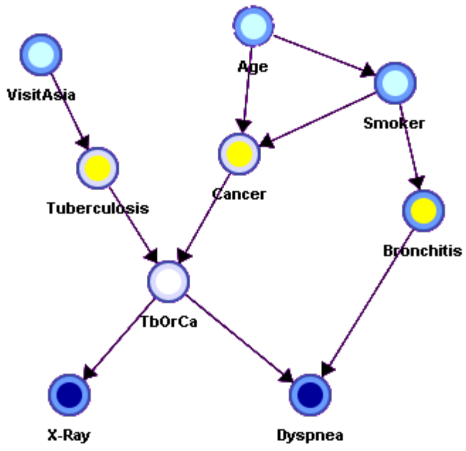


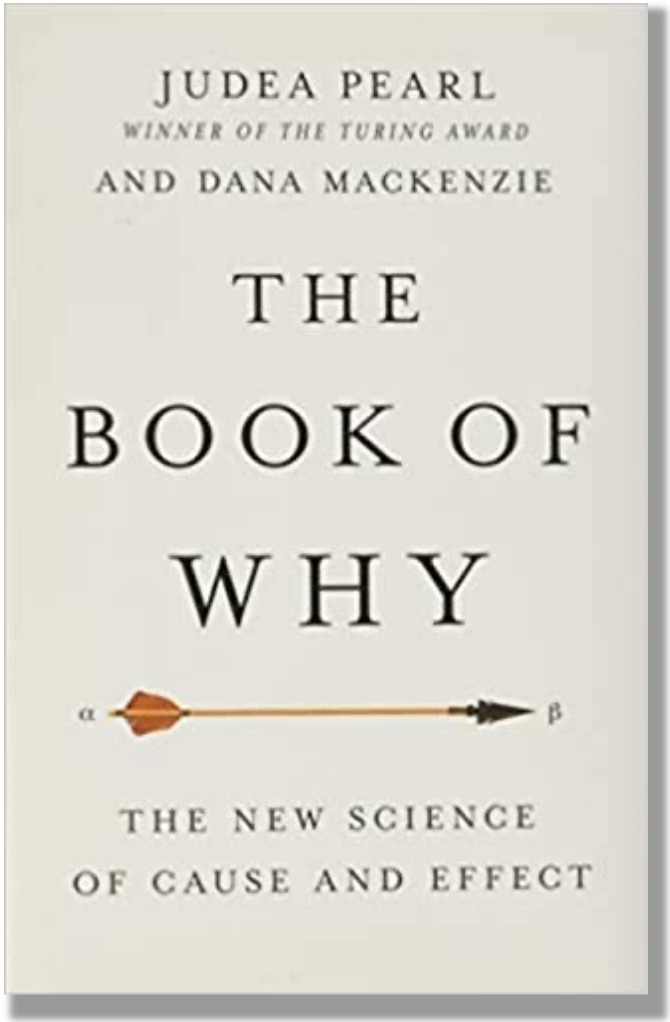
Reasoning with Bayesian Belief Networks

Overview

- Bayesian Belief Networks (BBNs) can reason with networks of propositions and associated probabilities
- BBNs encode causal associations between facts and events the propositions represent
- Useful for many AI problems
  - Diagnosis
  - Expert systems
  - Planning
  - Learning

Judea Pearl

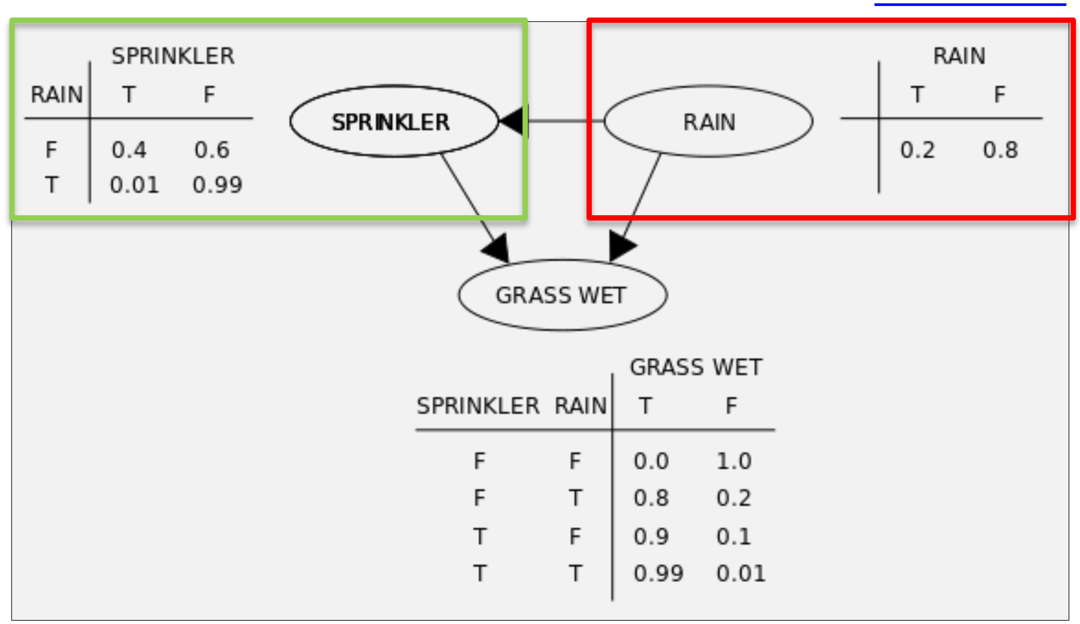
- UCLA CS professor
- Introduced Bayesian networks in the 1980s
- Pioneer of probabilistic approach to AI reasoning
- First to formalize causal modeling in empirical sciences
- Written many books on the topics, including the popular 2018 Book of Why



BBN Definition

- AKA Bayesian Network, Bayes Net
- A graphical model (as a DAG) of probabilistic relationships among a set of random variables
- Nodes are variables, links represent direct influence of one variable on another
- Nodes have **prior probabilities** or **conditional probability tables** (CPTs)

source



Recall Bayes Rule

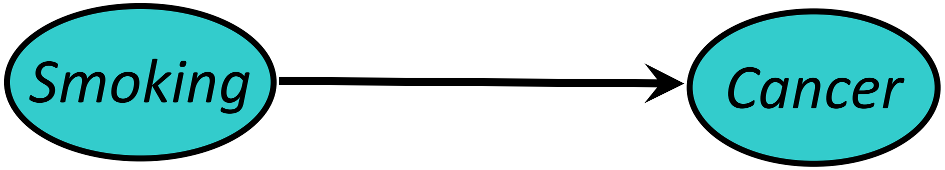
$$P(H,E) = P(H \mid E)P(E) = P(E \mid H)P(H)$$

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

Note symmetry: can compute probability of a ***hypothesis given its evidence*** as well as probability of ***evidence given hypothesis***

