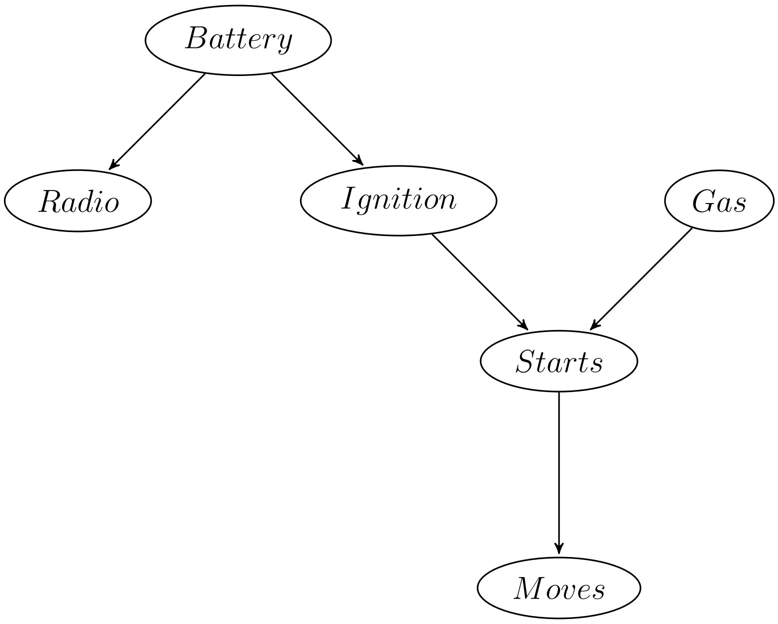
COSC1125/1127 Artificial Intelligence

School of Computer Science and IT
RMIT University
Semester 1, 2019

Tutorial Sheet 11
Bayesian Networks

1. Consider the network for car diagnosis shown in the figure below:

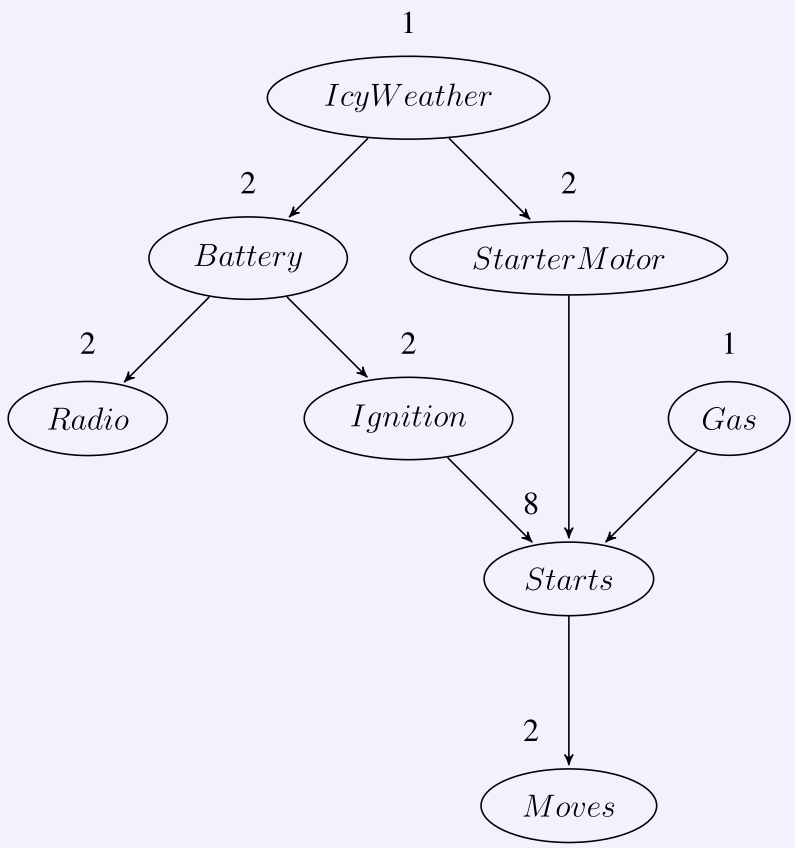


   (a) Extend the network with the Boolean variables *IcyWeather* and *StarterMotor*, by reasoning how components might affect each other. Being a design/modelling process, there may be many correct answers, just state your assumptions.

Answer

Adding variables to an existing net can be done in two ways. Formally speaking, one should insert the variable ordering and rerun the network construction process from the point where the first new variable appears. Informally speaking, one never really builds a network by a strict ordering. Instead, one asks what variables are direct causes or influences on what other ones, and builds local parent/child graphs that way. It is usually very easy to identify where in such a structure the new variable goes, but one must be very careful to check for possible induced dependencies downstream.

$IcyWeather$ is not caused by any of the car-related variables, so needs no parents. It directly affects the battery and starter moter. $StarterMotor$ is an additional precondition for $Starts$. The new network is shown below:



(b) Give reasonable conditional probability tables for all the nodes.

