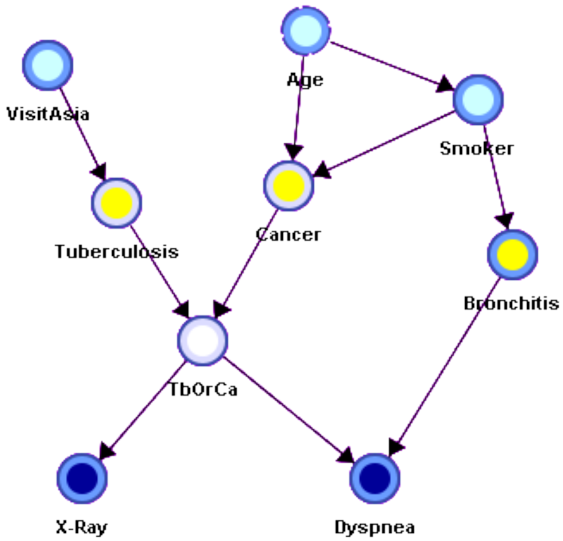
Reasoning with Bayesian Belief Networks

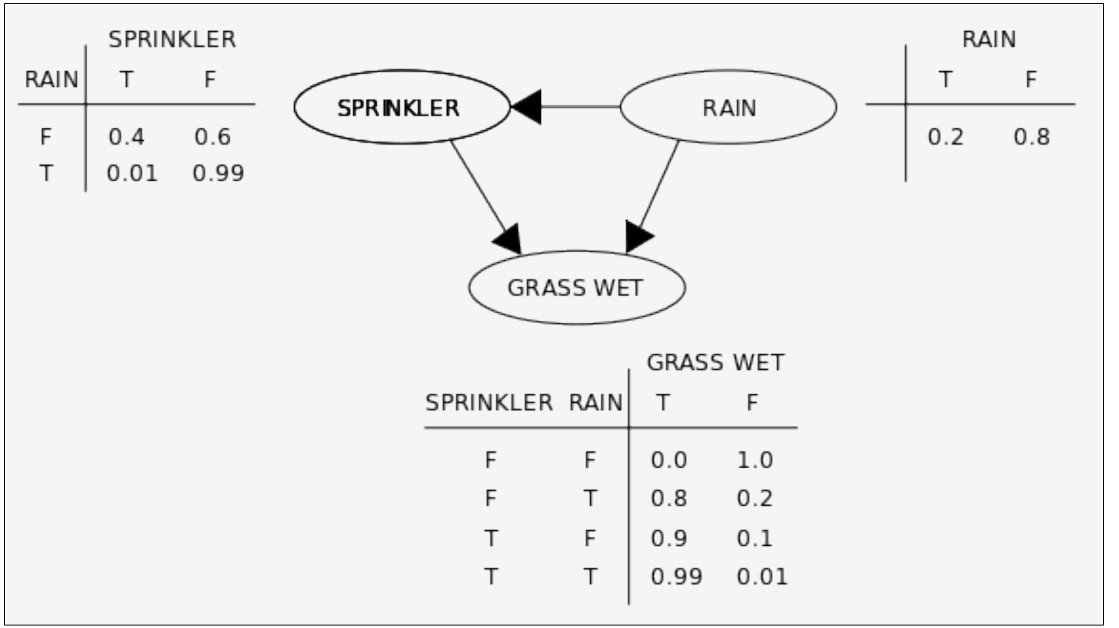


Overview

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

BBN Definition

- AKA Bayesian Network, Bayes Net
- A graphical model (as a DAG) of probabilistic relationships among a set of random variables
- Links represent direct influence of one variable on another



source

Recall Bayes Rule

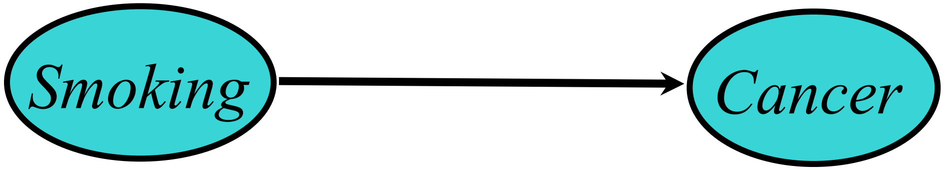
$$P(H,E) = P(H \mid E)P(E) = P(E \mid H)P(H)$$

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

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

Simple Bayesian Network

$S \in \{no, light, heavy\}$

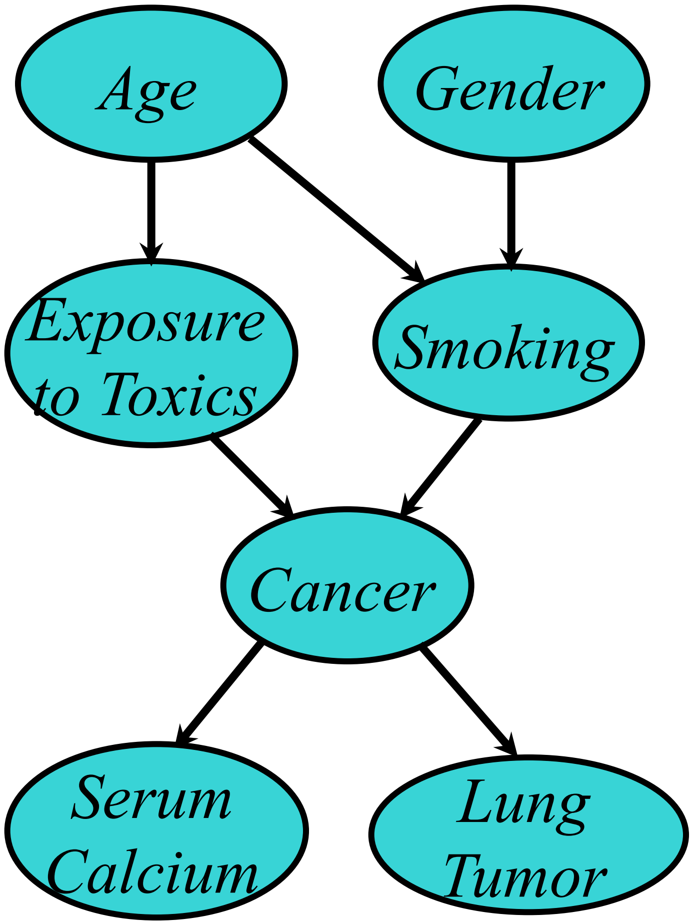


$C \in \{none, benign, malignant\}$

P(S=no)	0.80
P(S=light)	0.15
P(S=heavy)	0.05

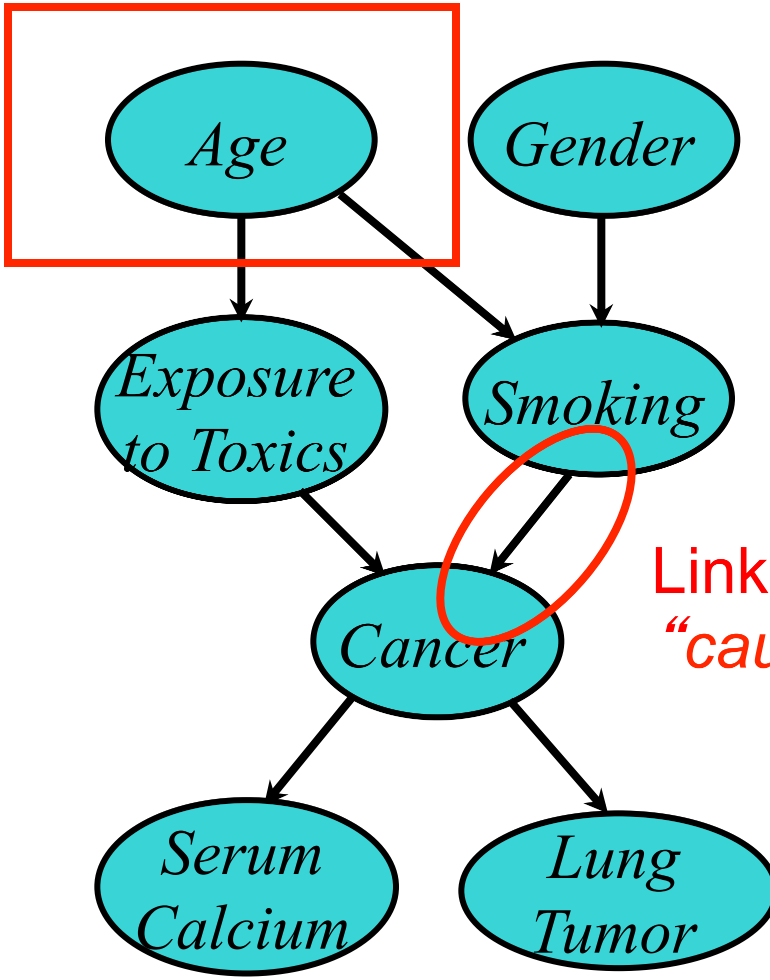
Smoking=	*no*	*light*	*heavy*
P(C=none)	0.96	0.88	0.60
P(C=benign)	0.03	0.08	0.25
P(C=malig)	0.01	0.04	0.15

