Escola Tècnica Superior d'Enginyeria
Universitat Rovira i Virgili

[DΣIM]

Approximate Reasoning

Aida Valls

Bayesian Networks

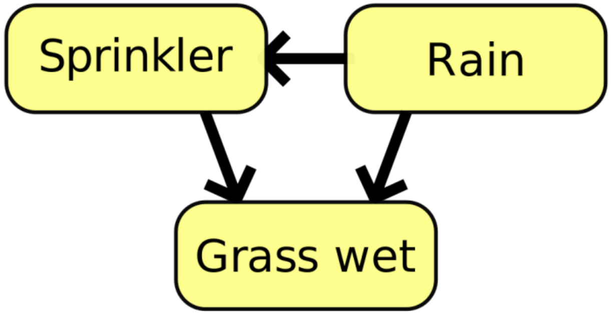
MESIIA – MIA

Introduction

- What are they?
  - Bayesian nets are a network-based framework for representing and analyzing models involving uncertainty
  - They follow a quasi-probabilistic model based on Bayes theorem
- Why are interesting?
  - Graphical interpretation which facilitates the understanding
  - Availability of easy to use commercial software
  - Growing number of creative applications

Introduction

- Bayesian Network is an acyclic graph
- Each node is an uncertain variable (it can have different possible values, states – it is not binary)
- The arcs indicate "causality" relations between the variables (conditional dependency)
- Nodes which are not connected indicate variables that are conditionally independent
- Any variable can be instantiated, if we know its value with certainty

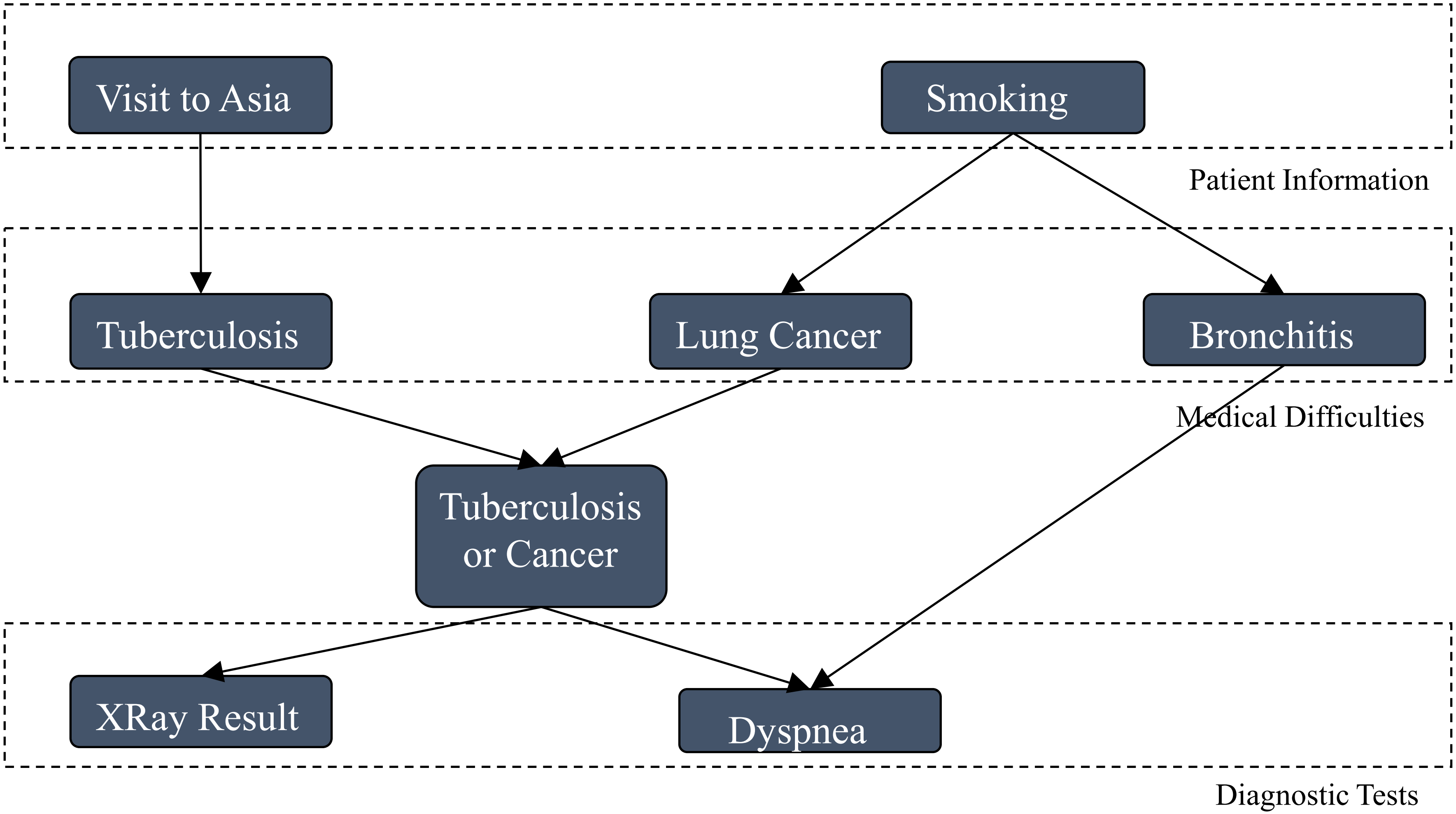


Bayesian Network

- Qualitative component :
  - nodes with some meaning

- Quantitative component :
  - Each node is associated with a probability function that takes as input a particular set of values for the node's parent variables and gives the probability of the variable represented by the node.
  - If the parents are m Boolean variables then the probability function could be represented by a table of 2^m entries