Simple Bayesian Network

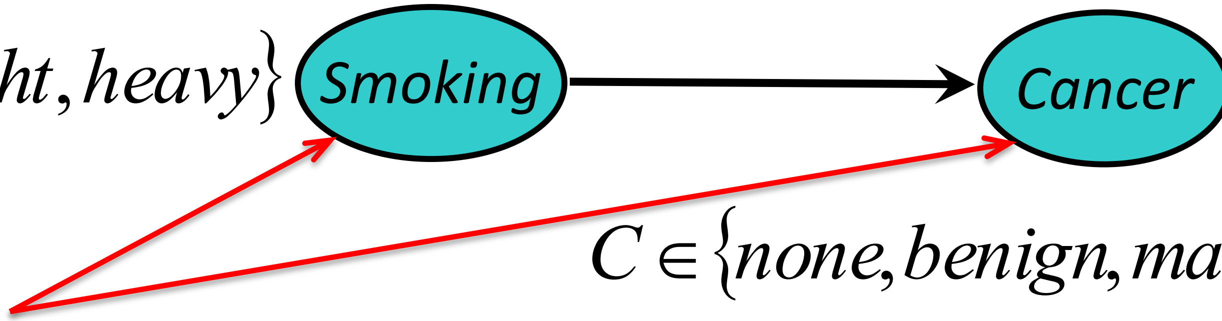
$S \in \{no, light, heavy\}$



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

Simple Bayesian Network

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



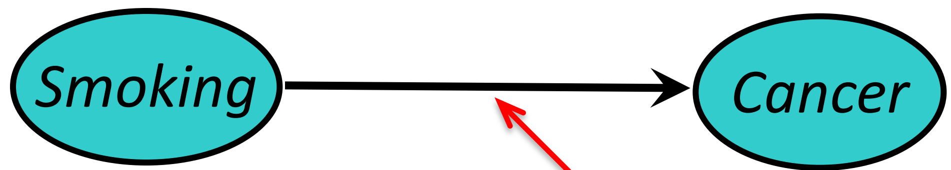
$$C \in \{none, benign, malignant\}$$

Nodes represent variables

- **Smoking** variable represents person's degree of smoking and has three possible values (no, light, heavy)
- **Cancer** variable represents person's cancer diagnosis and has three possible values (none, benign, malignant)

Simple Bayesian Network

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



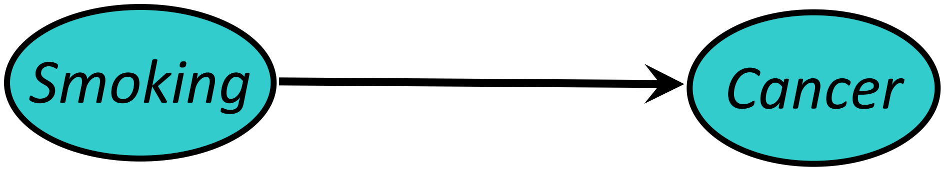
$C \in \{none, benign, malignant\}$

- **tl;dr:** smoking effects cancer
- **Smoking** behavior effects the probability of **cancer** outcome
- **Smoking** behavior considered evidence for whether a person is likely to have cancer or not

Directed links represent *"causal"* relations

Simple Bayesian Network

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



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

Prior probability of S

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

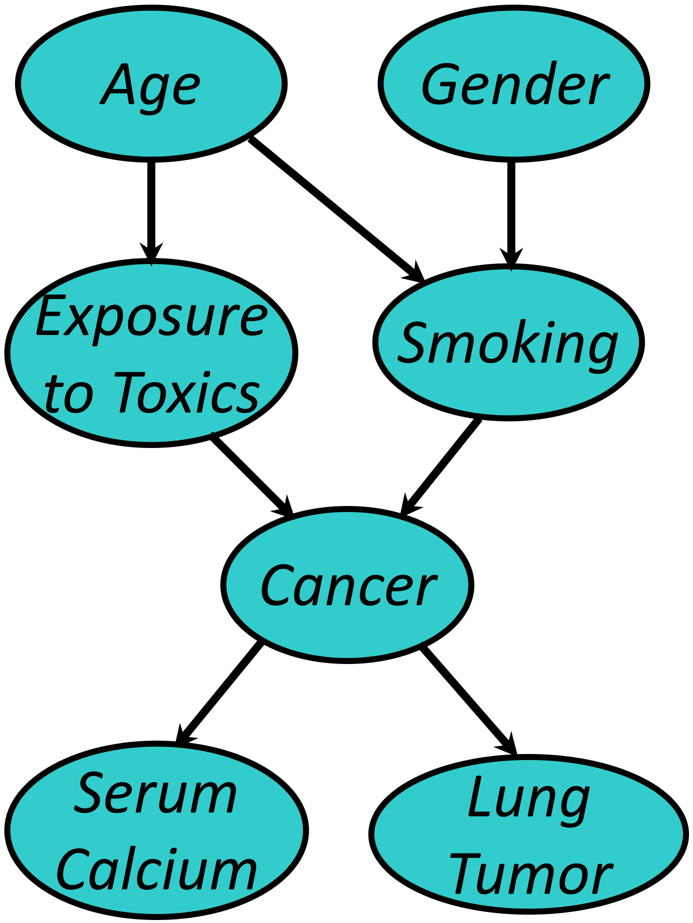
Nodes without in-links have **prior probabilities**

Joint distribution of S and C

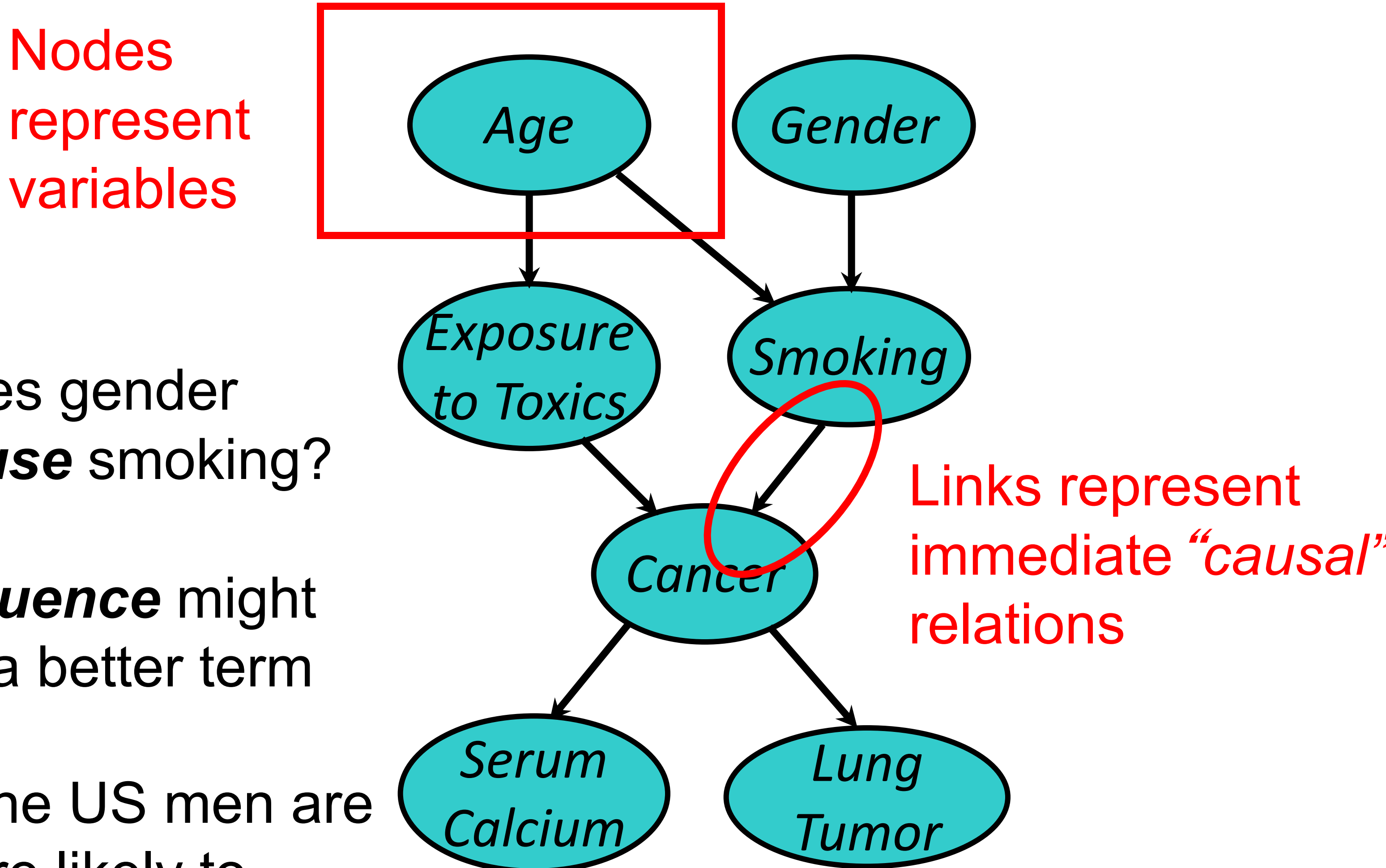
Smoking=	*no*	*light*	*heavy*
C=none	0.96	0.88	0.60
C=benign	0.03	0.08	0.25
C=malignant	0.01	0.04	0.15

