Heri Wijayanto
Information Technological Engineering, Engineering Faculty, University of Mataram
January, 15th, 2013

Articles of
BAYESIAN NETWORK

Bayes Theorem

Figure 1. Thomas Bayes

Bayesian Network is an extension of Bayes Theorem that was introduced by Thomas Bayes from London (1702-1761). His theory was written in *Essay towards solving a problem in the doctrine of chances* and it was published in the *Philosophical Transactions of the Royal Society of London* in 1764. The paper was sent to the Royal Society by Richard Price who was Thomas Bayes's best friend [http://www-history.mcs.st-and.ac.uk/Biographies/Bayes.html].

Bayes theorem concerns about a probability that is dependent to other probability of the previous occurrence. For example, someone who just came out from a train said that he had just had an interesting conversation with someone else (x), probability of x is a women is 50%; therefore, it can be written as P(W)=50%. From the next information is known that x had just gone home from a boutique. Therefore, this fact strengthens probability that x is a woman. If x is A and boutique is B, that probability can be written as P(A|B) and the equation of it is presented below.

$$P(A|B) = \frac{P(B|A)\,P(A)}{P(B)}$$

Prof. Gleen Tesler in his presentation [Tesler Glenn] explained it by Venn diagram which is shown an experiment of flipping a coin 3 times. The sample space is S = {HHH, HHT, HTH, HTT, THH, THT, TTH, TTT} and set of first flip is heads is written by A ={HHH, HHT,HTH, HTT} and B represents two flips are heads B = {HHT,HTH, THH}.

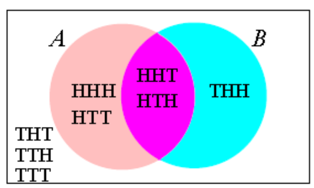

Figure 2. Venn Diagram of flipping a coin three times

From the figure 2 above, it is known that probability of A is 4/8 or 0.5, and B is 3/8. The probability of A that is dependent to the occurrence of B. Now, B becomes the universe of this probability; therefore it can be depicted in figure 3 below.

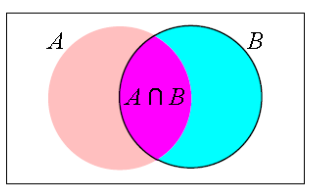

Figure 3. Venn Diagram of the probability of A that is dependent to B

The equation of figure 3 abova can be written as:

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}$$

By the same way, probability of P(B|A)= $\frac{P(A\cap B)}{P(B)}$; P($A \cap B) = P(B|A)P(B)$; therefore P(A|B)=$\frac{P(B|A)P(B)}{P(B)}$
and

$$P(A|B) = \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|A^c)P(A^c)}$$

Because

$$P(B) = P(B \cap A) + P(B \cap A^c) = P(B|A)P(A) + P(B|A^c)P(A^c)$$

Independency of two events is known if the knowledge of one event does not affect the probability of the other event

$$P(A|B) = P(A) \quad \Leftrightarrow \quad \frac{P(A \cap B)}{P(B)} = P(A) \quad \Leftrightarrow \quad P(A \cap B) = P(A)P(B)$$

Bayesian Network

Bayesian network is also known as Believe Network that is a member of probabilistic graphical model [Ruggeri F, et al, 2007]. This graphic structure is used for presenting knowledge about uncertain domain. Bayesian network in [Wong, et al] is a triplet (N, E,P). N is a set of nodes. Every node is labeled with random variable that is associated with space. Every single node is also unique; therefore, the term of node and variable can be exchanged. E is a set of edge; therefore, D=(N,E) that is a Directed Acyclic Graph (DAG). This edges show the existence of direct causal influence relation among variables that are connected. For every single node $A_i \in N$, the strength of casual influence of parent node π_i is determined by Conditional Probability Distribution $p(A_i|\pi_i)$ from A_i that are conditioned on the values of parent A_i. The base of assumption of independency that is attached in Bayesian Network is an independent variable from its non-descendant that is given by parent. Whereas, Joint Probability Distribution (JPD) of a Bayesian Network with n nodes, P can be formed by factorization form as $P=p(A_1 \ldots A_n)=\prod_{i=1}^{n} p(A_i|\pi_i)$. For instance, distribution of Bayesian Network probability that is shown in figure 1 can be written as:

$$p(x_1, x_2, x_3, x_4, x_5, x_6) = p(x_1)p(x_2|x_1)p(x_3|x_1)p(x_4|x_1, x_2)p(x_5|x_2, x_3)p(x_6|x_5)$$