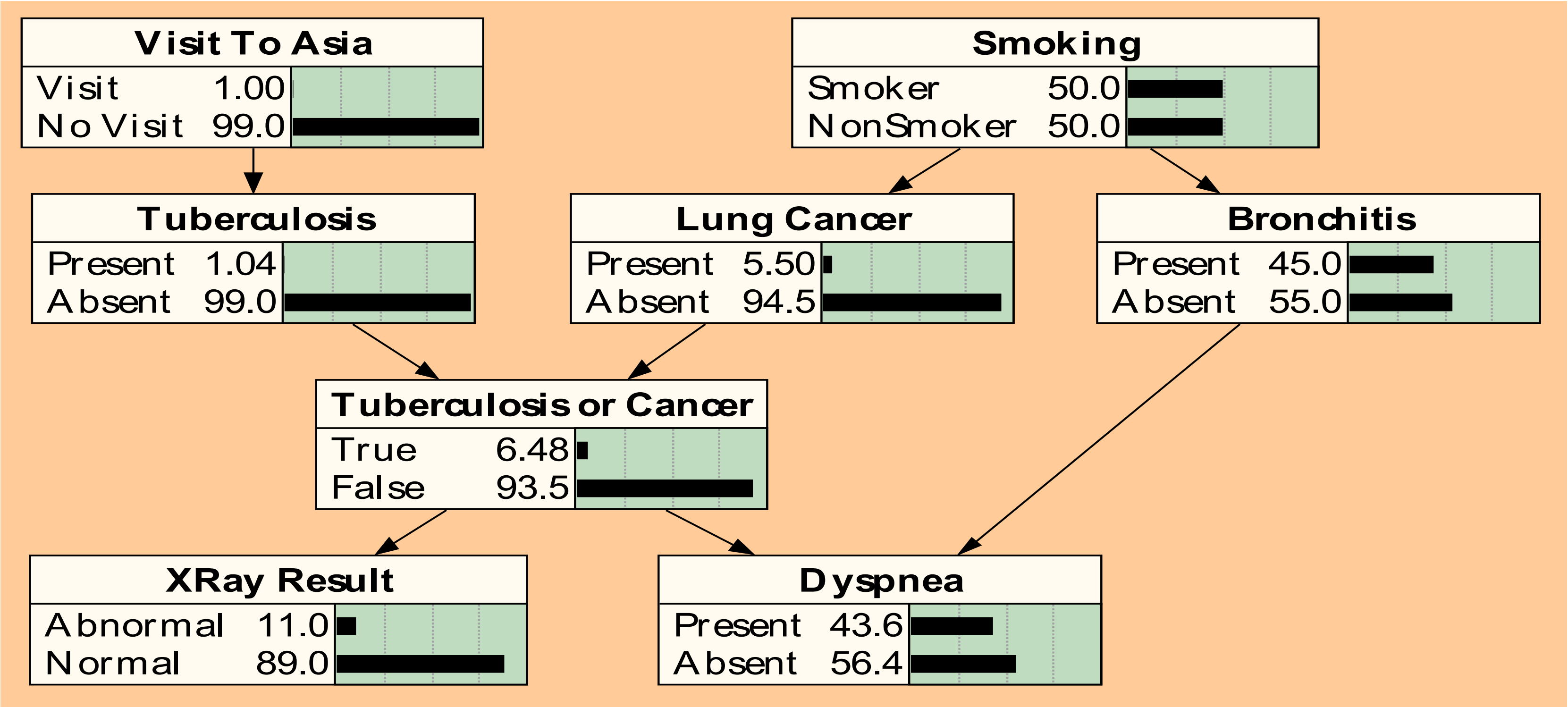
Example from Medical Diagnostics



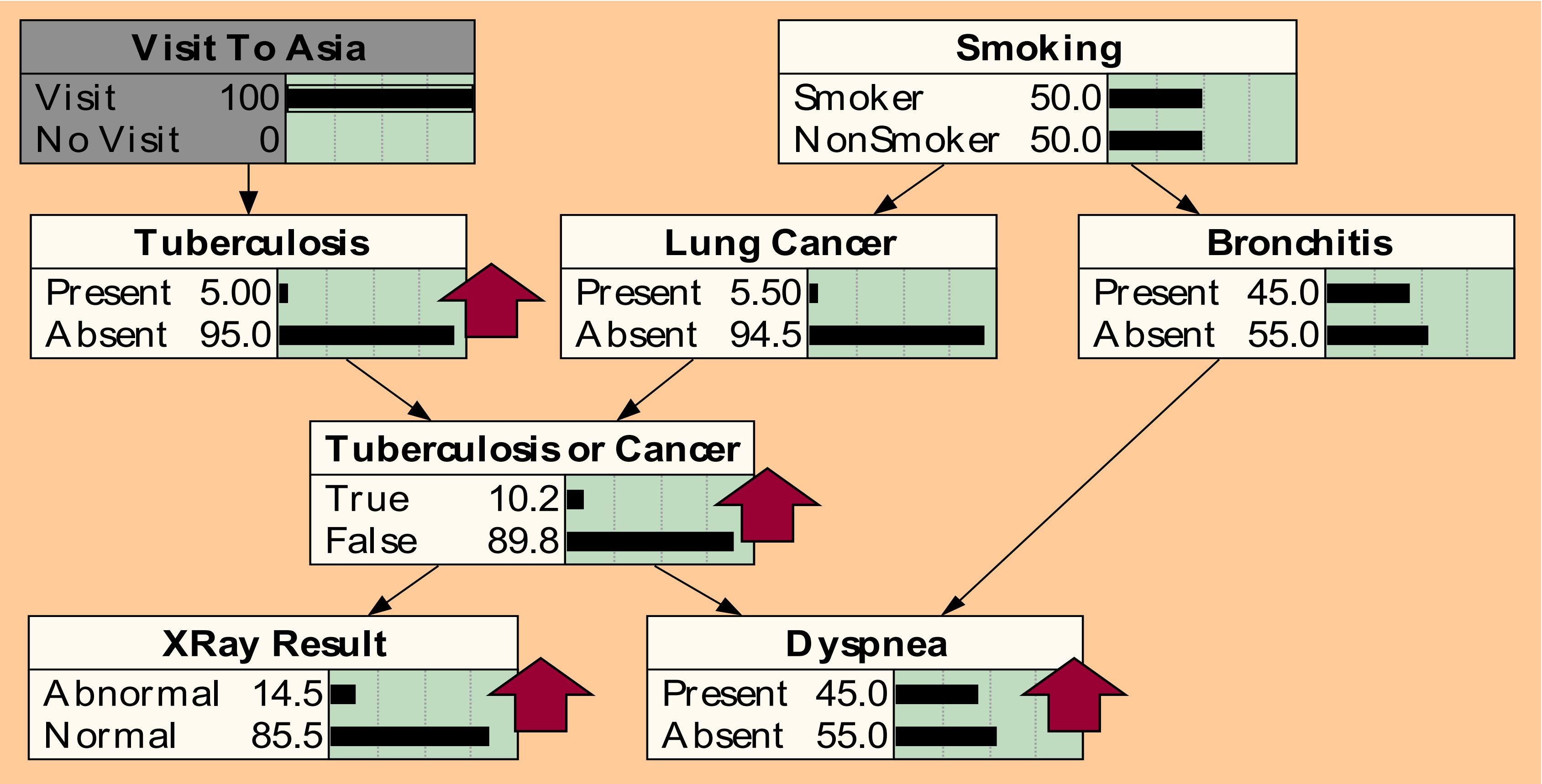
- Network represents a knowledge structure that models the relationship between medical difficulties, their causes and effects, patient information and diagnostic tests.

