Machine Learning

Bayesian Network

March Brag
cesarbrma91@gmail.com
@MarchBragagnini

Universidad Católica San Pablo

Centro de Investigación
e Innovación en
Ciencia Computación

Outline

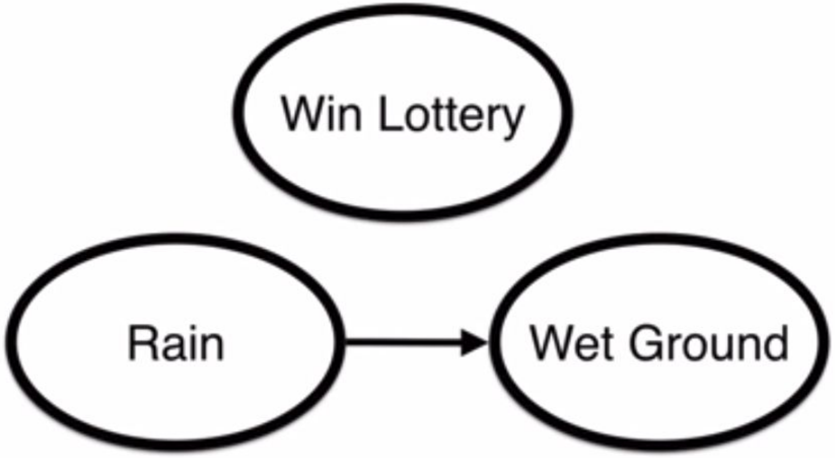
1. Bayes’ Theorem
2. Bayes Network
3. Conditional independencies in BN
4. Hybrid Bayesian Network
5. Exact Inference in Bayesian Networks
6. Approximate Inference in Bayesian Networks
7. Summary Bayesian Networks
8. Weka
9. References

Bayes' Theorem

$$p(Y|X) = \frac{p(X|Y)p(Y)}{p(X)}$$

- Comments:
  - Bayes' rule tells us how to 'invert' conditional probabilities, i.e. to find P(B|A) from P(A|B).

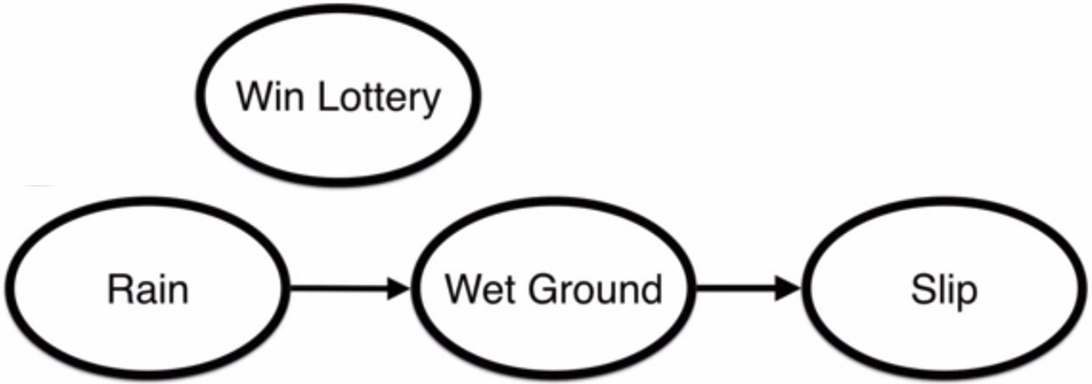
Bayes Network



$$P(L, R, W) =$$

[ ] https://www.youtube.com/watch?v=TuGDMj43ehw

Bayes Network

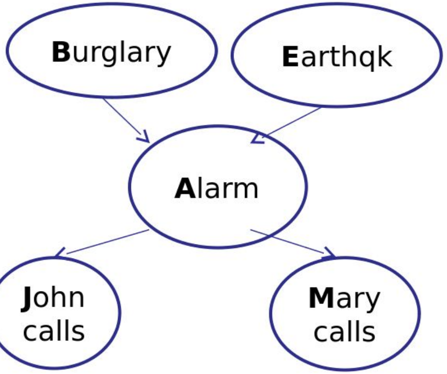


$$P(L, R, W, S) =$$

[ ] https://www.youtube.com/watch?v=TuGDMj43ehw

Bayes Network

B	P(B)
+b	0.001
-b	0.999



E	P(E)
+e	0.002
-e	0.998

B	E	A	P(A|B,E)
+b	+e	+a	0.95
+b	+e	-a	0.05
+b	-e	+a	0.94
+b	-e	-a	0.06
-b	+e	+a	0.29
-b	+e	-a	0.71
-b	-e	+a	0.001
-b	-e	-a	0.999

