# Reasoning with Bayesian Networks

## Overview

- Bayesian Belief Networks (BBNs) can reason with networks of propositions and associated probabilities
- Useful for many AI problems
  - Diagnosis
  - Expert systems
  - Planning
  - Learning

## Recall Bayes Rule

$$P(H,E) = P(H \mid E)P(E) = P(E \mid H)P(H)$$

$$\boxed{P(H \mid E) = \frac{P(E \mid H)P(H)}{P(E)}}$$

Note the symmetry: we can compute the probability of a hypothesis given its evidence and vice versa.

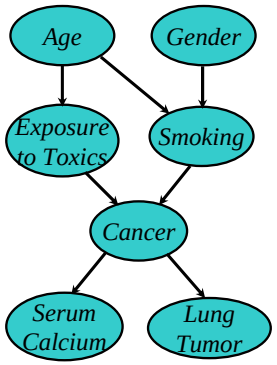
## Simple Bayesian Network


$C \in \{none, benign, malignant\}$

| | |
|---|---|
| *P(S=no)* | 0.80 |
| *P(S=light)* | 0.15 |
| *P(S=heavy)* | 0.05 |

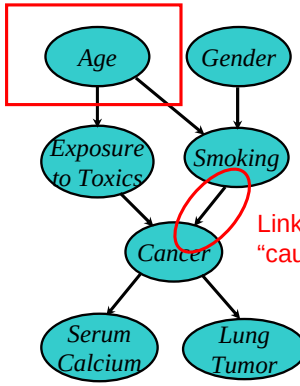
| *Smoking=* | *no* | *light* | *heavy* |
|---|---|---|---|
| *P(C=none)* | 0.96 | 0.88 | 0.60 |
| *P(C=benign)* | 0.03 | 0.08 | 0.25 |
| *P(C=malig)* | 0.01 | 0.04 | 0.15 |

## More Complex Bayesian Network




## More Complex Bayesian Network

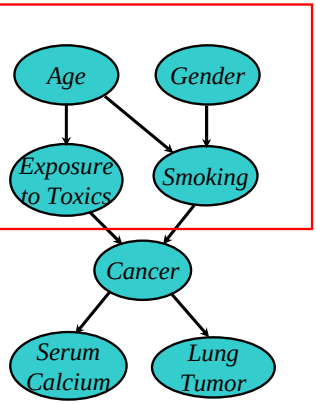
Nodes represent variables



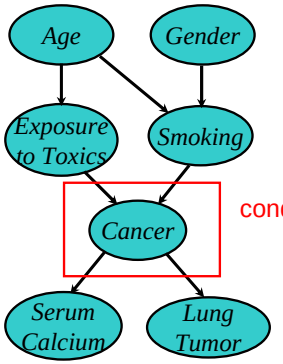

Links represent "causal" relations

## More Complex Bayesian Network

predispositions

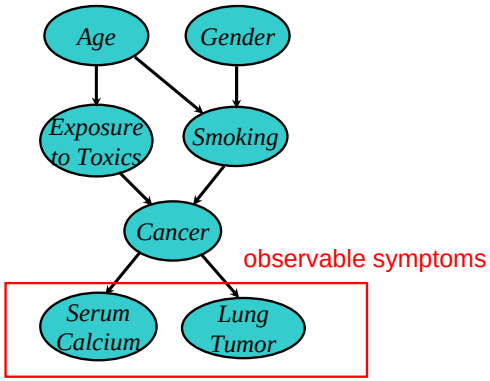



## More Complex Bayesian Network

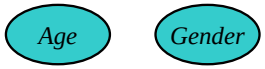



condition

## More Complex Bayesian Network




## Independence



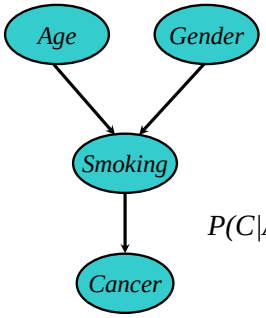

*Age* and *Gender* are independent.