Quantitative information in the nodes

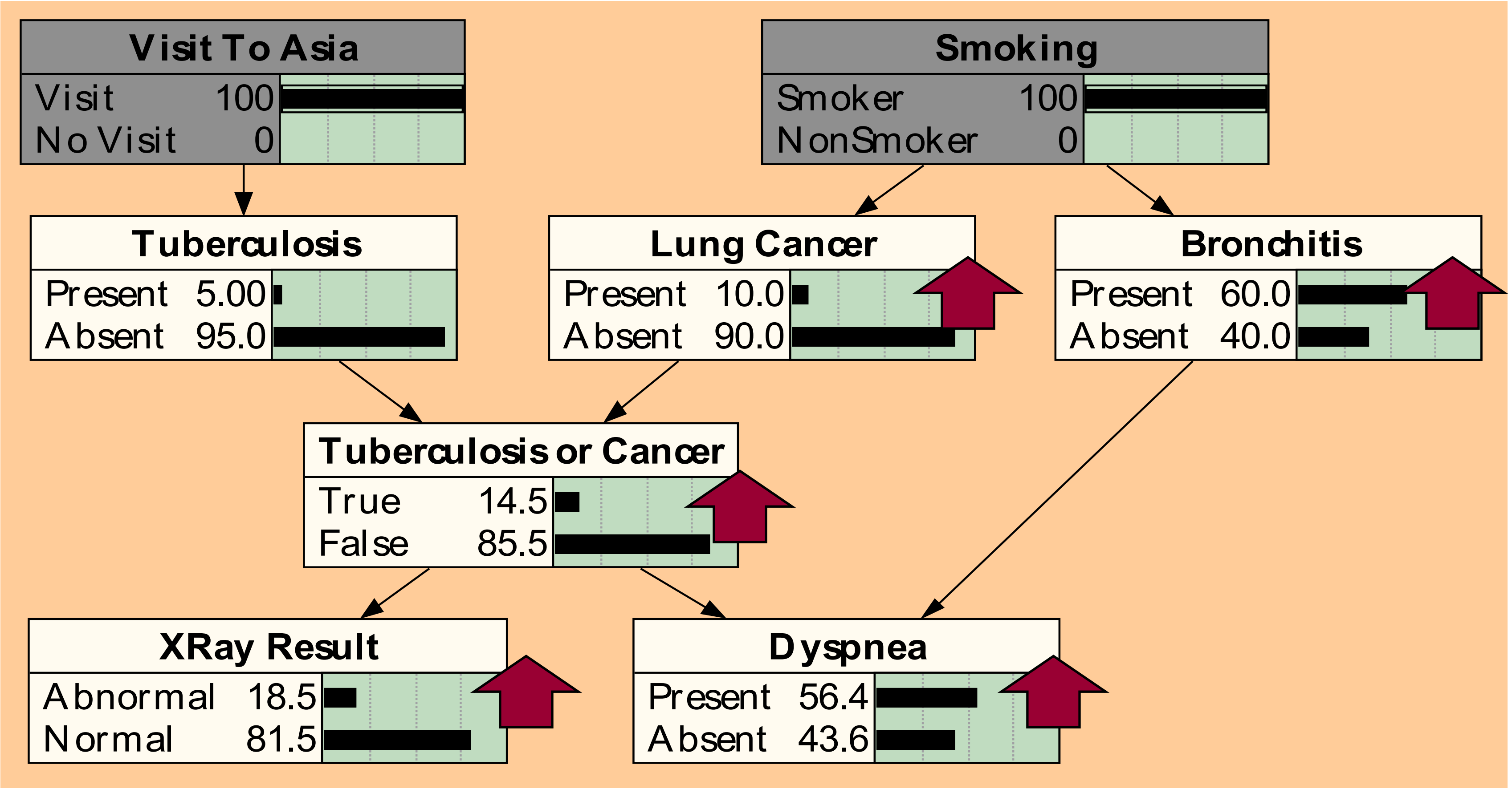
Visit to Asia	P(visit)=0,01
Tuberculosis	P(present | visit)=0,05, P(present | no visit)=0,01
Tuberculosis or Cancer	P(true | tuberc, cancer)=1, P(true | tuberc, no cancer)=1 P(true | no tuberc, cancer)=1, P(true | no tuberc, no cancer)=0
Smoking	P(present)=0,5
Lung Cancer	P(present | smoke)=0,1, P(present | no smoke)=0,01
Bronchitis	P(present | smoke)=0,6, P(present | no smoke)=0,3
X ray result	P(abnormal | tuberc)=0,98, P(abnormal | no tuberc)=0,05
Dyspnea	P(present | tubcan, bronc)=0,9, P(present | tubcan, no bronc)=0,7 P(present | no tubcan, bronc)=0,8, P(present | no tubcan, no bronc)=0,1