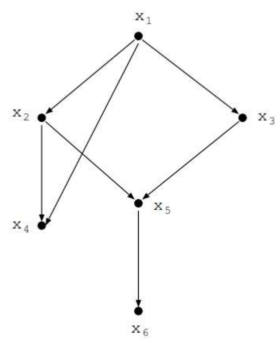

Figure 4. Example of *Bayesian Network* and its equation

Example of Simple BN

The easy way to calculate probability in Bayesian Network is by drawing Join Probability Distribution (JPD) of each node. For example, the probability of rain is 0.6 and the probability the grass to be wet if rain is 0.8 and the probability of grass to be wet if not rain is 0.3. Therefore, BN for this example that is simulated by Netica is shown in figure 5 below.

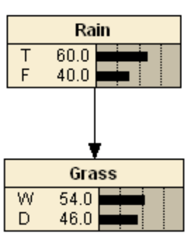

Figure 5. Simple BN

Suppose the JPD of Grass Node is shown below

Rain	**Grass**	**P(G|R) (%)**
T	W	80
T	D	20
F	W	15
F	D	85

Figure 6. JPD of Grass Node

$P(G^W|R) = P(G^W|R^T)\,P(R^T) + P(G^W|R^F)P(R^F)$
$= 0.8 \times 0.6 + 0.15 \times 0.4$
$= 0.48 + 0.06 =$ **0.54**
$\Rightarrow P(G^D) = 1 - 0.54 =$ **0.46**

If it is known that grass is wet, what is the probability of Rain is true

$P(R^T|G^W) = P(G^W|R^T)P(R^T)/P(G^W)$

$= 0.8 \times 0.6/0.54 = 0.48/0.54 = \mathbf{0.889}$

$P(R^F|G^W) = 0.15 \times 0.4/0.54 = 0.6/0.54 = \mathbf{0.111}$

or $1 - 0.889 = \mathbf{0.111}$

The simulation of the calculation above is shown in figure 7.

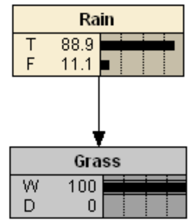

Figure 7. Probability of Rain if it is known that Grass is wet

Example of Simple Converging BN

Figure 8 shows the example of converging BN which has three nodes, sprinkler, rain, and grass. The grass can be wet by sprinkler is on or rain is true.

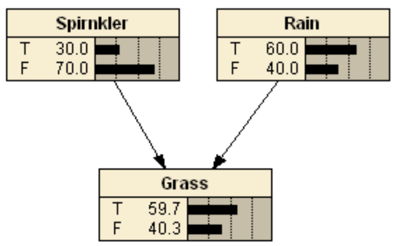

Figure 8. BN of Simple Converging Node

Figure 9 shows the JPD of the example in figure 8.

Sprinkler	Rain	Grass	P(G|S,R) (%)
T	T	T	90
T	T	F	10
T	F	T	65
T	F	F	35
F	T	T	75
F	T	F	25
F	F	T	15
F	F	F	85

Figure 9. JPD of figure 8.