$P(A,G) = P(G)P(A)$

$P(A|G) = P(A) \quad A \perp G$
$P(G|A) = P(G) \quad G \perp A$

$P(A,G) = P(G|A)\,P(A) = P(G)P(A)$
$P(A,G) = P(A|G)\,P(G) = P(A)P(G)$

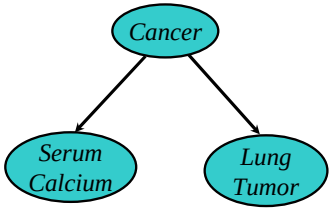
## Conditional Independence




*Cancer* is independent of *Age* and *Gender* given *Smoking*.

$P(C|A,G,S) = P(C|S) \quad C \perp A,G \mid S$

## Conditional Independence: Naïve Bayes



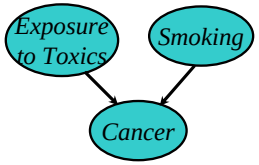

*Serum Calcium* and *Lung Tumor* are dependent

*Serum Calcium* is independent of *Lung Tumor*, given *Cancer*

$P(L|SC,C) = P(L|C)$

The naïve Bayes assumption is that all of the evidence (e.g., symptoms) is Independent given the disease. This make it easy to combine evidence

# Explaining Away




*Exposure to Toxics* and *Smoking* are independent

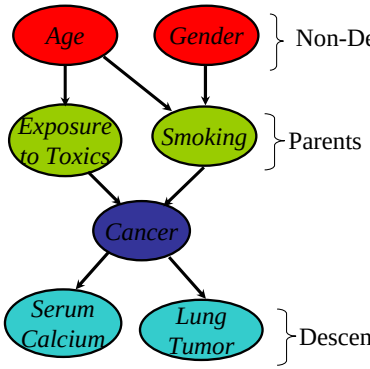
*Exposure to Toxics* is **dependent** on *Smoking*, given *Cancer*

$P(E = heavy \mid C = malignant) >$
$P(E = heavy \mid C = malignant, S=heavy)$

"Explaining away" is like abductive inference in that it moves from observation To possible causes or explanations.

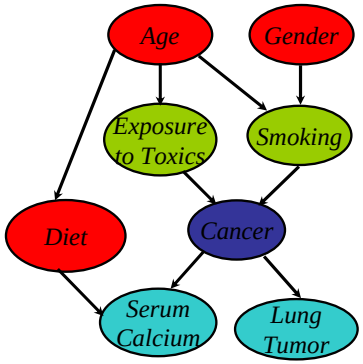
# Conditional Independence

A variable (node) is conditionally independent of its non-descendants given its parents.




*Cancer* is independent of *Age* and *Gender* given *Exposure to Toxics* and *Smoking*.

# Another non-descendant




*Cancer* is independent of *Diet* given *Exposure to Toxics* and *Smoking*.

# BBN Construction

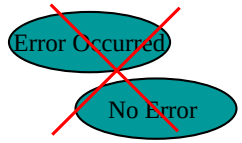
- The knowledge acquisition process for a BBN involves three steps
  - Choosing appropriate variables
  - Deciding on the network structure
  - Obtaining data for the conditional probability tables

## (1) Choosing variables

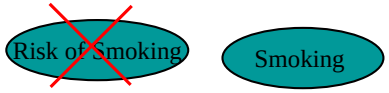
- Variables should be collectively exhaustive, mutually exclusive values

$$x_1 \vee x_2 \vee x_3 \vee x_4$$

$$\neg(x_i \wedge x_j) \quad i \neq j$$




- They should be values versus and not probabilities

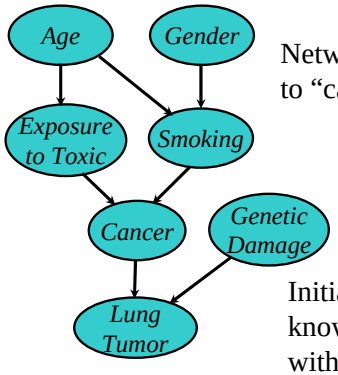



## Heuristic: Knowable in Principle

Example of good variables

- Weather {Sunny, Cloudy, Rain, Snow}
- Gasoline: Cents per gallon
- Temperature { ≥ 100F , < 100F}
- User needs help on Excel Charting {Yes, No}
- User's personality {dominant, submissive}

## (2) Structuring

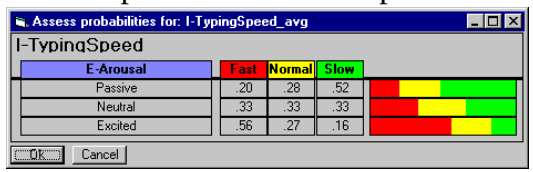



Network structure corresponding to "causality" is usually good.