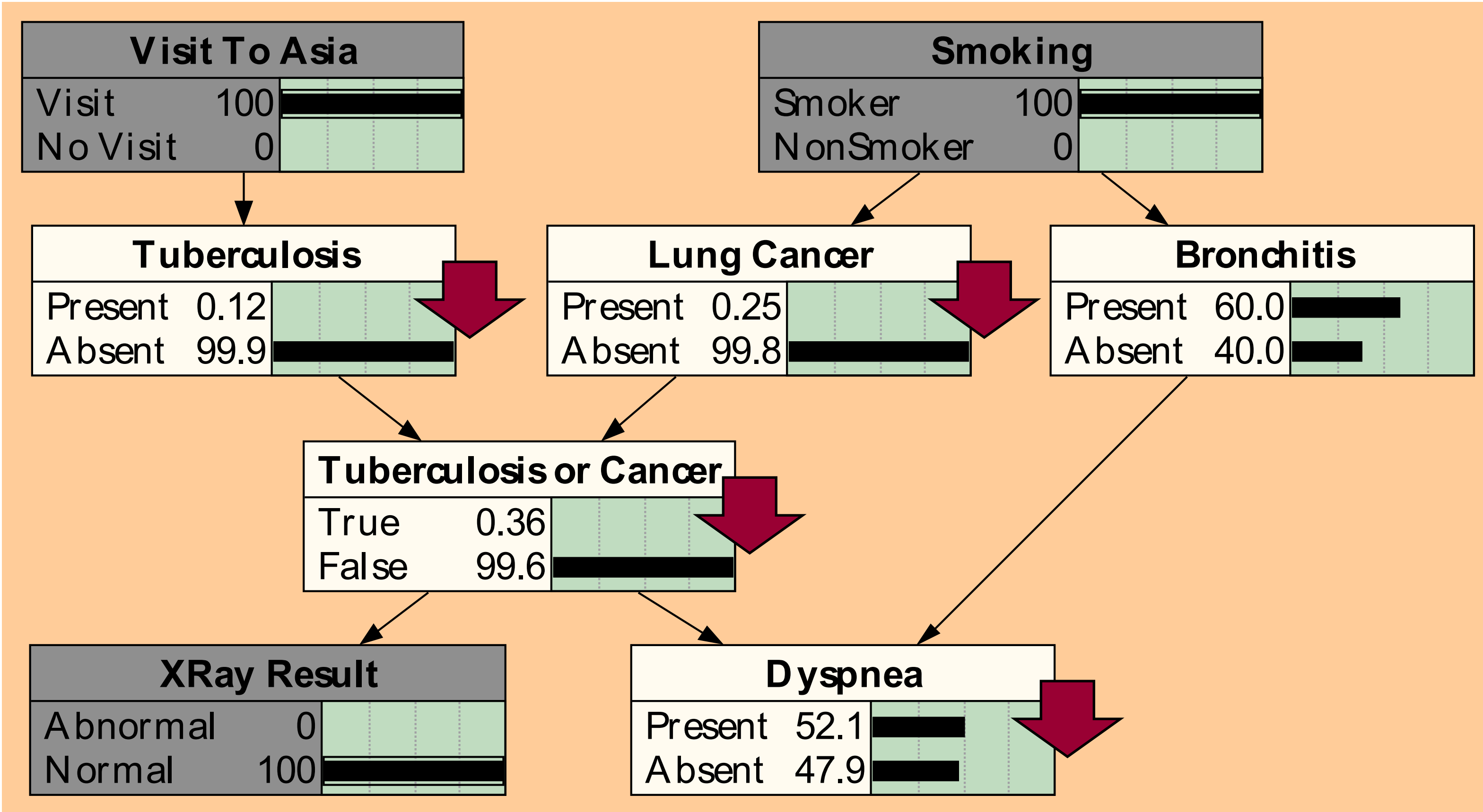
- Propagation algorithm processes relationship information to provide an unconditional or marginal probability distribution for each node
- The unconditional or marginal probability distribution is frequently called the belief function of that node



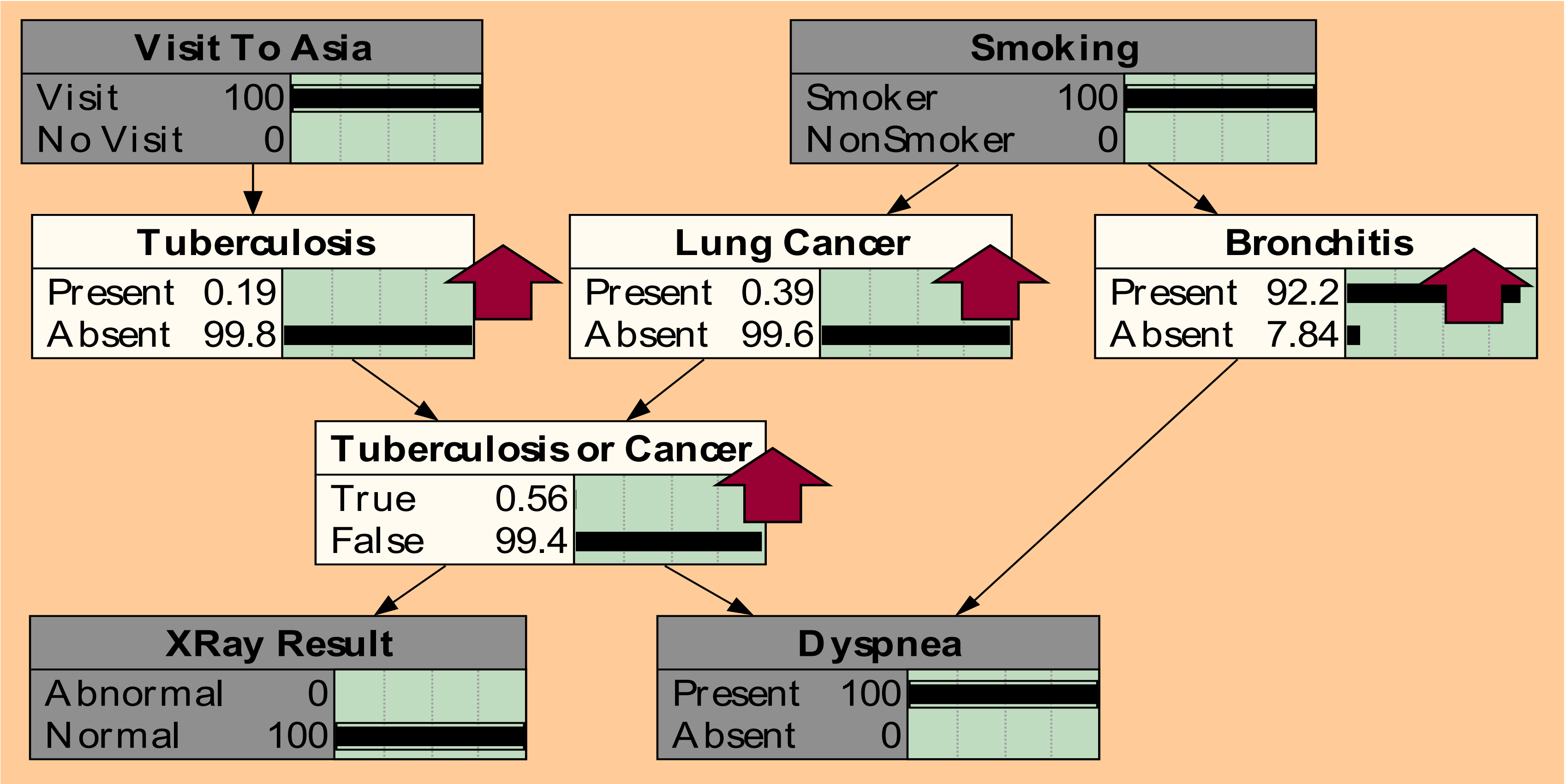
- As a finding is entered (grey node), the propagation algorithm updates the beliefs attached to each relevant node in the network
- Interviewing the patient produces the information that "Visit to Asia" is "Visit"
- This finding propagates through the network and the belief functions of several nodes are updated