Nodes with in-links have **joint probability distributions**

More Complex Bayesian Network

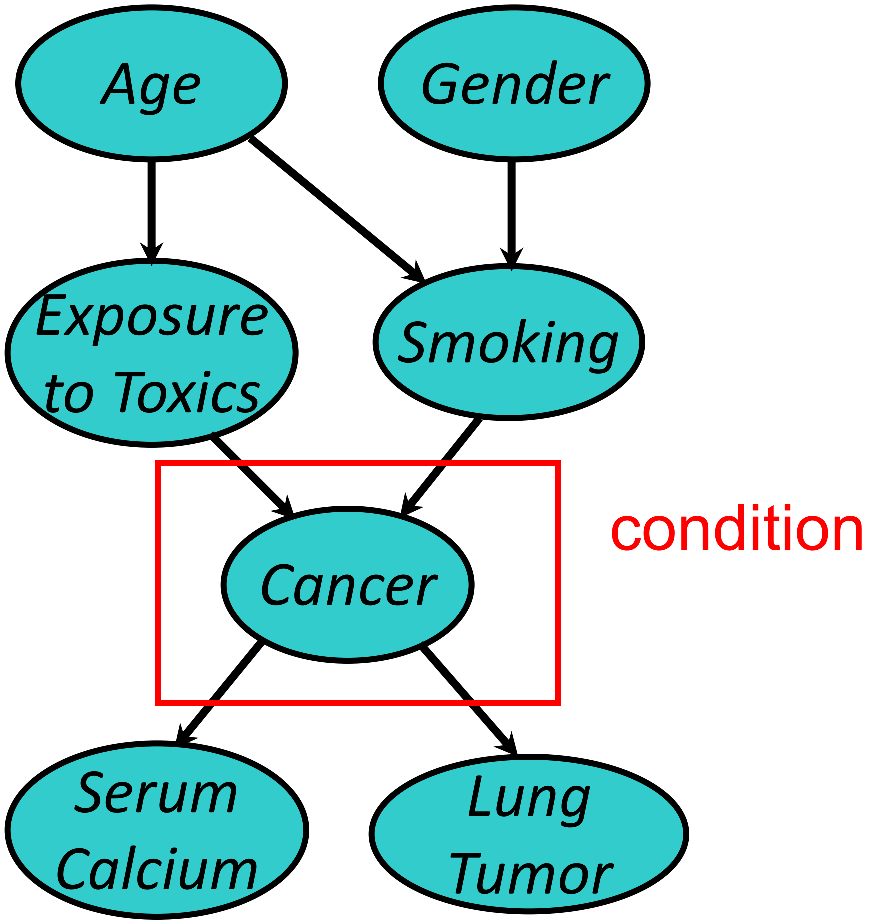


More Complex Bayesian Network

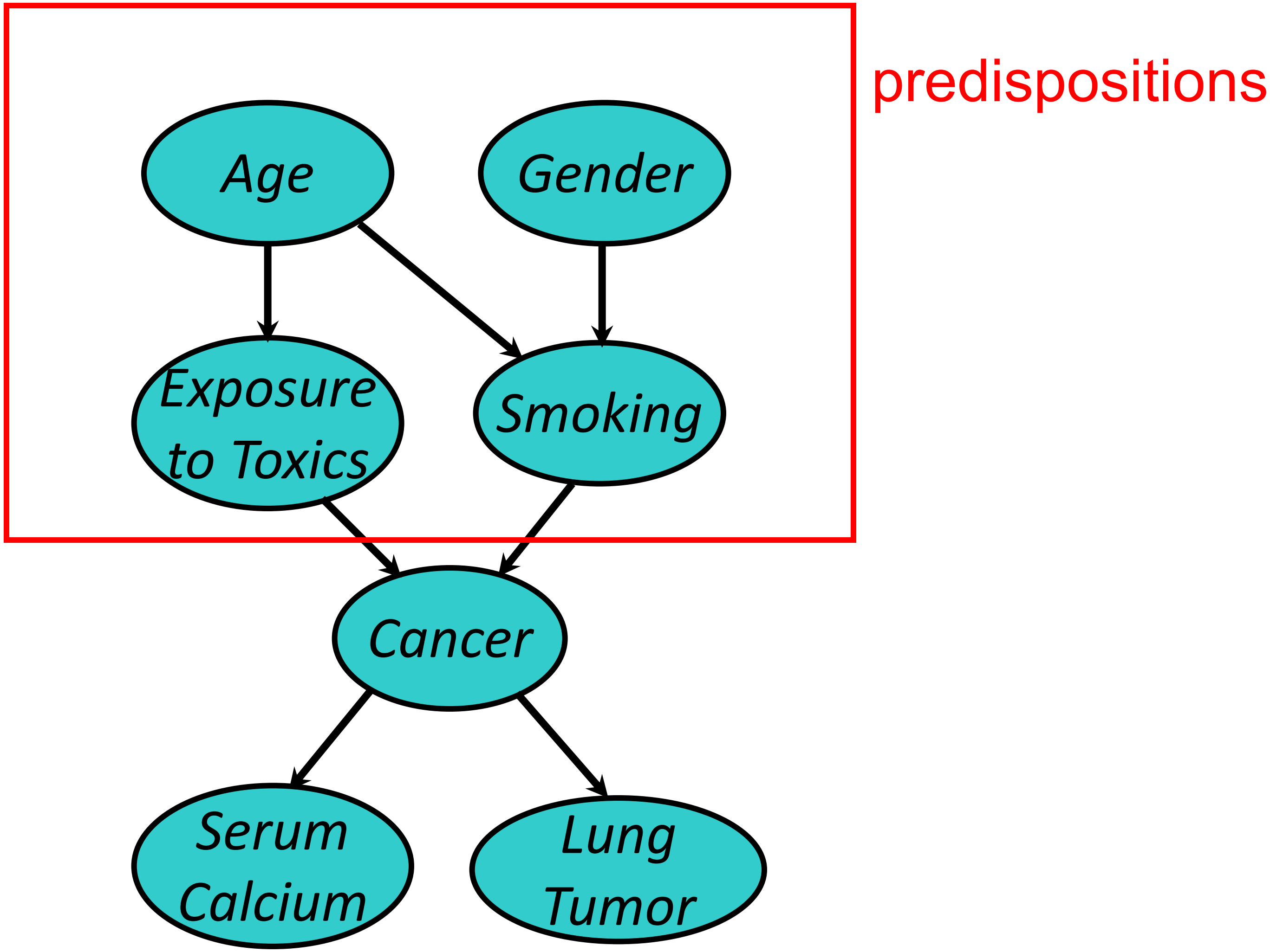


- Does gender ***cause*** smoking?
- ***Influence*** might be a better term
- In the US men are more likely to smoke