More Complex Bayesian Network



More Complex Bayesian Network

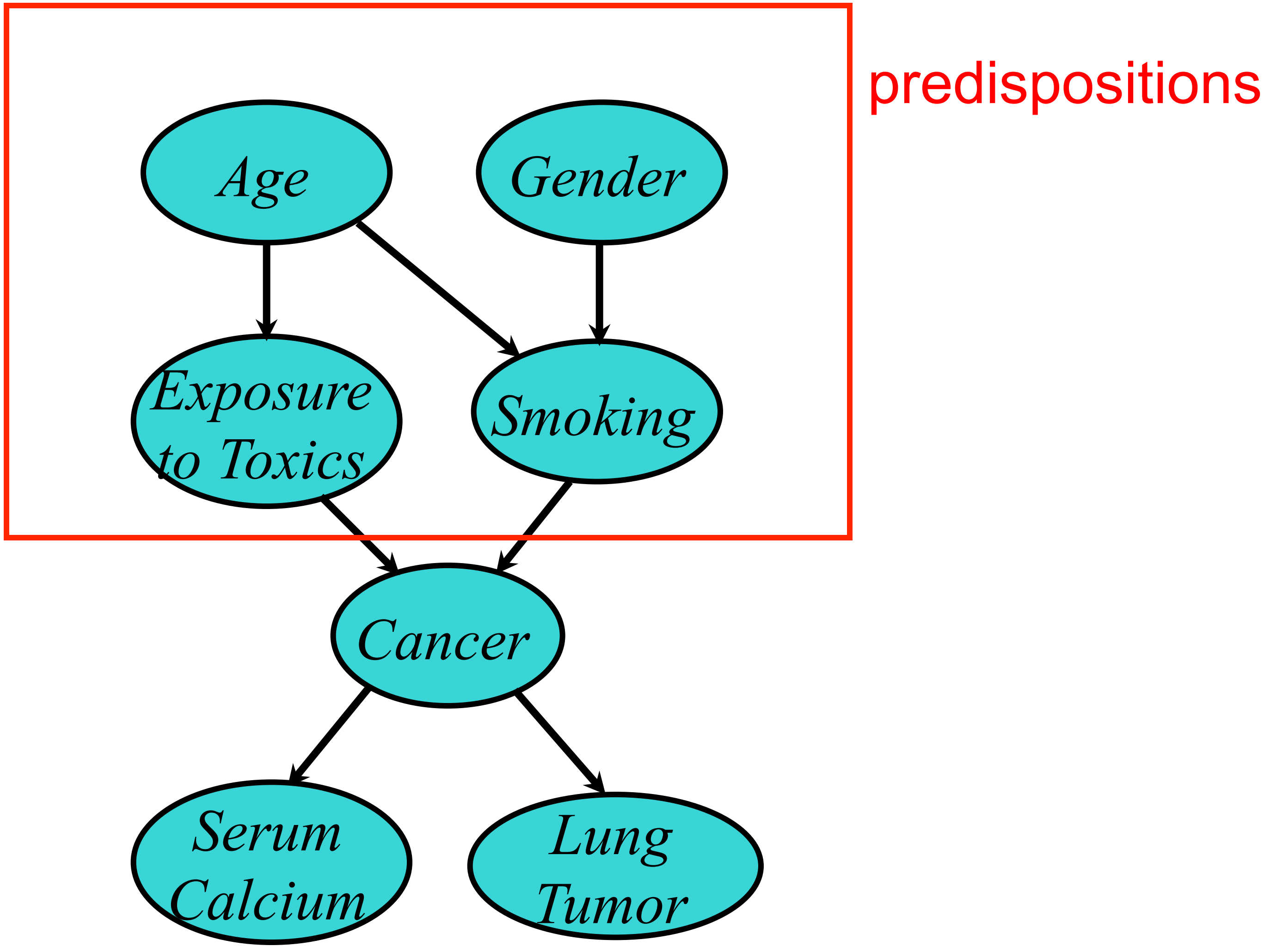
Nodes represent variables



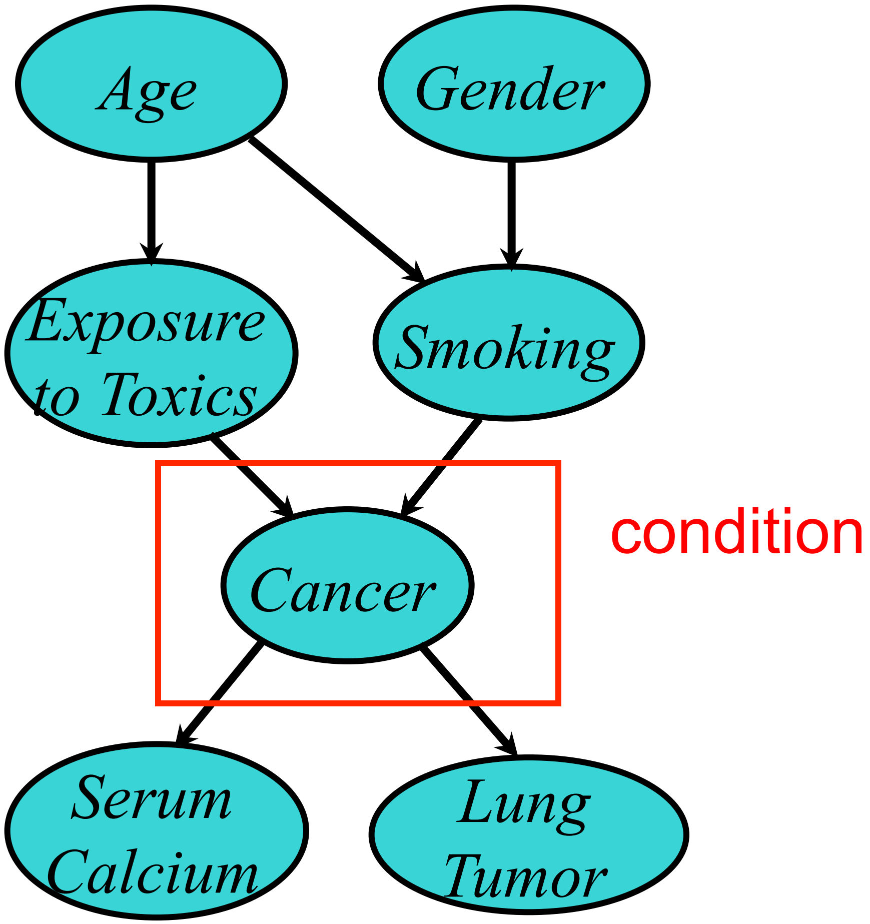
Links represent *"causal"* relations

- Does gender cause smoking?
- Influence might be a more appropriate term

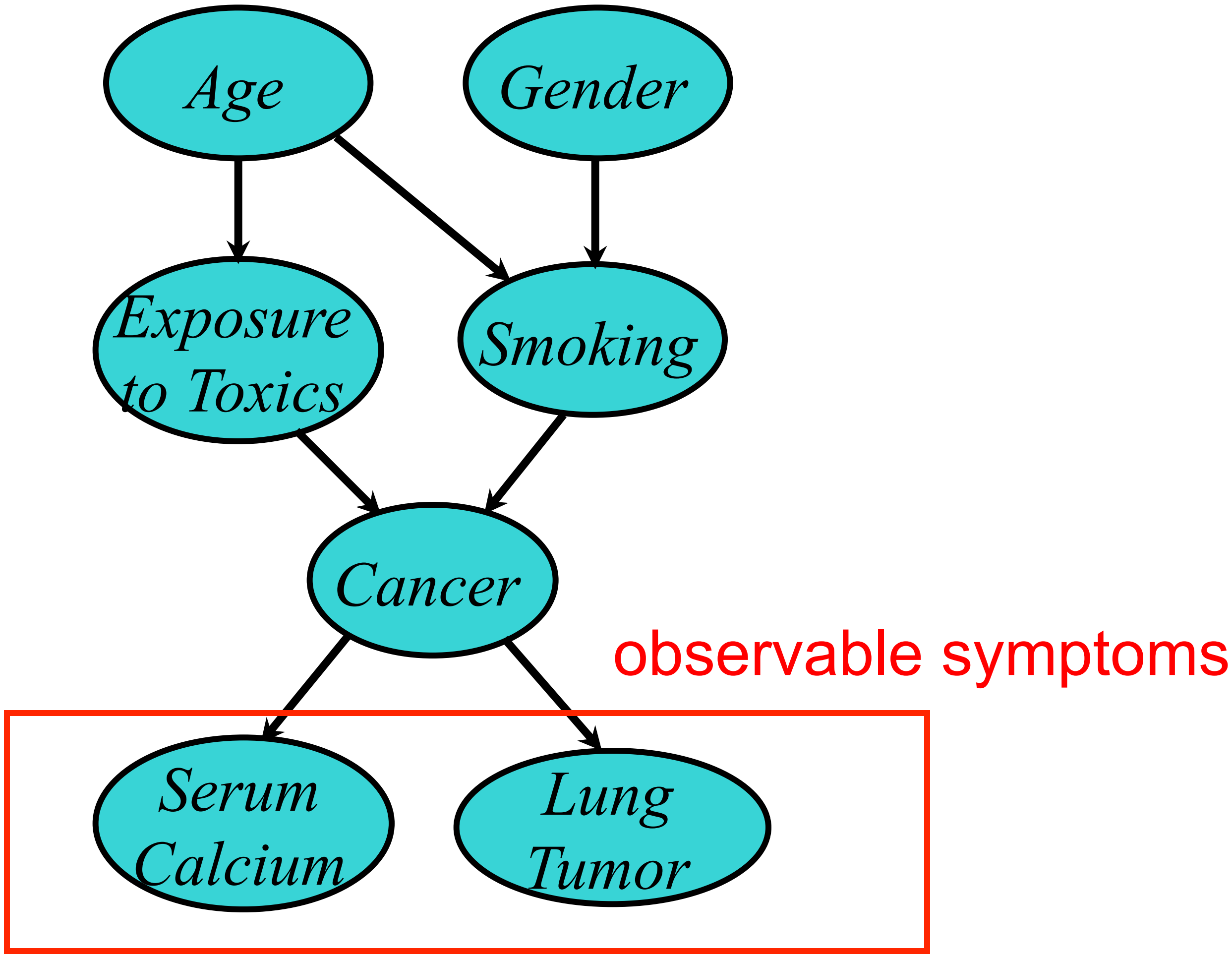
More Complex Bayesian Network



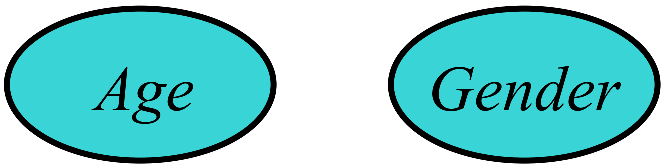
More Complex Bayesian Network



More Complex Bayesian Network



Independence



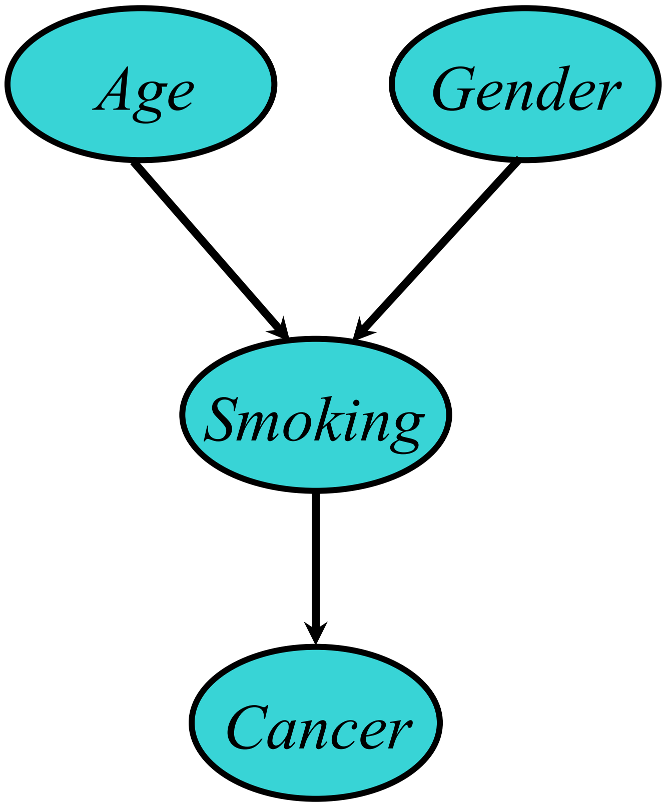
Age and *Gender* are independent.