- Further interviewing of the patient produces the finding “Smoking” is “Smoker”
- This information propagates through the network

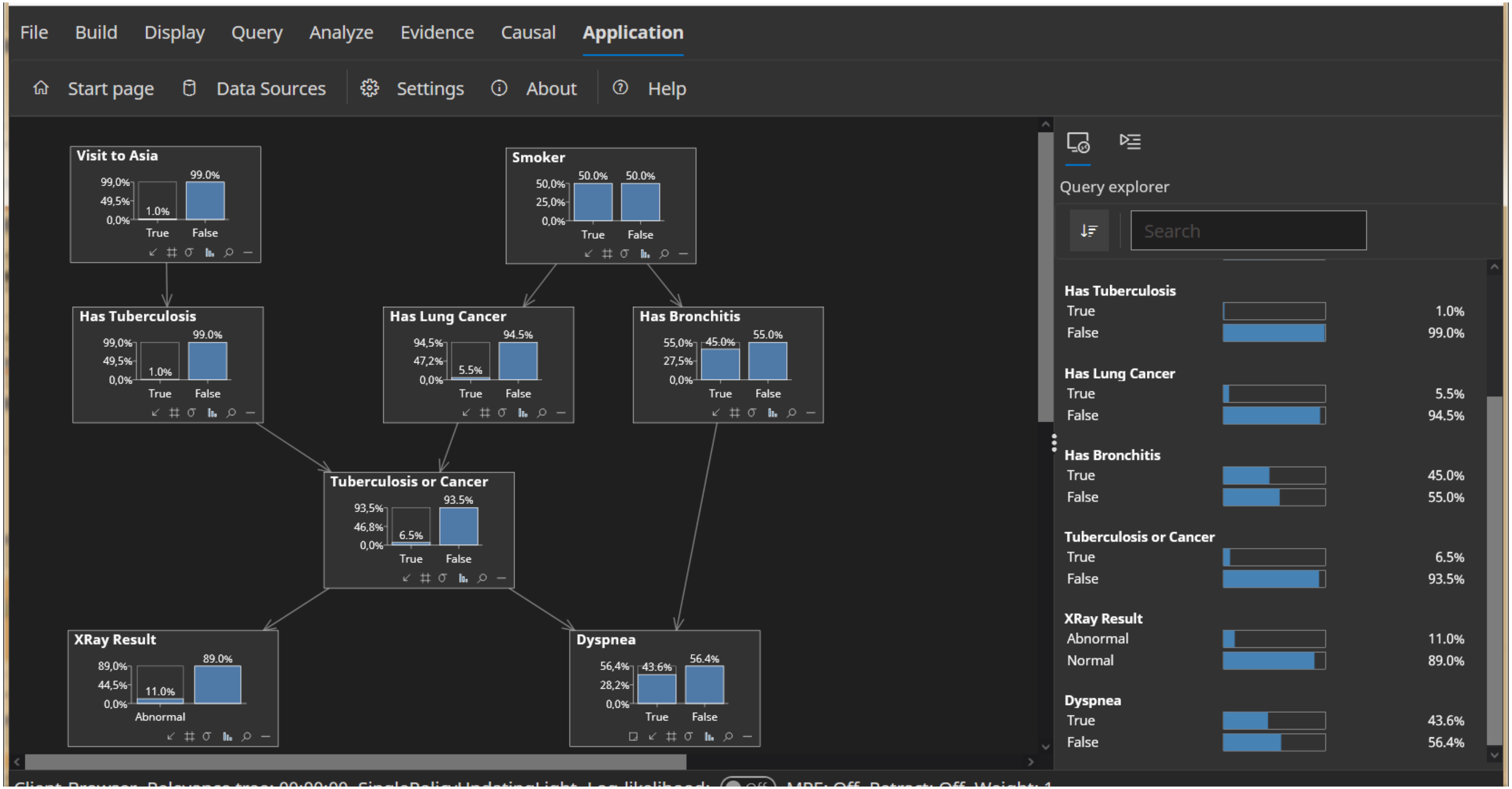


- Finished with interviewing the patient, the physician begins examination
- The physician now moves to specific diagnostic tests such as an X-Ray, which results in a “Normal” finding which propagates through the network
- Note that the information from this finding propagates backward and forward through the arcs