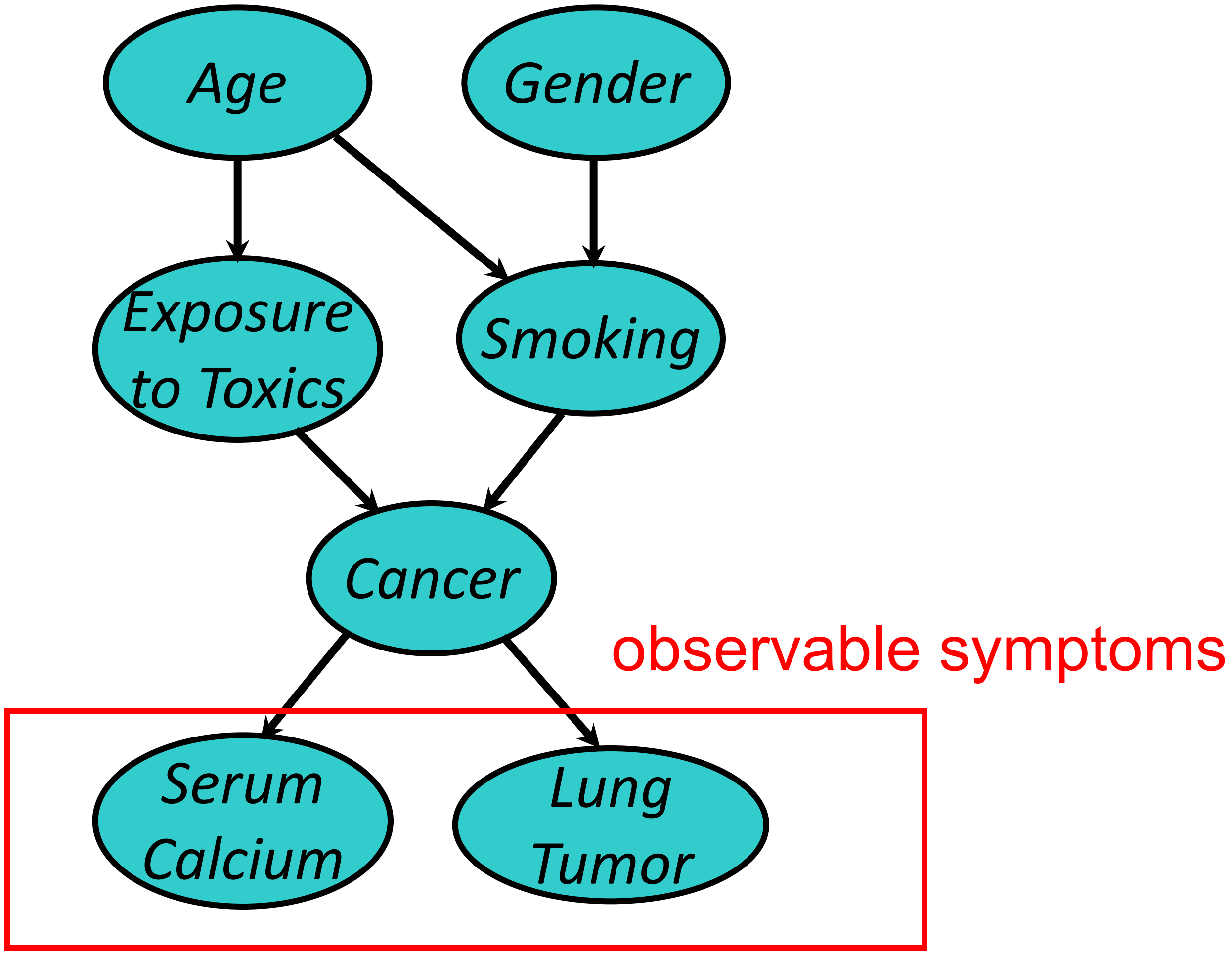
More Complex Bayesian Network



More Complex Bayesian Network

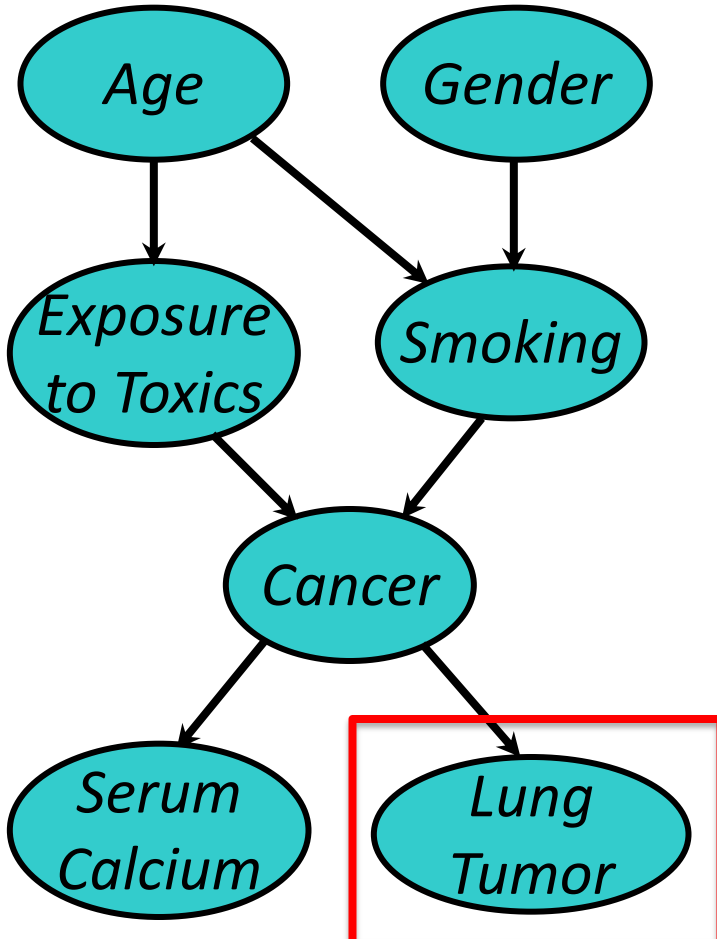


More Complex Bayesian Network



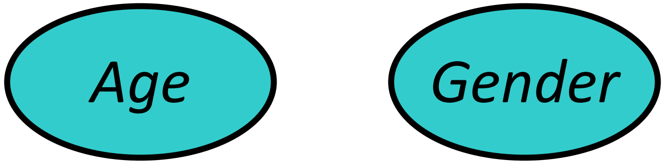
More Complex Bayesian Network

Can we predict likelihood of **lung tumor** given values of other 6 variables?