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

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

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

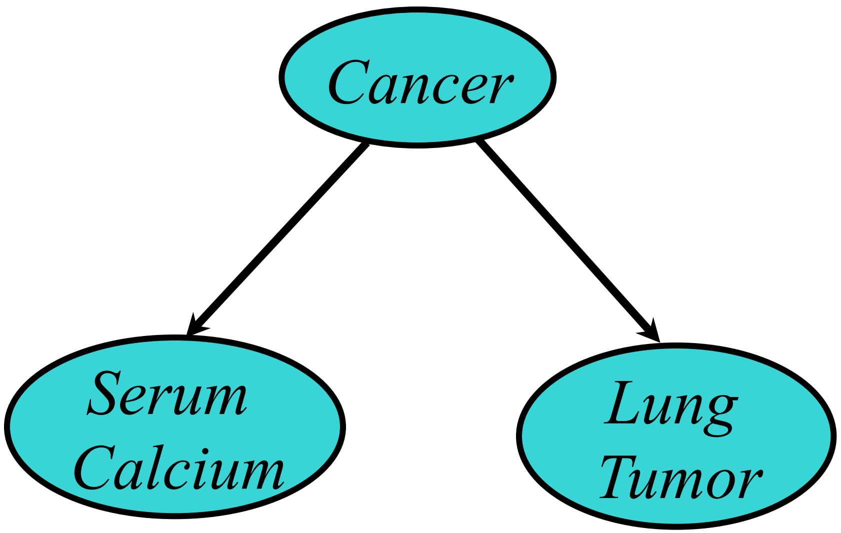
Conditional Independence



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

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

Conditional Independence: Naïve Bayes



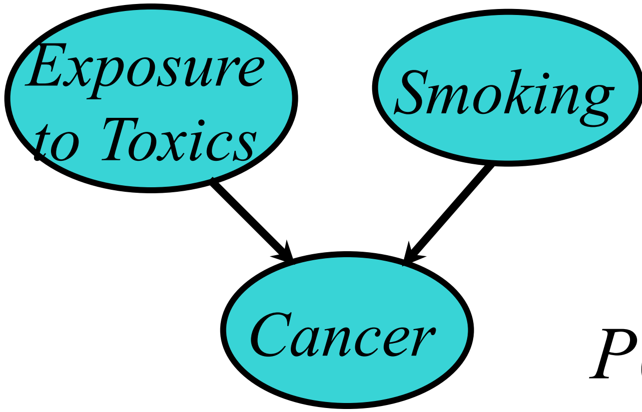
Serum Calcium and *Lung Tumor* are dependent

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

$$P(L \mid SC, C) = P(L \mid C)$$
$$P(SC \mid L, C) = P(SC \mid C)$$

Naïve Bayes assumption: 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

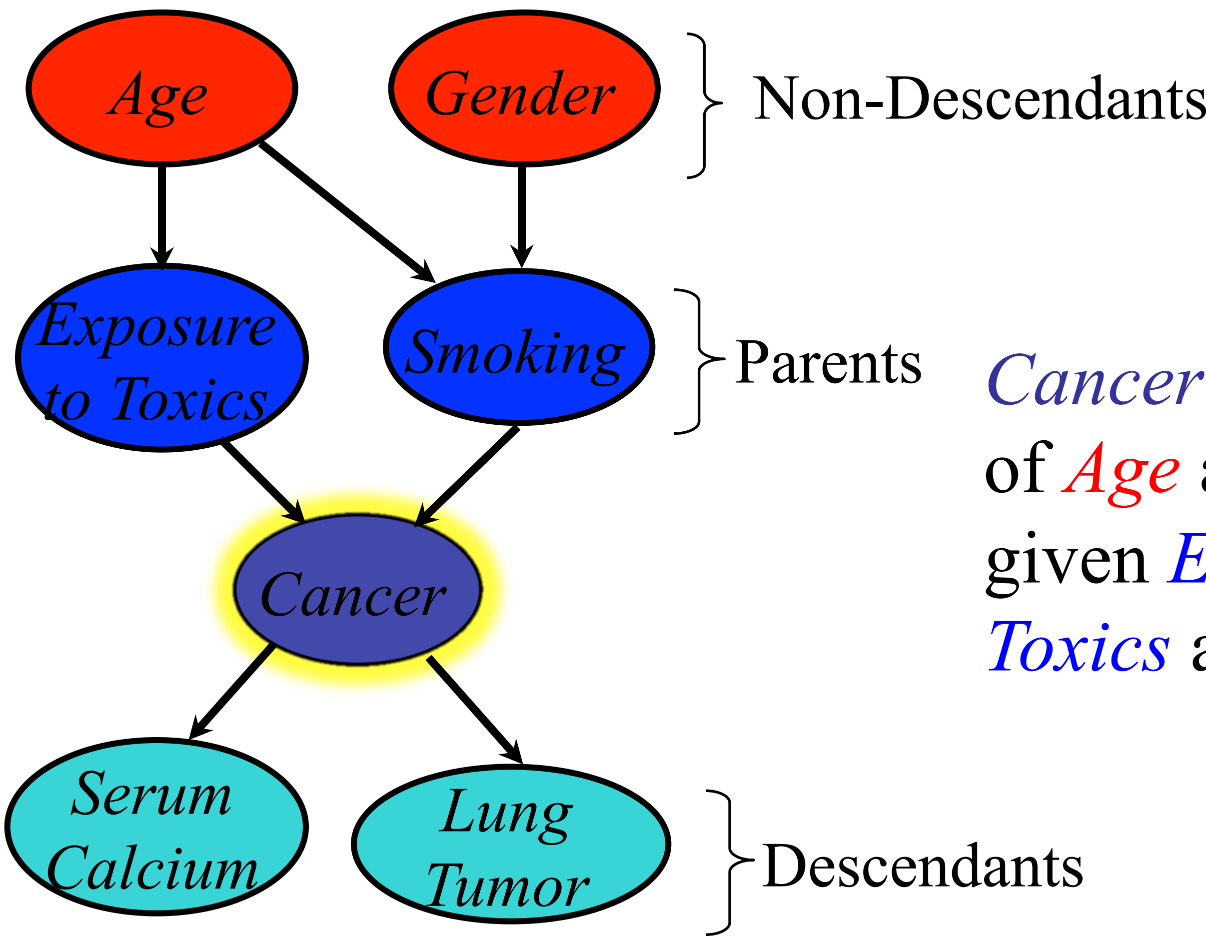
Exposure to Toxics is **dependent** on *Smoking*, given *Cancer*

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

- *Explaining away:* reasoning pattern where confirmation of one causereduces need to invoke alternatives
- Essence of Occam's Razor (prefer hypothesis with fewest assumptions)
- Relies on independence of causes

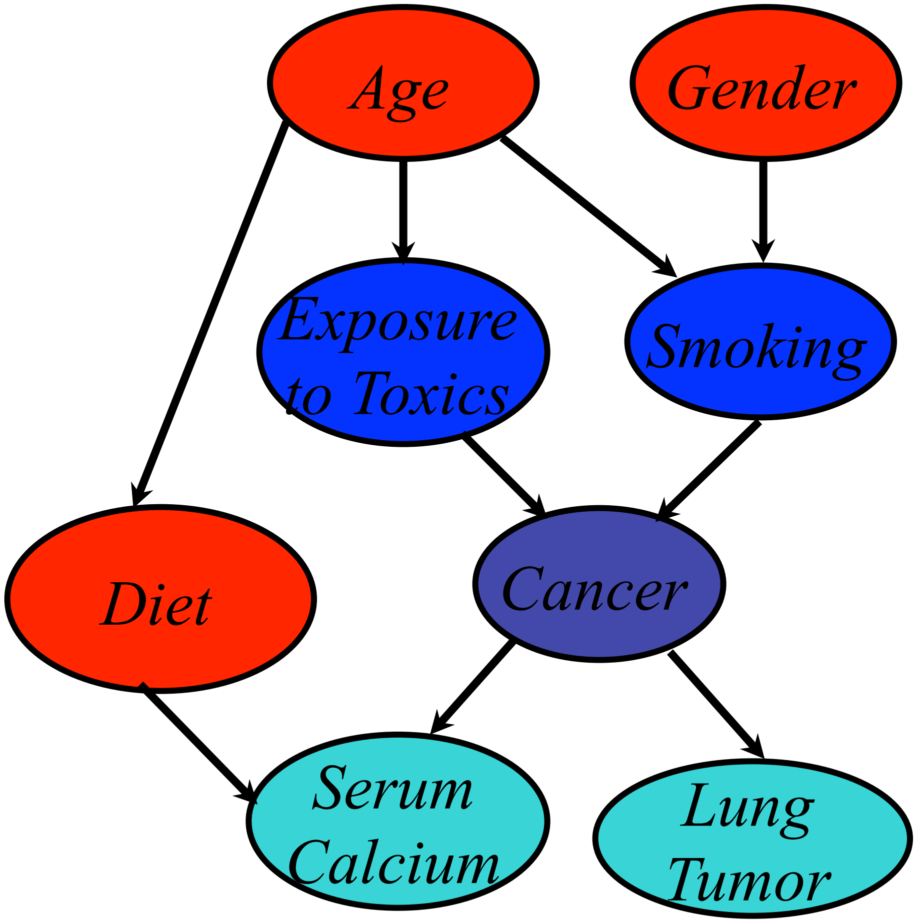
Conditional Independence

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



A variable is conditionally independent of its non-descendants given its parents

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

BBN Construction

The **knowledge acquisition** process for a BBN involves three steps

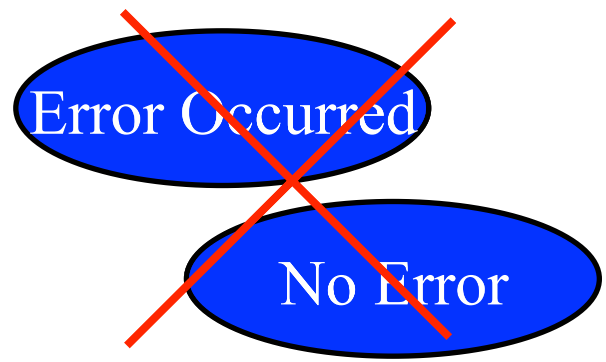
KA1: Choosing appropriate variables

KA2: Deciding on the network structure

KA3: Obtaining data for the conditional probability tables

KA1: Choosing variables

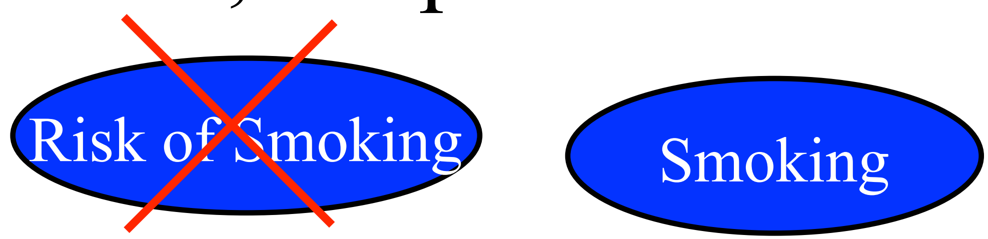
Variables should be collectively exhaustive, mutually exclusive values



