CS5011 Artificial Intelligence

Assignment 4: Reasoning with Uncertainty – Bayesian Networks

University of
St Andrews

Reasoning with Uncertainty – Bayesian Networks

1 Finished part

(1) Part 1

(2) Part 2

2 Literature Review

2.1 Bayesian networks

2.1.1 Brief introduction of Bayesian networks

A Bayesian network (BN) is a directed acyclic graphical model that encodes probabilistic relationships within distinctions of interest when dealing with uncertainty [1]. Bayesian networks, also called belief networks, is a well-known graphical model applied in expressing and reasoning uncertain knowledge, which is the expansion of Bayesian approach method [1]. The topology of a BN is a directed acyclic composed of variable nodes and directed lines connecting these nodes, of which nodes are random variables, and directed lines represent the relationships with nodes (parent nodes pointing to their children by an arrow, i.e. conditional dependence)[1]. Conditional probability is used to express how strong the relationships between nodes are. If a node has no a parent node, information can be represented by priori probability [1]. Variable nodes are abstractions of problems in real world. For instance, the temperature of a device, the gender of a patient, a feature of an object and the occurrence of an event[2].

2.1.3 Three types of structures in Bayesian network

Directed acyclic graphs are widely used to interpret causal relationships which is often used in Bayesian network [3]. There are three fundamental types of structures in Bayesian network.

(1) Head to head

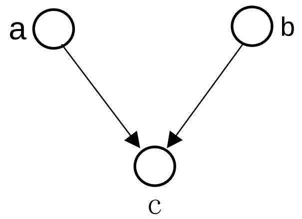


Figure 1

As is shown in Figure 1, when the c is unknown, a and b is blocked and independent, which is called head-to-head conditional independence. This can be described as:

$$P(a,b,c)=P(a)*P(b)*P(c|a,b)$$

➢ $P(a,b,c)=P(a,b)=P(a)*P(b)$ [3]

(2) Tail to tail

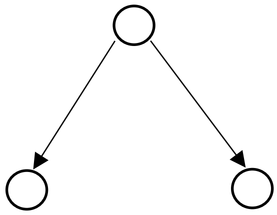


Figure 2

- As is shown in Figure 3, when c is unknown, this can be described as:

$$P(a,b,c)=P(a)*P(c|a)*P(b|c)$$

But it can be transferred to P(a,b)=P(a)P(b), namely, a and b are not independent when c is unknown.

- When c is known, this can be described as:

$$P(a,b|c)=P(a,b,c)/P(c) \text{ and } P(a,b,c)=P(a)*P(c|a)*P(b|c)$$

  - $P(a,b|c)=P(a,b,c)/P(c)$

$$= P(c)*P(a|c)*P(b|c)/ P(c)$$

$$= P(a|c) *P(b|c)$$

Thus, when c is known, a and b are blocked and independent, which is called tail-to-tail conditional independence [3].

(3) Head to tail

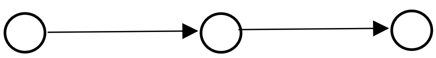


Figure 3

- As is shown in Figure 3, when c is unknown, this can be described as:

$$P(a,b,c)=P(a)*P(c|a)*P(b|c)$$

But it can be transferred to P(a,b)=P(a)P(b), namely, a and b are not independent when c is unknown.

- When c is known, this can be described as:

$$P(a,b|c)=P(a,b,c)/P(c) \text{ and } P(a,c)=P(a)*P(c|a)=P(c)*P(a|c)$$

  - $P(a,b|c)=P(a,b,c)/P(c)$

$$= P(a)*P(c|a)*P(b|c)/ P(c)$$

$$= P(a,c) *P(b|c)/ P(c)$$

$$= P(a|c) *P(b|c)$$

Thus, when c is known, a and b are blocked and independent, which is called head-to-tail conditional independence [3].

2.2 Current application of Bayesian network

Bayesian network is widely used in many different fields including protein structure, gene expression analysis and gene regulatory networks[4], medicine[5], document classification, information retrieval[6], semantic search[7], biomonitoring[8], decision support systems[9], image processing, data fusion, engineering, sports betting[10][11], gaming, geophysics and volcano monitoring [12], law[13][14], study design[15] and risk analysis[16], financial and marketing informatics[17].

Recently, Netflix apply to Bayesian network to business and optimize personalized recommendation system. The target of this system is to forward right content at the right time to each member. To implement this, Netflix uses contextual bandit counterparts with Thompson Sampling, LinUCB, or Bayesian methods that intelligently balance making the best prediction with data exploration [18].

3 Part 1

3.1 Design the Bayesian Network for car faults

In this car faults model, random variables and their corresponding finite countable domains are revealed as follows:

(1) Battery Age (BA)={new, old, very_old};

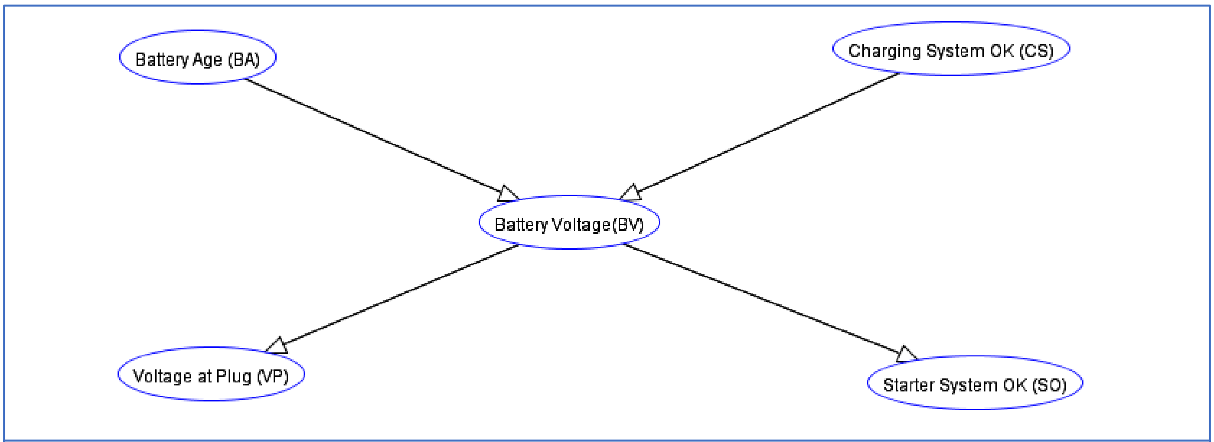
(2) Charging System OK (CS)={true, false};

(3) Battery Voltage (BV)={strong, weak, dead};

(4) Voltage at Plug (VP)={high, low};

(5) Starter System OK (SO)={true, false};

The figure 1 shows the topology of Bayesian Network (BN) of car faults model and conditional distributions for each node.



	P(Battery Age (BA)=new)	P(Battery Age (BA)=old)	P(Battery Age (BA)=very_old)
Prior Probability	0.4	0.4	0.2

	P(Charging System OK (CS)=true)	P(Charging System OK (CS)=false)
Prior Probability	0.99	0.01

Battery Age (BA)	Charging System OK (CS)	P(Battery Voltage(BV)=strong)	P(Battery Voltage(BV)=weak)	P(Battery Voltage(BV)=dead)
new	true	0.99	0.005	0.005
new	false	0.8	0.15	0.05
old	true	0.8	0.01	0.19
old	false	0.05	0.3	0.65
very_old	true	0.6	0.3	0.1
very_old	false	0.002	0.1	0.898

Battery Voltage(BV)	P(Voltage at Plug (VP)=high)	P(Voltage at Plug (VP)=low)
strong	0.9	0.1
weak	0.2	0.8
dead	0.0	1.0

Battery Voltage(BV)	P(Starter System OK (SO)=true)	P(Starter System OK (SO)=false)
strong	0.99	0.01
weak	0.7	0.3
dead	0.0	1.0

Figure 1 BN of car fault model

3.2 Analyse the relationships between variables in car faults BN

(1) BA and CS are cause of effect (BV), so BA and CS directly affect BV;

(2) When either BA is true or CS is new, BV is more likely to be strong;

(3) BV is cause of effects (VP and SO), so BV directly affects VP and SO;

(4) When BV is strong, VP is more likely to be high;

(5) When BV is week, VP is more likely to be low;

(6) When BV is dead, VP must be low;

(7) When BV is strong or weak, SO is more likely to be true;

(8) When BV is dead, SO must be false.

(9) BA and CS are independent, VP and SO are conditionally independent given by BV;

(10) P(BA) and P(CS) are prior probabilities. P(BV),P(VP) and P(SO) are posterior probabilities.

3.3 Evaluate results when observations are changed in car faults BN

3.3.1 Predict results when observations are changed in car faults BN

Note: when talking about increase or decrease, it compares with initial probability distribution in car faults BN.

(a) Diagnostic: the car starter system is not OK

(i) Predict BV

SO is influenced directly by BV. When BV becomes weaker, P(SO) being false increases. On the contrary, when SO is observed as false, it means the effect is caused by P(BV) getting weaker. In addition, even if BV is weak, SO still have higher likelihood to be true than to be false, which means if BV is false, it is more likely affected by BV becoming dead. Therefore, when the agent observes that the car starter system is not OK:

- P(BV=weak) increases slightly;
- P(BV=dead) increases dramatically;
- P(BV=strong) decreases;

(ii) Predict VP

VP is influenced directly by BV. When BV becomes weaker, P(SO) being low increases. After observation, as my prediction that BV becomes weaker, and P(BV) being dead increases dramatically, I predict that:

- P(VP=high) decreases dramatically;
- P(VP=low) increases dramatically;

(iii) Predict BA

BA directly influence BV, When BA becomes older, P(BV) being weak increases. On the contrary, when P(BV) being weak and dead increases, it means the effect is caused by P(BA) getting older. In addition, as initial probabilities shown in figure 1, it is apparent to see that when BA is new, P(BV) being dead is near to 0. Thus, If BV is dead, it P(BA) being new approaches to 0. I predict that:

- P(BA=new) decreases significantly;
- P(BA=old) increases;
- P(BA=very_old) increases;

(iv)Predict CS

CS directly influence BV, When CS becomes false, P(BV) being weak increases. On the contrary, when P(CS) being weak and dead increases, it means the effect is caused by P(CS) becoming false. In addition, as initial probabilities shown in figure 1, it is apparent to see that when CS is true, P(BV) being dead is very low. Thus, If BV is dead, P(CS=true) approaches to 0. I predict that:

- P(CS) decreases significantly;
- P(¬CS) increases significantly;

(b) Diagnostic: the car starter system is not OK, the voltage at plug is low

(i) Predict BV

SO and VP are influenced directly by BV. When BV becomes weaker, P(SO=false) or P(VP=low) increase. On the contrary, when SO and VP are observed as false and low respectively, it means the effects are caused by P(BV) getting weaker. In addition, even if BV is weak, SO still have higher likelihood to be true than to be false, and VP still have higher likelihood to be high than to be low, which means if BV is false or VP is low, it is more likely affected by BV becoming dead. Furthermore, in this case both VP and SO are observed to be negative states, it reveals that they are greatly affected by BV. It means P(BV=dead) is higher than when only SO is observed as false. Therefore, when the agent observes that the car SO is false and VP is low:

- P(BV=weak) increases slightly (higher than the probability predicted in (a) (i));
- P(BV=dead) increases dramatically (higher than the probability predicted in (a) (i));
- P(BV=strong) decreases (lower than the probability predicted in (a) (i));

(ii) Predict BA

Based on the analysis in 3.3.1 (a) (iii) and the prediction in 3.3.1(b) (i), I predict that:

- P(BA=new) decreases significantly (lower than the prediction in 3.3.1(a) (iii));
- P(BA=old) increases (higher than the probability predicted in 3.3.1(a) (iii));

- P(BA=very_old) increases (higher than the probability predicted in 3.3.1(a) (iii));

(iii) Predict CS

CS directly influence BV, When CS becomes to be false, P(BV=weak) increases. On the contrary, when P(CS=weak) and P(CS=dead) increases, it means the effect is caused by P(CS=false). In addition, CS is an independent variable. Initial data shown in figure 1 reveals that the probability of CS being true is naturally extremely high, which means if BV becomes weak or dead, it more likely to be affected by a change from the probabilities of the other variable (i.e. BA). Combining with prediction from (b)(i), I predict that:

- P(CS=true) decreases slightly (lower than the probability predicted in 3.3.1(a) (iv));
- P(CS=false) increases slightly (higher than the probability predicted in 3.3.1 (a) (iv));

(c) Profiling: Determine the general characteristics of a car and what issues it might have depending on its battery age.

(i) Assume BA is new

- BA and CS are independent variables as I have mentioned in 3.2 (9), so when BA is new, it do not influence CS. Namely the probability distribution of CS will not change.
- According to the topology of BN of this model, it is easy to see that BA is one cause which can affect BV. How BA affects BV is explained in 3.2 (2) and 3.3.1

(a)(iii), thus P(BV=strong) increases, P(BV=weak) decreases and P(BV=dead) decreases.

- How BV affects VP and SO are explained in 3.2 (3)-(8), 3.3.1 (a)(ii) and 3.3.1 (b)(i), thus P(VP=high) increases, P(VP=low) decreases, P(SO=true) increases and P(SO=false) decreases.

(ii) Assume BA is old

- BA and CS are independent variables as I have mentioned in 3.2 (9), so when BA is old, it do not influence CS. Namely the probability distribution of CS will not change.
- According to the topology of BN of this model, it is easy to see that BA is one cause which can affect BV. How BA affects BV is explained in 3.2 (2) and 3.3.1 (a)(iii), thus P(BV=strong) decreases, P(BV=weak) increases and P(BV=dead) increases.
- How BV affects VP and SO are explained in 3.2 (3)-(8), 3.3.1 (a)(ii) and 3.3.1 (b)(i), thus P(VP=high) decreases, P(VP=low) increases, P(SO=true) decreases and P(SO=false) increases.

(iii) Assume BA is very old

- BA and CS are independent variables as I have mentioned in 3.2 (9), so when BA is old, it do not influence CS. Namely the probability distribution of CS will not change.
- According to the topology of BN of this model, it is easy to see that BA is one cause which can affect BV. How BA affects BV is explained in 3.2 (2) and 3.3.1

(a)(iii), thus P(BV=strong) decreases continually, P(BV=weak) increases continually and P(BV=dead) increases continually.

- How BV affects VP and SO are explained in 3.2 (3)-(8), 3.3.1 (a)(ii) and 3.3.1 (b)(i), thus P(VP=high) decreases continually, P(VP=low) increases continually, P(SO=true) decreases continually and P(SO=false) increases continually.

(iii) Assume which depends on its battery age

As it shown in figure 1, when BA changes from new to old to very old (CS remaining to be true), BV are more likely to be weak than to be dead. When BV is weak, it has big influence on VP. Once BV becomes weak, VP are probably to be low.

<u>(d) Predictive: Both car starter system and charging system are not OK. Predict the situation of voltage at plug.</u>

When CS is false, the probability of BV being dead increase dramatically unless BA can ensure to be new. Based on my analysis in 3.3.1 (a) (i), when SO is false, it is probably that BV becomes weaker, especially becomes dead. Further, both CS and BV are false, therefore the probability of BV being strong is very low.