- Model has 7 variables
- Complete joint probability distribution has 7 dimensions!
- **Too much data required** ☹
- BBN simplifies: nodes have a CPT with data on itself & parents in graph

Independence



Age and *Gender* are independent*

No path between them in the graph

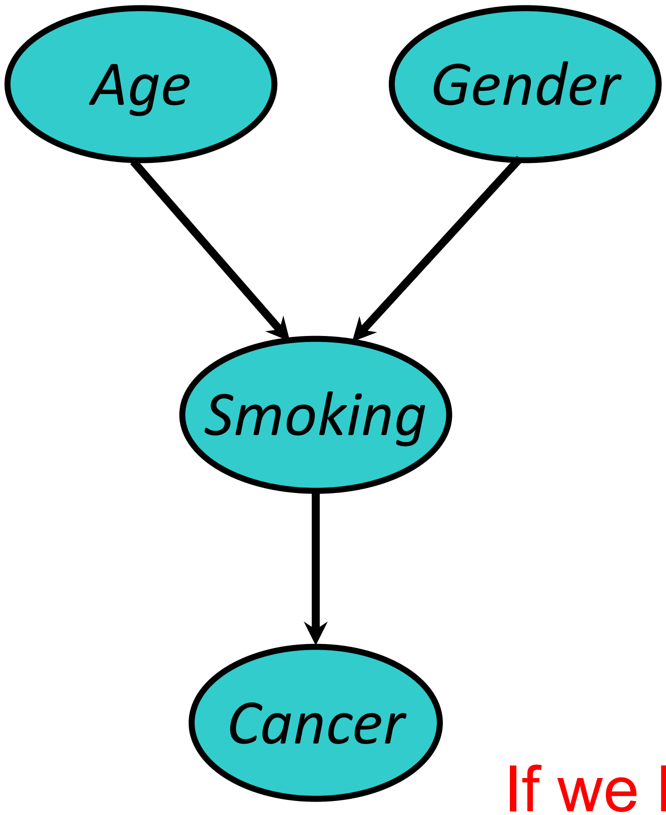
$$P(A,G) = P(G) * P(A)$$

$$P(A \mid G) = P(A)$$
$$P(G \mid A) = P(G)$$

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

* Not strictly true, but a good approximation

Conditional Independence

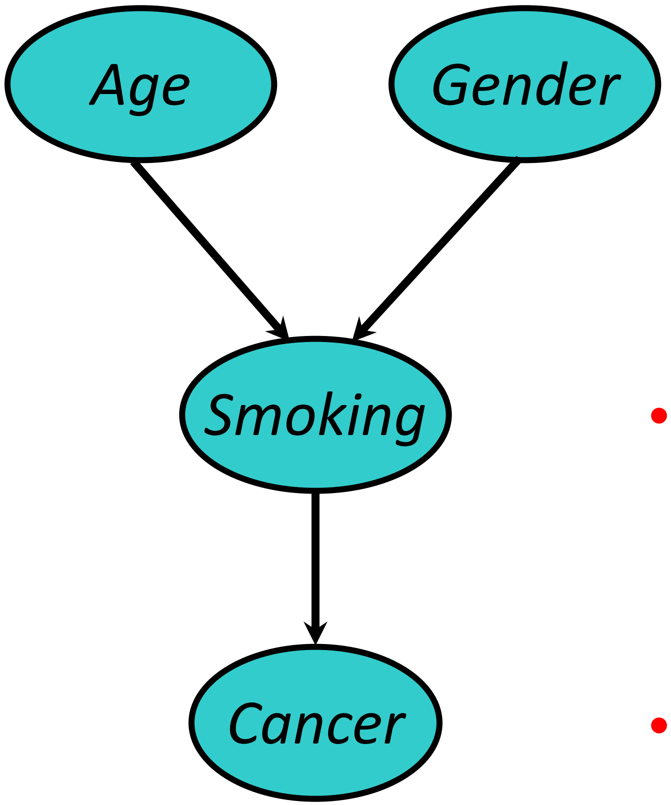


Cancer is independent of *Age* and *Gender* given *Smoking*

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

If we know value of smoking, there is no need to know values of age or gender

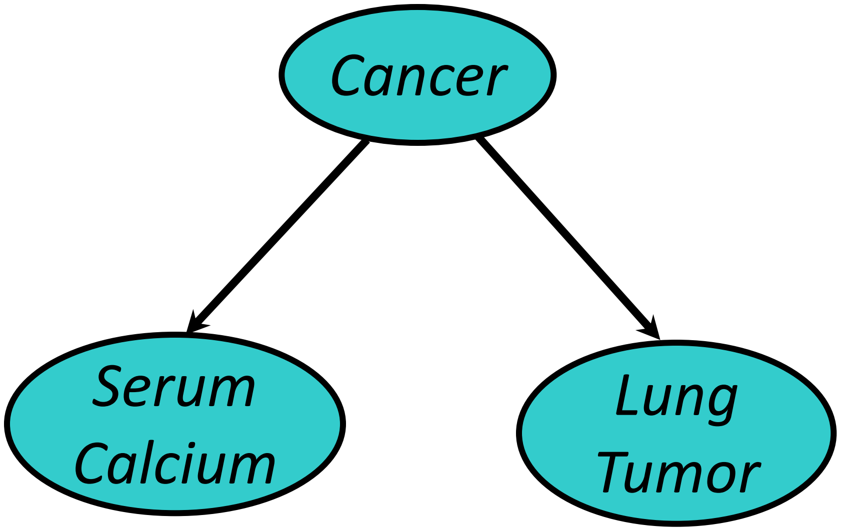
Conditional Independence



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

- Instead of one big CPT with 4 variables, we have two smaller CPTs with 3 and 2 variables
- If all variables binary: 12 models ($2^3 + 2^2$) rather than 16 (2^4)

Conditional Independence: Naïve Bayes



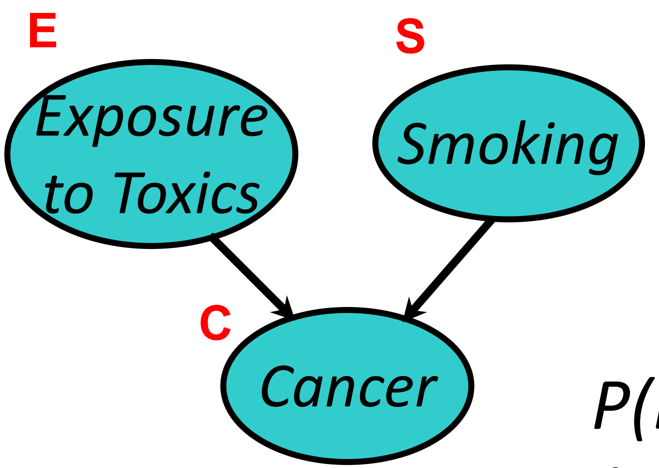
Serum Calcium and *Lung Tumor* are dependent (their presence is correlated)