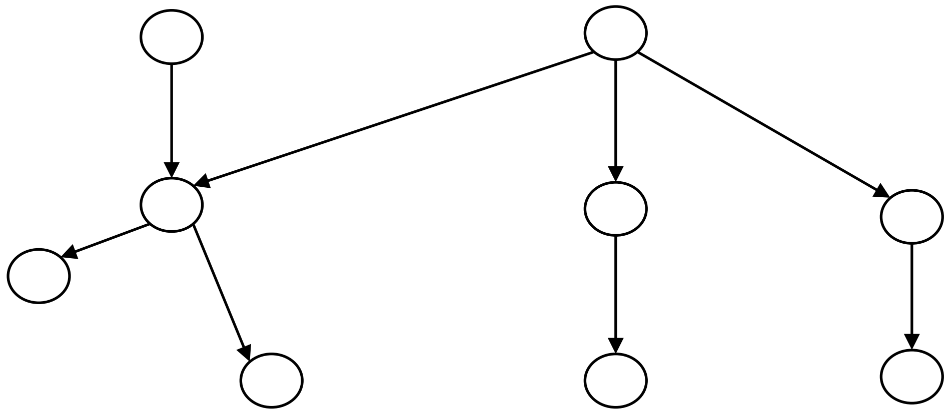
- The physician also determines that the patient is having difficulty breathing, the finding "Present" is entered for "Dyspnea" and is propagated through the network
- The doctor might now conclude that the patient has bronchitis and does not have tuberculosis or lung cancer

- Online tool: https://online.bayesserver.com/
- You can test this example if you open Asia network

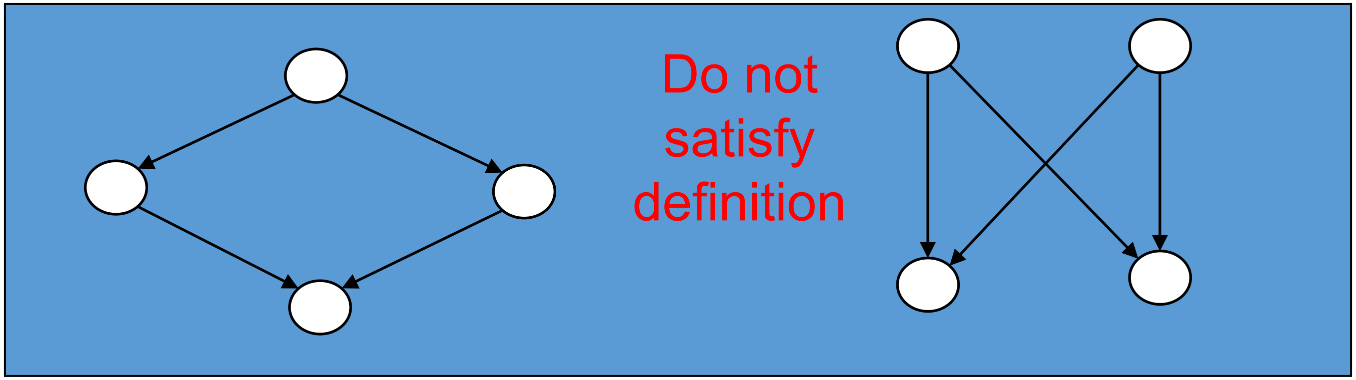


Inference in Bayesian Networks

- Polytree: is an acyclic graph where there is only a single path connecting each pair of nodes.