Because of the probabilities of BV being weak and being dead increase dramatically, based on the analysis in 3.2 (4)-(6) and 3.3.1(a)(ii), I predict that

- P(VP=high) decreases dramatically;
- P(VP=low) increases dramatically;

3.2.2 Show results and obtain conclusions in car faults BN

(a) Figure 1 and 2 shows initial probabilities. The result after observation (SO is false) is shown in figure 3. My predictions of the characteristics are very similar to that of the results. However P(CS=true) decreases slightly, and P(CS=false) increases slightly. This is because CS is an independent variable. Initial data shown in figure 1 reveals that the probability of CS being true is naturally extremely high, which means if BV becomes weak or dead, it more likely to be affected by a change from the probabilities of the other variable (i.e. BA). According to diagnostic performed, BA is more likely to make SO become false.

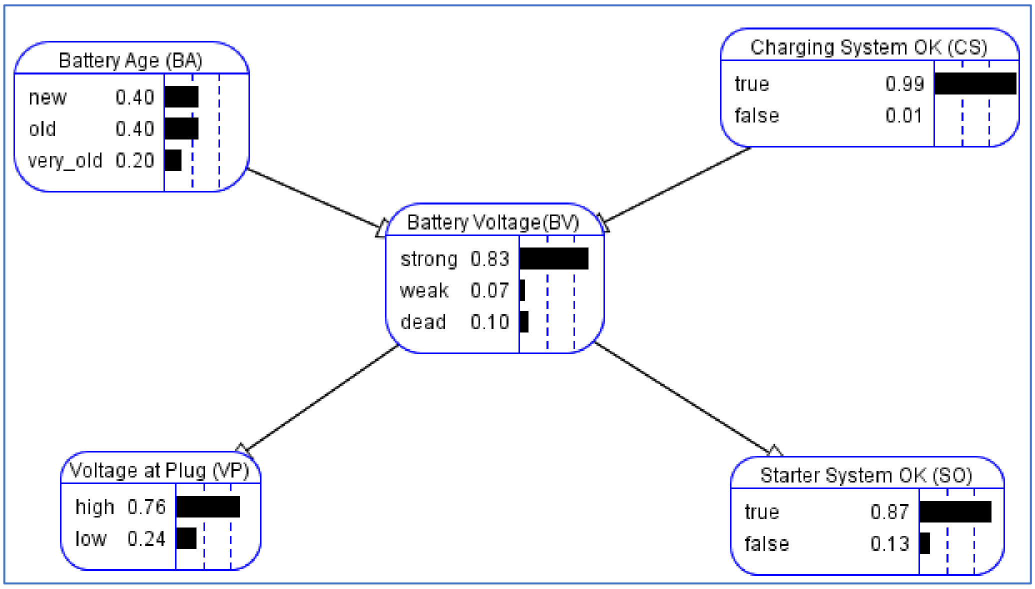


Figure 2 before the car starter becomes not OK

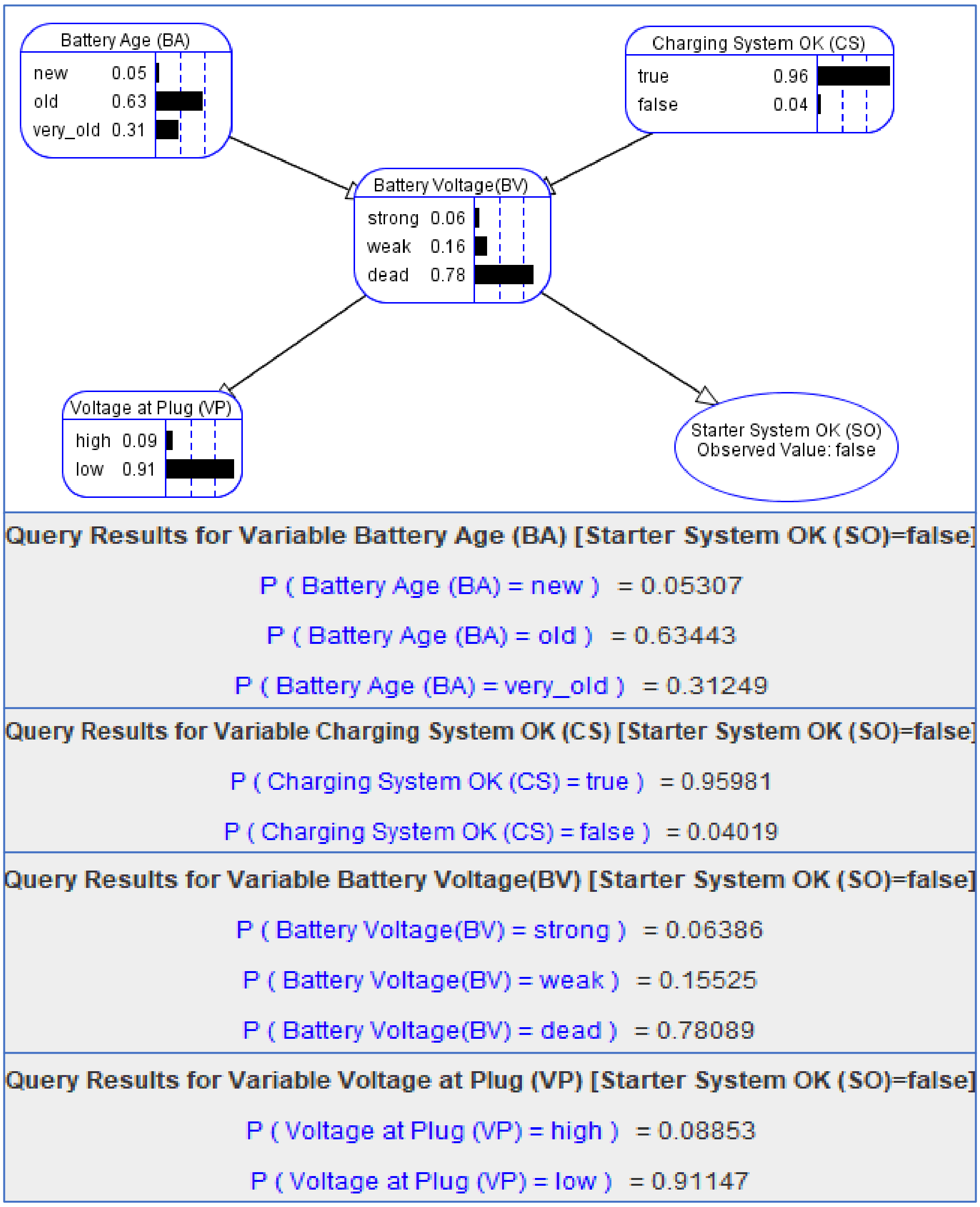


Figure 3 after the car starter becomes not OK

(b) Figure 1 and 2 shows initial probabilities. The result after observation (SO is false, VP is low) is shown in figure 3. My predictions of the characteristics are similar to that of the results. However P(BA=very_old) in (b) is smaller than that in (a). P(BV=weak) in (b) is smaller than that in (a). The first prediction error implies when the probability of BV being dead is very high and over a certain percentage, it mainly caused by CS becoming false. The second prediction error implies I underestimate the great impact of BV being dead on VP being low or SO being false. According to diagnostic performed, CS is more likely to make VP low.

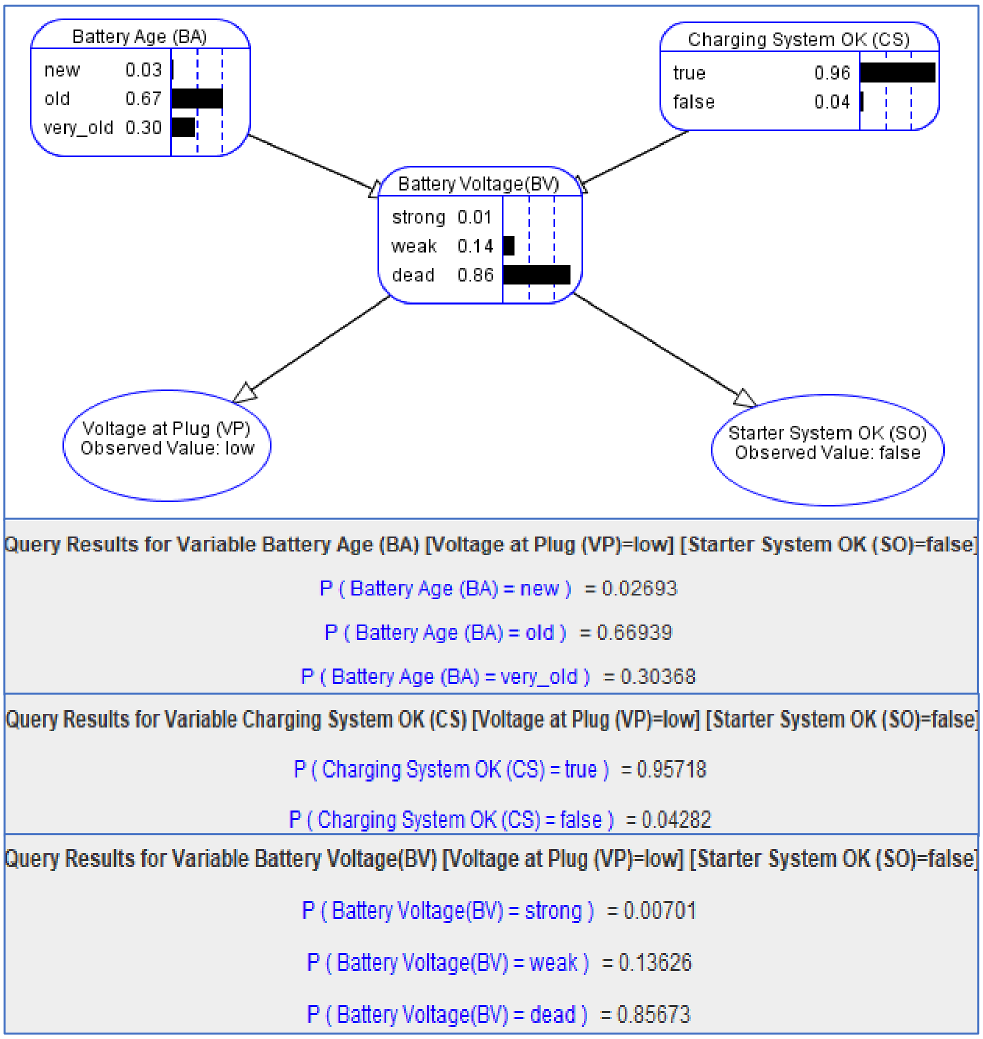
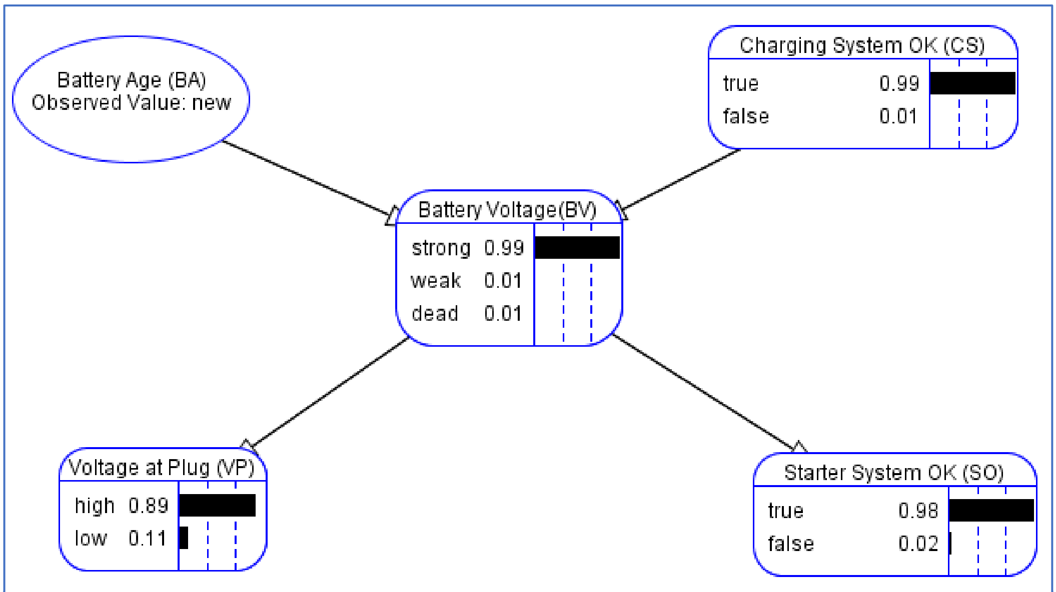


Figure 4 Figure 3 after SO becomes false and VP becomes low

(c) Figure 1 and 2 shows initial probabilities.

(i) Assume BA is new

My predictions of the characteristics are as that of the results. Besides, the result shows that in this case, P(BV=dead) approaches to 0. Because of that, P(VP=high) and P(SO=true) increase markedly. Results are shown in figure 5.



Query Results for Variable Charging System OK (CS) [Battery Age (BA)=new]

P (Charging System OK (CS) = true) = 0.99

P (Charging System OK (CS) = false) = 0.01

Query Results for Variable Battery Voltage(BV) [Battery Age (BA)=new]

P (Battery Voltage(BV) = strong) = 0.9881

P (Battery Voltage(BV) = weak) = 0.00645

P (Battery Voltage(BV) = dead) = 0.00545

Query Results for Variable Voltage at Plug (VP) [Battery Age (BA)=new]

P (Voltage at Plug (VP) = high) = 0.89058

P (Voltage at Plug (VP) = low) = 0.10942

Query Results for Variable Starter System OK (SO) [Battery Age (BA)=new

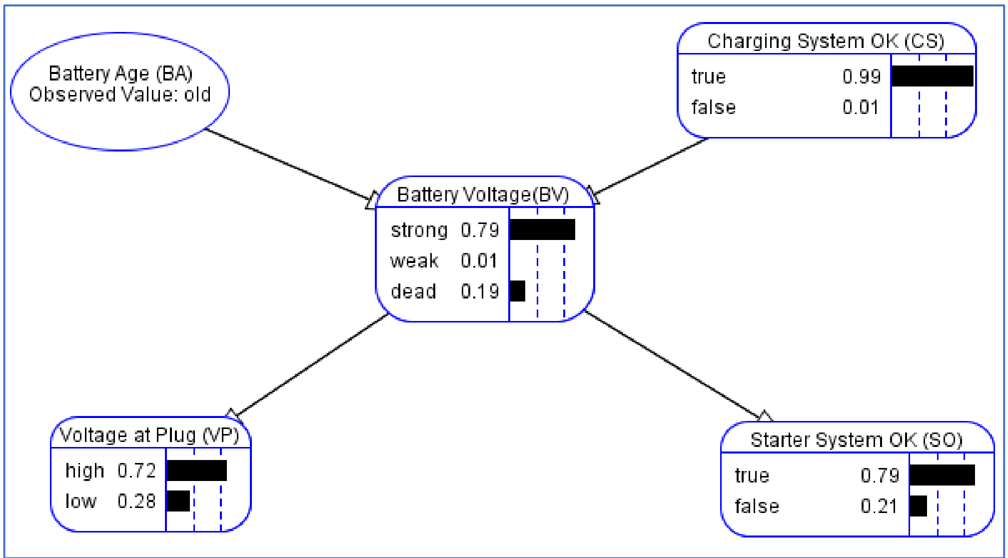
P (Starter System OK (SO) = true) = 0.98273

P (Starter System OK (SO) = false) = 0.01727

Figure 5 assume BA is new

(ii) Assume BA is old

My predictions of the characteristics are similar to that of the results. However, P(BV=weak) decreases but P(BV=dead) increases, which means when BA gets older, BV is more likely to dead than weak. Besides, the result shows that in this case, P(SO=true) decreases more markedly than P(VP=high), which means BV becoming dead has more influence on SO.