The probability of grass is wet can be calculated by summing of all condition that grass is wet there are

$P(G^T) = P(G^T|S^T,R^T)P(S^T)P(R^T) + P(G^T|S^T,R^F)P(S^T)P(R^F) + P(G^T|S^F,R^T)P(S^F)P(R^T) + P(G^T|S^F,R^F)P(S^F)P(R^F)$

$= 0.9 \times 0.3 \times 0.6 + 0.65 \times 0.3 \times 0.4 + 0.75 \times 0.7 \times 0.6 + 0.15 \times 0.7 \times 0.4$

$= 0.162 + 0.078 + 0.315 + 0.042 = \mathbf{0.597}$

The probability of grass is wet if sprinkler is on is as follows

$P(G^T|S^T,R)$

$= P(G^T|S^T,R^T)P(S^T)P(R^T) + P(G^T|S^T,R^F)P(S^T)P(R^F) + P(G^T|S^F,R^T)P(S^F)P(R^T) + P(G^T|S^F,R^F)P(S^F)P(R^F)$

$= 0.9 \times 1 \times 0.6 + 0.65 \times 1 \times 0.4 + 0.75 \times 0 \times 0.6 + 0.15 \times 0 \times 0.4$

$= 0.54 + 0.26 + 0 + 0$

$= \mathbf{0.8}$

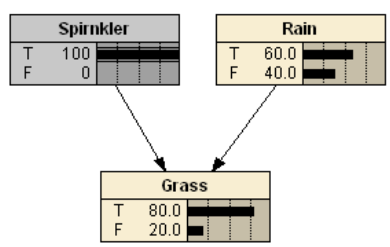

Figure 10. The result of simulation if Spinkler is on

The probability of sprinkler is on and the rain is true if the grass is wet is as follow (Sprinkler and Rain are independent)

Heri Wijayanto
Information Technological Engineering, Engineering Faculty, University of Mataram

Bayes Theorem

$P(S^T,R|G^T) = (P(G^T|S^T,R^T).P(S^T)P(R^T) + P(G^T/S^T,R^F)P(S^T)P(R^F))/P(G)$
$= (0.9x0.3x0.6 + 0.65.0.3.0.4) /0.597$
$= (0.162 + 0.078)/0.597 = \mathbf{0.402}$

$P(S,R^T|G^T) = (P(G^T|S^T,R^T).P(S^T)P(R^T) + P(G^T|S^F,R^T)P(S^F)P(R^T))/P(G)$
$= (0.9x0.3x0.6 + 0.75.0.7.0.6) /0.597$
$= (0.162 + 0.315)/0.597 = \mathbf{0.799}$

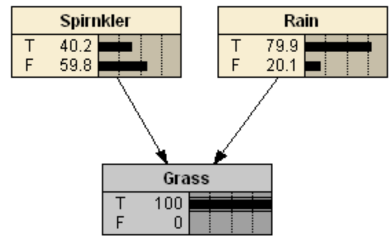

Figure 11. The result of simulation if Grass is wet

The probability of Rain if sprinkler is on and grass is wet can be reckoned by the probability of Grass if sprinkler is on such as in figure 10, and then the probability of Rain can be calculated as follows.

$P(S^1,R^T|(G^T|S^1))$
$= (P(G^T|S^1,R^T).P(S^1)P(R^T)+ P(G^T/S^0,R^T)P(S^o)P(R^T))/P(G)$
$= (0.9x1x0.6 + 0.75x0x0.6) /0.8$
$= (0.54 + 0)/0.8 = \mathbf{0.675}$

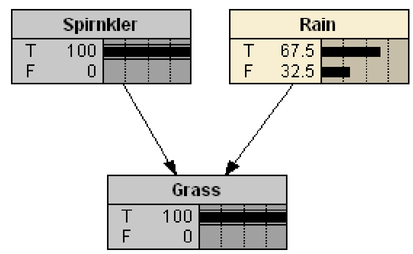

Figure 12. Probability when sprinkler is on and the grass is wet

Example of Simple Diverging BN

Figure 13 shows the example of diverging BN that contains of 3 node, there are rain, storm and flooding. Probability of flooding and storm are dependent to the probability of Rain, but node flooding and storm are independent to each other. Calculating probability of flooding P(F) and storm P(S) is similar to the calculating of example in figure 7 that is applied twice for P(F) and P(S).