A	J	P(J|A)
+a	+j	0.9
+a	-j	0.1
-a	+j	0.05
-a	-j	0.95

A	M	P(M|A)
+a	+m	0.7
+a	-m	0.3
-a	+m	0.01
-a	-m	0.99

$$P(+b, -e, +a, -j, +m) =$$

Conditional independencies in BN

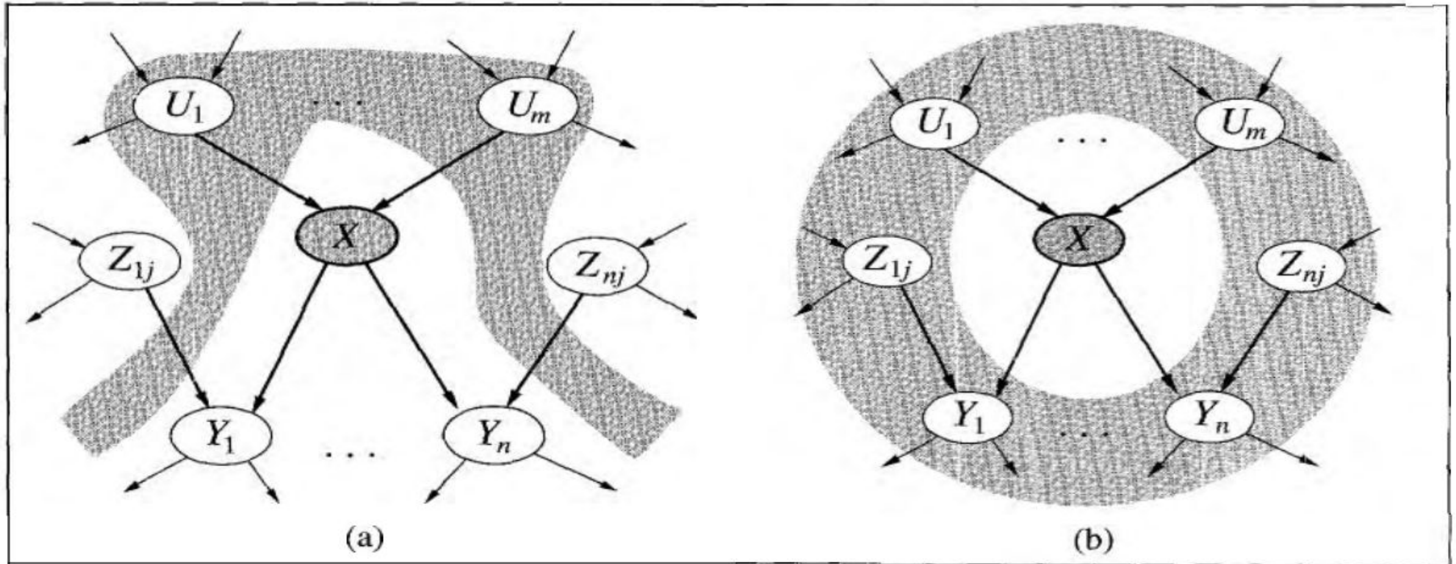
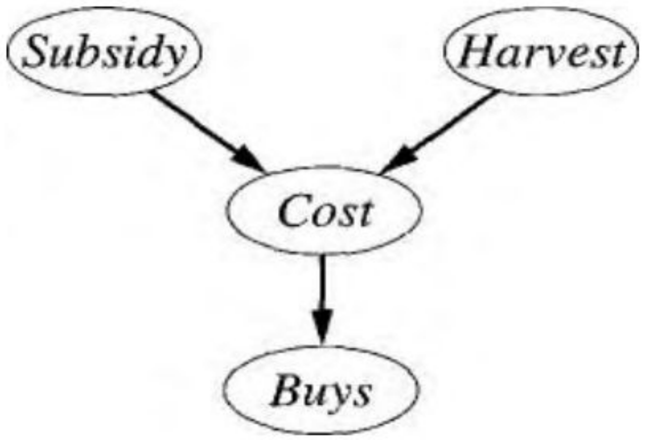