Query Results for Variable Charging System OK (CS) [Battery Age (BA)=old]

P (Charging System OK (CS) = true) = 0.99

P (Charging System OK (CS) = false) = 0.01

Query Results for Variable Battery Voltage(BV) [Battery Age (BA)=old]

P (Battery Voltage(BV) = strong) = 0.7925

P (Battery Voltage(BV) = weak) = 0.0129

P (Battery Voltage(BV) = dead) = 0.1946

Query Results for Variable Voltage at Plug (VP) [Battery Age (BA)=old]

P (Voltage at Plug (VP) = high) = 0.71583

P (Voltage at Plug (VP) = low) = 0.28417

Query Results for Variable Starter System OK (SO) [Battery Age (BA)=old]

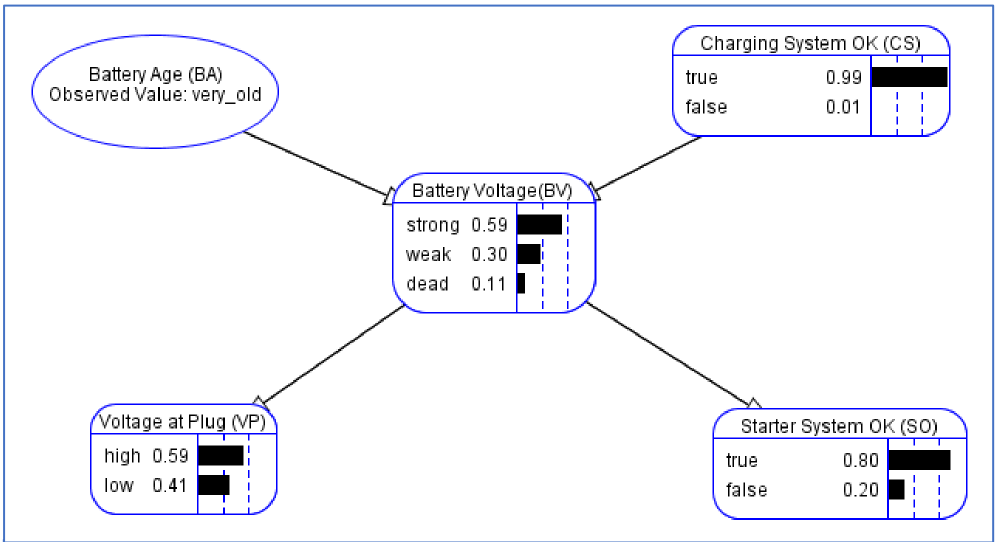
P (Starter System OK (SO) = true) = 0.79361

P (Starter System OK (SO) = false) = 0.2064

Figure 6 assume BA is old

(iii) Assume BA is very old

My predictions of the characteristics are similar to that of the results. P(BV=weak) decreases much more markedly than the increase of P(BV=dead), which means when BA becomes very old, BV is more likely to weak than dead. Besides, the result shows that in this case, P(VP=high) decreases much more markedly than the decrease of P(SO=true), which means BV becoming weak has more influence on VP.



Query Results for Variable Charging System OK (CS) [Battery Age (BA)=very_old]

P (Charging System OK (CS) = true) = 0.99

P (Charging System OK (CS) = false) = 0.01

Query Results for Variable Battery Voltage(BV) [Battery Age (BA)=very_old]

P (Battery Voltage(BV) = strong) = 0.59402

P (Battery Voltage(BV) = weak) = 0.298

P (Battery Voltage(BV) = dead) = 0.10798

Query Results for Variable Voltage at Plug (VP) [Battery Age (BA)=very_old]

P (Voltage at Plug (VP) = high) = 0.59422

P (Voltage at Plug (VP) = low) = 0.40578

Query Results for Variable Starter System OK (SO) [Battery Age (BA)=very_old]

P (Starter System OK (SO) = true) = 0.79668

P (Starter System OK (SO) = false) = 0.20332

Figure 7 assume BA is very old

(d) Figure 1 and 2 shows initial probabilities. The results after CS becoming false and SO becoming false. My predictions of the characteristics are same to that of the results. Results also shows that P(VP=high) approaches to 0. The results are shown in figure 8.

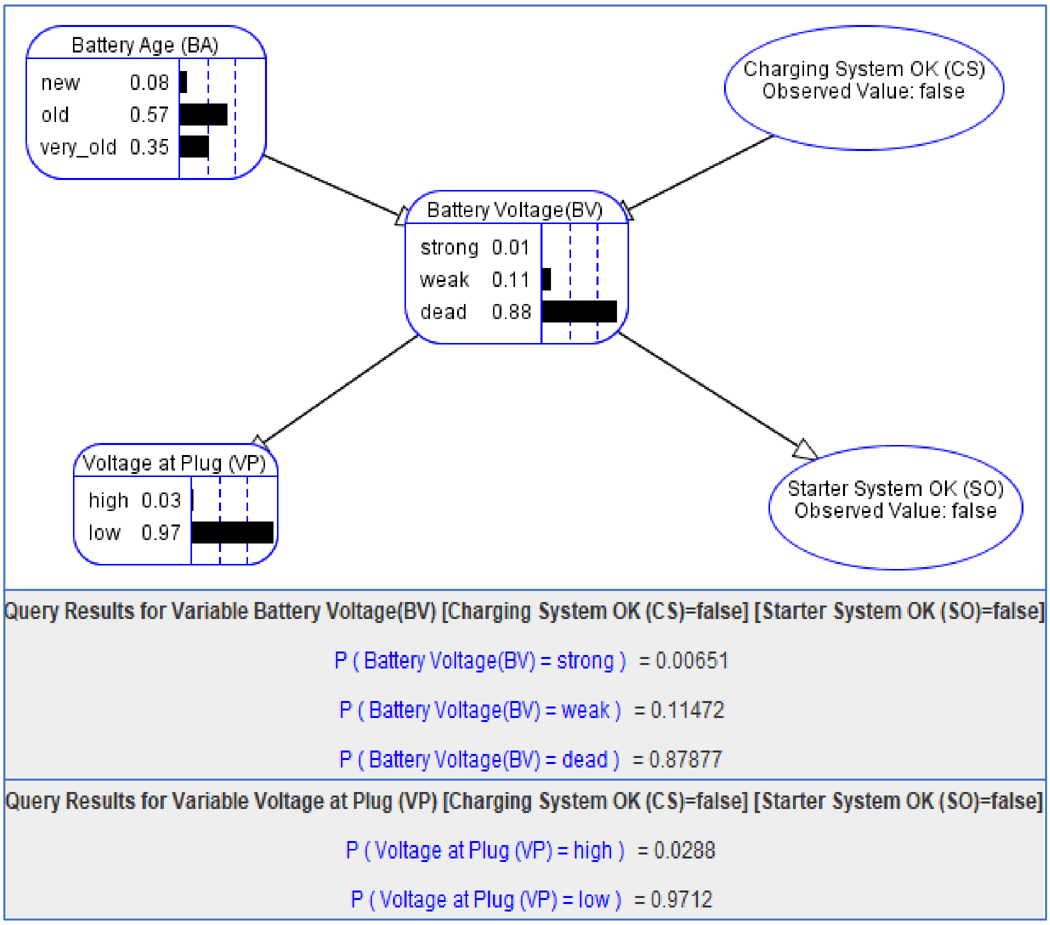


Figure 8 CS is false and SO is false

3.4 Analyse the relationships between variables in medical diagnosis BN

Figure 9 shows BN for medical diagnosis

(i) In this medical diagnosis model, random variables and their corresponding finite countable domains are revealed as follows:

(1) Influenza={T, F};

(2) Smokes={T, F};

(3) Sore Throat={T, F};

(4) Fever={T, F};

(5) Bronchitis={T, F};

(6) Coughing={T, F};

(7) Wheezing={T, F};

(ii) Relationships between variables are analysed as follows:

(1) Influenza and smokes are causes of effect (bronchitis), so influenza and smokes directly affect bronchitis;

(2) When either influenza is true or smokes is true, bronchitis has very high likely to be true;

(3) When influenza changes from true to false (Smokes remaining to be false), P(Bronchitis=T) changes from 0.9 to 0.0001. It means if influenza being true has a huge impact on if bronchitis being true;

(4) When smokes changes from true to false (influenza remaining to be false), P(Bronchitis=T) changes from 0.7 to 0.0001. It means if smokes being true has a huge impact on if bronchitis being true;

(5) When both influenza and smokes are true, the probability of bronchitis being true is almost 100%, which means these two variables are almost whole causes of making bronchitis true.

(6) When both influenza and smokes are false, the probability of bronchitis being true is almost 0, which means in this case, bronchitis is almost impossible to be true.

(7) Bronchitis is cause of effects (coughing and wheezing), so bronchitis directly affects coughing and wheezing;

(8) When bronchitis is true, coughing is 80% likely to be true;

(9) When bronchitis is false, coughing is 93% likely to be false, which means once bronchitis is false, coughing almost will not happen;

(10) When bronchitis is true, wheezing is 60% likely to be true and 40% to be false, which also reveals that it still has a quit high probability for wheezing to be true;

(11) When bronchitis is false, wheezing is 99.9% likely to be false, which means once wheezing is false, basically it can be asserted that coughing almost will not happen;

(12) Influenza is also a cause of effect (sore throat, fever), so influenza directly affect sore throat, fever too;

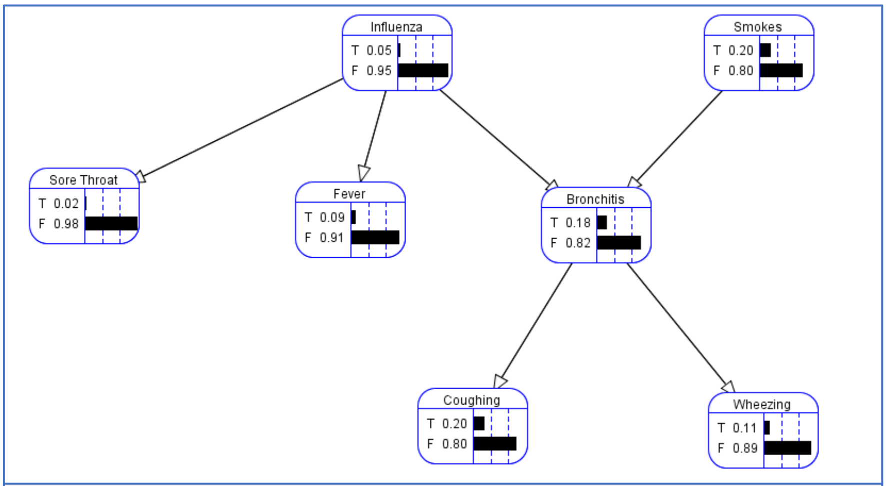
(13) When influenza is true, throat has 30% likelihood to be sore. When influenza is false, sore throat is 99.9% likely to be false, which means once influenza is false, basically it can be asserted that sore throat almost will not happen; Actually, based on probability distribution of

sore throat given by influenza, it is apparently that whether influenza is true or not, sore throat is more likely to be false.

(14) When influenza is true, fever has 90% likelihood to be true. When influenza is false, fever is 95% likely to be false, which means if influenza is true has huge influence on if fever is true. Influenza is true, fever is extremely likely to true and vice versa.

(15) Influenza and smokes are independent, coughing and wheezing are conditionally independent given by bronchitis, and sore throat and fever are conditionally independent given by influenza;

(16) The probabilities of influenza and smokes are prior probabilities. The probabilities of bronchitis, coughing, wheezing, sore throat and fever are posterior probabilities.



	P(Smokes=T)	P(Smokes=F)
Prior Probability	0.2	0.8

	P(Influenza=T)	P(Influenza=F)
Prior Probability	0.05	0.95

Influenza	Smokes	P(Bronchitis=T)	P(Bronchitis=F)
T	T	0.99	0.01
T	F	0.9	0.1
F	T	0.7	0.3
F	F	1.0E-4	0.9999

Bronchitis	P(Coughing=T)	P(Coughing=F)
T	0.8	0.2
F	0.07	0.93

Bronchitis	P(Wheezing=T)	P(Wheezing=F)
T	0.6	0.4
F	0.001	0.999

Influenza	P(Sore Throat=T)	P(Sore Throat=F)
T	0.3	0.7
F	0.001	0.999

Influenza	P(Fever=T)	P(Fever=F)
T	0.9	0.1
F	0.05	0.95

Figure 9 BN for medical diagnosis

3.5 Evaluate results when observations are changed in medical diagnosis BN

3.5.1 Predict results when observations are changed in medical diagnosis BN

Note: when talking about increase or decrease, it compares with initial probability distribution in car faults BN.

(a) Diagnostic: The agent observes that the patient is wheezing, how likely is that the patient smokes.

As the analysis shown in 3.4 (ii)(7)(10), wheezing is influenced by bronchitis. Hence when wheezing is observed as true, it is because the probability of bronchitis being true increases significantly. Combined with the analysis shown in 3.4(ii)(1)(2)(4), when bronchitis becomes true, it is more likely due to the reason that smokes becomes true. Therefore, I predict that:

- P(Smokes=T) increases markedly;
- P(Smokes=F) decreases markedly;

(b) Predictive: The agent observes that the patient is a smoker, how likely is the patient will develop bronchitis? And that they will have sore throat?

As the analysis shown in 3.4 (ii)(1)(2)(4), when the patient is a smoker, the patient will probably to develop bronchitis; Because of that and combined the analysis in 3.4 (ii)(1)(3), the increase of probability of bronchitis being true is caused by the increase of probability of influenza being true. Additionally, the variable smokes is dependant with the variables sore throat. Hence, if a patient is a smoker will not influence if throat is sore. Therefore, I predict that:

- P(Bronchitis=T) increases significantly;
- P(Bronchitis=F) decreases significantly;
- P(Sore throat=T) will not change;

- P(Sore throat=F) will not change;

(c) Predictive: The agent observes that the patient is a smoker, and has influenza, how likely is the patient will develop have sore throat?

As the analysis shown in 3.4 (ii)(12)(13), influence has small influence on sore throat, but when P(Influenza=true) is bigger, throat will be more likely to be sore than when P(Influenza=false) is bigger. As if the patient is smoker is not the cause of if throat is sore, namely the former have no impact on the latter, Therefore, I predict that:

- P(sore throat=T)=0.3;
- P(sore throat=F)=0.7;

(d) Predictive: The agent observes that the patient is not a smoker, and has influenza, how likely is the patient will be coughing?

By checking initial probability distributions shown in figure 9, when influenza is true, and smokes is false, the probability of bronchitis being true or false is determined.

- P(Bronchitis=T)=0.9;
- P(Bronchitis=F)=0.1;

Based on the analysis in 3.4(ii)(7)(8), the increase of probability of bronchitis being true causes the increase of probability of coughing being true. Therefore, I predict that:

- P(Coughing=T) increases dramatically;
- P(Coughing=F) increases dramatically;

(e) Intercausal: The agent observes that the patient is not a smoker and, has got bronchitis. How likely is that they have influenza?