Initially this uses the designer's knowledge but can be checked with data

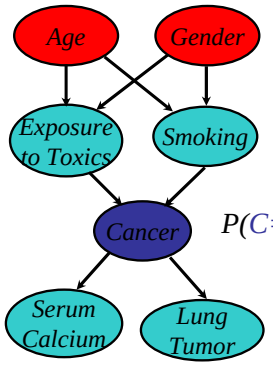
## (3) The numbers

- Second decimal usually doesn't matter
- Relative probabilities are important




- Zeros and ones are often enough
- Order of magnitude is typical: $10^{-9}$ vs $10^{-6}$
- Sensitivity analysis can be used to decide accuracy needed

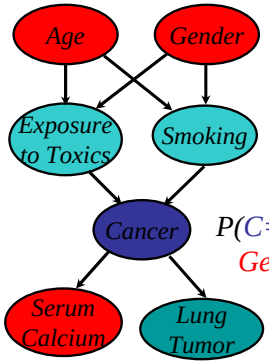
## Predictive Inference




How likely are elderly males to get malignant cancer?

*P(C=malignant | Age>60, Gender= male)*

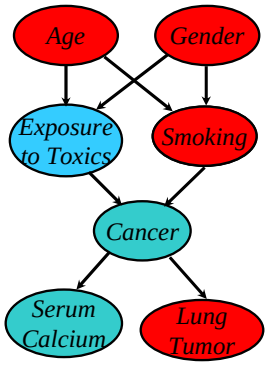
## Predictive and diagnostic combined




How likely is an elderly male patient with high Serum Calcium to have malignant cancer?

*P(C=malignant | Age>60, Gender= male, Serum Calcium = high)*

## Explaining away



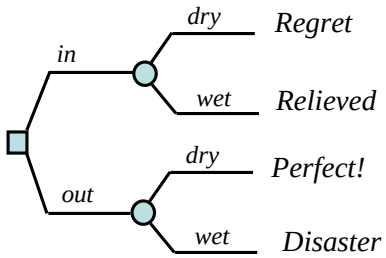

- If we see a lung tumor, the probability of heavy smoking and of exposure to toxics both go up.
- If we then observe heavy smoking, the probability of exposure to toxics goes back down.

## Decision making

- Decision - an irrevocable allocation of domain resources
- Decision should be made so as to maximize expected utility.
- View decision making in terms of
  - Beliefs/Uncertainties
  - Alternatives/Decisions
  - Objectives/Utilities

# A Decision Problem

Should I have my party inside or outside?



