Section 1. Uncertainty

U1

(2p) Which of the following statements are true?

- [x] If $P(a|b,c) = P(b|a,c)$, then $P(a|c) = P(b|c)$
- [ ] If $P(a|b,c) = P(a)$, then $P(b|c) = P(b)$
- [ ] If $P(a|b) = P(a)$, then $P(a|b,c) = P(a|c)$

U2

(2p) We have the following joint distribution for the Toothache, Cavity, Catch world:

	toothache		**¬ toothache**	
	catch	**¬ catch**	**catch**	**¬ catch**
cavity	0.108	0.012	0.072	0.008
¬ cavity	0.016	0.064	0.144	0.576

Compute the following probabilities. Write the probabilities using exactly 2 decimals and rounding with the floor function.

Example:

- 0.3 => Write: 0.30
- 0.789 => Write: 0.78
- 0.781 => Write: 0.78

$P(toothache) =$

0.20

$P(cavity) =$

0.20

$P(\neg toothache|cavity) =$

0.40

$P(\neg cavity|toothache \vee catch) =$

0.53

Section 2. Decision Trees

DT1

(2p)

Suppose we generate a training set from a decision tree and then apply decision-tree learning to that training set. Is it the case that the learning algorithm will eventually return the correct tree as the training-set size goes to infinity and if the method for generating examples eventually generates all possible combinations of input attributes?

- ○ Yes, the learning algorithm will return the exact initial tree.
- ◉ No, the trees may be different, but they necessarily are logically equivalent
- ○ No, the trees may be different and there is also no guarantee that the trees will be logically equivalent

DT2

(3.5p) We are given the following table:

..Other attributes..	F = Favourite Film Genre	C = Likes Classical Musical
...	**thriller**	**no**
...	**comedy**	**no**
...	**thriller**	**yes**
...	**drama**	**yes**
...	**comedy**	**no**
...	**comedy**	**yes**
...	**drama**	**yes**

Using the formulas from the course slides calculate the following. For the values of logarithms you can use wolframalpha.com. Write the values with exactly 2 decimals precision, using the floor function. Example:

- 0 => 0.00
- 1.3 => Write 1.30
- 23.789 => Write 23.78
- 23.781 => Write 23.78

The entropy of the F variable, H(F) =

1.55

The specific conditional entropies of variable C, given specific values of variable F:

H(C | F = thriller) =

1.00

H(C | F = comedy) =

0.91

H (C | F = drama) =

0.00

The general Conditional Entropy:

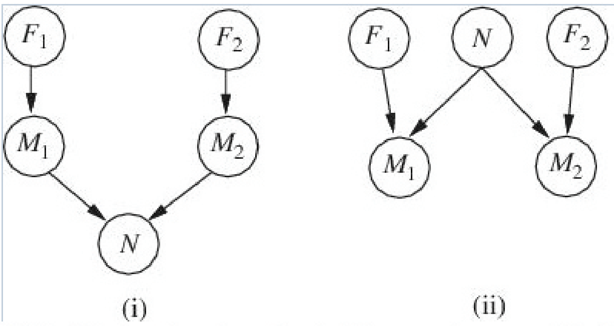
H (C | F) =

0.67

Section 3. Bayesian Belief Networks

BBN1

(3.5p) Two astronomers in different parts of the world make measurements M1 and M2 of the number of start N in some small region of the sky, using their telescopes. Normally, there is a small possibility *e* of error by up to one star in each direction. Each telescope can also (with a much smaller probability *f*) be badly out of focus (events F1 and F2), in which case the scientist will undercount by three or more stars (or if N is less than 3, fail to detect any stars at all). Consider the two networks in the figure:

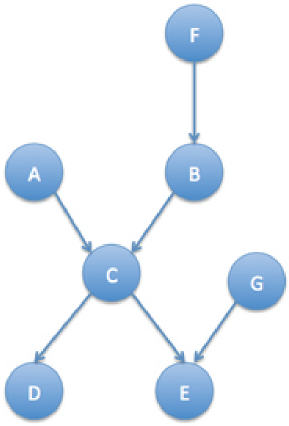


Which of the two Bayesian networks is/are correct representation(s) of the preceding information? Think about whether the independence and conditional independence and dependence rules that you can substract from the Bayesian nets are equivalent to what you should be able to deduce from the problem formulation. You can use D-separation (see Exercise the session slides).

- [ ] (i)
- [x] (ii)

BBN2

(3.5p) Take the following Bayesian network and construct its moral graph.