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

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

They should be values, not probabilities

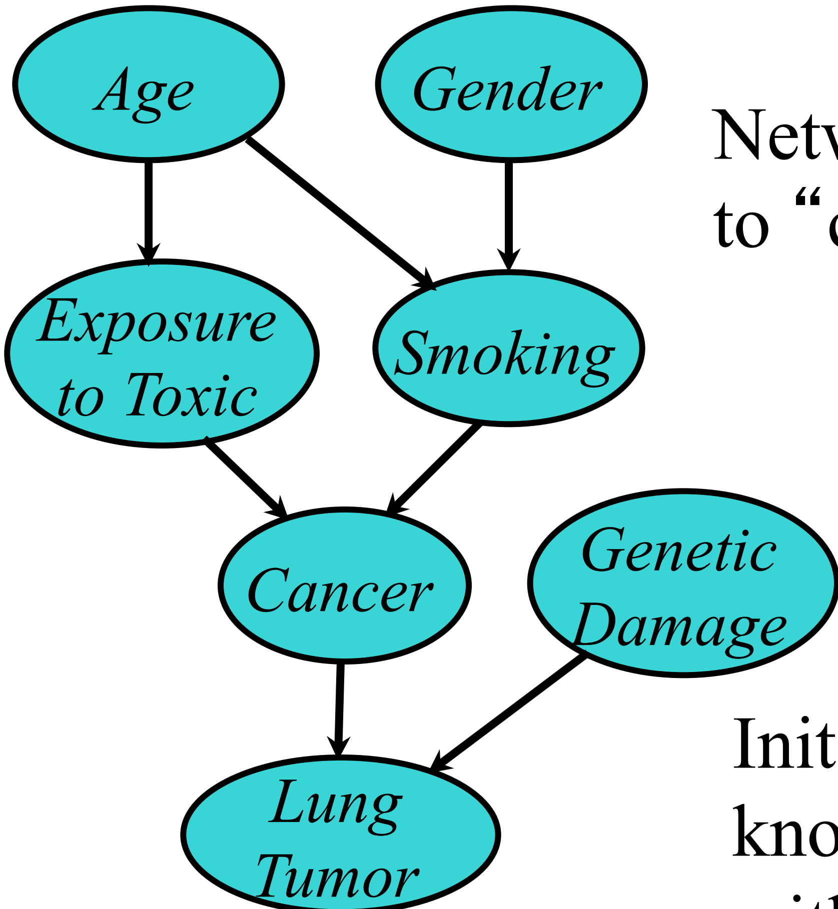


Heuristic: Knowable in Principle

Example of good variables

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

KA2: Structuring



Network structure corresponding to "causality" is usually good.

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

KA3: The Numbers

- For each variable we have a table of probability of its value for values of its parents
- For variables w/o parents, we have prior probabilities

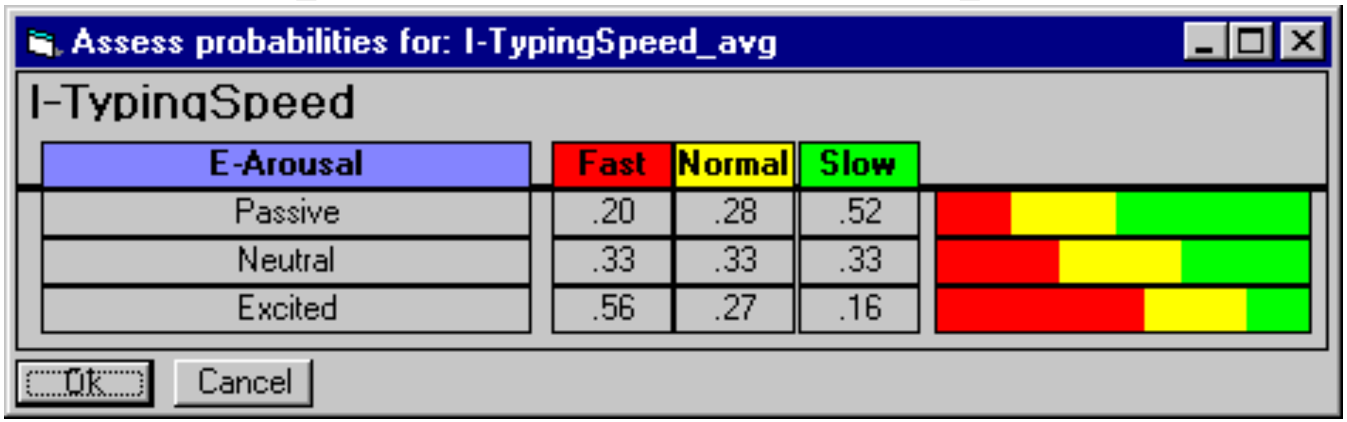
$$S \in \{no, light, heavy\}$$
$$C \in \{none, benign, malignant\}$$

smoking priors	
no	0.80
light	0.15
heavy	0.05

	smoking		
cancer	**no**	**light**	**heavy**
none	0.96	0.88	0.60
benign	0.03	0.08	0.25
malignant	0.01	0.04	0.15

KA3: 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

Three kinds of reasoning

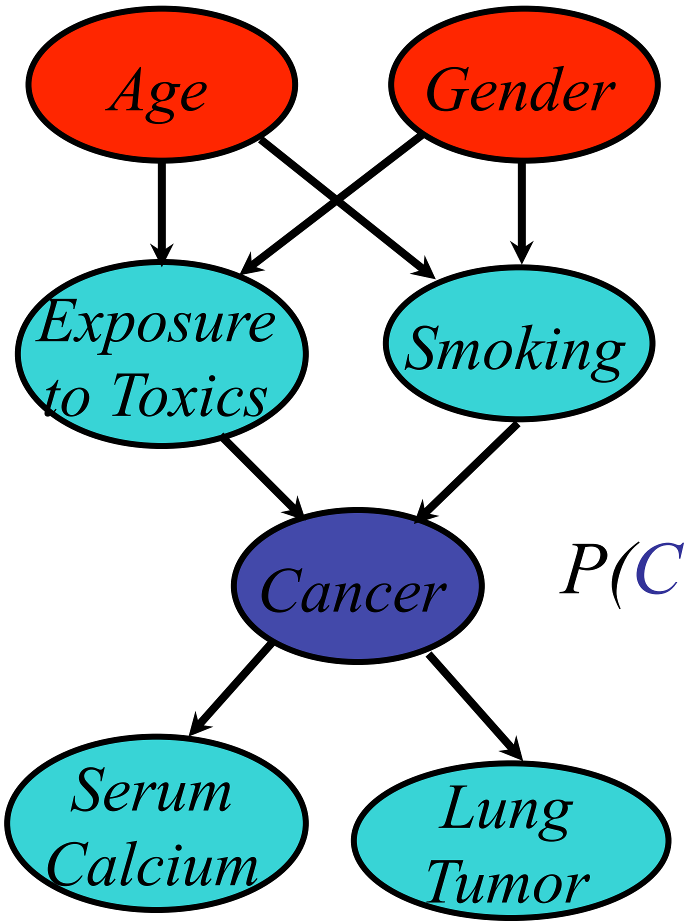
BBNs support three main kinds of reasoning:

- **Predicting** conditions given predispositions
- **Diagnosing** conditions given symptoms (and predisposing)
- **Explaining** a condition by one or more predispositions

To which we can add a fourth:

- **Deciding** on an action based on probabilities of the conditions

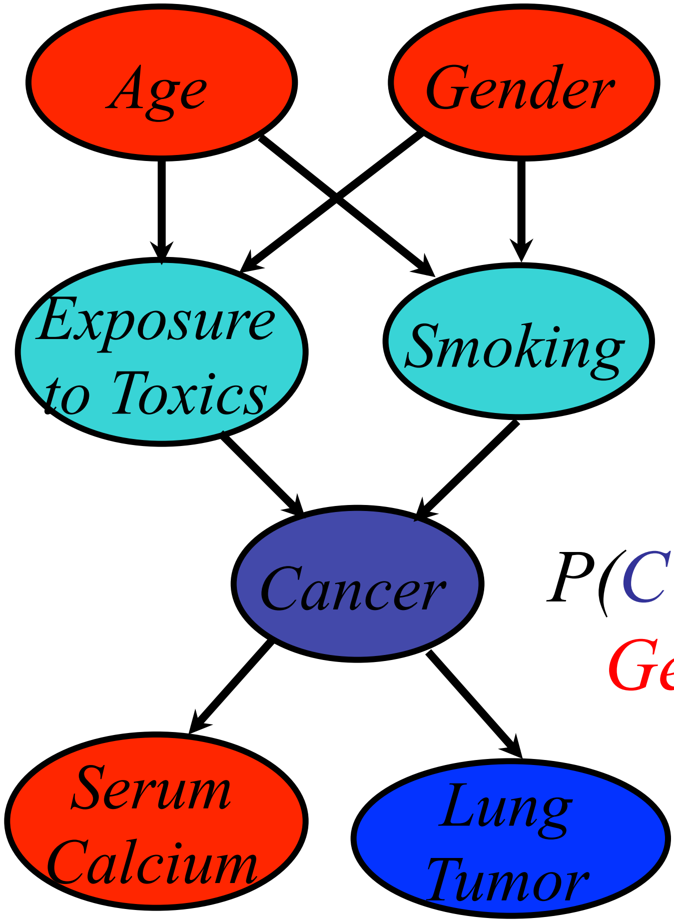
Predictive Inference



How likely are elderly males to get malignant cancer?