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

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

Naïve Bayes assumption: evidence (e.g., symptoms) independent given disease; easy to combine evidence

Explaining Away



Exposure to Toxics and *Smoking* are independent

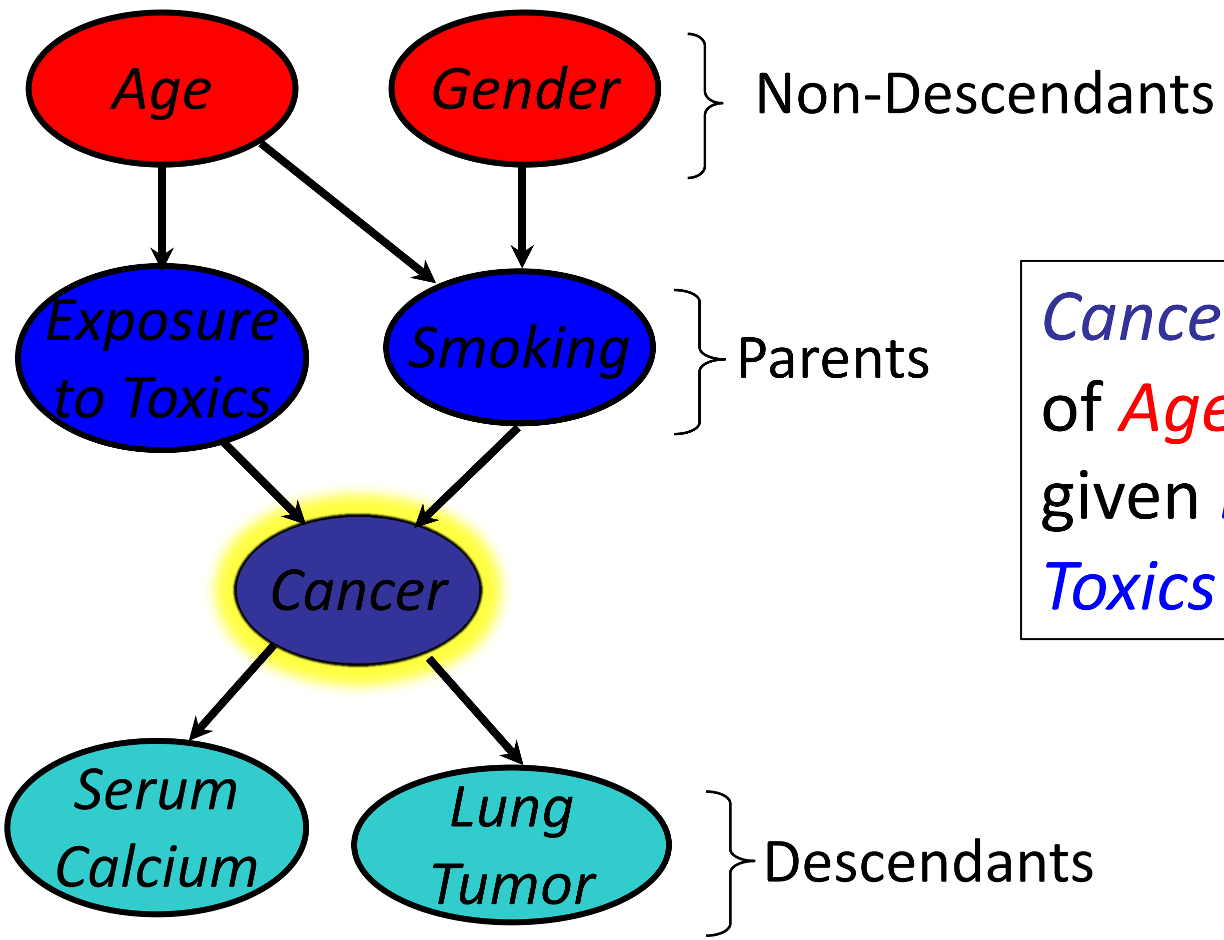
Exposure to Toxics is **dependent** on *Smoking*, given *Cancer*

P(E=heavy | C=malignant) > P(E=heavy | C=malignant, S=heavy)

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

Conditional Independence

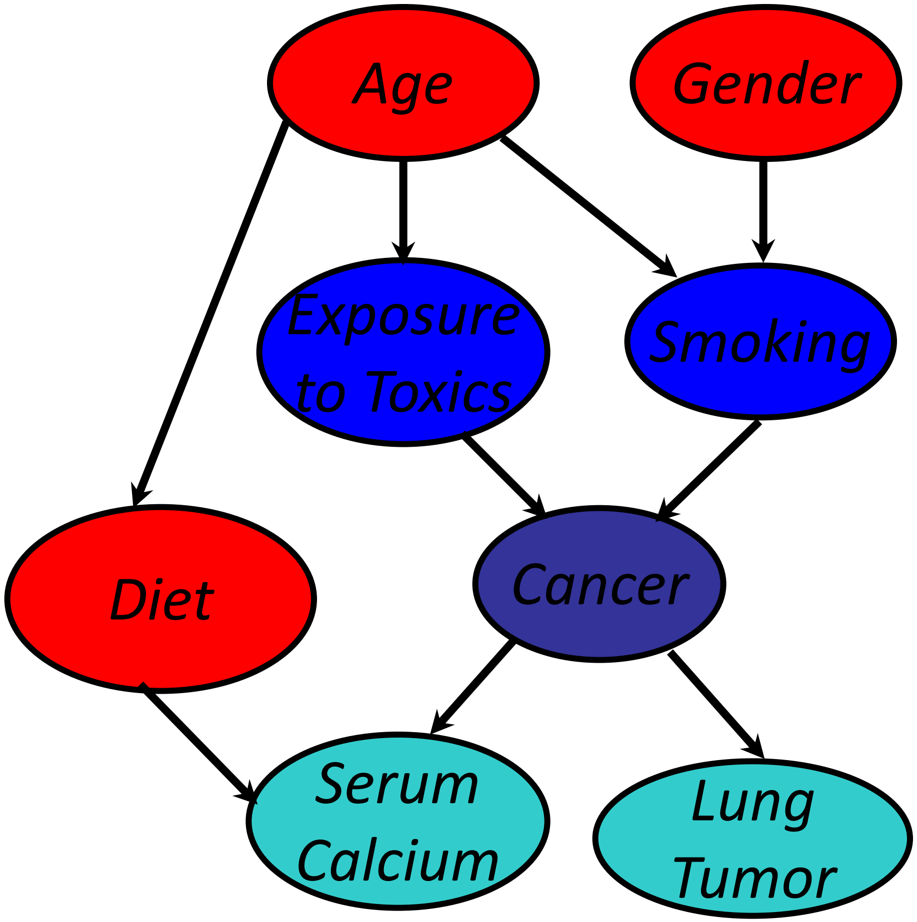
A variable (node) is conditionally independent of its non-descendants given its parents



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

The major benefit of the BBN model !