Based on the analysis in 3.4(ii)(1)-(5), if the agent only observes that the patient has got bronchitis, it must be caused by either influenza being true or smokes being true, or both of them being true. However another evidence is that the patient is not smoker, so the probability of influenza must be really big, and even approaches to 1.

- P(Influenza=T) increases and approaches to 1;
- P(Influenza=F) increases and approaches to 0;

(f) Profiling: Determine the characteristics of patients that smoke and compare those against patients that do not smoke.

(i) Assume smokes is false

Based on the analysis in 3.4 (ii)(1)(2)(4), the increase of the probability of smokes being false causes the increase of the probability of bronchitis being false. And from figure 9, the probability of influenza being true is close to 0.05, as another cause of making bronchitis be true. I predict that

- P(Bronchitis=T) decrease and approaches to 0;
- P(Bronchitis=F) increases and approaches to 1;

As the analysis in 3.4 (ii)(7)-(11), the increase of the probability of bronchitis being false causes the increase of the probability of coughing being false and the increase of the probability of wheezing being false. I predict that

- P(Coughing= T) decrease and approaches to 0;

- P(Coughing=F) increases and approaches to 1;
- P(Wheezing=T) decrease and approaches to 0;
- P(Wheezing=F) increases and approaches to 1;

(ii) Assume smokes is true

Similarly, when smoke is false, I predict that

- P(Bronchitis=T) increase markedly;
- P(Bronchitis=F) decrease markedly;
- P(Coughing=T) increase markedly;
- P(Coughing=F) decreases markedly;

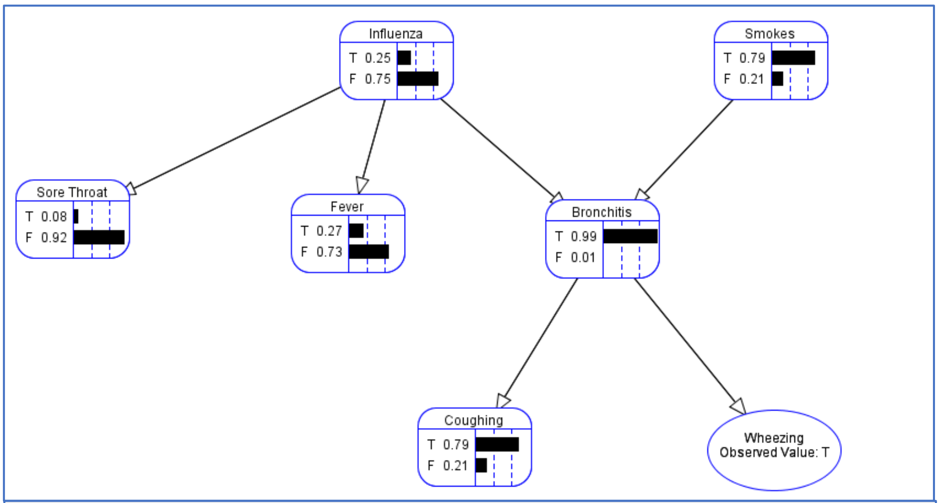
Based on 3.4 (ii)(10), the probability of bronchitis being true has small impact on the probability of wheezing being true.

- P(Wheezing=T) increases slightly;
- P(Wheezing=F) decreases slightly;

Besides, the variable smokes is dependant with the variables sore throat and fever. Hence, if a patient is a smoker will not influence if throat is sore and if a patient get fever.

3.5.2 Show results and obtain conclusions in simple diagnostic BN

(a) Figure 10 shows the result of when the agent observes that the patient is wheezing, which is same as my prediction, and the patient is 79% likely to smoke.



Query Results for Variable Bronchitis [Wheezing=T]

P (Bronchitis = T) = 0.99241

P (Bronchitis = F) = 0.00759

Query Results for Variable Smokes [Wheezing=T]

P (Smokes = T) = 0.7929

P (Smokes = F) = 0.2071

Figure 10 the patient is wheezing

(b) Figure 11 shows the result of when the agent observes that the patient is a smoker, which is same as my prediction. The patient is 71% likely to develop bronchitis and 2% likely to have sore throat.

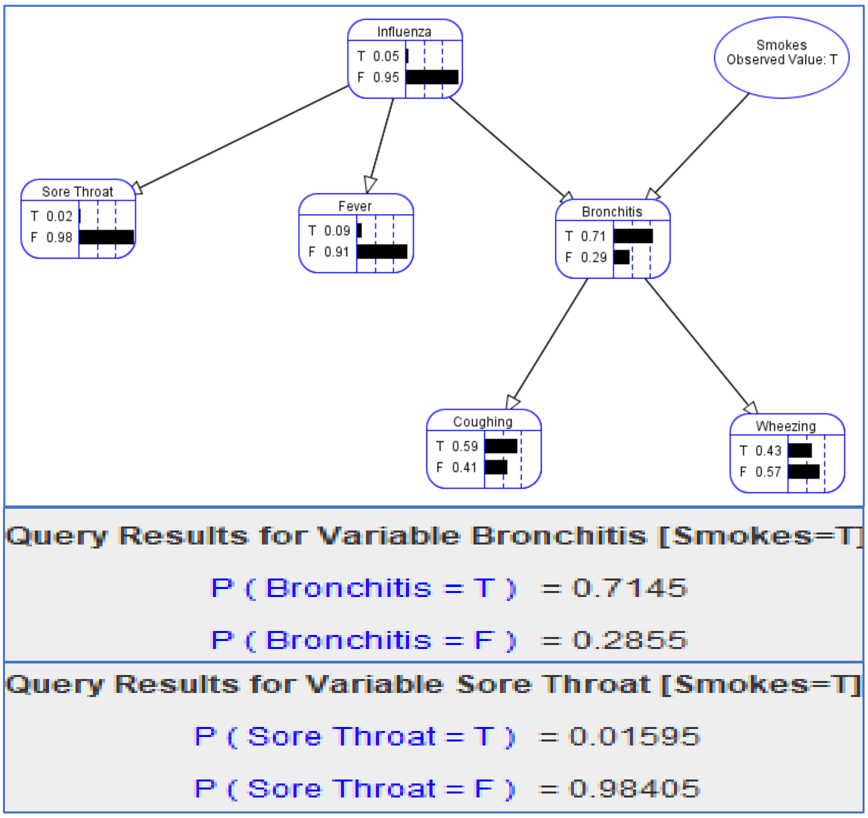


Figure 11 patient is a smoker

(c) Figure 12 shows the result of when the agent observes that the patient is a smoker and has influenza, which is same as my prediction. The patient is 30% likely to develop sore throat.

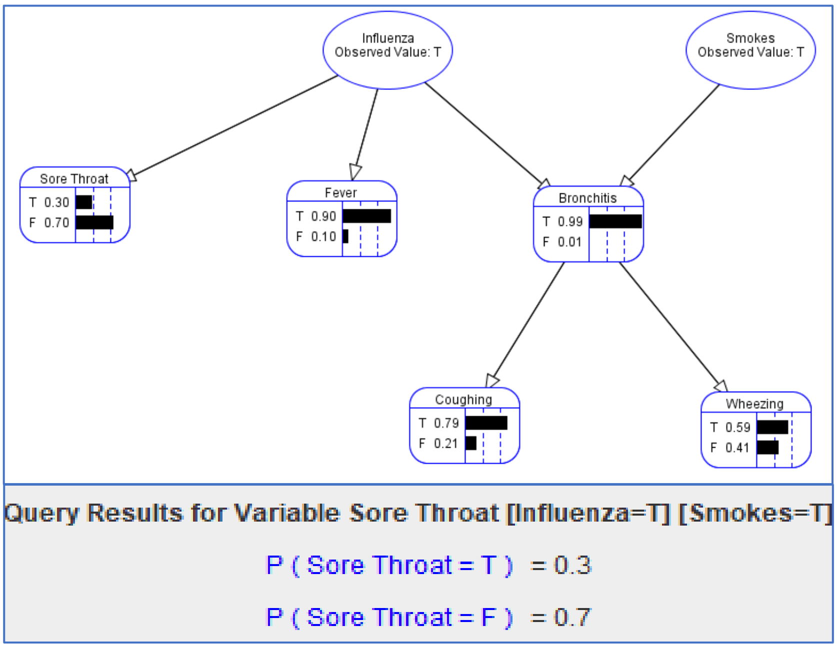


Figure 12 the patient is a smoker and has influenza

(d) Figure 13 shows the result of when the agent observes that the patient is not a smoker and has influenza, which is same as my prediction. The patient is 73% likely to have coughing.

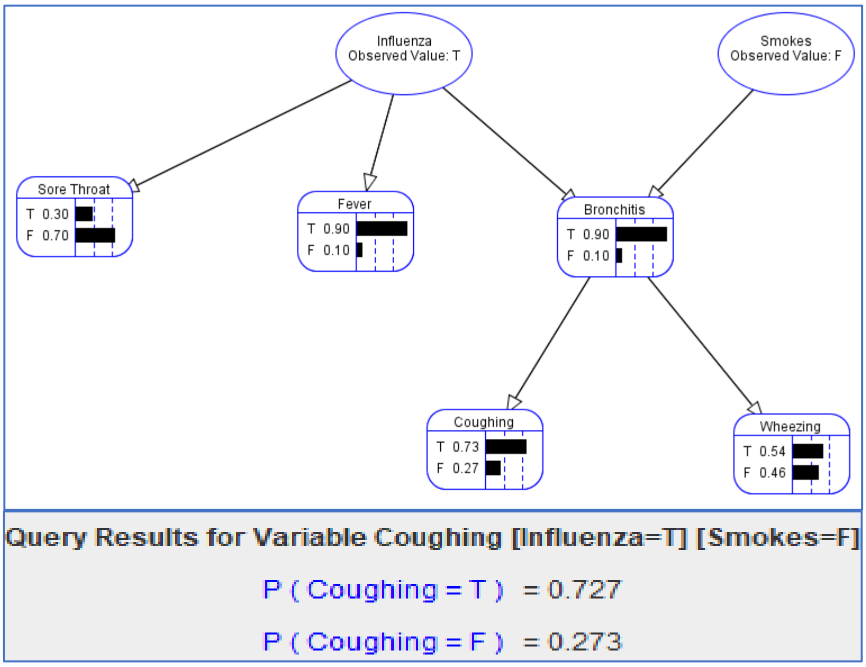


Figure 13 the patient is not a smoker and has influenza

(e) Figure 14 shows the result of when the agent observes that the patient is not a smoker and has influenza, which is same as my prediction. The patient is 99.789% likely to have influenza.

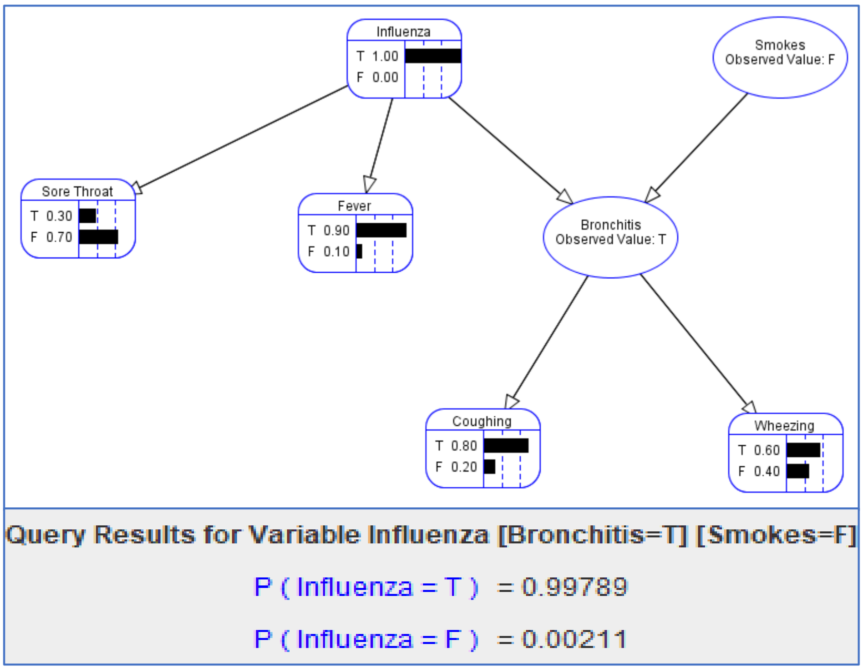
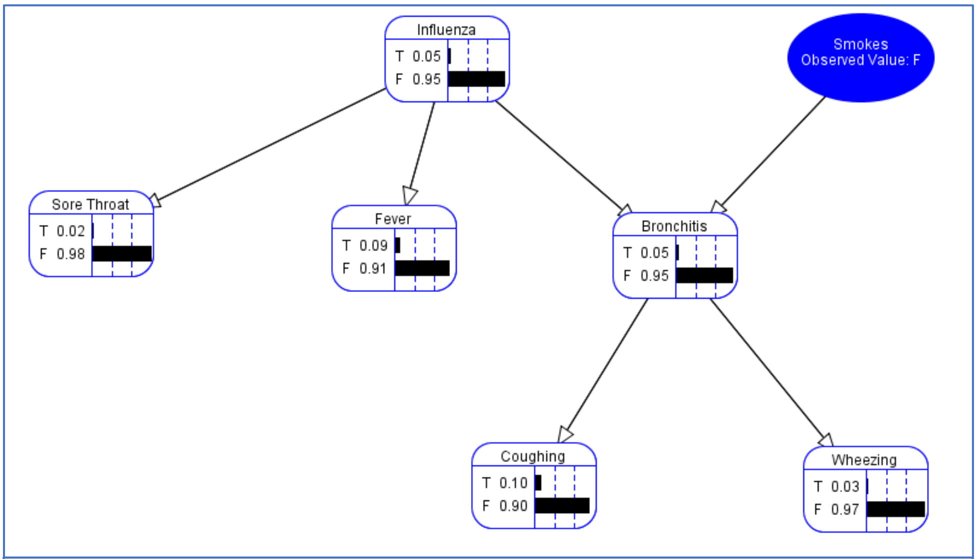


Figure 14 the patient is not a smoker and has got bronchitis

(f) Figure 15 and Figure 16 shows the results of when smokes is assumed to be false and true respectively, which is same as my prediction. The detail information about the differences of characteristics between those who smoke and not smoke are shown in figure 15 and figure 16 too.