$P(C{=}malignant \mid Age{>}60, Gender{=}male)$

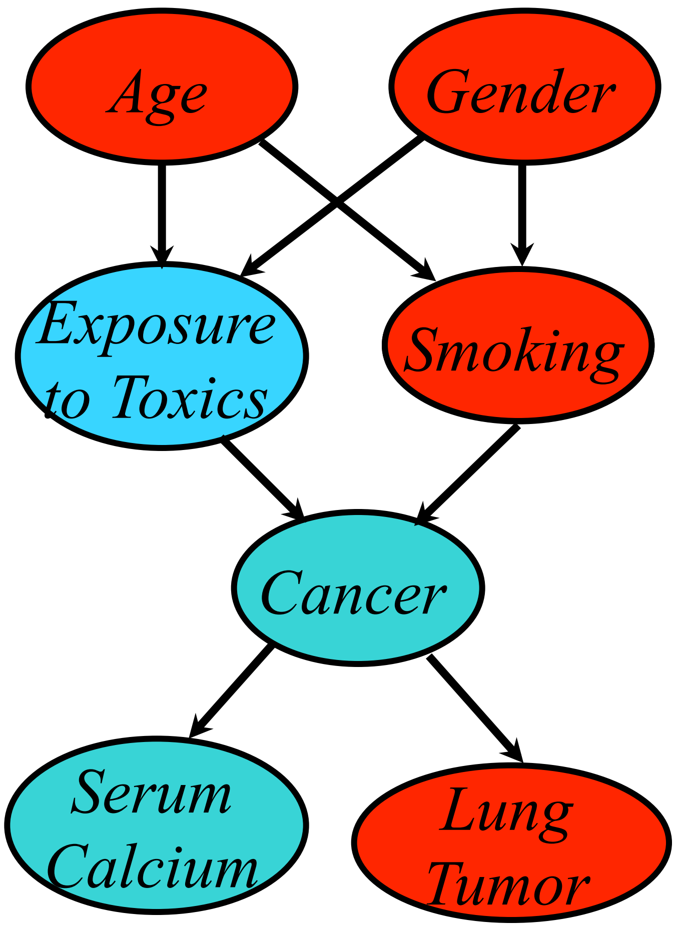
Predictive and diagnostic combined



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

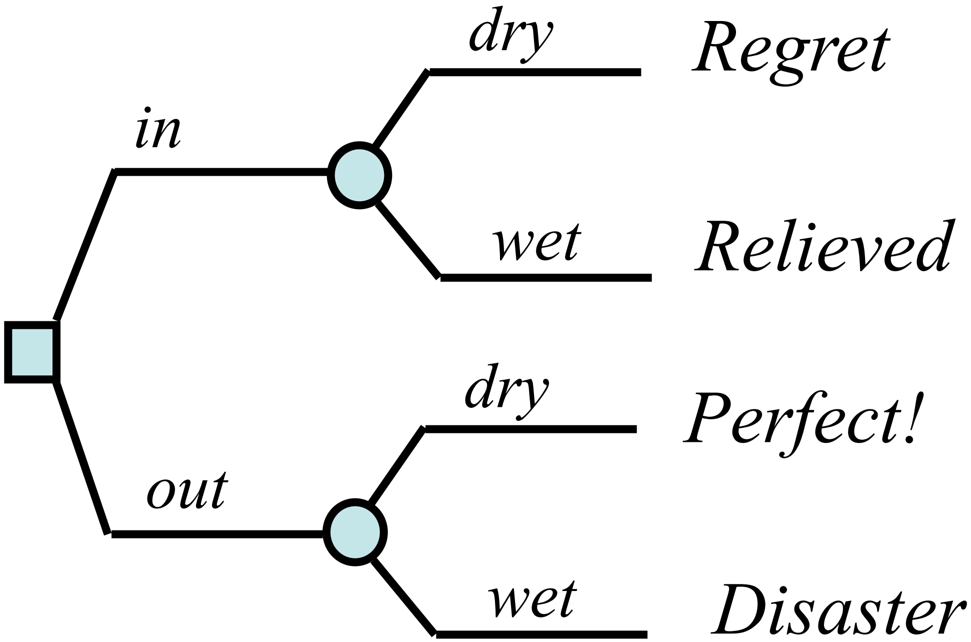
- If we then observe heavy smoking, the probability of exposure to toxics goes back down

Decision making

- A decision is a medical domain might be a choice of treatment (e.g., radiation or chemotherapy)
- Decisions should be made 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?



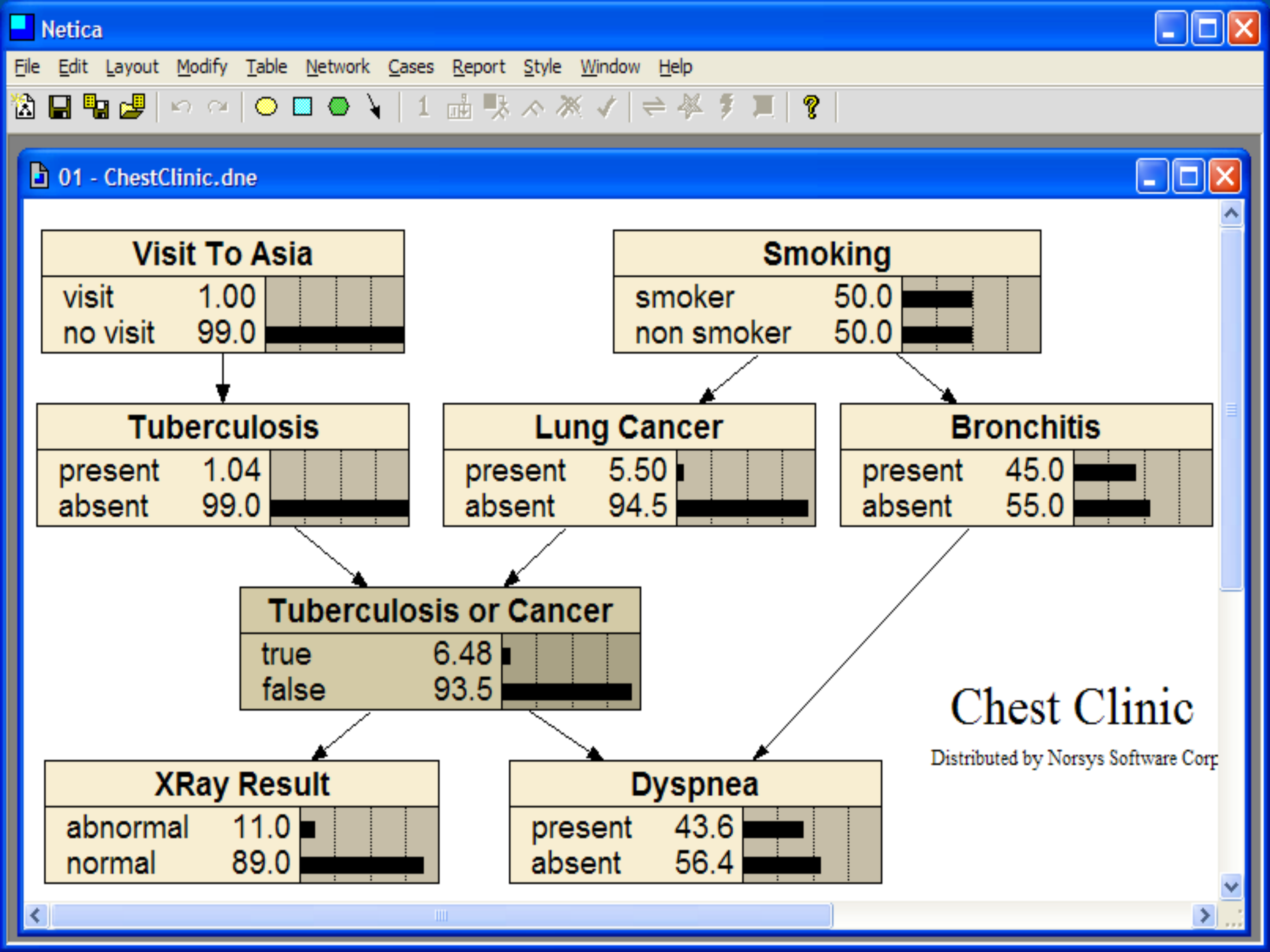
Value Function

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

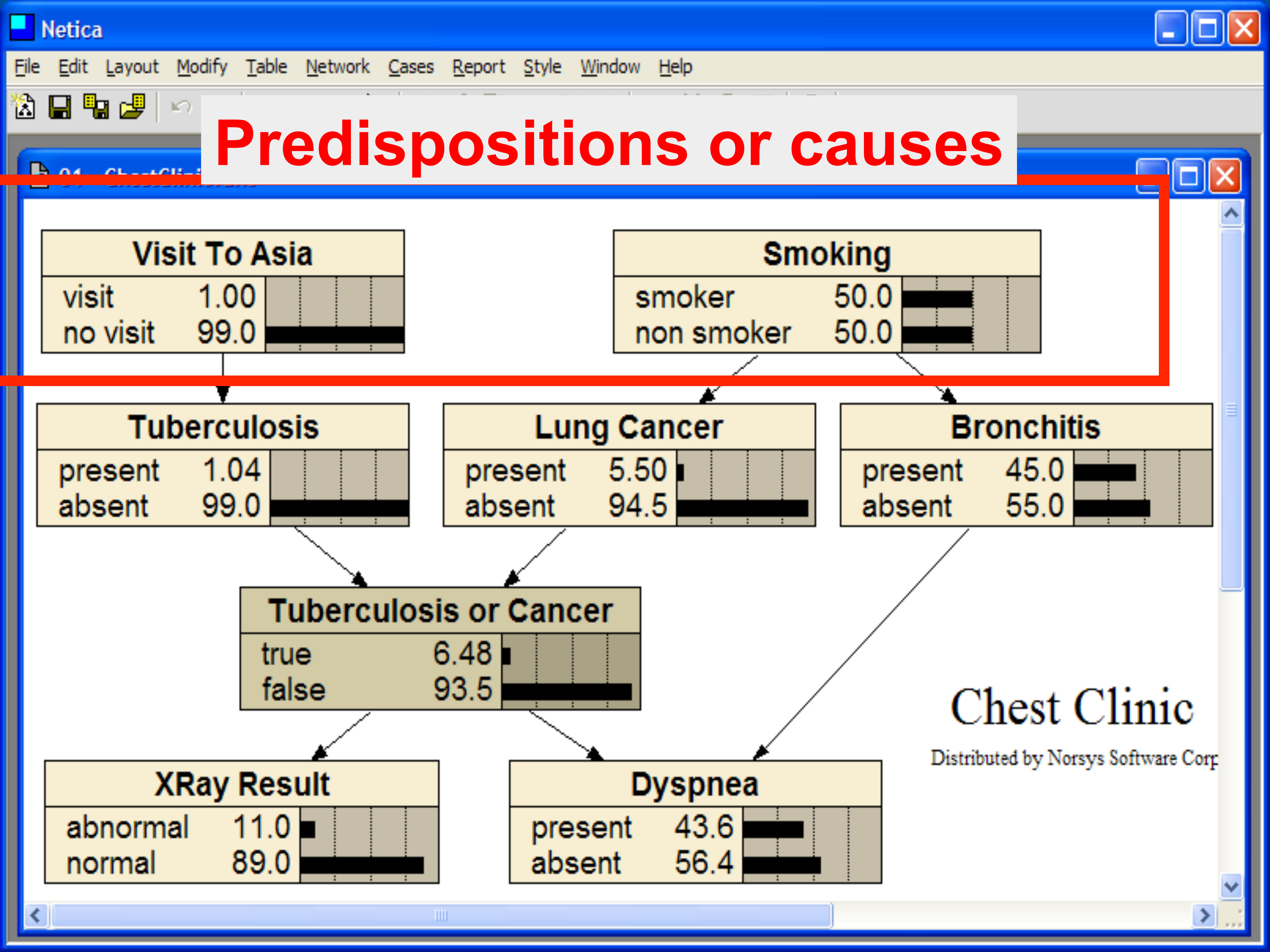
Location?	Weather?	Value
in	dry	$50
in	wet	$60
out	dry	$100
out	wet	$0