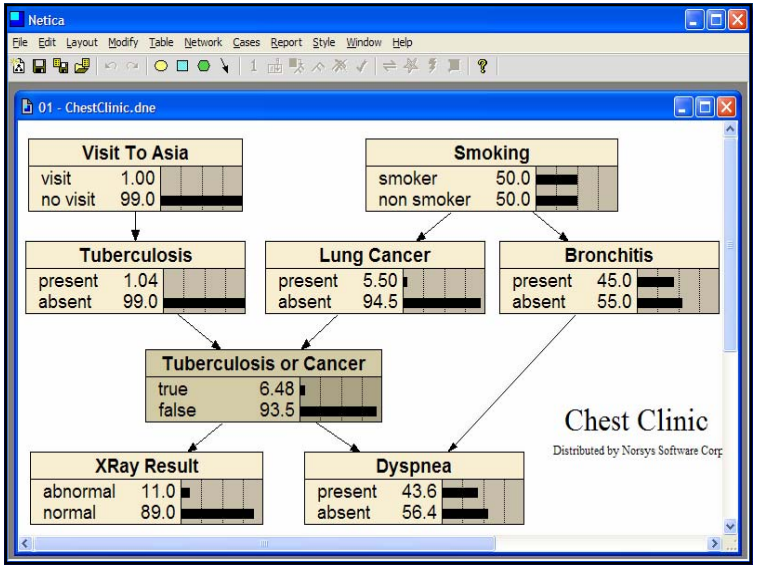

# Value Function

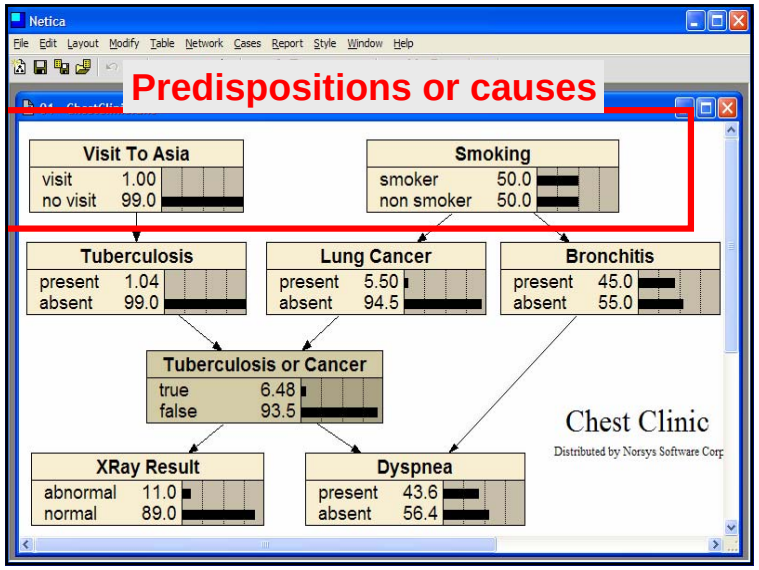
A numerical score over all possible states of the world allows BBN to be used to make decisions

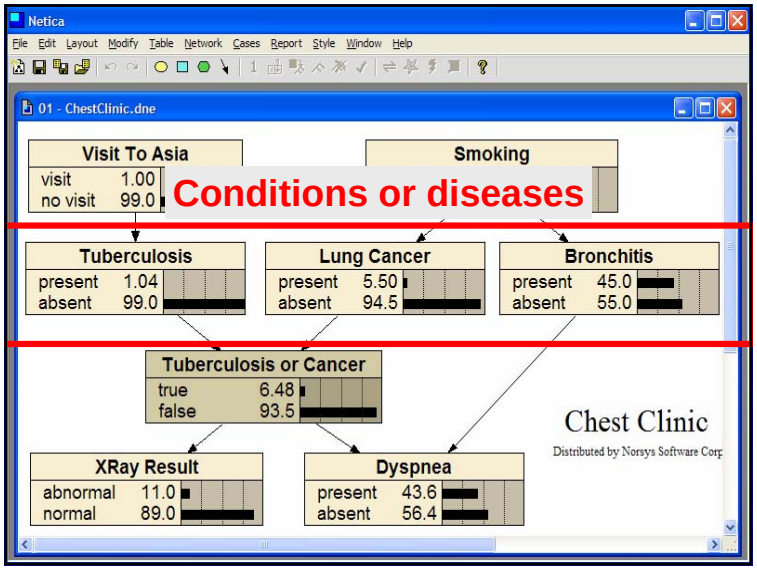
| Location? | Weather? | Value |
|---|---|---|
| in | dry | $50 |
| in | wet | $60 |
| out | dry | $100 |
| out | wet | $0 |