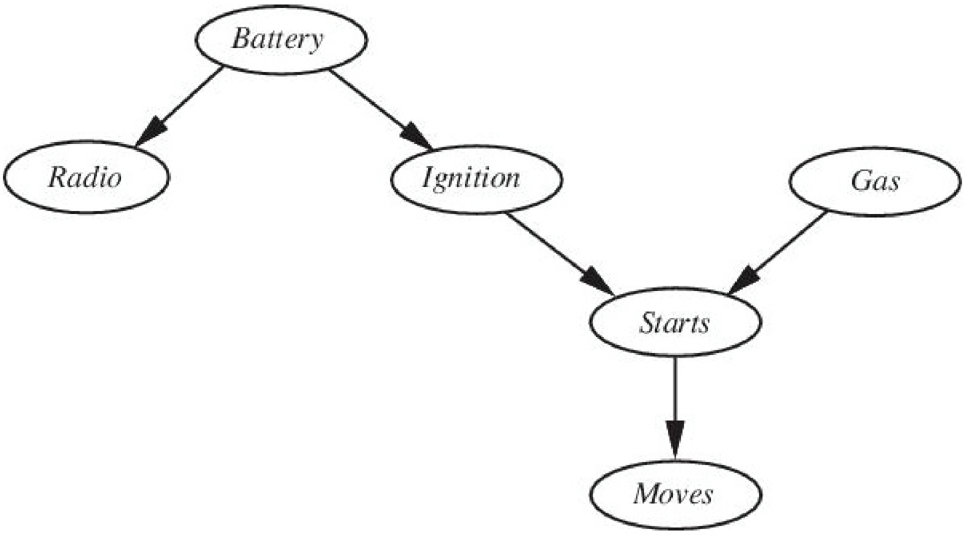


Which of the following are true? You can use the D-separation rules from the Exercise class slides

- [x] $A \perp G \mid C$ (A is independent of G given C)
- [x] In the moral graph, if we remove C, A and G will get disconnected
- [ ] $A \perp B \mid C$
- [ ] In the moral graph, if we remove C, A and B will get disconnected
- [x] $A \perp B$
- [ ] In the moral graph, A and B are disconnected
- [x] If we build a moral graph only for A and B and their ancestors, A and B will be disconnected

BBN3

(2.5p) What is the Markov blanket of the Ignition variable in the following network?

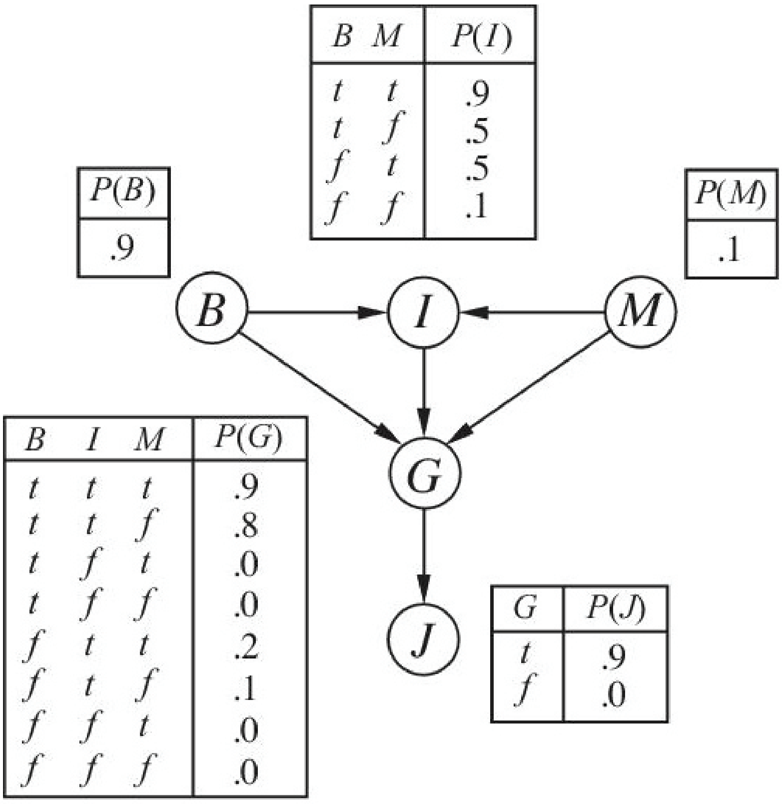


Check all the variables that are inside the Markov blanket.

- [x] Battery
- [ ] Radio
- [x] Gas
- [x] Starts
- [ ] Moves

BBN4

(2p) Consider the following simple Bayes network with Boolean variables: B = *BrokenElectionLaw*, I = *Indicted*, M =*PolliticallyMotivatedProsecutor*, G = *FoundGuilty*, J = *Jailed*.



$P(B)$
.9

B	M	$P(I)$
t	t	.9
t	f	.5
f	t	.5
f	f	.1

$P(M)$
.1

B	I	M	$P(G)$
t	t	t	.9
t	t	f	.8
t	f	t	.0
t	f	f	.0
f	t	t	.2
f	t	f	.1
f	f	t	.0
f	f	f	.0

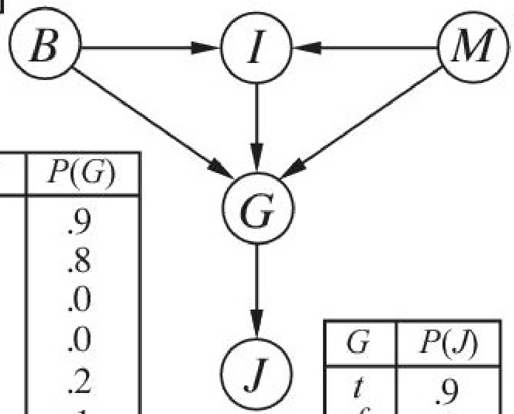
G	$P(J)$
t	.9
f	.0

Which, if any, of the following are asserted by the network structure (**ignore the tables** for now)?