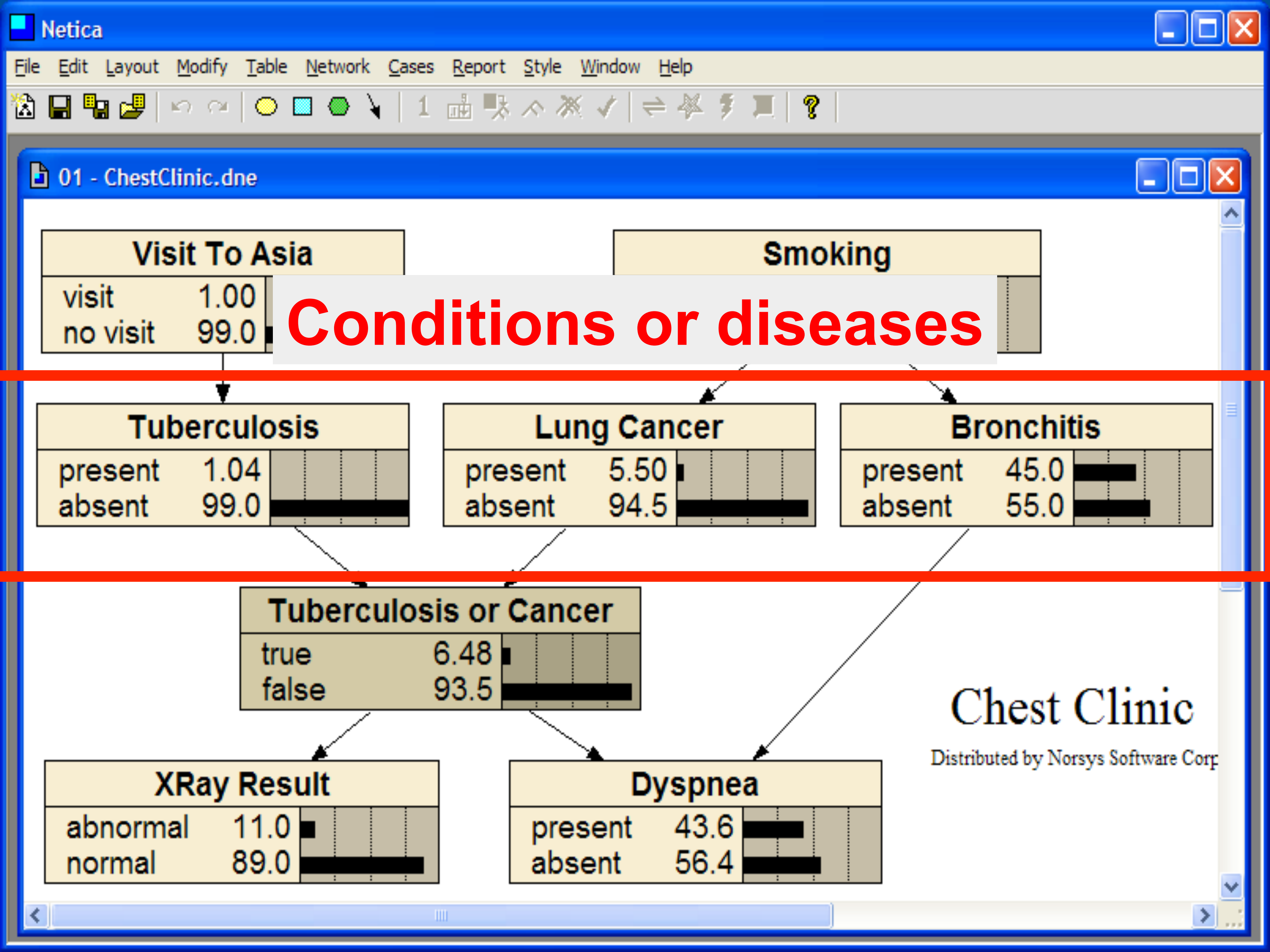
Two software tools

- Netica: Windows app 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.
- Samiam: Java system for modeling and reasoning with Bayesian networks
  - Includes a GUI and reasoning engine

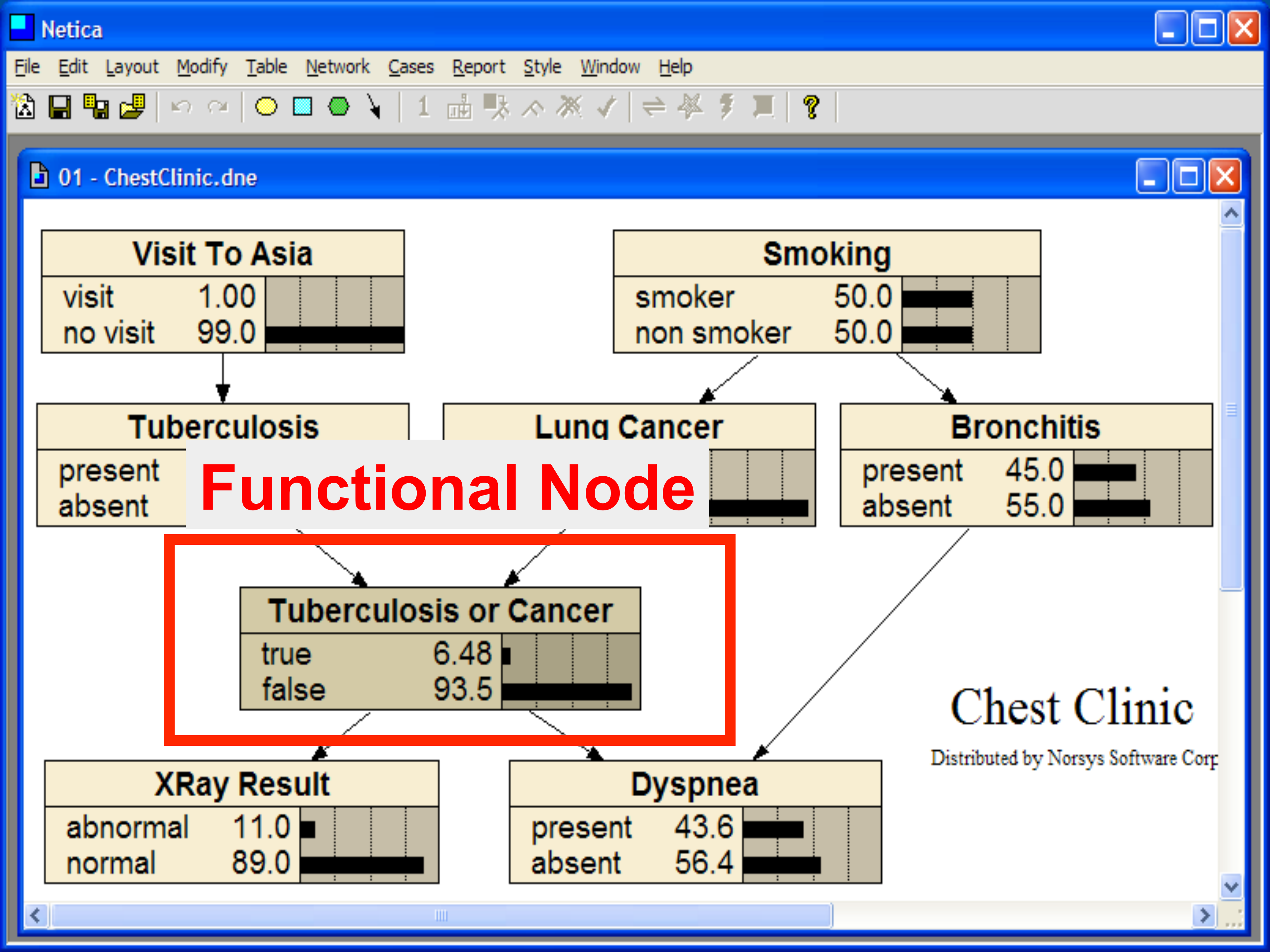

Netica
File Edit Layout Modify Table Network Cases Report Style Window Help
01 - ChestClinic.dne
Visit To Asia
visit 1.00
no visit 99.0
Smoking
smoker 50.0
non smoker 50.0
Tuberculosis
present 1.04
absent 99.0
Lung Cancer
present 5.50
absent 94.5
Bronchitis
present 45.0
absent 55.0
Tuberculosis or Cancer
true 6.48
false 93.5
Chest Clinic
Distributed by Norsys Software Corp
XRay Result
abnormal 11.0
normal 89.0
Dyspnea
present 43.6
absent 56.4

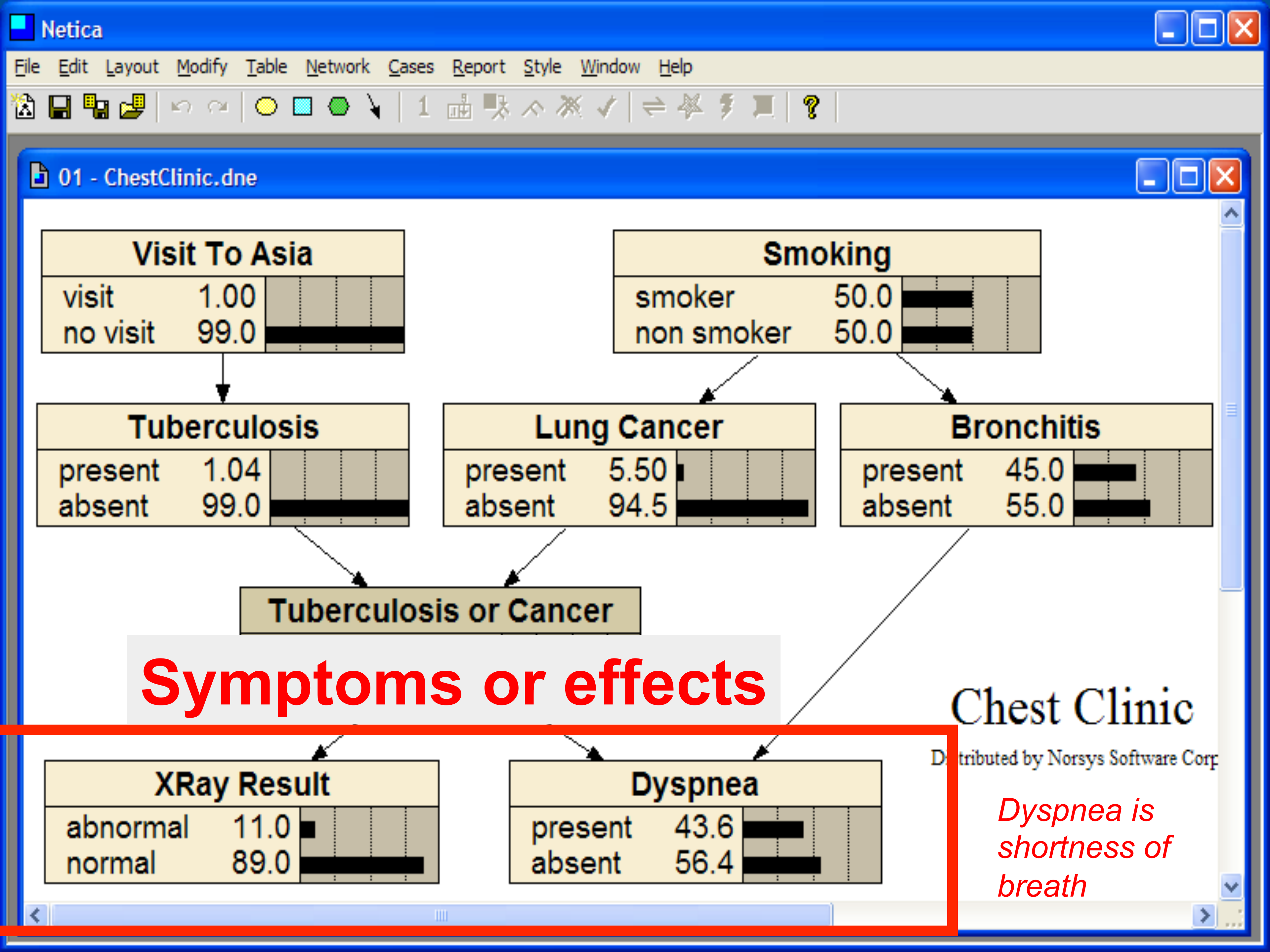
Predispositions or causes





Conditions or diseases
Netica
File Edit Layout Modify Table Network Cases Report Style Window Help
01 - ChestClinic.dne
Visit To Asia
visit 1.00
no visit 99.0
Smoking
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
File Edit Layout Modify Table Network Cases Report Style Window Help
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
Bronchitis
present 45.0
absent 55.0
Functional Node
Tuberculosis or Cancer
true 6.48
false 93.5
Chest Clinic
Distributed by Norsys Software Corp
XRay Result
abnormal 11.0
normal 89.0
Dyspnea
present 43.6
absent 56.4