Figure 14.4 (a) A node X is conditionally independent of its non-descendants (e.g., the Z_{ij}s) given its parents (the U_is shown in the gray area). (b) A node X is conditionally independent of all other nodes in the network given its Markov blanket (the gray area).

[ ] Artificial Intelligence A Modern Approach, 2nd edition, Russel & Norvig

Hybrid Bayesian Network

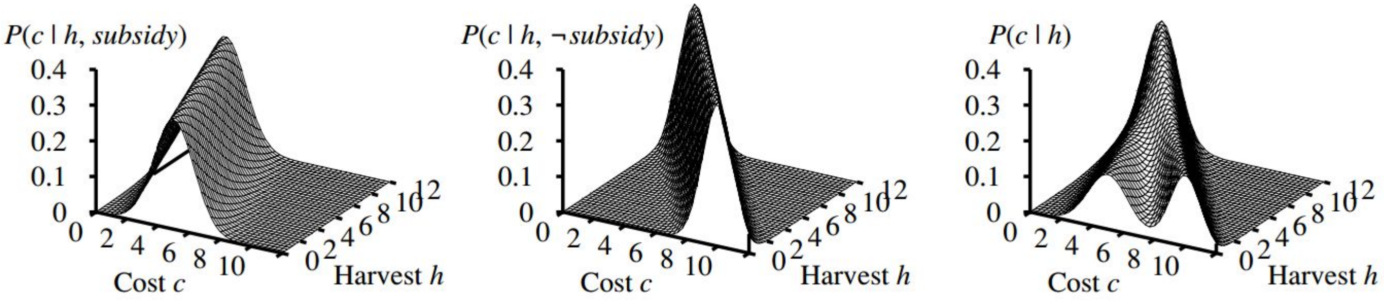
A network with both discrete and continuous variables is called a hybrid Bayesian network. To specify a hybrid network, we have to specify two new kinds of distributions: the conditional distribution for a continuous variable given discrete or continuous parents; and the conditional distribution for a discrete variable given continuous parents.



A simple network with discrete variables(Subsidy and Buys)and continuous variables (Harvest and Cost).

[ ] **Artificial Intelligence A Modern Approach, 2nd edition, Russel & Norvig**

Hybrid Bayesian Network



(a) (b) (c)

The graphs in (a) and (b) show the probability distribution over Cost as a function of Harvest size, with Subsidy true and false, respectively. Graph (c) shows the distribution P(Cost | Harvest), obtained by summing over the two subsidy cases.

$$P(c\,|\,h, subsidy) \;=\; N(a_t h + b_t, \sigma_t^2)(c) = \frac{1}{\sigma_t\sqrt{2\pi}}\, e^{-\frac{1}{2}\left(\frac{c-(a_t h+b_t)}{\sigma_t}\right)^2}$$

$$P(c\,|\,h, \neg subsidy) \;=\; N(a_f h + b_f, \sigma_f^2)(c) = \frac{1}{\sigma_f\sqrt{2\pi}}\, e^{-\frac{1}{2}\left(\frac{c-(a_f h+b_f)}{\sigma_f}\right)^2}$$

[ ] Artificial Intelligence A Modern Approach, 2nd edition, Russel & Norvig

Exact Inference in Bayesian Networks

Considering the following equation:

$$\mathbf{P}(X \mid \mathbf{e}) = \alpha\, \mathbf{P}(X, \mathbf{e}) = \alpha \sum_{\mathbf{y}} \mathbf{P}(X, \mathbf{e}, \mathbf{y})$$