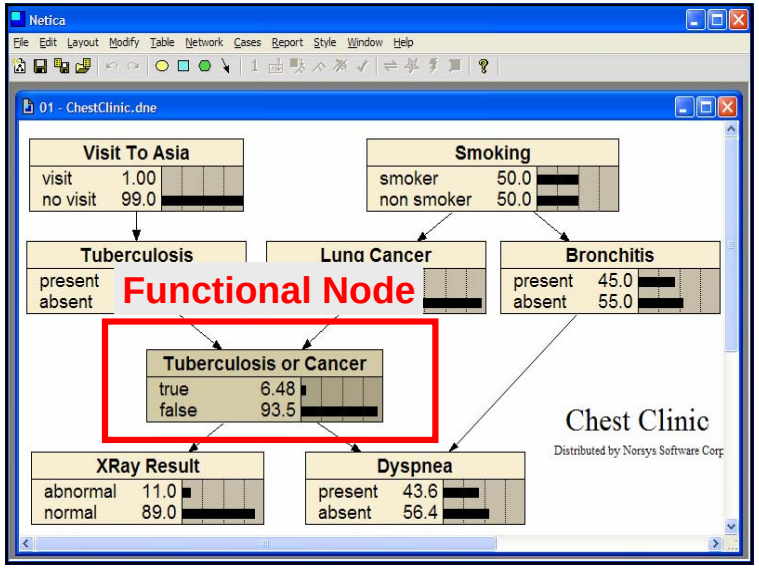
# Netica

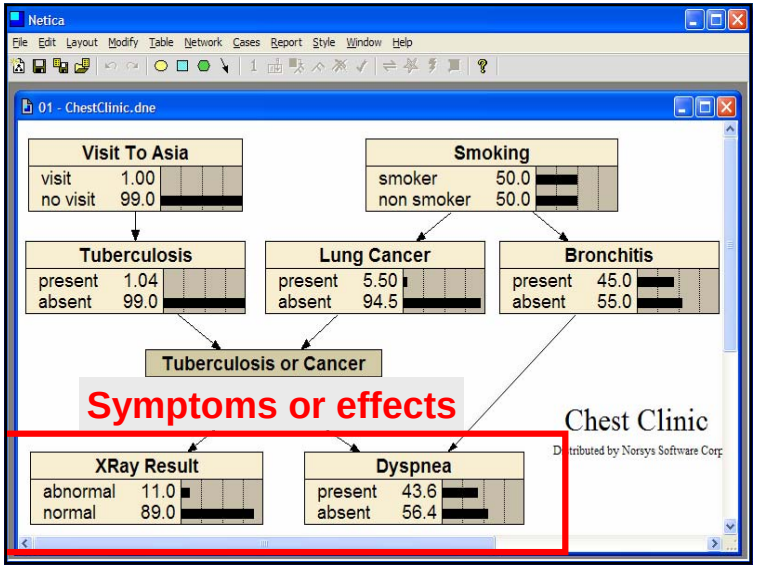
- Software for working with Bayesian belief networks and influence diagrams
- A commercial product but free for small networks
- Includes a graphical editor, compiler, inference engine, etc.
- http://www.norsys.com/