Query Results for Variable Bronchitis [Smokes=F]

P (Bronchitis = T) = 0.0451

P (Bronchitis = F) = 0.95491

Query Results for Variable Coughing [Smokes=F]

P (Coughing = T) = 0.10292

P (Coughing = F) = 0.89708

Query Results for Variable Wheezing [Smokes=F]

P (Wheezing = T) = 0.02801

P (Wheezing = F) = 0.97199

Query Results for Variable Sore Throat [Smokes=F]

P (Sore Throat = T) = 0.01595

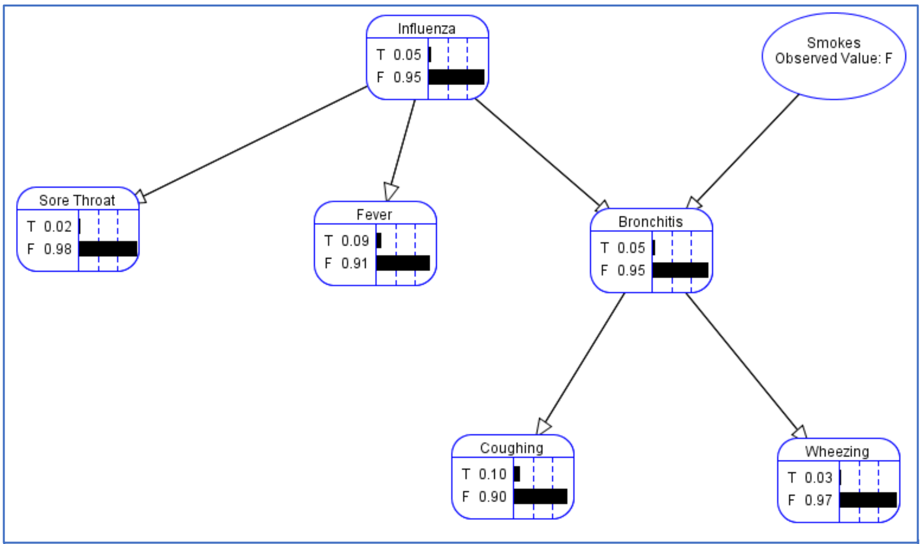
P (Sore Throat = F) = 0.98405

Query Results for Variable Fever [Smokes=F]

P (Fever = T) = 0.0925

P (Fever = F) = 0.9075

Figure 15 assume smokes is false



Query Results for Variable Bronchitis [Smokes=F]

P (Bronchitis = T) = 0.0451

P (Bronchitis = F) = 0.95491

Query Results for Variable Coughing [Smokes=F]

P (Coughing = T) = 0.10292

P (Coughing = F) = 0.89708

Query Results for Variable Wheezing [Smokes=F]

P (Wheezing = T) = 0.02801

P (Wheezing = F) = 0.97199

Query Results for Variable Influenza [Smokes=F]

P (Influenza = T) = 0.05

P (Influenza = F) = 0.95

Query Results for Variable Sore Throat [Smokes=F]

P (Sore Throat = T) = 0.01595

P (Sore Throat = F) = 0.98405

Query Results for Variable Fever [Smokes=F]

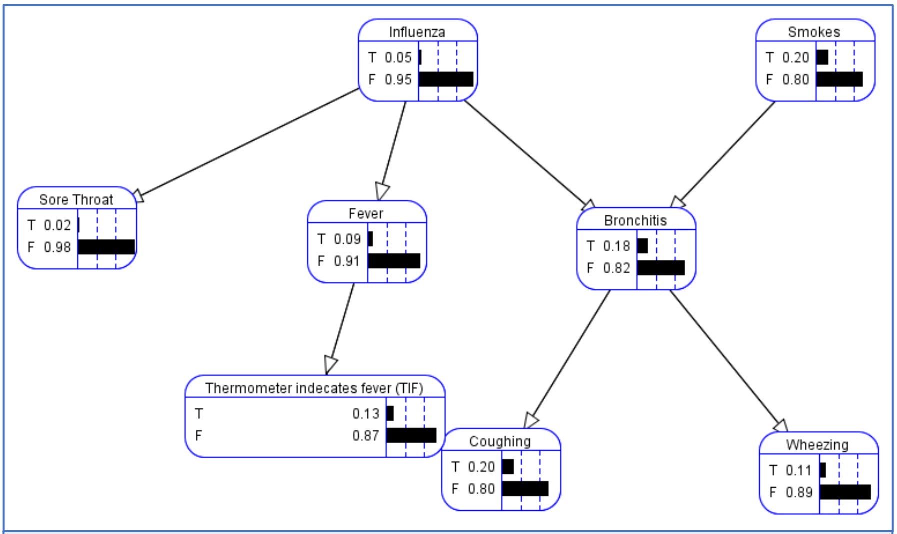
P (Fever = T) = 0.0925

P (Fever = F) = 0.9075

Figure 16 assume smokes is true

3.6 The development of BN for simple diagnostic example

3.6.1 Construct BN



Fever	P(Thermometer indecates fever (TIF)=T)	P(Thermometer indecates fever (TIF)=F)
T	0.9	0.1
F	0.05	0.95

Figure 17 The developed version of BN for simple diagnostic example

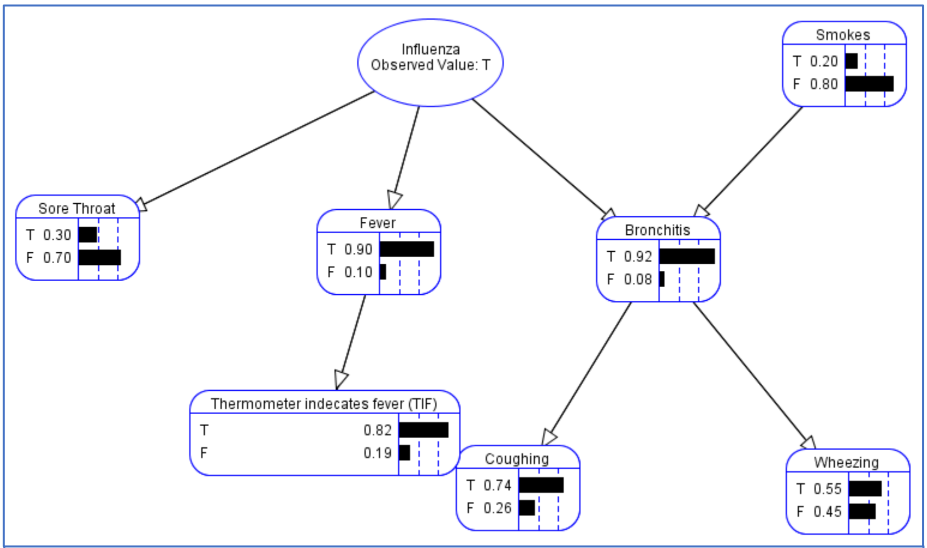
3.6.2 Predict that if the patient has influenza, how likely is that the thermometer will indicate fever?

Based on 3.4 (ii)(12)(14), when the patient has influenza, the patient has very high probability to get fever. Fever can be seen as a cause of effect (thermometer indicating fever), and when a patient has influenza, thermometer are very likely to indicate fever. Therefore, I predict that

- P(TIP) increases significantly;
- P(¬TIP) decreases significantly;

3.6.3 Result of the changes and the likelihood of thermometer indicating fever.

Figure 18 shows the result when the patient has influenza, which is same as my prediction. Thermometer will be 82% likely to indicate fever.



Query Results for Variable Fever [Influenza=T]

P (Fever = T) = 0.9

P (Fever = F) = 0.1

Query Results for Variable Thermometer indecates fever (TIF) [Influenza=T]

P (Thermometer indecates fever (TIF) = T) = 0.815

P (Thermometer indecates fever (TIF) = F) = 0.185

Figure 18 the probability distribution of TIF when the patient has influenza

4 Part 2

4.1 Construct a BN to represent an alert system

(i) Construct a BN to represent an alert system

The topology of BN for an alert system is shown in graph 19.

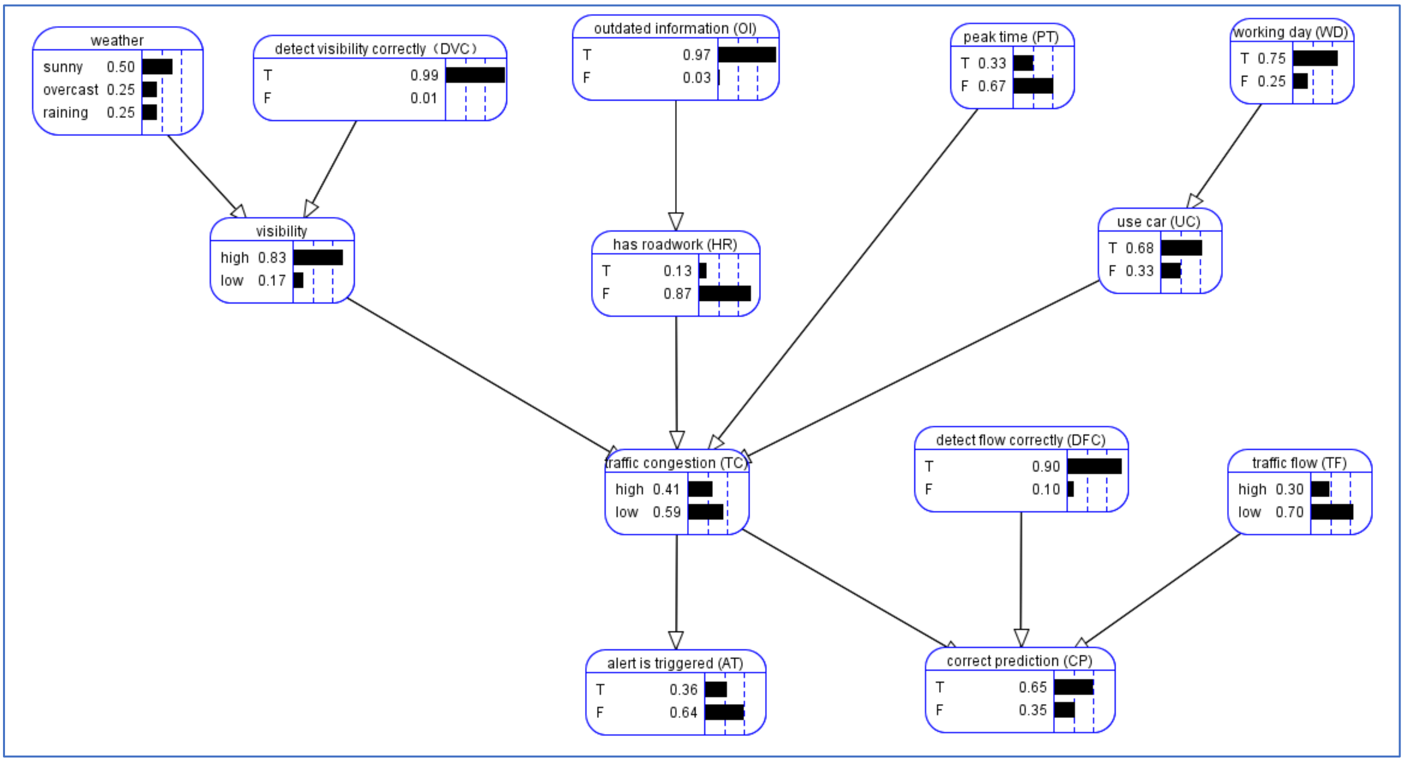


Figure 19 BN for an alarm system

(ii) Domain variables

- There are 13 variables;
- weather={sunny, overcast, raining};
- detect visibility correctly(DVC)={T,F};
- Visibility={high, low};
- outdated information (OI)={T, F};

- has roadwork (HR)={T,F};
- peak time (PT)={T,F};
- working day (WD)={T,F};
- use car (UC)={T,F};
- traffic congestion (TC)={high, low};
- detect flow correctly (DFC)={T, F};
- traffic flow (TF)={high, low};
- correct prediction (CP)={T,F};
- alert is triggered (AT)={T,F}

(iii) Relationships between different domain variables

(1)Weather, DVC, OI, PT, WD, DFC and TF are independent, and their probabilities are prior probability.

(2)Visibility is the child of weather and DVC. If visibility is high is influenced by weather and DVC. When DVC is true, visibility decreases from high to low as the weather turns from sunny to overcast to raining. However when DVC is false, results change in opposite way.

(3)HR is the child of OI. If HR is true is influenced by OI. Data updating late make the roadwork be detected late, which means real P(HR=T) should be higher than 10%. Namely, P(HR=T) increases as OI turns from false to true.

(4)UC is the child of WD. If UC is true is influenced by WD. People in working day are more likely to drive than in holiday. Namely, P(HR=T) increases as WD turns from false to true and vice versa.

(5)CP is the child of TC, DFC and TF. If CP is true is influenced by TC, DFC and TF. When DFC is true, both TF and TC are high or low, P(CP=T) is 1. When

DFC is true, the value of TF and TC are opposite, P(CP=T) decreases significantly. When DFC is false, both TF and TC are high or low, P(CP=T) decreases significantly. When DFC is false, the value of TF and TC are opposite, P(CP=T) is 1.

(6)TC is the child of visibility, HR, PT and UC. If TC is high is influenced by visibility HR, PT and UC. P(TC=high) increases as visibility turns from P(visibility=low) increases ,or P(HR=T) increases, or P(PT=T)increases or P(UC=T) increases and vice versa.

(7)AT is the child of TC. If AT is high is influenced by TC. P(AT=T) increases as P(TC=high) increases.

(iv) How I choose values for those probabilities are vacant

(1) Probabilities of weather and DFC are prior probabilities, thus their probability distributions (shown in figure 20) are determined by life experience. Further it is more rationale that the probability distribution of TF is similar to that of TC.

weather	detect visibility correctly（DVC）	P(visibility=high)	P(visibility=low)
sunny	T	1.0	0.0
sunny	F	0.0	1.0
overcast	T	0.65	0.35
overcast	F	0.35	0.65
raining	T	0.7	0.3
raining	F	0.3	0.7

	P(traffic flow (TF)=high)	P(traffic flow (TF)=low)
Prior Probability	0.3	0.7

Figure 20 probability tables for visibility and traffic flow

(2)Probabilities of visibility are conditional probabilities. The relationship between visibility and weather and DVC are described in 4.1(iii)(2), which is reasons why I set probabilities as showing in figure 21.

weather	detect visibility correctly（DVC）	P(visibility=high)	P(visibility=low)
sunny	T	1.0	0.0
sunny	F	0.0	1.0
overcast	T	0.65	0.35
overcast	F	0.35	0.65
raining	T	0.7	0.3
raining	F	0.3	0.7

Figure 21 the probability distribution of visibility

(3)Probabilities of HR are conditional probabilities. The relationship between HR and OI are described in 4.1(iii)(3).), which is reasons why I set probabilities as showing in figure 22.

outdated information (OI)	P(has roadwork (HR)=T)	P(has roadwork (HR)=F)
T	0.13	0.87
F	0.1	0.9

Figure 22 the probability distribution of HR

(4)Probabilities of CP are conditional probabilities. The relationship between CP and DFC, TF and TC are described in 4.1(iii)(5) , which is reasons why I set probabilities as showing in figure 23.

detect flow correctly (DFC)	traffic flow (TF)	traffic congestion (TC)	P(correct prediction (CP)=T)	P(correct prediction (CP)=F)
T	high	high	1.0	0.0
T	high	low	0.25	0.75
T	low	high	0.25	0.75
T	low	low	1.0	0.0
F	high	high	0.25	0.75
F	high	low	1.0	0.0
F	low	high	1.0	0.0
F	low	low	0.25	0.75

Figure 23 the probability distribution of CP

(5)Probabilities of TC are conditional probabilities. The relationship between TC and visibility, HR, PT and UC are described in 4.1(iii)(6), which is reasons why I set probabilities as showing in figure 24.

visibility	has roadwork (HR)	peak time (PT)	use car (UC)	P(traffic congestion (TC)=high)	P(traffic congestion (TC)=low)
high	T	T	T	0.85	0.15
high	T	T	F	0.8	0.2
high	T	F	T	0.5	0.5
high	T	F	F	0.3	0.7
high	F	T	T	0.65	0.35
high	F	T	F	0.5	0.5
high	F	F	T	0.25	0.75
high	F	F	F	0.1	0.9
low	T	T	T	0.95	0.05
low	T	T	F	0.9	0.1
low	T	F	T	0.75	0.25
low	T	F	F	0.7	0.3
low	F	T	T	0.75	0.25
low	F	T	F	0.7	0.3
low	F	F	T	0.6	0.4
low	F	F	F	0.45	0.55

Figure 24 the probability distribution of TC

(6)Probabilities of AT are conditional probabilities. The relationship between AT and TC are described in 4.1(iii)(7), and combining with given information that an alert is triggered if traffic congestion has 80% probability to be high. I set probabilities as showing in figure 25.

traffic congestion (TC)	P(alert is triggered (AT)=T)	P(alert is triggered (AT)=F)
high	0.8	0.2
low	0.05	0.95

Figure 25 the probability distribution of AT

4.2 Prediction and evaluation based on BN for the initial alarm system

(a) Predictive: The agent observes that weather is overcast, how likely that traffic congestion is high.