Calculate

$$\mathbf{P}(Burglary \mid JohnCalls = true, MaryCalls = true)$$

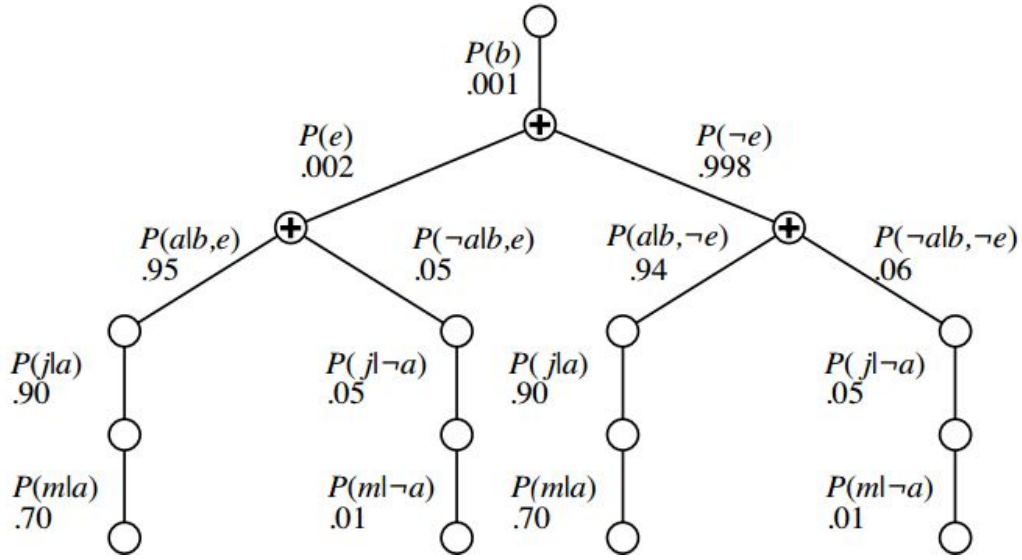
$$\mathbf{P}(B \mid j, m) = \alpha\, \mathbf{P}(B, j, m) = \alpha \sum_{e} \sum_{a} \mathbf{P}(B, j, m, e, a,)$$

[ ] Artificial Intelligence A Modern Approach, 2nd edition, Russel & Norvig

Exact Inference in Bayesian Networks

$$P(b \mid j, m) = \alpha \sum_e \sum_a P(b)P(e)P(a \mid b, e)P(j \mid a)P(m \mid a)$$

$$P(b \mid j, m) = \alpha\, P(b) \sum_e P(e) \sum_a P(a \mid b, e)P(j \mid a)P(m \mid a)$$