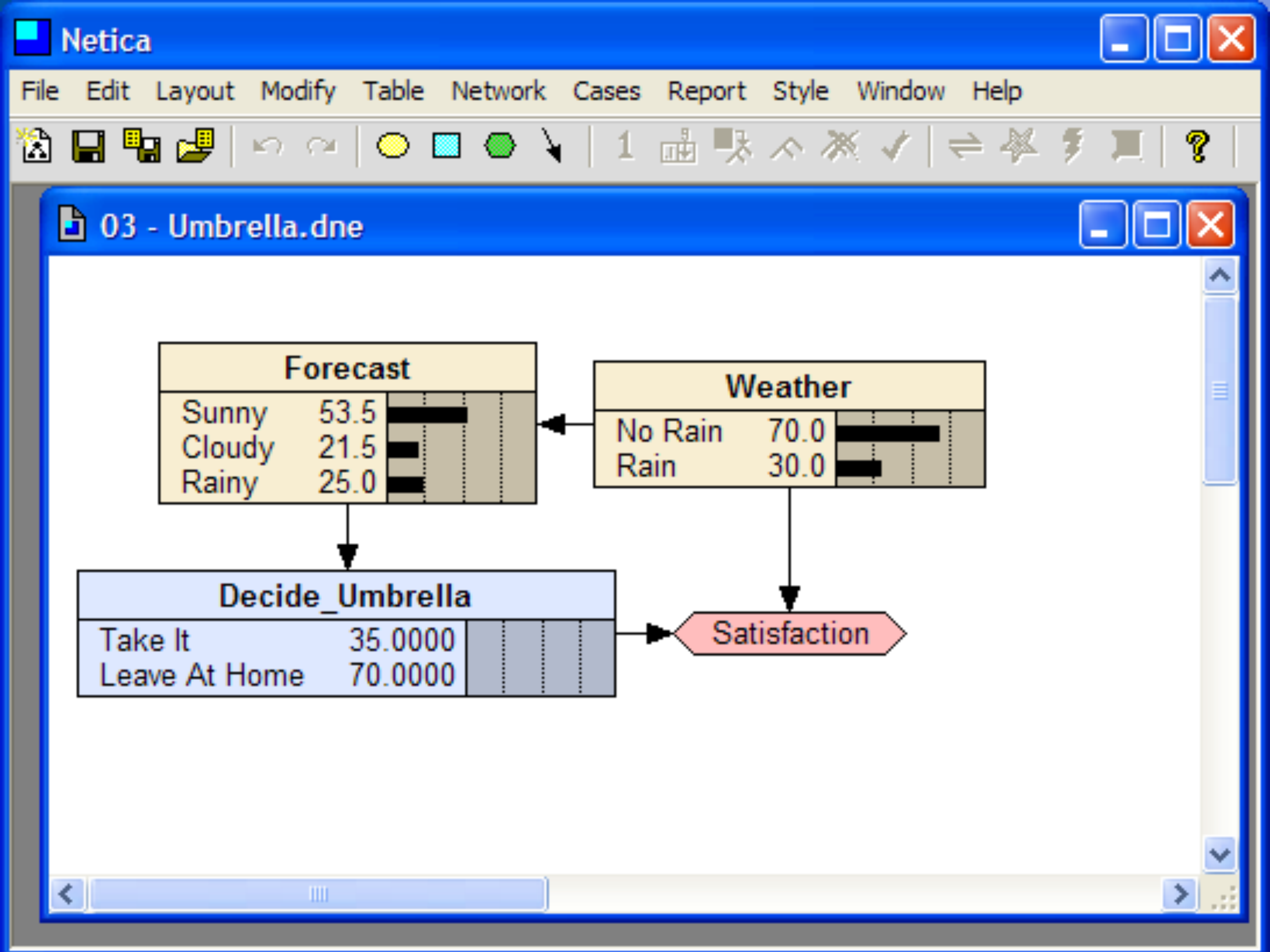

Netica
File Edit Layout Modify Table Network Cases Report Style Window Help
01 - ChestClinic.dne
Visit To Asia
visit 1.00
no visit 99.0
Smoking
smoker 50.0
non smoker 50.0
Tuberculosis
present 1.04
absent 99.0
Lung Cancer
present 5.50
absent 94.5
Bronchitis
present 45.0
absent 55.0
Tuberculosis or Cancer
Symptoms or effects
Chest Clinic
XRay Result
abnormal 11.0
normal 89.0
Dyspnea
present 43.6
absent 56.4
Dyspnea is shortness of breath