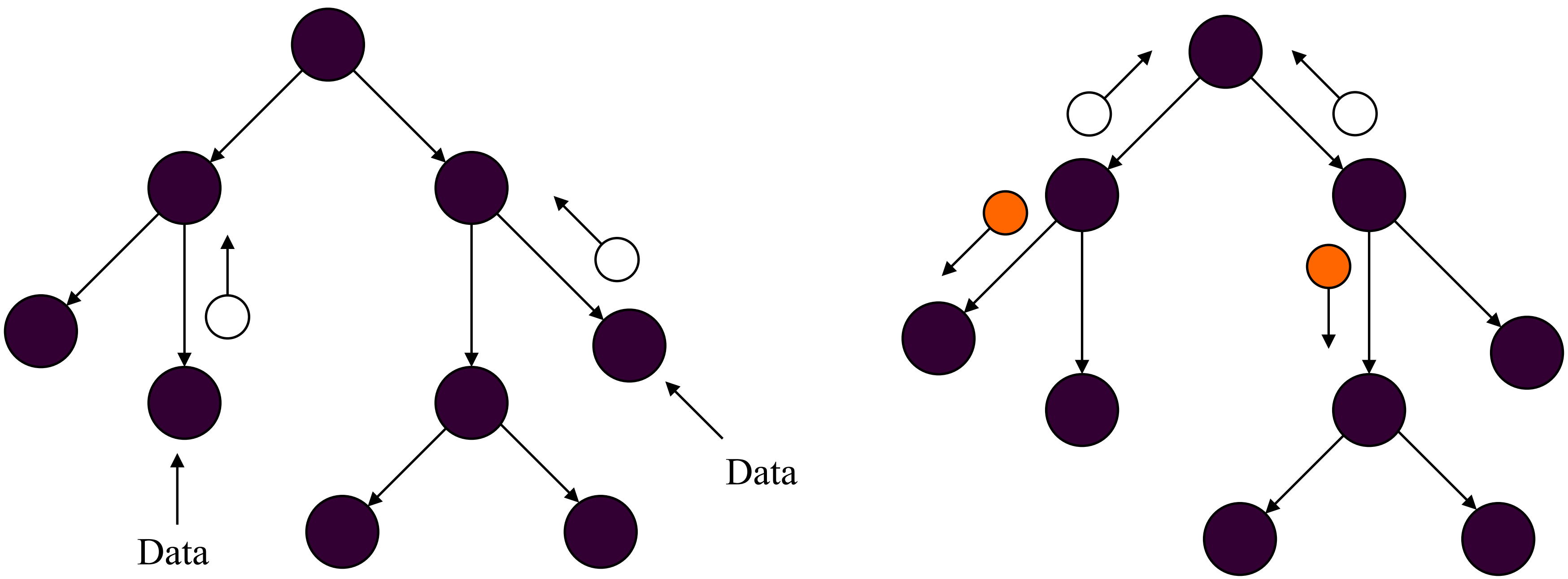


Multiple parents and/or multiple children



Inference in Bayesian Networks

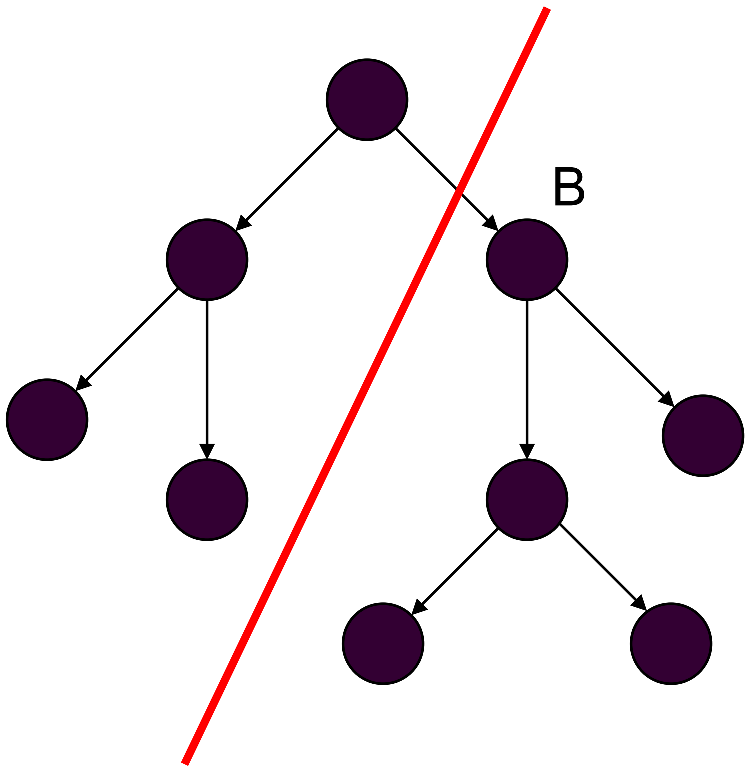
- Propagation algorithm: calculate the probabilities of the variables a posteriori, given certain evidences of some other variables.



Pearl propagation algorithm

- We distinguish two sets with respect to the node we are recalculating the belief, node B.

E+: nodes not descendents of B



E-: B and its descendents

Pearl propagation algorithm

$$p(B|E) = \frac{p(E-, E+|B) \cdot p(B)}{p(E)}$$

$$p(B|E) = \alpha \cdot \underbrace{\boxed{p(B|E+)}}_{\pi(B)} \cdot \underbrace{\boxed{p(E-|B)}}_{\lambda(B)}$$

Normalizing factor (value in 0..1)

Propagated to the immediate nodes in E- (sons)

Propagated to the upper node in E+ (father)

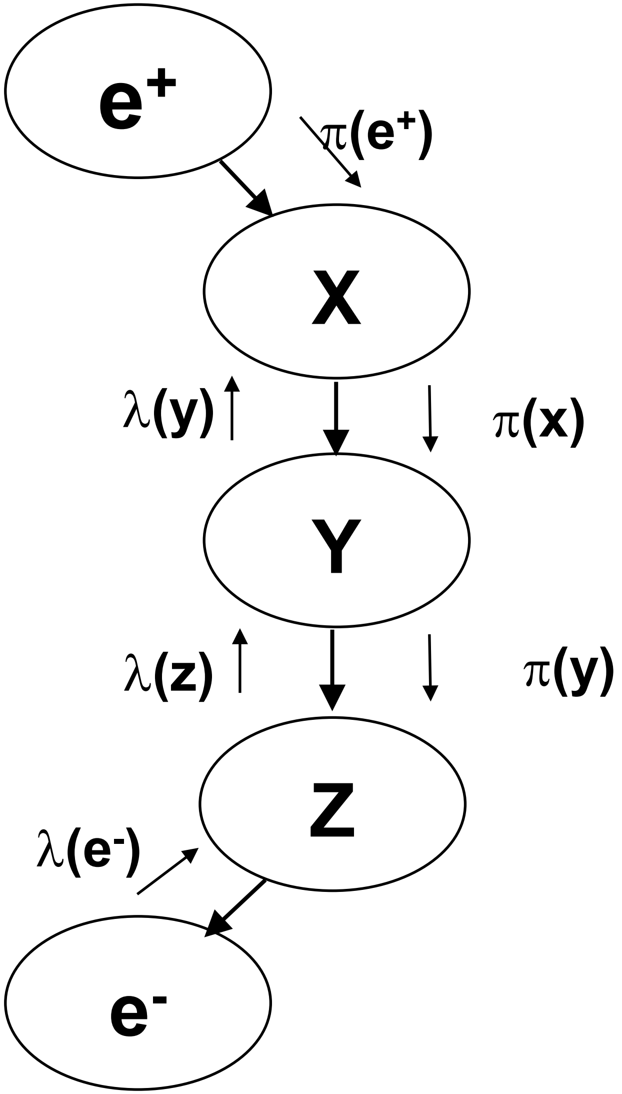
Pearl propagation algorithm

We will study the case of propagation in a single branch.

Each node stores the values of π and λ

The probability of B is calculated as:

$$p(B|E) = \alpha \cdot \pi(B) \cdot \lambda(B)$$



Pearl propagation algorithm

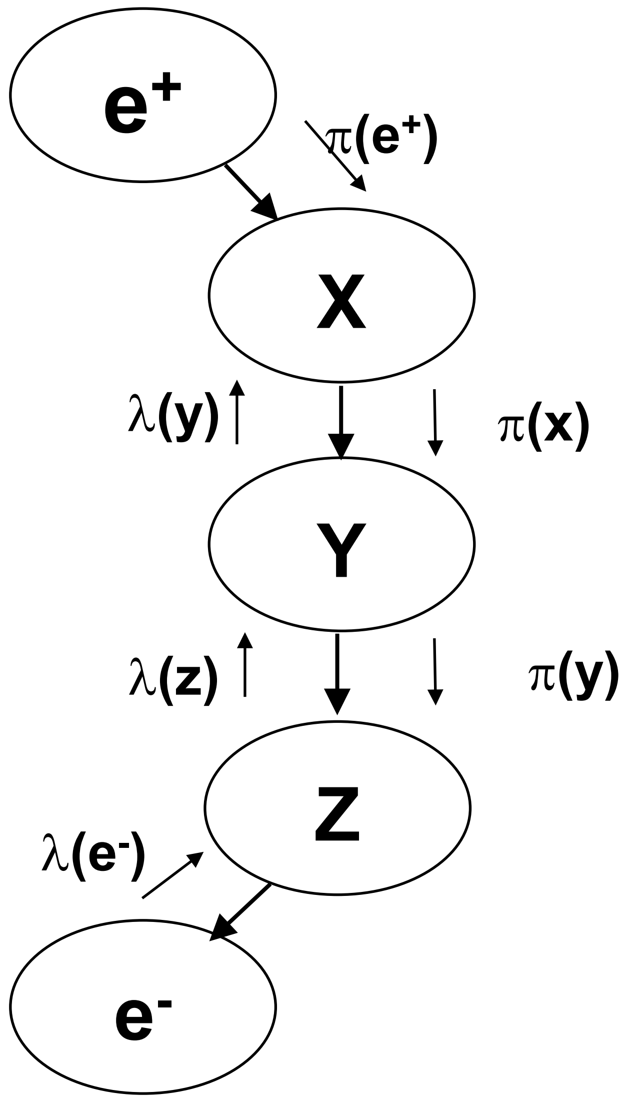
For each node we store the π and λ values, as well as the belief *B* for each of the posible values.

Initialization:

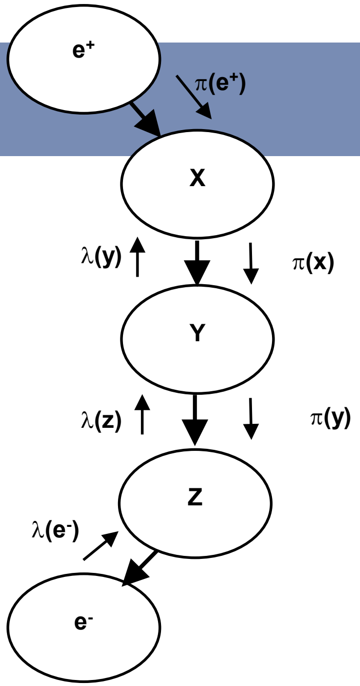
π takes the a priori probability values in the node at the top

λ is 1on all the nodes of the network

B is equal to π



Pearl propagation algorithm



For each node we store the π and λ values, as well as the belief *B* for each of the posible values.

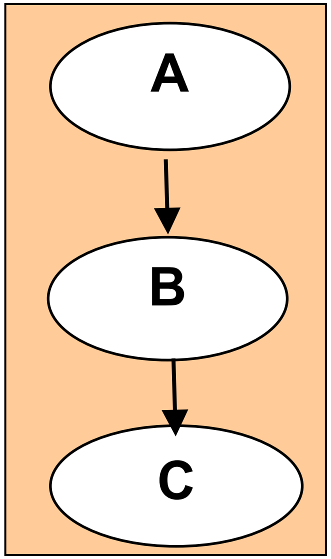
Propagation in the network:

π is propagated downwards << New evidence from the top e+

λ is propagated upwards << New evidence from the bottom e-

B is recalculated each time a new value of π or λ arrives

Pearl propagation algorithm



$$M_{B|A} = \begin{array}{c} \begin{array}{cc} a1 & a2 \end{array} \\ \begin{bmatrix} .8 & .1 \\ .2 & .9 \end{bmatrix} \end{array} \begin{array}{c} b1 \\ b2 \end{array}$$

$$M_{C|B} = \begin{array}{c} \begin{array}{cc} b1 & b2 \end{array} \\ \begin{bmatrix} .5 & .1 \\ .4 & .3 \\ .1 & .6 \end{bmatrix} \end{array} \begin{array}{c} c1 \\ c2 \\ c3 \end{array}$$

	π(A)	Bel(A)	λ(A)
a1	0.8	0.8	1.0
a2	0.2	0.2	1.0

Propagation down: $\pi(B) = \pi(A)\ M_{B|A}$

	π(B)	Bel(B)	λ(B)
b1	0.66	0.66	0.71
b2	0.34	0.34	0.71

Propagation up: $\lambda(B) = M_{C|B}\ \lambda(C)$

	π(C)	Bel(C)	λ(C)
c1	0.36	0.25	0.5
c2	0.37	0.52	1.0
c3	0.27	0.23	0.6

Example

- In an Excel file in Moodle you will find an example of propagation of evidence in a Bayesian network.
- The example is a simplification of the Medical Diagnosis network presented before.
- Equations can be found in each sheet.

	Asia	
Tuberculosis	visit	no visit
yes	0,1	0,01
no	0,9	0,99

	Tuberculosis	
X-ray	yes	no
abnormal	0,98	0,05
normal	0,02	0,95

	X-ray	
Dyspnea	abnormal	normal
yes	0,6	0,1
no	0,4	0,9

Asia	pi	Baux	B	lambda
yes	0,01	0,01	0,01	1
no	0,99	0,99	0,99	1

Tuberculosis	pi	Baux	B	lambda
yes				
no				

X-ray	pi	Baux	B	lambda
abnormal				
normal				

Dyspnea	pi	Baux	B	lambda
yes				
no				

Learning in Bayesian Networks

- Learn the parameters
  - Estimate the marginal and conditional probabilities from examples
- Learn the structure of the network
  - Find dependencies between variables
- Learn dynamic networks
  - Some of the relations between states are temporal. These temporal relations can be learned automatically.

References

- Inteligencia Artificial. Técnicas, métodos y aplicaciones. José T. Palma Méndez, Roque Marín Morales, Ed. Mc-Graw Hill, 2008 (004.8 Int)
- Artificial Intelligence. Elaine Rich & Kevin Knight. Ed. Mc-Graw Hill, 1991 (004.8 Ric)
- Bayesian networks : with examples in R, Marco Scutari & Jean-Baptiste Denis, 2014 (URV online)

Acknowledgements:

This material has been prepared using the previous work Buede, Tatman and Bresnik (Tutorial 66th MORS symphosium) and materials of Dr. Godo (UOC).