(i) Prediction

As the analysis shown in 4.1(iii)(2), when weather is overcast, P(visibility=low) increases. Then according to 4.1(iii)(6), when P(visibility=low) increases, I predict that:

- P(TC=high) increases;
- P(TC=low) decreases;

***(ii)** Results*

Figure 26 shows results, which is same as my prediction. TC will be 46% likely to be high.

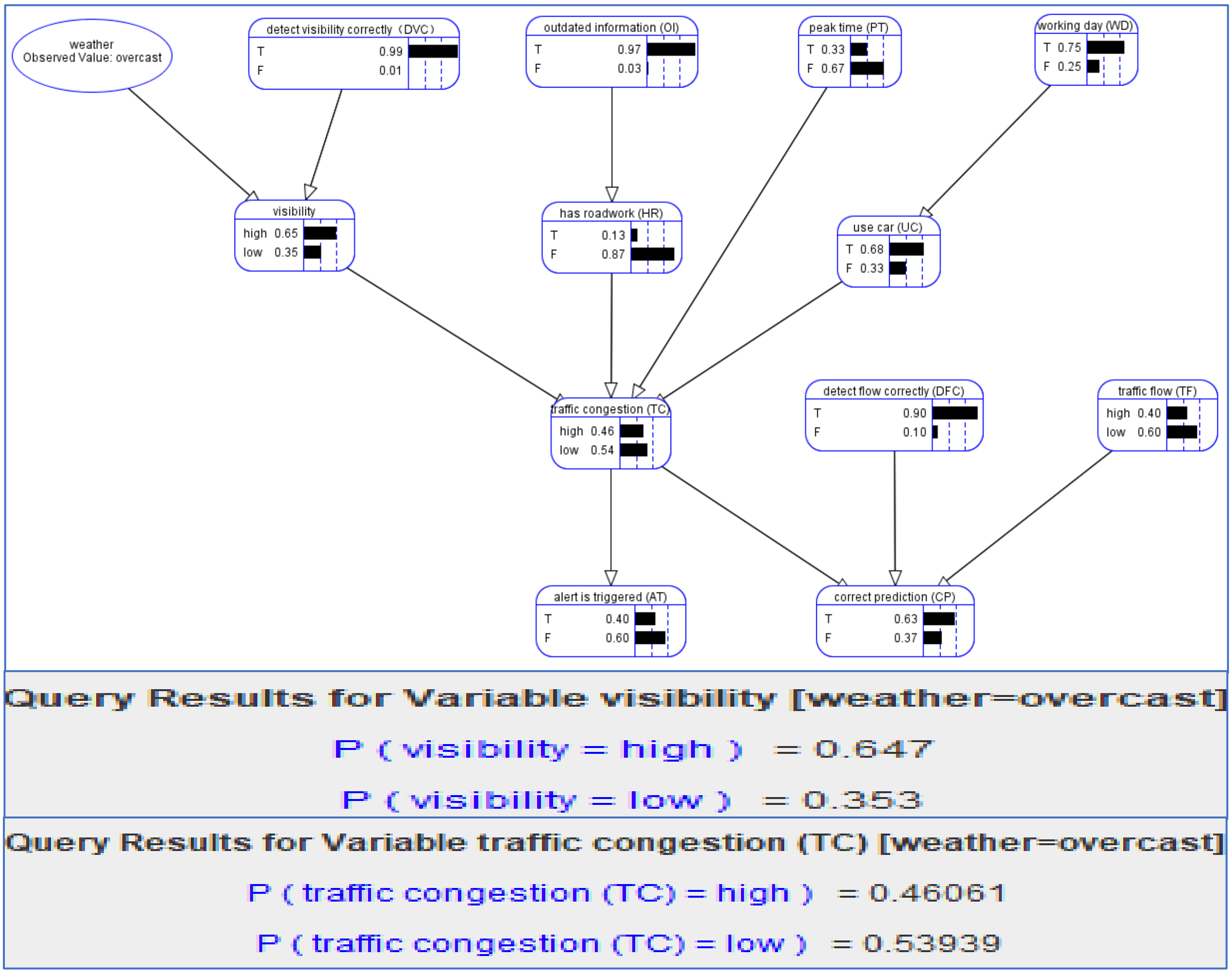


Figure 26 weather is overcast

<u>(b) Predictive: The agent observes that the day is a holiday, how likely that traffic congestion is high.</u>

***(i)** Prediction*

As the analysis shown in 4.1(iii)(4), when working day is false, P(UC=T) decreases. Then according to 4.1(iii)(6), when P(UC=T) decreases, I predict that:

- P(TC=high) decreases;
- P(TC=low) increases;

(ii) Results

Figure 27 shows results, which is same as my prediction. TC will be 40% likely to be high.

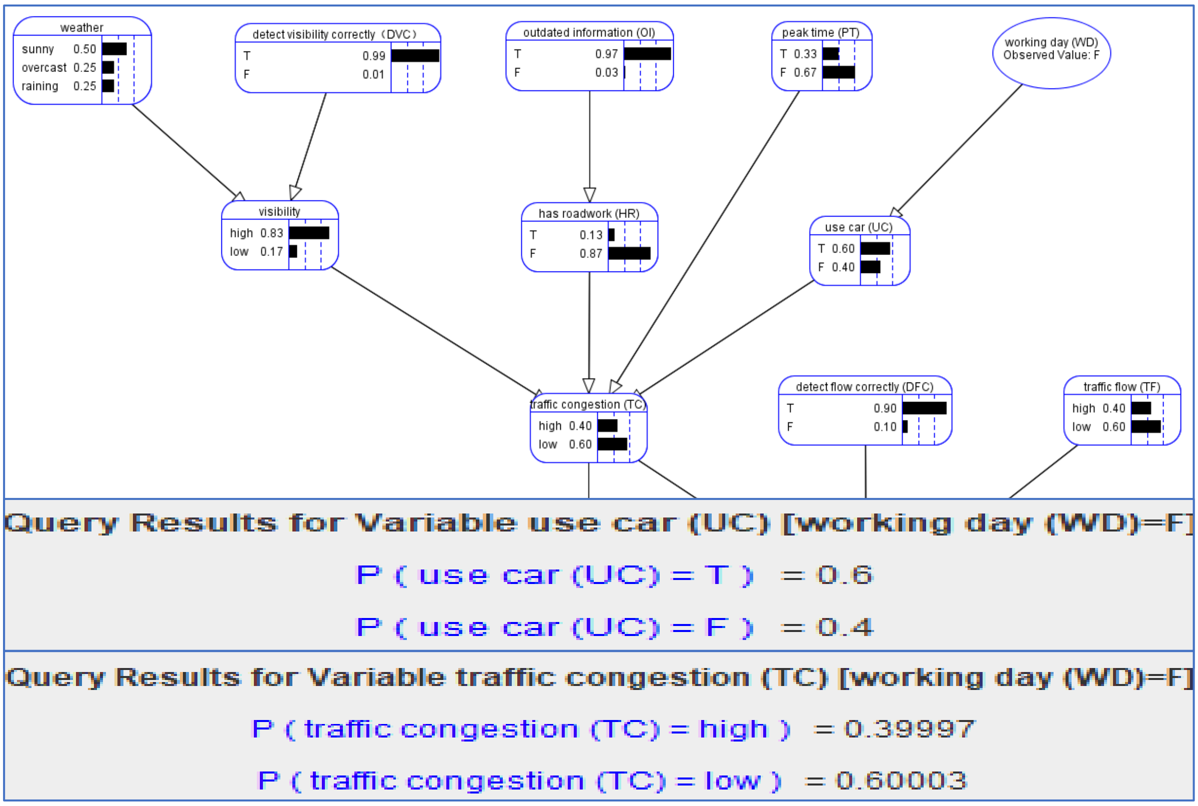


Figure 27 the day is a holiday

(c) Predictive: The agent observes that traffic flow is high, and it is detected correctly, how likely that prediction is true?

(i) Prediction

As the analysis shown in 4.1(iii)(5), when traffic flow is high, and it is detected correctly, the key in this case is that if P(TC=high) is bigger than P(TC=low). If it is, P(CP=T) increases. However, initial probability distribution of TC shows P(TC=high) is smaller than P(TC=low). Therefore, I predict that:

- P(CP=T) decreases;
- P(CP=F) increases;

(ii) Results

Figure 28 shows results, which is same as my prediction. CP will be 56% likely to be high.

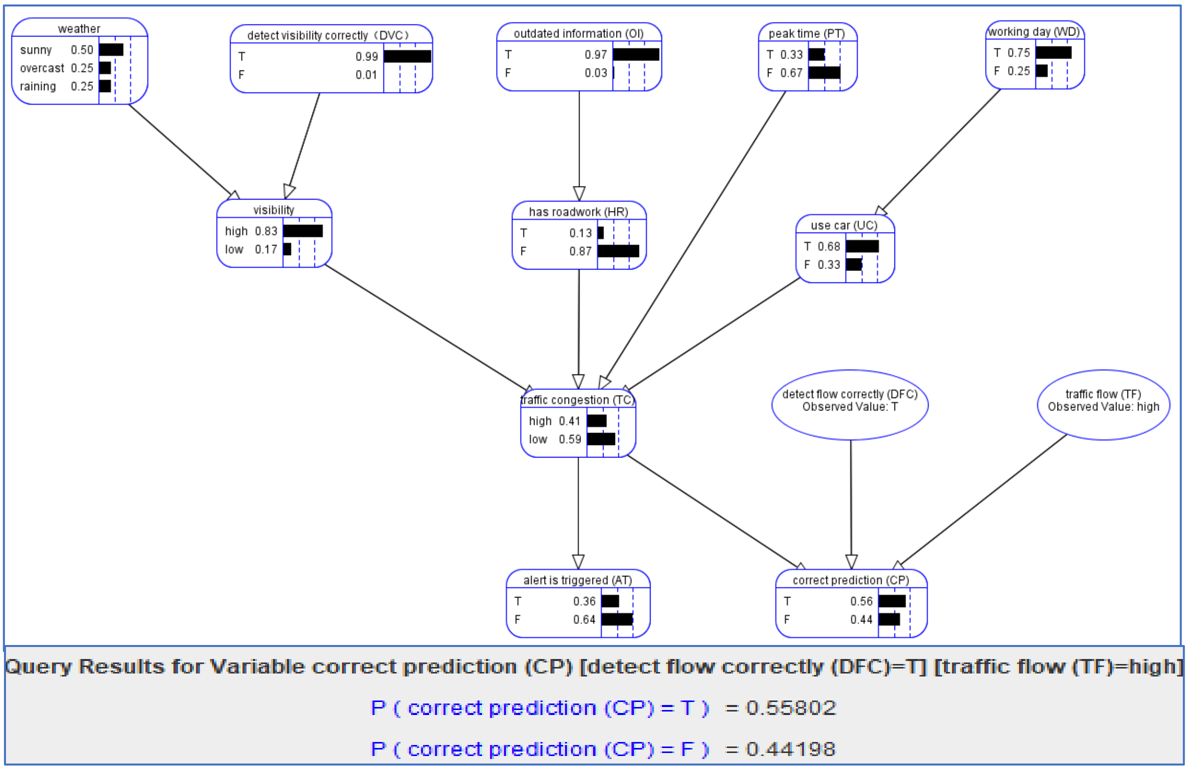


Figure 28 TF is high, and DFC is true

(d) Diagnostic: The agent observes that alert is triggered, how likely that it is in peak time?

(i) Prediction

As the analysis shown in 4.1(iii)(7), when alert is triggered, it reveals that it is very likely to be caused by TC. Thus, P(TC=T) increases. Then according to 4.1(iii)(6), when P(TC=T) increases, it means that PT as a cause of making TC high has higher likelihood to be true than original state. Therefore, I predict that:

- P(PT=T) increases;
- P(PT=F) decreases;

(ii) Results

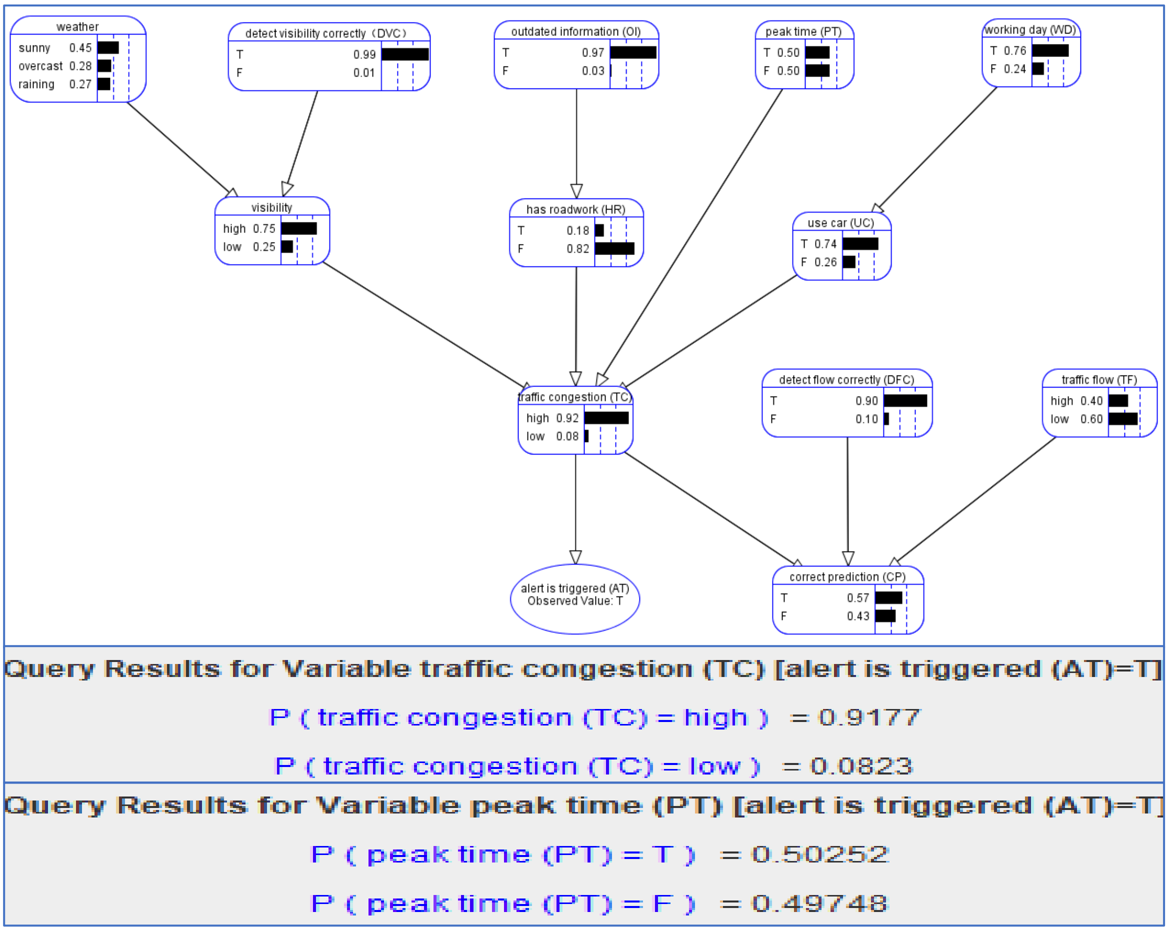


Figure 29 alert is triggered

Figure 29 shows results, which is same as my prediction. It will be 50% likely to be peak time.