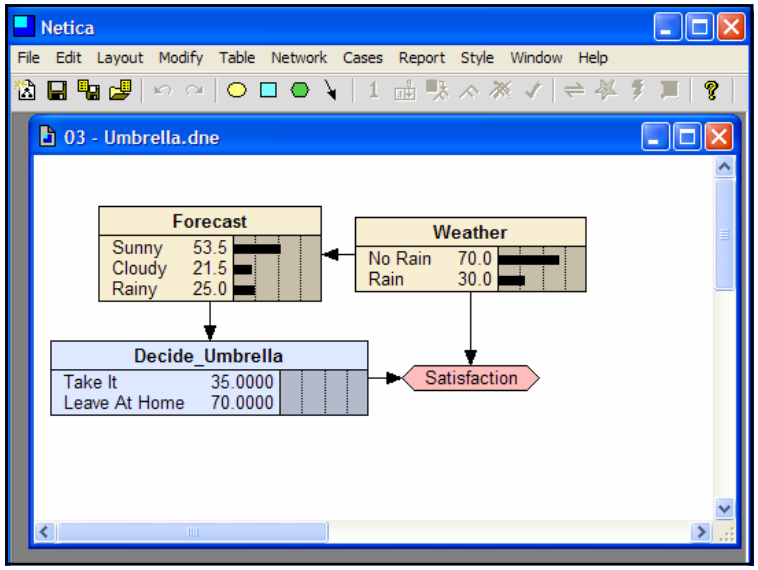


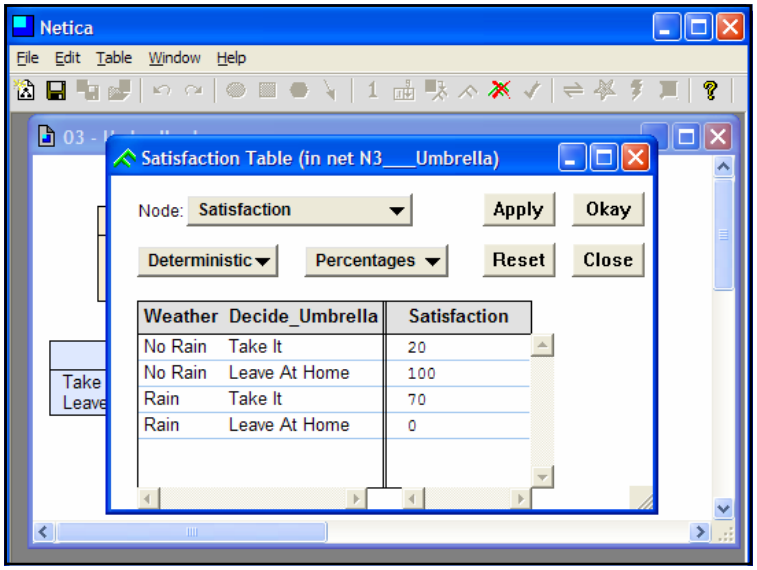

Netica
Predispositions or causes
Visit To Asia
visit 1.00
no visit 99.0
Smoking
smoker 50.0
non smoker 50.0
Tuberculosis
present 1.04
absent 99.0
Lung Cancer
present 5.50
absent 94.5
Bronchitis
present 45.0
absent 55.0
Tuberculosis or Cancer
true 6.48
false 93.5
XRay Result
abnormal 11.0
normal 89.0
Dyspnea
present 43.6
absent 56.4
Chest Clinic
Distributed by Norsys Software Corp

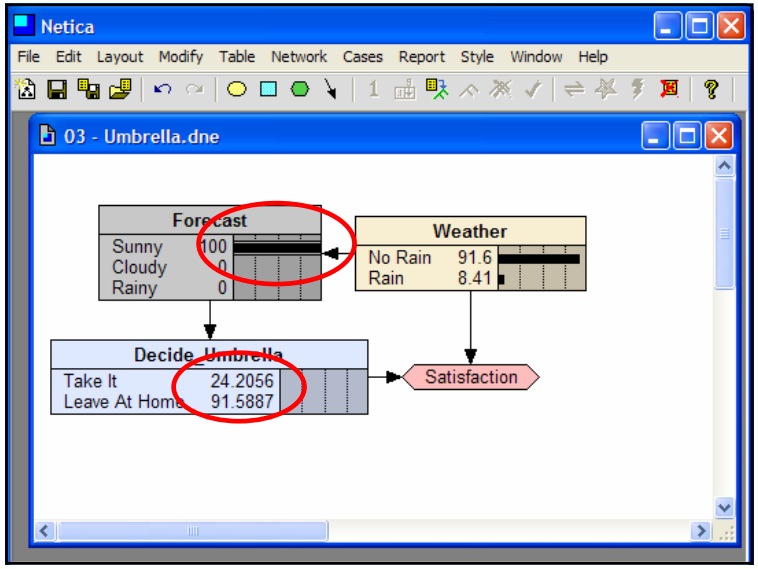

Netica
01 - ChestClinic.dne
Visit To Asia
visit 1.00
no visit 99.0
Smoking
Conditions or diseases
Tuberculosis
present 1.04
absent 99.0
Lung Cancer
present 5.50
absent 94.5
Bronchitis
present 45.0
absent 55.0
Tuberculosis or Cancer
true 6.48
false 93.5
XRay Result
abnormal 11.0
normal 89.0
Dyspnea
present 43.6
absent 56.4
Chest Clinic
Distributed by Norsys Software Corp