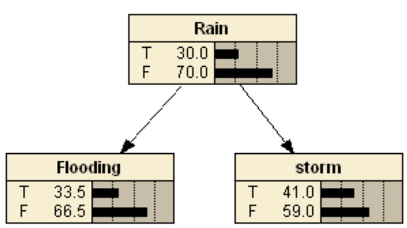

Figure 13. Example of simple diverging BN

Rain	Flooding	P(F|R) (%)
T	T	65
T	F	35
F	T	20
F	F	80

Figure 14. JPD of Flooding Node

Rain	Storm	P(S|R) (%)
T	T	55
T	F	45
F	T	35
F	F	65

Figure 15. JPD of Storm Node

In case flooding is true, it will change the probability of Rain, and also storm. F and S are independent; therefore

$P(R^T|F^T)$ = $P(F^T|R^T)P(R^T) / P(F)$
= $(0.65x0.3)/0.335$
= $0.195/0.335$
= **0.582**

Even more, node Strom has new probability:

$P(S^T|R)$ = $P(S^T|R^T).P(R^T) + P(S^T|R^F).P(R^F)$
= $0.55x0.582 + 0.35x0.418$
= $0.3201 + 0.1463$ = **0.466**

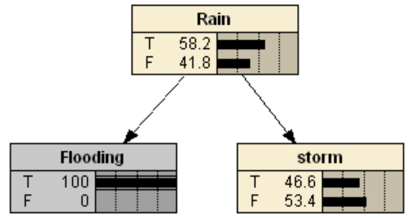

Figure 16. Probability after simulating that Rain is true

If flooding and storm are happened simultaneously, the probability of rain is $P(R|F,S) = P(F,S|R)P(R) / P(F,S)$ to calculate this equation, $P(F,S|R)$ and $P(F,S)$ must be calculated first. Because F and S are **independent (use independency rule)**, therefore

$P(F,S|R^T) = P(F|R^T).P(S|R^T) = 0.65x0.55 = 0.36$
$P(F,S|R^F) = P(F|R^F).P(S|R^F) = 0.2x0.35 = 0.07$
$P(F,S) = P(F,S|R^T)P(R^T) + P(F,S|R^F)P(R^F)$
= $0.36x0.3 + 0.07x0.7$
= $0.108 + 0.049 = 0.157$
$P(R|F,S) = P(F,S|R)P(R) / P(F,S)$
= $0.108/0.157$ = **68.7**

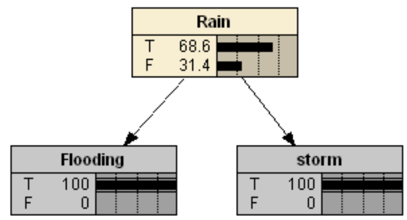

Figure 16. Probability after simulating that Rain is true and Storm is true

Example of Simple Linier BN

To calculate probability in Flooding and Epidemic are:

$P(F)$ = $P(F|R^T).P(R^T) + P(F|R^F).P(F^F)$
= $0.65x0.3 + 0.25x0.7$ = **0.37**

$P(E)$ = $P(E|F^T).P(F^T) + P(E|F^F).P(F^F)$
= $0.85x0.37 + 0.3x0.63$ = **0.503**

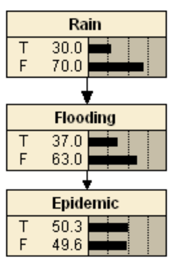

Figure 17. Simple linier BN

Rain	Flooding	P(F|R) (%)
T	T	65
T	F	35
F	T	25
F	F	75

Figure 18. JPD of Flooding

Flooding	Epidemic	P(E|F) (%)
T	T	85
T	F	15
F	T	30
F	F	70

Figure 19. JPD of Epidemic

Because F is initiated, R and E are independent

$P(R|F)$ = $P(F|R).P(R)/P(F)$
= $0.65.0.3/0.37$=**0.527**

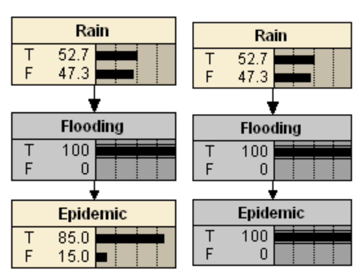

Figure 21 and 22. Simulation of Flooding is known

Because P(R)=true ➔ $P(F^T)=0.65$ and $P(F^T)=0.35$
New $P(E) = P(E|F^T).P(F^T) + P(E|F^F).P(F^F)$
$= 0.85x0.65 + 0.3x0.35$
$= 0.5525 + 0.105 = 0.6575$
$P(F|E) = P(E|F).P(F)/P(E)$
$= 0.85x0.65/0.6575 =$ **0.84**

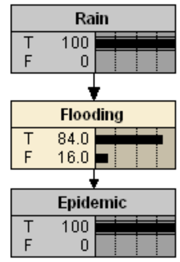

Figure 23. Simulation of Rain and Epidemic are known

If epidemic is known, probability of flooding and rain become
$P(F|E) = P(E|F).P(F)/P(E)$
$= 0.85x0.37 / 0.503 =$ **0.625**

To find probability in Rain,
chain rule [Jianguo Ding, 2010]is used.

To infer the probability in Rain by chain rule, it is easy if JPD in figure 18 and 19 are used.

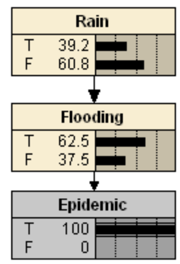

Figure 20. Simulation of Epidemic is known

$P(R^T|FE^T) = P(FE^T|R^T).P(R^T)/P(FE^T)$
$= P(\mathbf{F}|R^T).P(E^T|\mathbf{F}).P(R^T)/ P(\mathbf{F}E^T)$
Because F is not instantiated; therefore Rain and Epidemic are dependent
$P(\mathbf{F}|R^T).P(E^T|\mathbf{F}).P(R^T) = P(R^T).P(F^T|R^T).P(E^T|F^T) + P(R^T).P(F^F|R^T).P(E^T|F^F)$
$= 0.3x0.65x0.85 + 0.3x0.35x0.3 = 0.19725$

Rain	P(R) (%)
T	30
F	70

Rain	Flooding	P(F|R) (%)
T	T	65
T	F	35
F	T	25
F	F	75

Flooding	Epidemic	P(E|F) (%)
T	T	85
T	F	15
F	T	30
F	F	70

$P(\mathbf{FE}^T) = P(R^T).P(F^T|R^T).P(E^T|F^T) + P(R^T).P(F^F|R^T).P(E^T|F^F) + P(R^F).P(F^T|R^F).P(E^T|F^T) + P(R^F).P(F^F|R^F).P(E^T|F^F)$
$= 0.3x0.65x0.85 + 0.3x0.35x0.3 + 0.7x0.25x0.85 + 0.7x0.75x0.3 = 0.5035$

$P(R^T|FE^T) = 0.19725/0.5035 =$ **39.1757**

References

Tesler Glenn; "*Conditional Probability and Bayes' Theorem (2.4) Independence (2.5)*"; University of California, San Diego; http://www.math.ucsd.edu/~gptesler/186/slides/bayesthm_11-handout.pdf
Wong S.K.M, Xiang. Y and Nie. X;"*Representation of Bayesian Networks as Relational Databases*"; Department of Computer Science University of Regina Regina, Saskatchewan, Canada S4S 0A2.
Ruggeri F., Faltin F., Kenett R; "*Bayesian Networks*"; Encyclopedia of Statistics in Quality & Reliability, Wiley & Sons; 2007.

Heri Wijayanto
Information Technological Engineering, Engineering Faculty, University of Mataram
January, 15th, 2013

Jianguo Ding. Probabilistic Inferences in Bayesian Networks, Bayesian Network, Ahmed Rebai (Ed.), ISBN: 978-953-307-124-4, InTech, Available from: http://www.intechopen.com/books/bayesiannetwork/probabilistic-reasoning-inferences-in-bayesian-networks; 2010.