(e) Diagnostic: The agent observes that traffic congestion is high, how likely is that information is outdated?

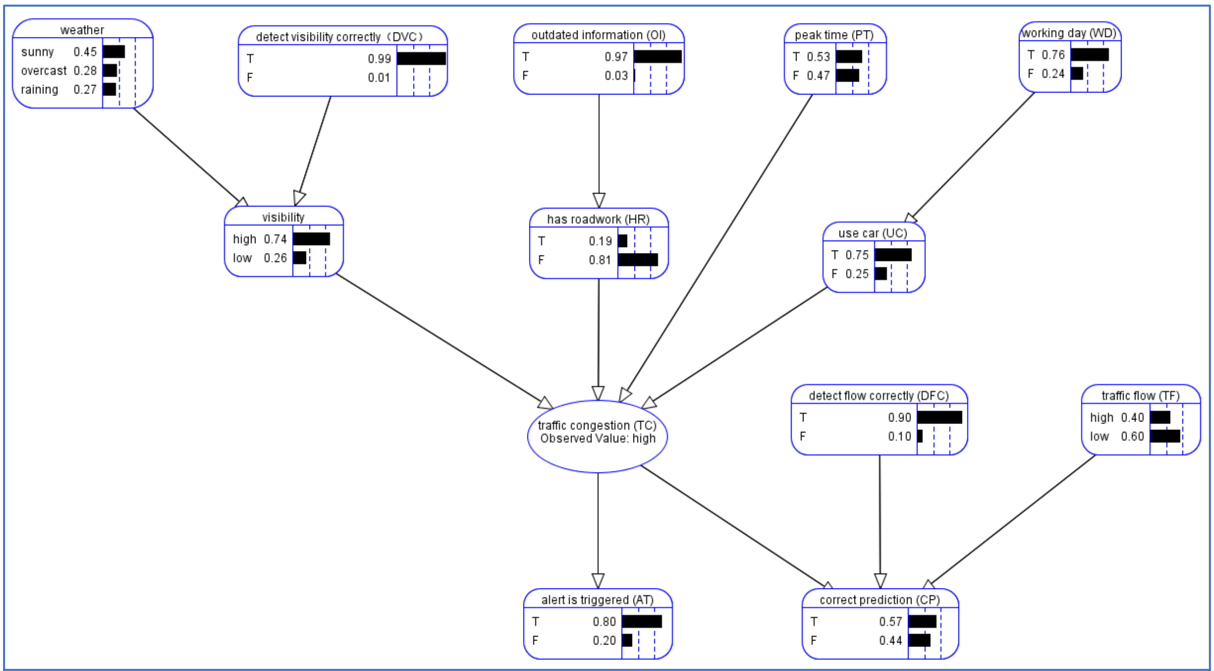
(i) *Prediction*

As the analysis shown in 4.1(iii)(7), when traffic congestion is high, it means that HR as a cause of making TC high has higher likelihood to be true than original state. According to 4.1(iii)(3), The increase of P(HR=T) is caused by the increase of P(OI=T). Therefore, I predict that:

- P(OI=T) increases;
- P(PT=F) decreases;

(ii) *Results*

Figure 30 shows results, which is same as my prediction. It will be 97.048% likely to be peak time. P(PT=T) has very slightly increase.



Query Results for Variable has roadwork (HR) [traffic congestion (TC)=high]

P (has roadwork (HR) = T) = 0.19089

P (has roadwork (HR) = F) = 0.80911

Query Results for Variable outdated information (OI) [traffic congestion (TC)=high]

P (outdated information (OI) = T) = 0.97048

P (outdated information (OI) = F) = 0.02952

Figure 30 traffic congestion is high

4.3 BN for new version alarm system

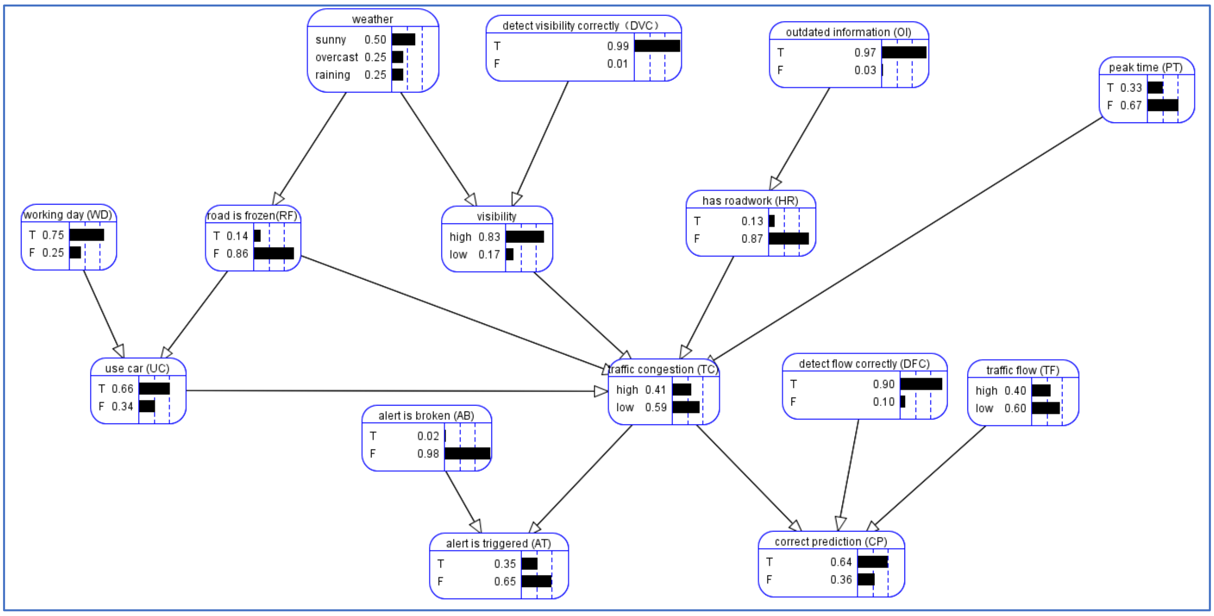


Figure 31 new version alarm system

In new version alarm system, I add two new variables. One is "road is frozen(RF)={T,F}", directly influence UC and TC. Another is "alert is broken", which directly influence AT. "alert is broken" describe the state of physical devices or codes depending on what is alert here. Figure 31 shows this new version system, and figure 32 shows updated probabilities. When weather is sunny, the road still may be frozen because of winter.

weather	P(road is frozen(RF)=T)	P(road is frozen(RF)=F)
sunny	0.1	0.9
overcast	0.15	0.85
raining	0.2	0.8

working day (WD)	road is frozen	P(use car (UC)=T)	P(use car (UC)=F)
T	T	0.6	0.4
T	F	0.7	0.3
F	T	0.5	0.5
F	F	0.6	0.4

	P(alert is broken (AB)=T)	P(alert is broken (AB)=F)
Prior Probability	0.02	0.98

traffic congestion (TC)	alert is broken	P(alert is triggered (AT)=T)	P(alert is triggered (AT)=F)
high	T	0.0	1.0
high	F	0.8	0.2
low	T	0.0	1.0
low	F	0.05	0.95

Figure 32 set new probabilities for new version alarm system

4.4 Prediction and evaluation based on BN for the new alarm system

(a) Predictive: The agent observes that weather is rain, how likely that traffic congestion is high.

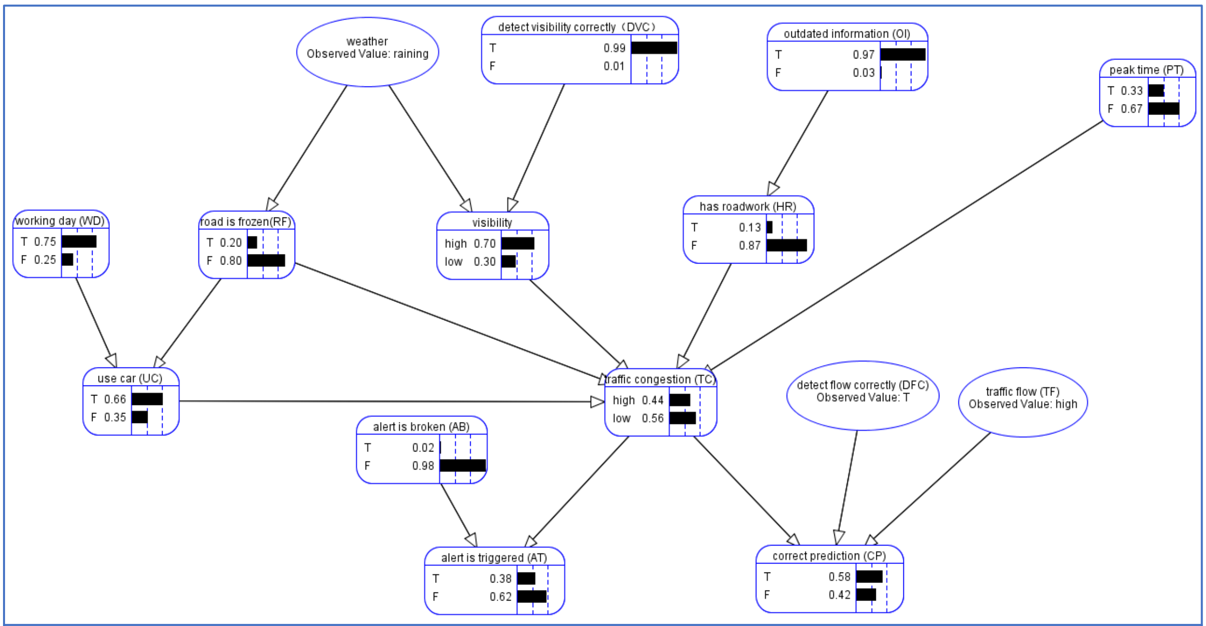
(i) *Prediction*

Firstly, as the analysis shown in 4.1(iii)(2), when weather is overcast, P(visibility=low) increases. Secondly, weather also affects if the road is frozen. When the weather is raining, P(RF=T) increases, which makes the amount of people who use cars increase, namely P(UC=T) decreases. The increase of P(RF=T) stimulates the increase P(TC=high), but the decrease of P(UC=T) reduces P(TC=high). In this case, the influence of RF is bigger than UC. Therefore I predict that:

- P(traffic congestion=high) increases slightly;
- P(traffic congestion=low) decreases slightly;

(ii) *Results*

Figure 33 shows results, which is same as my prediction. TC will be 44% likely to be high.



Query Results for Variable road is frozen(RF) [weather=raining] [detect flow correctly (DFC)=T] [traffic flow (TF)=high]

P (road is frozen(RF) = T) = 0.2

P (road is frozen(RF) = F) = 0.8

Query Results for Variable use car (UC) [weather=raining] [detect flow correctly (DFC)=T] [traffic flow (TF)=high]

P (use car (UC) = T) = 0.655

P (use car (UC) = F) = 0.345

Query Results for Variable traffic congestion (TC) [weather=raining] [detect flow correctly (DFC)=T] [traffic flow (TF)=high]

P (traffic congestion (TC) = high) = 0.44455

P (traffic congestion (TC) = low) = 0.55545

Figure 33 weather is rain

(b) Predictive: The agent observes that the day is a holiday, how likely that traffic congestion is high.

(i) *Prediction*

As the analysis shown in 4.1(iii)(4), when working day is false, P(UC=T) decreases. Further, in this new version BN, the emerge of RF causes P(UC =T) decrease more. Then according to 4.1(iii)(6), when P(UC=T) decreases, I predict that:

- P(TC=high) decreases;
- P(TC=low) increases;

(ii) Results

Figure 34 shows results, which is same as my prediction. TC will be 40% likely to be high.

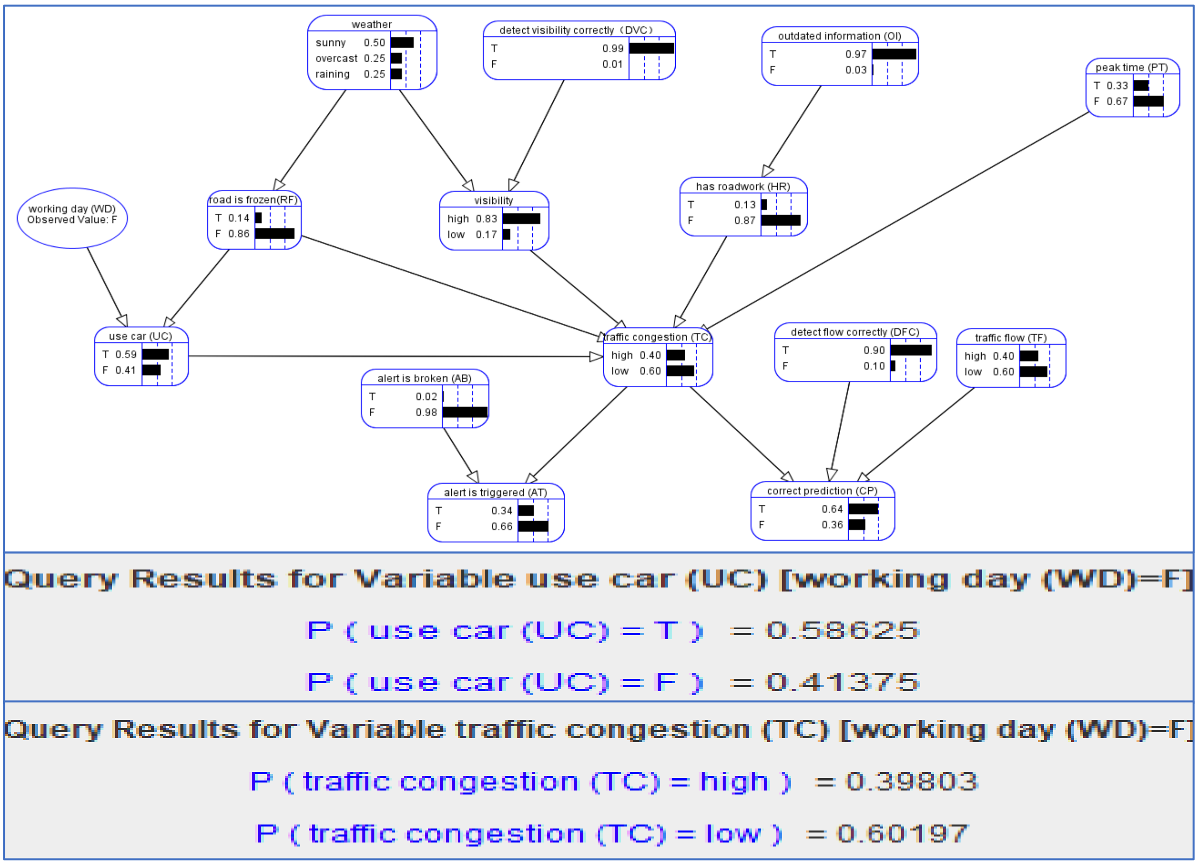


Figure 34 the day is a holiday

<u>(c) Predictive: The agent observes that traffic flow is high, and it is detected correctly, how likely that prediction is true?</u>

(i) Prediction

This part in new version BN is same as in old version BN. As the analysis shown in 4.1(iii)(5), when traffic flow is high, and it is detected correctly, the key in this case is that if P(TC=high) is bigger than P(TC=low). If it is, P(CP=T) increases. However, initial probability distribution of TC shows P(TC=high) is smaller than P(TC=low). Therefore, I predict that:

- P(CP=T) decreases;
- P(CP=F) increases;

***(ii)** Results*

Figure 35 shows results, which is same as my prediction. CP will be 56% likely to be high.