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**

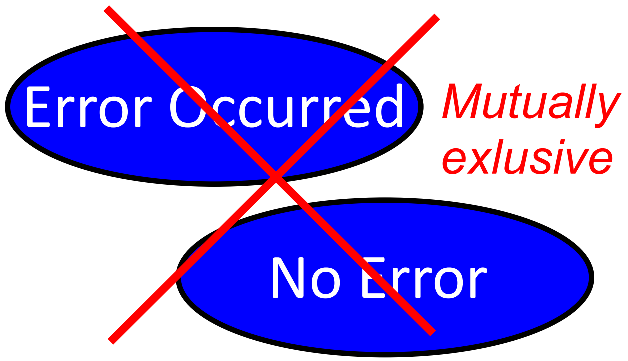
KA3: Obtaining the **conditional probability table data**

KA1: Choosing variables

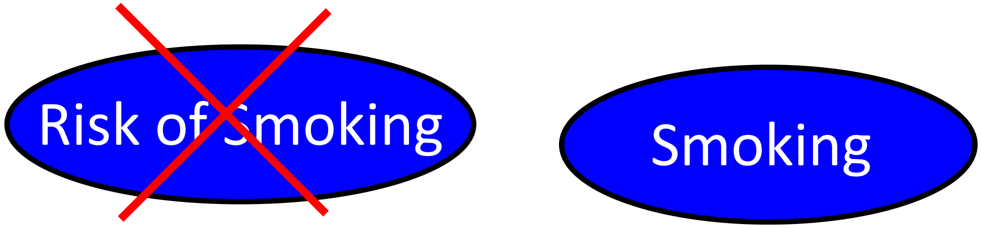
- Variable values: integers, reals or enumerations
- Variable should have collectively *exhaustive*, *mutually exclusive* values

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

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



- They should be values, not probabilities

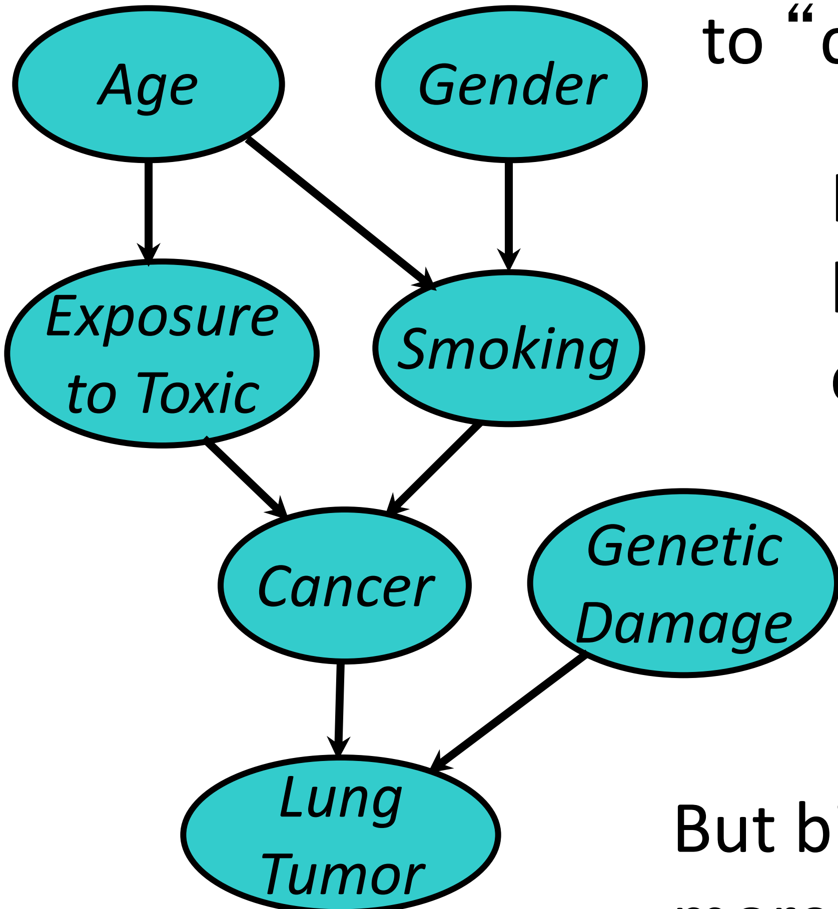


Heuristic: Knowable in Principle

Example of good variables

- Weather: {Sunny, Cloudy, Rain, Snow}
- Gasoline: $ per gallon {<2, 2-3, 3-4, >4}
- Temperature: { $\geq$ 100 F , < 100 F}
- User needs help on Excel Charts: {Yes, No}
- User's personality: {dominant, submissive}

KA2: Structuring



Network structure corresponding to "causality" is usually good.

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

May be better to add suspected links than to leave out

But bigger CPT tables mean more joint data must be collected

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**

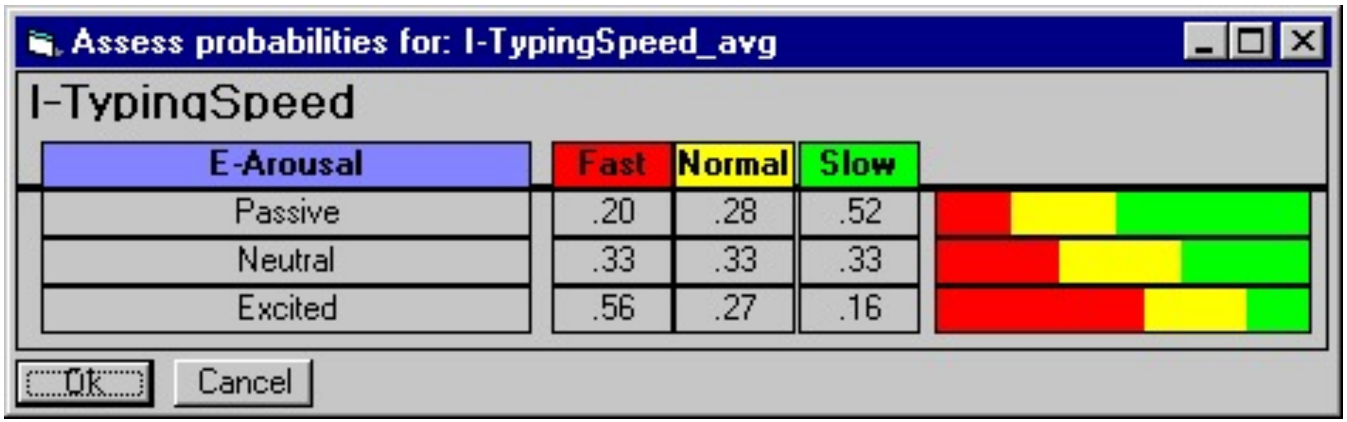
$$S \in \{no, light, heavy\}$$
$$C \in \{none, benign, malignant\}$$

