes is that the values of attributes are independent of each other.

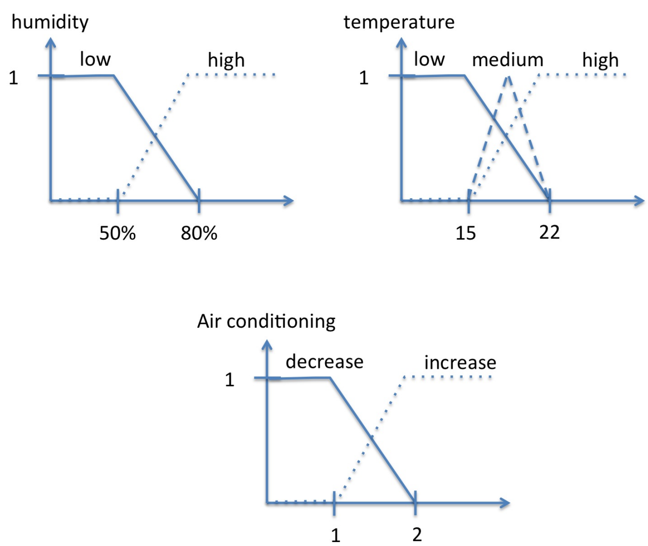


Figure 2: Examples of fuzzy membership functions.