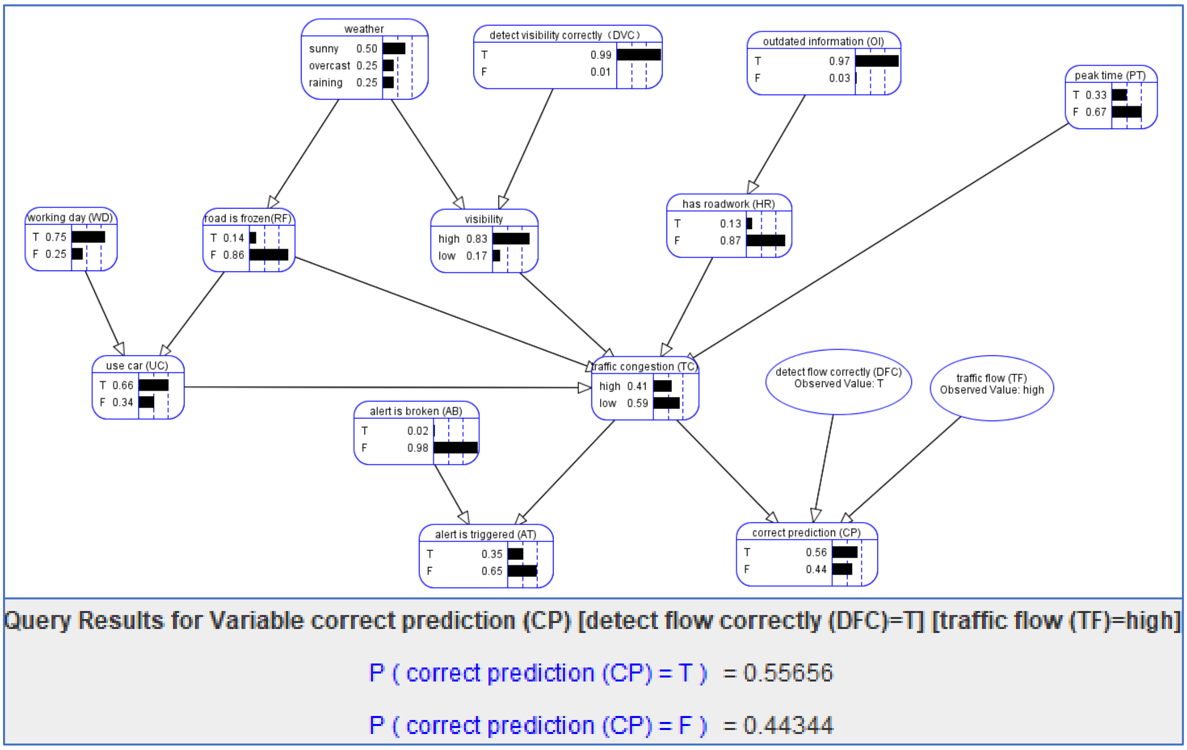


Figure 35 TF is high, and DFC is true

(d) Diagnostic: The agent observes that alert is triggered, how likely that it is in peak time?

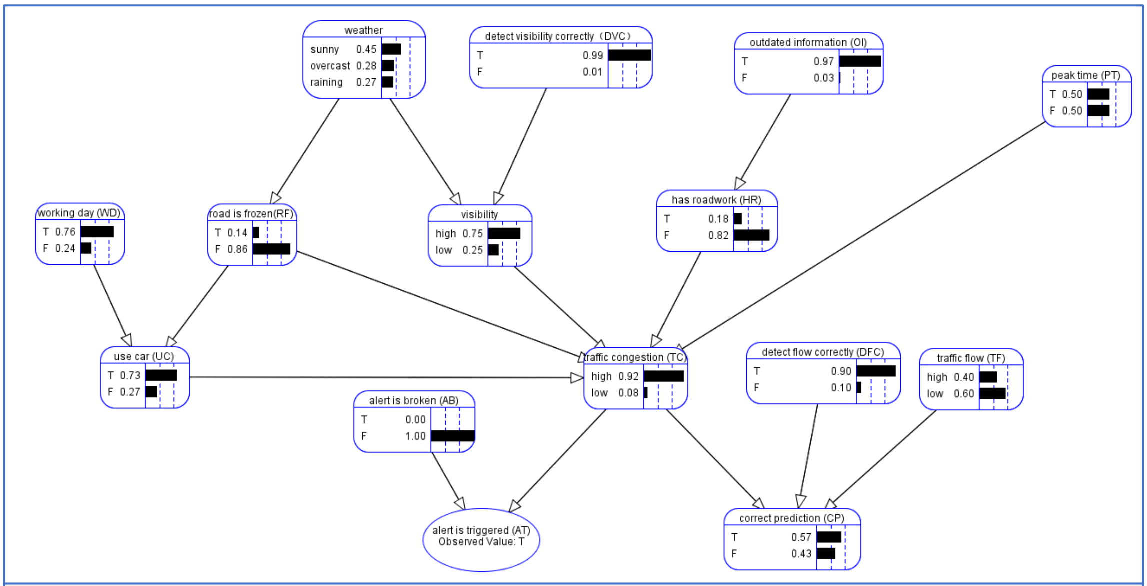
(i) Prediction

As the analysis shown in 4.1(iii)(7), when alert is triggered, it reveals that it is very likely to be caused by TC. Thus, P(TC=T) increases. Then according to 4.1(iii)(6), when P(TC=T) increases, it means that PT as a cause of making TC high has higher likelihood to be true than original state. Therefore, I predict that:

- P(PT=T) increases;
- P(PT=F) decreases;

(ii) Results

Figure 36 shows results, which is same as my prediction. It will be 50% likely to be peak time.



Query Results for Variable traffic congestion (TC) [alert is triggered (AT)=T]

P (traffic congestion (TC) = high) = 0.91709

P (traffic congestion (TC) = low) = 0.08291

Query Results for Variable peak time (PT) [alert is triggered (AT)=T]

P (peak time (PT) = T) = 0.50343

P (peak time (PT) = F) = 0.49657

Figure 36 alert is triggered

<u>(e) Diagnostic: The agent observes that traffic congestion is high, how likely is that information is outdated?</u>

(i) Prediction

As the analysis shown in 4.1(iii)(7), when traffic congestion is high, it means that HR as a cause of making TC high has higher likelihood to be true than original state. According to 4.1(iii)(3), The increase of P(HR=T) is caused by the increase of P(OI=T). Therefore, I predict that:

- P(OI=T) increases;
- P(PT=F) decreases;

(ii) Results

Figure 37 shows results, which is same as my prediction. It will be 97.048% likely to be peak time. P(PT=T) has very slightly increase.

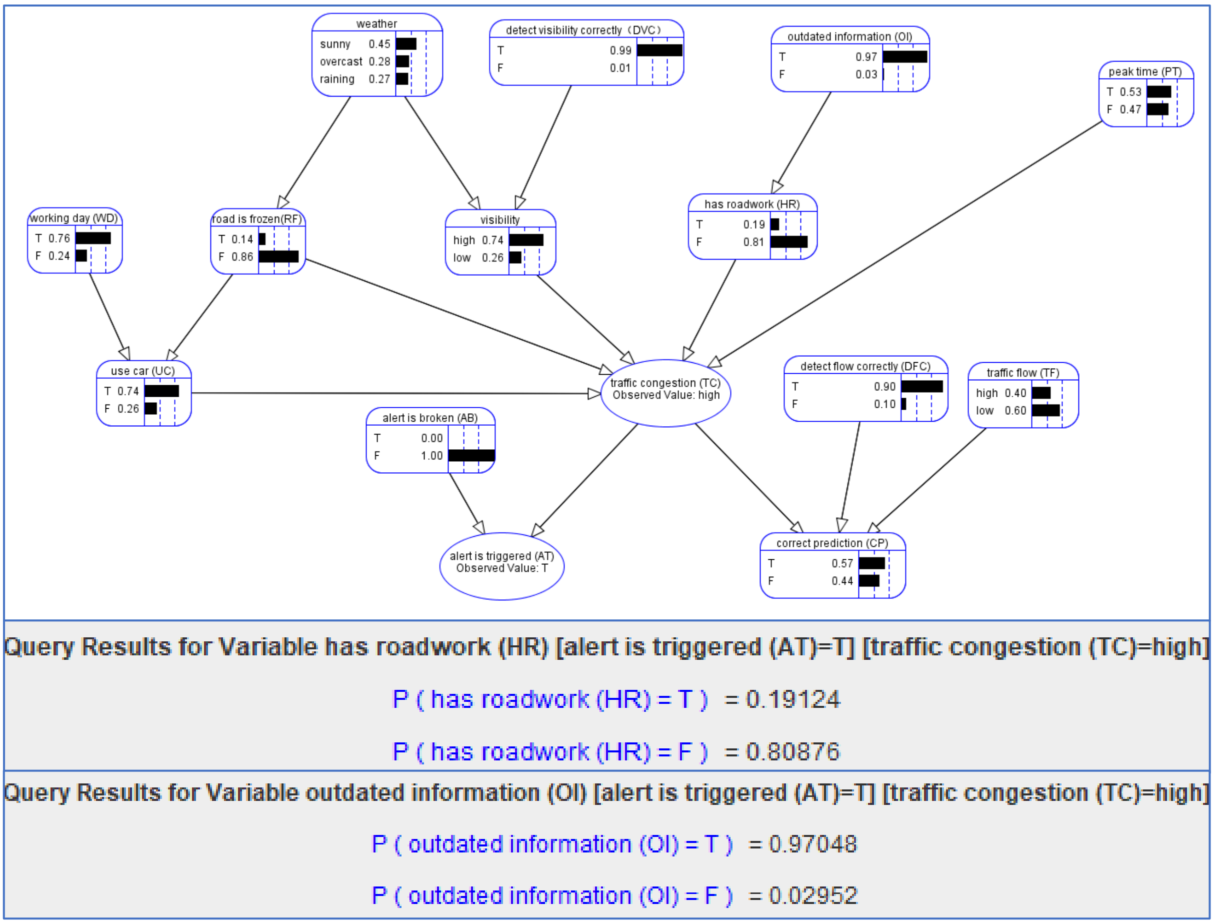


Figure 37 traffic congestion is high

5 General evaluation

Uncertainty is a very important part of artificial intelligence because it might solve problems that other methods cannot solve. Even though, the evidence is not enough or we do not know everything of environment, using uncertainty still help us to give a most rationale results based on calculating probabilities. Bayesian Network is a very common and useful method when reasoning with uncertainty. It can clearly show the relationships between different variables. Building a proper BN for a specific problem is the key to resolve this problem. The more reasonable the network is, the more easy and efficient the problem is to be solved.

Reference

1. Pearl J. Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference. Morgan Kaufmann Publishers, Inc. San Francisco. 1988.
2. Pearl J. Bayesian networks. Department of Statistics, UCLA. 2011 Oct 25.
3. Koski T, Noble J. Bayesian networks: an introduction. John Wiley & Sons; 2011 Aug 26.
4. Friedman N, Linial M, Nachman I, Pe'er D. Using Bayesian networks to analyze expression data. Journal of computational biology. 2000 Aug 1;7(3-4):601-20.
5. Uebersax JS. Breast cancer risk modeling: an application of Bayes nets. Methodology Divison Research Report. Raven Pack International SL. 2004 Aug 2.
6. de Campos LM, Fernández-Luna JM, Huete JF. Bayesian networks and information retrieval: an introduction to the special issue. Information processing & management. 2004 Sep 30;40(5):727-33.
7. Koumenides CL, Shadbolt NR. Combining link and content-based information in a Bayesian inference model for entity search. InProceedings of the 1st Joint International Workshop on Entity-Oriented and Semantic Search 2012 Aug 12 (p. 3). ACM.
8. Jiang X, Cooper GF. A Bayesian spatio-temporal method for disease outbreak detection. Journal of the American Medical Informatics Association. 2010 Jul 1;17(4):462-71.
9. Díez FJ, Mira J, Iturralde E, Zubillaga S. DIAVAL, a Bayesian expert system for echocardiography. Artificial Intelligence in Medicine. 1997 May 1;10(1):59-73.
10. Constantinou AC, Fenton NE, Neil M. pi-football: A Bayesian network model for forecasting Association Football match outcomes. Knowledge-Based Systems. 2012 Dec 31;36:322-39.
11. Constantinou AC, Fenton NE, Neil M. Profiting from an inefficient Association Football gambling market: Prediction, Risk and Uncertainty using Bayesian networks. Knowledge-Based Systems. 2013 Sep 30;50:60-86.
12. Cannavò F, Cannata A, Cassisi C, Di Grazia G, Montalto P, Prestifilippo M, Privitera E, Coltelli M, Gambino S. A multivariate probabilistic graphical model for real-time volcano monitoring on Mount Etna. Journal of Geophysical Research: Solid Earth. 2017 Jan 1.
13. Davis GA. Bayesian reconstruction of traffic accidents. Law, Probability and Risk. 2003 Jun 1;2(2):69-89.

14. Pourret O, Naïm P, Marcot B, editors. Bayesian networks: a practical guide to applications. John Wiley & Sons; 2008 Apr 30.
15. Karvanen J. Study design in causal models. Scandinavian Journal of Statistics. 2015 Jun 1;42(2):361-77.
16. Trucco P, Cagno E, Ruggeri F, Grande O. A Bayesian Belief Network modelling of organisational factors in risk analysis: A case study in maritime transportation. Reliability Engineering & System Safety. 2008 Jun 30;93(6):845-56.
17. Shmilovici A, Kahiri Y, Ben-Gal I, Hauser S. Measuring the efficiency of the intraday forex market with a universal data compression algorithm. Computational Economics. 2009 Mar 1;33(2):131-54.
18. https://medium.com/netflix-techblog/artwork-personalization-c589f074ad76