Answer

i. A reasonable prior for $IcyWeather$ might be 0.05 (perhaps, depending on the location and season).

ii. $P(Battery|IcyWeather) = 0.95$,
$P(Battery|\neg IcyWeather) = 0.997$

iii. $P(StarterMotor|IcyWeather) = 0.98$,
$P(StarterMotor|\neg IcyWeather) = 0.999$

iv. $P(Radio|Battery) = 0.9999$,
$P(Radio|\neg Battery) = 0.05$

v. $P(Ignition|Battery) = 0.998$,
$P(Ignition|\neg Battery) = 0.01$

vi. $P(Gas) = 0.995$

vii. $P(Starts|Ignition, StarterMotor, Gas) = 0.9999$

viii. $P(Moves|Starts) = 0.998$

(c) How many independent values are contained in the joint probability distribution for eight Boolean nodes, assuming that no conditional independence relations are known to hold among them?

Answer

With 8 Boolean variables, the joint has $2^8 - 1 = 255$ independent entries. In general there are 2^n entries in a joint probability distribution; however, only $2^n - 1$ *independent* entries. If the probabilities of all of the independent entries are summed up, the probability of the final entry must be $1 - \sum P(independent)$, making its value *dependent* on the rest.

(d) How many independent probability values do your network tables contain?

Answer

Given the topology shown in the figure from question 1a, the total number of independent CPT entries is $1 + 2 + 2 + 2 + 2 + 1 + 8 + 2 = 20$

(e) Reconstruct, showing your workings, five instances of the full joint probability distribution (e.g., reconstruct $P(\neg iW, b, r, i, g, s, sM, \neg m)$).

Answer

Here is one example. The probabilities are taken from question 1(b).

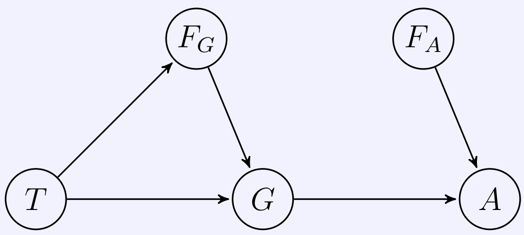
$$
\begin{aligned}
&P(\neg iW, b, r, i, g, s, sM, \neg m)\\
&= P(\neg iW) \cdot P(b \mid \neg iW) \cdot P(r|b) \cdot P(i \mid b) \cdot P(g) \cdot P(s \mid i, sM, g) \cdot P(\neg m \mid s)\\
&= (1 - P(iW)) \cdot P(b \mid \neg iW) \cdot P(r|b) \cdot P(i \mid b) \cdot P(g) \cdot P(s \mid i, sM, g) \cdot (1 - P(m \mid s))\\
&= 0.95 \cdot 0.997 \cdot 0.9999 \cdot 0.998 \cdot 0.995 \cdot 0.9999 \cdot 0.002\\
&= 0.00188
\end{aligned}
$$

2. In your local nuclear power station, there is an alarm that senses when a temperature gauge exceeds a given threshold. The gauge measures the temperature of the core. Consider the Boolean variable A (alarm sounds), F_A (alarm is faulty), and F_G (gauge is faulty) and the multivalued nodes G (gauge reading) and T (actual core temperature).

   (a) Draw a Bayesian network for this domain, given that the gauge is more likely to fail when the core temperature gets too high.

Answer

A suitable network is shown in the figure below. The key aspects are: the failure nodes are parents of the sensor nodes, and the temperature node is a parent of both the gauge and the gauge failure node. It is exactly this kind of correlation that makes it difficult for humans to understand what is happening in complex systems with unreliable sensors.



   (b) Is your network a polytree?

Answer

No, it is not a polytree, since if we replace the directed edges with undirected edges, we obtain an undirected graph which is cyclic. Therefore, no matter which way the network is drawn, it should not be a polytree because of the fact that the temperature influences the gauge in two ways.

   (c) Suppose there are just two possible actual and measured temperatures, normal and high; the probability that the gauge gives the correct temperature is x when it is working, but y when it is faulty. Give the conditional probability table associated with G.

Answer

The CPT for G is shown below. The wording of the question is a little tricky because x and y are defined in terms of "incorrect" rather than "correct".

	$T = Normal$		$T = High$	
	F_G	$\neg F_G$	F_G	$\neg F_G$
$G = High$	$1 - y$	$1 - x$	y	x
$G = Normal$	y	x	$1 - y$	$1 - x$

(d) Suppose the alarm works correctly unless it is faulty, in which case it never sounds. Give the conditional probability table associated with A.

Answer

The CPT for A is as follows:

	$G = Normal$		$G = High$	
	F_A	$\neg F_A$	F_A	$\neg F_A$
A	0	0	0	1
$\neg A$	1	1	1	0