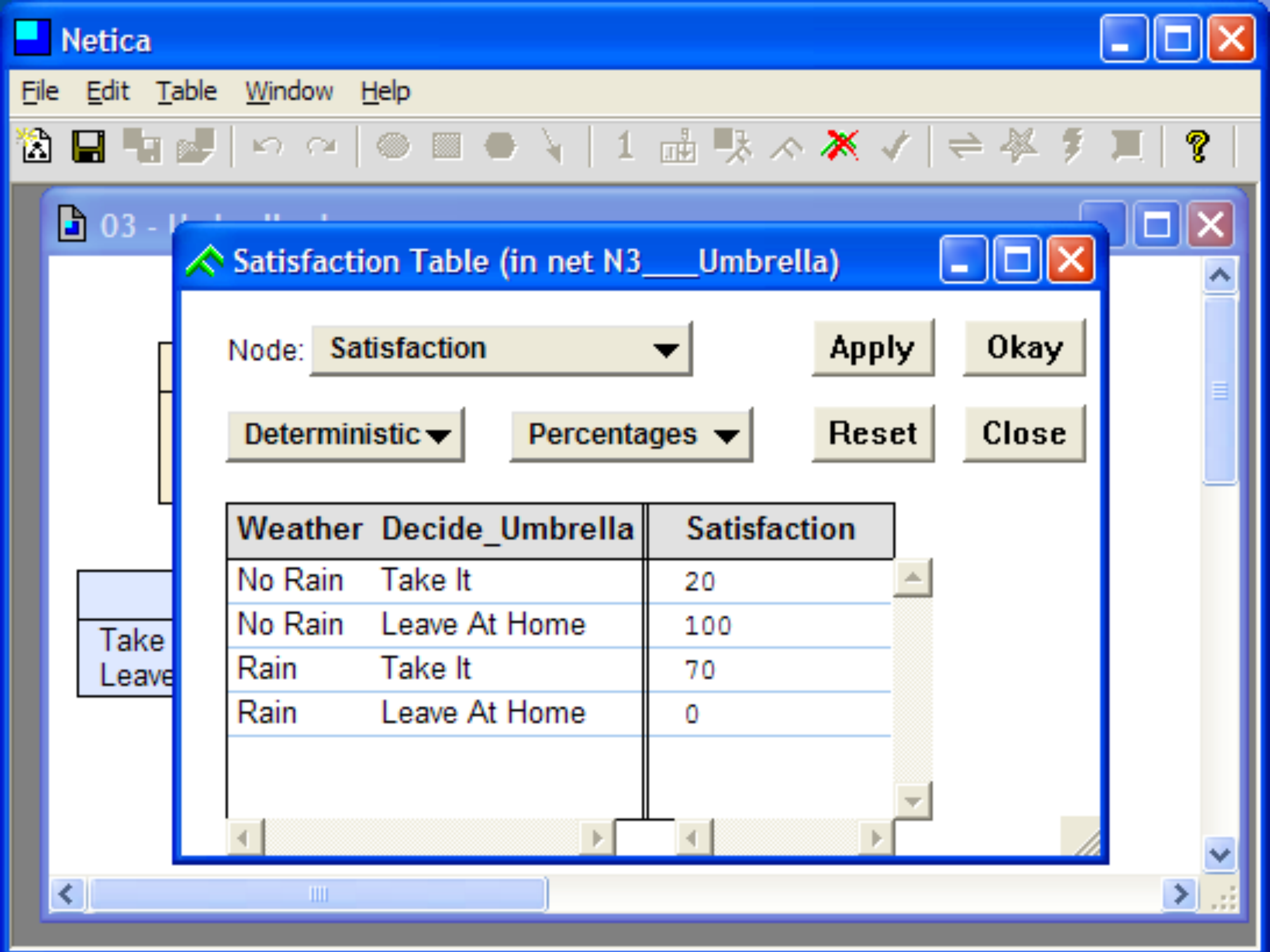
Decision Making with BBNs

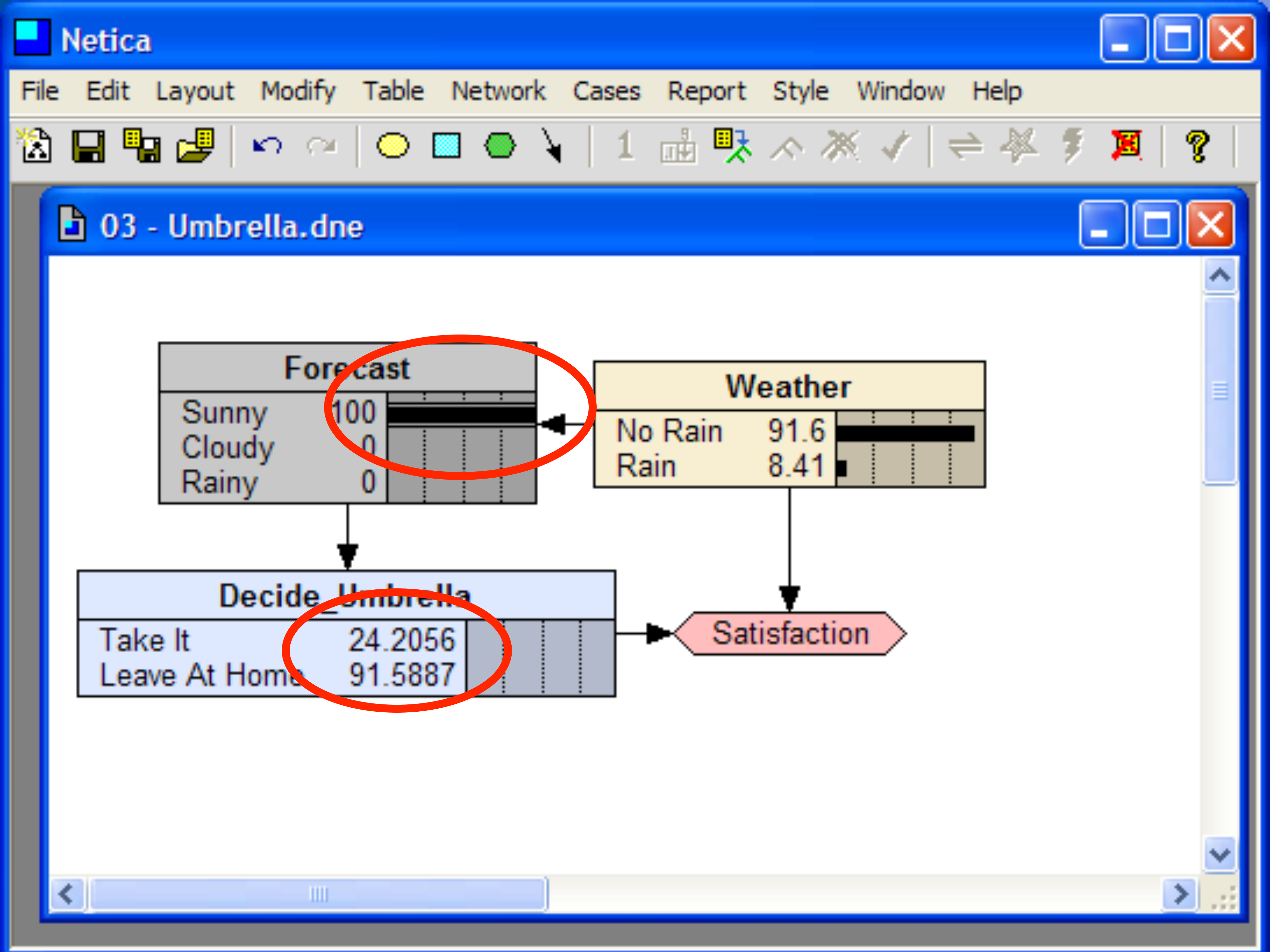
- Today's weather forecast might be either sunny, cloudy or rainy
- Should you take an umbrella when you leave?
- Your decision depends only on the forecast
  - The forecast "depends on" the actual weather
- Your satisfaction depends on your decision and the weather
  - Assign a utility to each of four situations: (rain|no rain) x (umbrella, no umbrella)

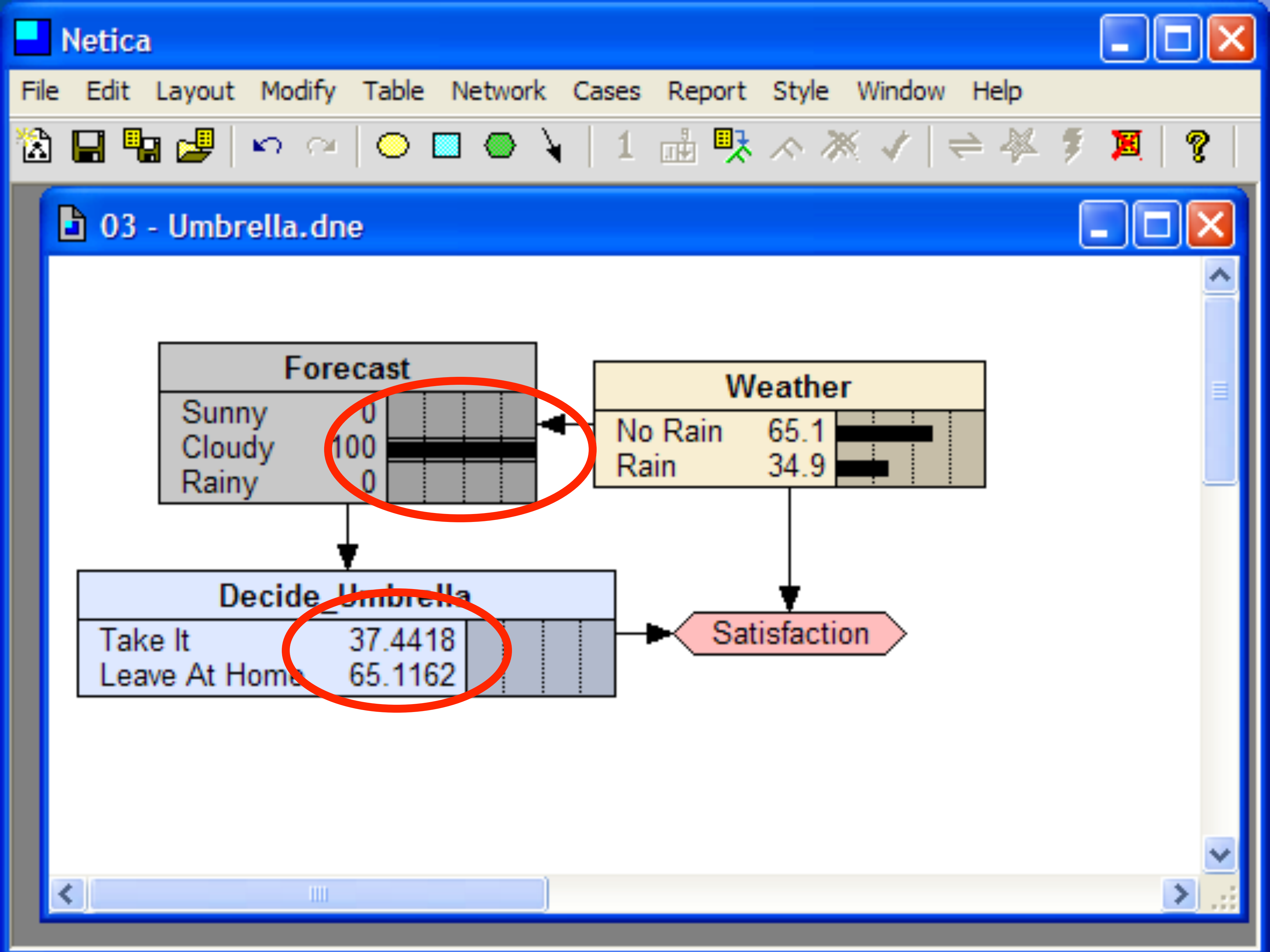
Decision Making with BBNs

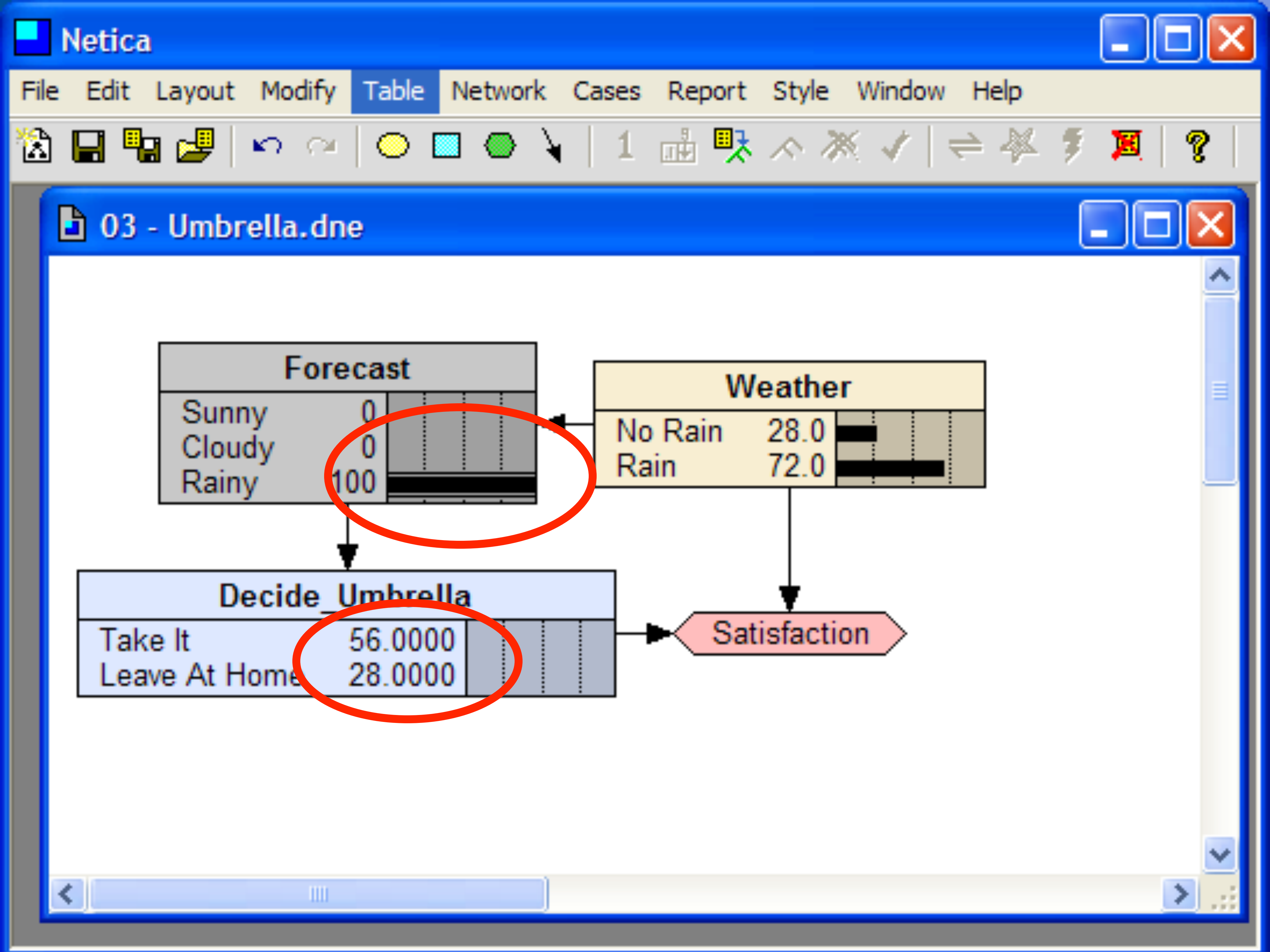
- Extend the BBN framework to include two new kinds of nodes: Decision and Utility
- A **Decision** node computes the expected utility of a decision given its parent(s), e.g., forecast, an a valuation
- A **Utility** node computes a utility value given its parents, e.g. a decision and weather
  - We can assign a utility to each of four situations: (rain|no rain) x (umbrella, no umbrella)
  - The value assigned to each is probably subjective


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
Take
Leave


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