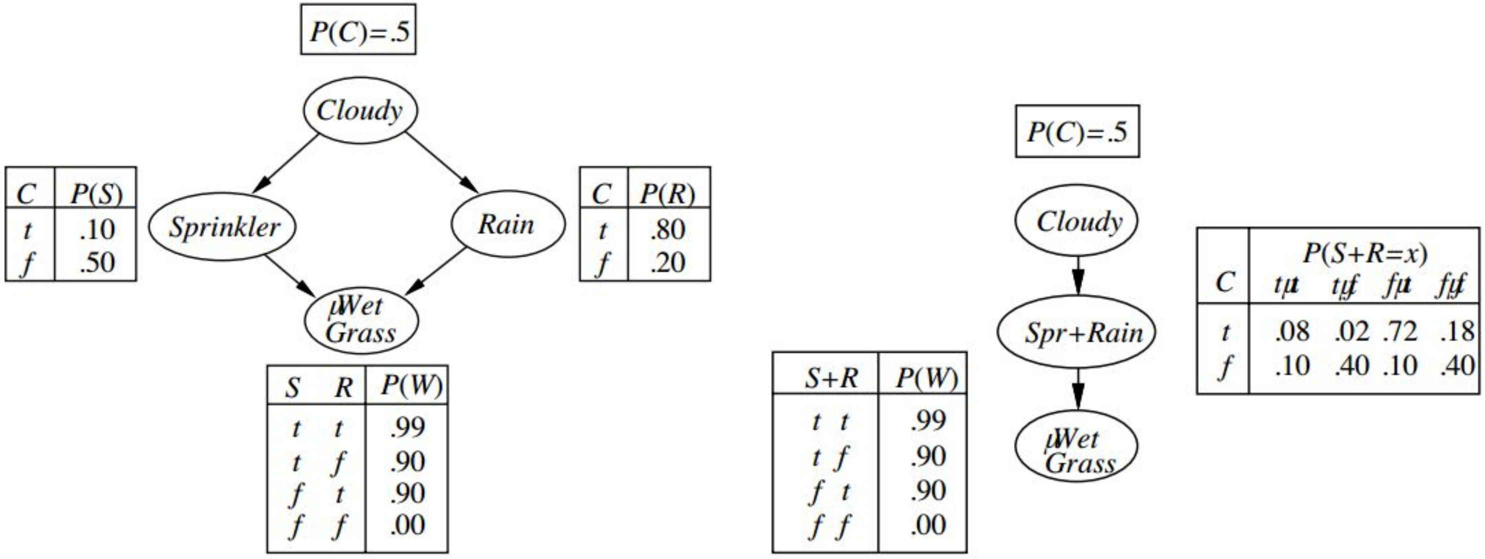


[ ] **Artificial Intelligence A Modern Approach, 2nd edition, Russel & Norvig**

Exact Inference in Bayesian Networks

Complexity: NP Hard (generally)
One method: For example, in a polytree network, one would need to issue O(n) queries CLUSTERING costing O(n) each, for a total of 0 (n 2) time. Using clustering algorithms (also known as JOIN TREE join tree algorithms), the time can be reduced to O(n).

Clustering:



(a) (b)

[ ] **Artificial Intelligence A Modern Approach, 2nd edition, Russel & Norvig**

Approximate Inference in Bayesian Networks

1. Monte Carlo algorithms.
    - Markov Chain Monte Carlo
        - Gibbs Sampler

2. Likelihood Weighting

[ ] Artificial Intelligence A Modern Approach, 2nd edition, Russel & Norvig

Summary Bayesian Networks

1. A Bayesian network is a directed acyclic graph whose nodes correspond to random variables; each node has a conditional distribution for the node, given its parents.
2. Bayesian networks provide a concise way to represent conditional independence relationships in the domain.
3. A Bayesian network specifies a full joint distribution; each joint entry is defined as the product of the corresponding entries in the local conditional distributions. A Bayesian network is often exponentially smaller than the full joint distribution.
4. Stochastic approximation techniques such as likelihood weighting and Markov chain Monte Carlo can give reasonable estimates of the true posterior probabilities in a network and can cope with much larger networks than can exact algorithms.

[ ] Artificial Intelligence A Modern Approach, 2nd edition, Russel & Norvig

Weka

Download & commands

- https://www.cs.waikato.ac.nz/ml/weka/downloading.html

```
$ cd weka*
$ java -jar weka.jar
$ curl https://archive.ics.uci.edu/ml/machine-learning-databases/iris/iris.data --output iris.csv
```

References

- https://machinelearningmastery.com/load-csv-machine-learning-data-weka/
- https://www.youtube.com/watch?v=tpH905jiBZ0
- http://web.ydu.edu.tw/~alan9956/docu/refer/BayesWEKA.pdf
- https://www.youtube.com/watch?v=TuGDMj43ehw
- Artificial Intelligence: A Modern Approach http://aima.cs.berkeley.edu/
- CS 5804: Introduction to Artificial Intelligence http://courses.cs.vt.edu/cs4804/Fall16/
- UC Berkeley CS188 Intro to AI -- Course Materials http://ai.berkeley.edu/lecture_slides.html
- JavaBayes - https://www.cs.cmu.edu/~javabayes/Home/node3.html

Machine Learning

Bayesian Network

March Brag

cesarbrma91@gmail.com

@MarchBragagnini

Universidad Católica San Pablo

Centro de Investigación e Innovación en Ciencia Computación