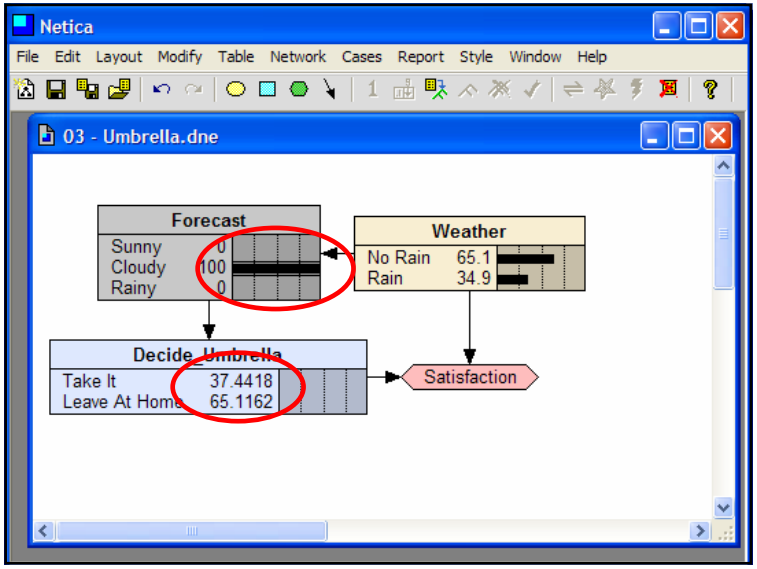

Netica
01 - ChestClinic.dne
Visit To Asia
visit 1.00
no visit 99.0
Smoking
smoker 50.0
non smoker 50.0
Tuberculosis
present
absent
Lung Cancer
Functional Node
Bronchitis
present 45.0
absent 55.0
Tuberculosis or Cancer
true 6.48
false 93.5
XRay Result
abnormal 11.0
normal 89.0
Dyspnea
present 43.6
absent 56.4
Chest Clinic
Distributed by Norsys Software Corp

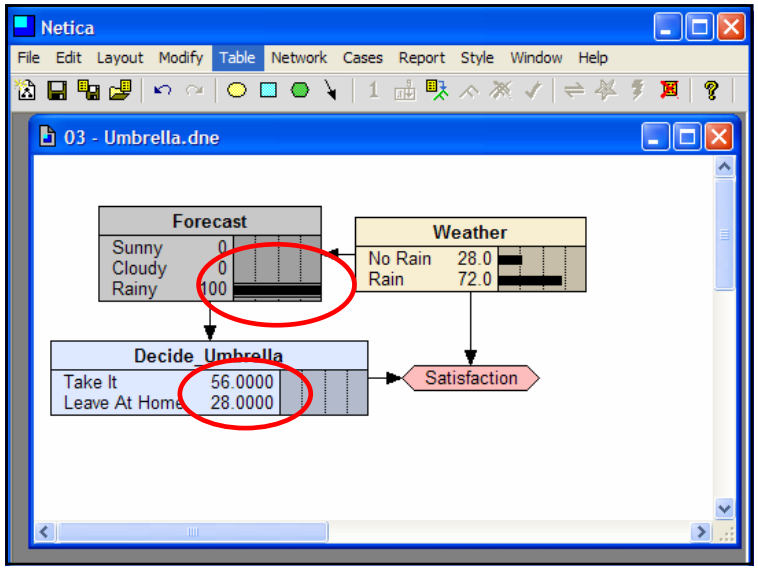

Netica
01 - ChestClinic.dne
Visit To Asia
visit 1.00
no visit 99.0
Smoking
smoker 50.0
non smoker 50.0
Tuberculosis
present 1.04
absent 99.0
Lung Cancer
present 5.50
absent 94.5
Bronchitis
present 45.0
absent 55.0
Tuberculosis or Cancer
Symptoms or effects
XRay Result
abnormal 11.0
normal 89.0
Dyspnea
present 43.6
absent 56.4
Chest Clinic


Netica
File Edit Layout Modify Table Network Cases Report Style Window Help
03 - Umbrella.dne
Forecast
Sunny 53.5
Cloudy 21.5
Rainy 25.0
Weather
No Rain 70.0
Rain 30.0
Decide_Umbrella
Take It 35.0000
Leave At Home 70.0000
Satisfaction


Netica
File Edit Table Window Help
Satisfaction Table (in net N3___Umbrella)
Node: Satisfaction
Apply
Okay
Deterministic
Percentages
Reset
Close
Weather Decide_Umbrella Satisfaction
No Rain Take It 20
No Rain Leave At Home 100
Rain Take It 70
Rain Leave At Home 0


Netica
File Edit Layout Modify Table Network Cases Report Style Window Help
03 - Umbrella.dne
Forecast
Sunny 100
Cloudy 0
Rainy 0
Weather
No Rain 91.6
Rain 8.41
Decide_Umbrella
Take It 24.2056
Leave At Home 91.5887
Satisfaction


Netica
File Edit Layout Modify Table Network Cases Report Style Window Help
03 - Umbrella.dne
Forecast
Sunny 0
Cloudy 100
Rainy 0
Weather
No Rain 65.1
Rain 34.9
Decide_Umbrella
Take It 37.4418
Leave At Home 65.1162
Satisfaction


Netica
File Edit Layout Modify Table Network Cases Report Style Window Help
03 - Umbrella.dne
Forecast
Sunny 0
Cloudy 0
Rainy 100
Weather
No Rain 28.0
Rain 72.0
Decide_Umbrella
Take It 56.0000
Leave At Home 28.0000
Satisfaction