- [ ] **P**(B, I, M) = **P**(B)**P**(I)**P**(M)
- [x] **P**(J|G) = **P**(J|G,I)
- [x] **P**(M|G, B, I) = **P**(M| G, B, I, J)

BBN5

(2p) Consider the following simple Bayes network with Boolean variables: B = *BrokenElectionLaw*, I = *Indicted*, M =*PolliticallyMotivatedProsecutor*, G = *FoundGuilty*, J = *Jailed*.



$P(B)$
.9

B	M	$P(I)$
t	t	.9
t	f	.5
f	t	.5
f	f	.1

$P(M)$
.1

B	I	M	$P(G)$
t	t	t	.9
t	t	f	.8
t	f	t	.0
t	f	f	.0
f	t	t	.2
f	t	f	.1
f	f	t	.0
f	f	f	.0

G	$P(J)$
t	.9
f	.0

Calculate the following probabilities. Write the probabilities using exactly 2 decimals and rounding with the floor function.

Example:

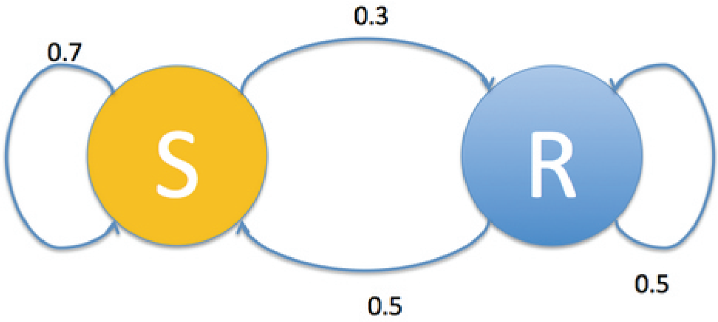
- 0.3 => Write: 0.30
- 0.789 => Write: 0.78
- 0.781 => Write: 0.78

P(b, i, ¬m, g, j) = 0.29 The probability that someone goes to jail **given** that they broke the law, have been indicted, and face a politically motivated prosecutor. 0.81

Section 4. Hidden Markov Models

HMM1

(2.5p) Suppose we have only sunny and rainy weather and we have the following transition diagram:



So, if the previous day it was sunny, it will stay sunny with 0.7 probability. If the previous day was rainy, it will stay rainy with 0.5 probability.

We don't know for sure how it was yesterday and we trust our friend's observation that it was sunny only 0.9. So, **P**(S0) = 0.9.

Calculate the following probabilities.

Write the probabilities using exactly **3 decimals!** and rounding with the floor function.

Example:

- 0.3 => Write: 0.300
- 0.64 => Write: 0.640
- 0.1789 => Write: 0.178
- 0.1781 => Write: 0.178

The probability that today it will be sunny.

P(S1) =

0.680

The probability that tomorrow it will be sunny

P(S2) =

0.636

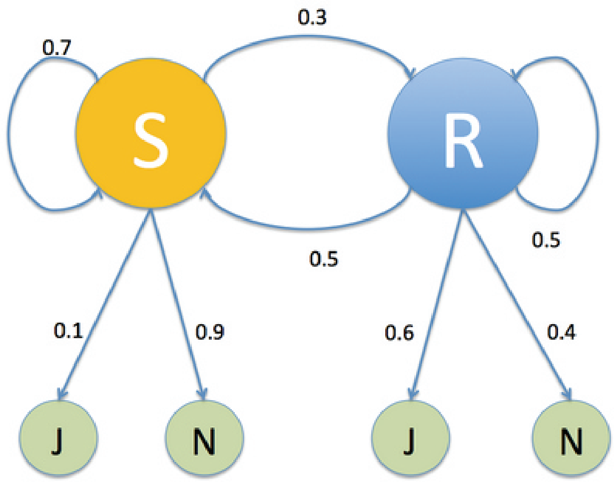
The stationary distribution for being sunny.

P(S∞) =

0.625

HMM2

(3.5p) Now, we add up the following observations:



If today is sunny, this morning our friend suffered from joint pain with probability **P**(Ji|Si) = 0.1 and had no joint pain with probability **P**(Ni|Si) = 0.9.

If today is rainy, then this morning he had joint pain with probability **P**(Ji|Ri) = 0.6 and no joint pain with probability **P**(Ni|Ri) = 0.4.

Given that yesterday was for sure sunny **P**(S0) = 1, and the observation that our friend had joint pain this morning, what is the probability that today it will be raining?

Write the probabilities using exactly **3 decimals!** and rounding with the floor function.

Example:

- 0.3 => Write: 0.300
- 0.64 => Write: 0.640
- 0.1789 => Write: 0.178
- 0.1781 => Write: 0.178

P(R1|J1) =

0.720