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

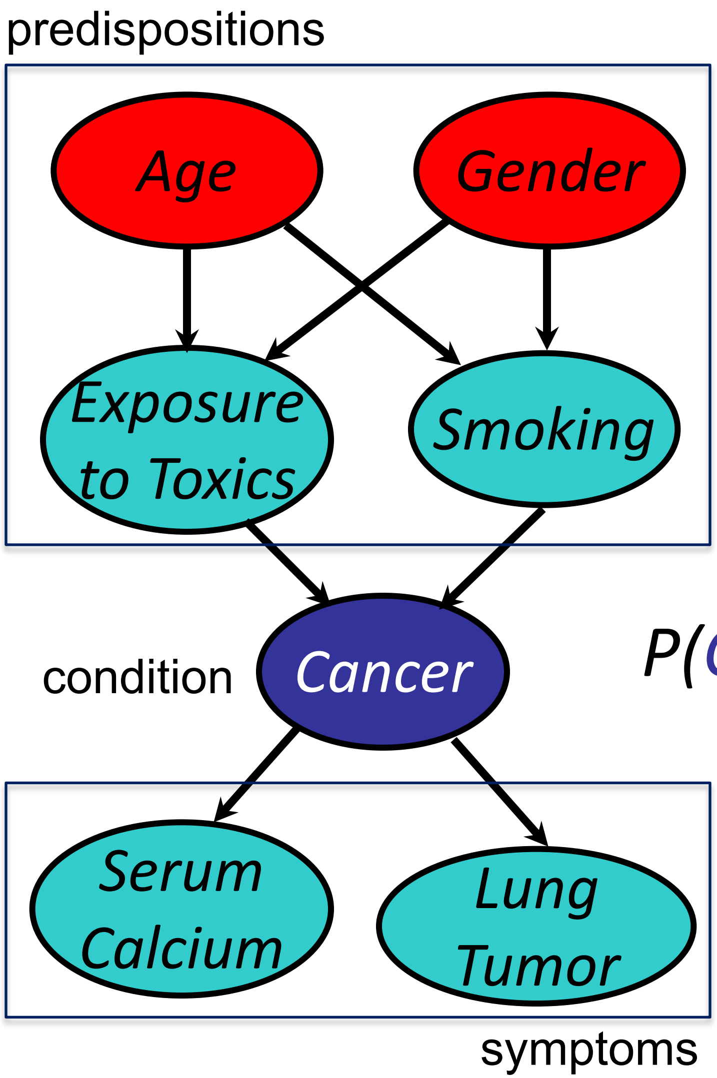
BBNs support three main kinds of reasoning:

- **Predicting** conditions given predispositions
  - "You are likely to get cancer since you are a heavy smoker"
- **Diagnosing** conditions given symptoms
  - "You're likely to have cancer given your high serum calicium level"
- **Explaining** a condition by predispositions
  - "Your cancer was probably caused by your exposure to lead"

To which we can add a fourth:

- **Deciding** on an action based on condition probabilities
  - "We should remove the lung tumor which might be cancerous"

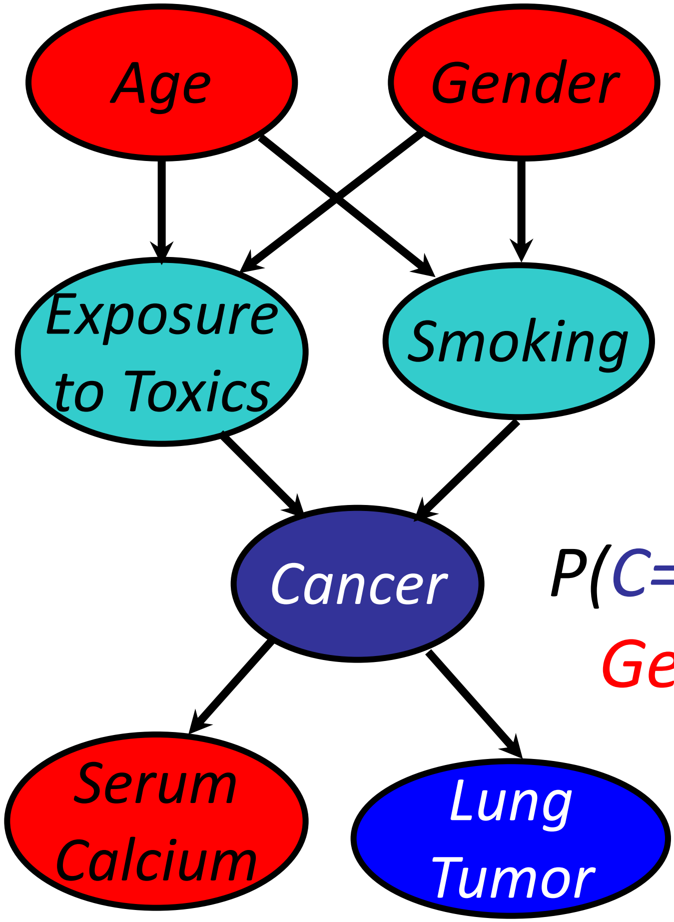
Predictive Inference



- From predispositions, predict condition
- How likely are elderly males to get malignant cancer?

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

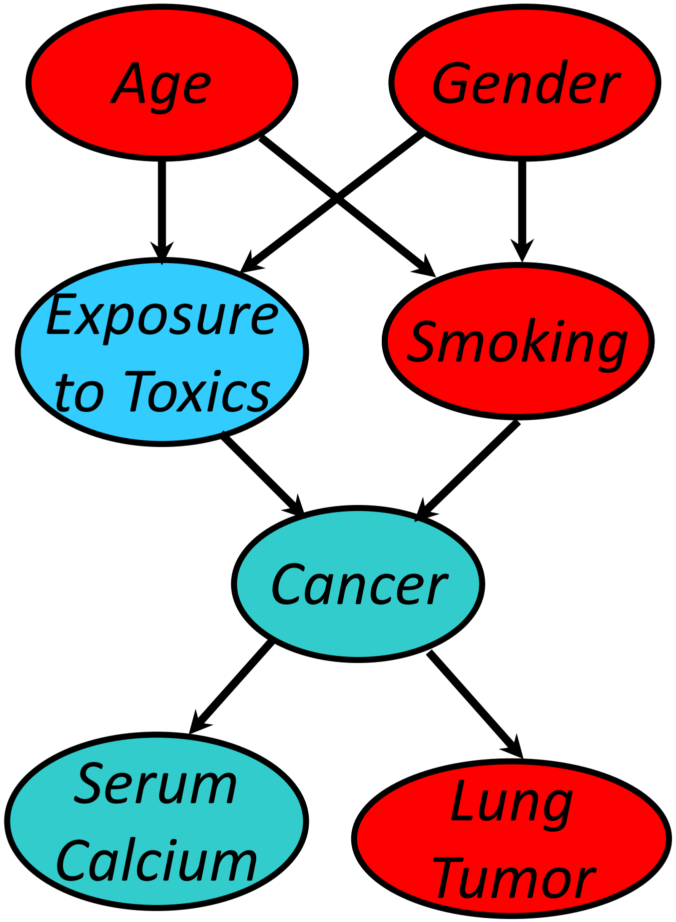
Predictive and diagnostic combined



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

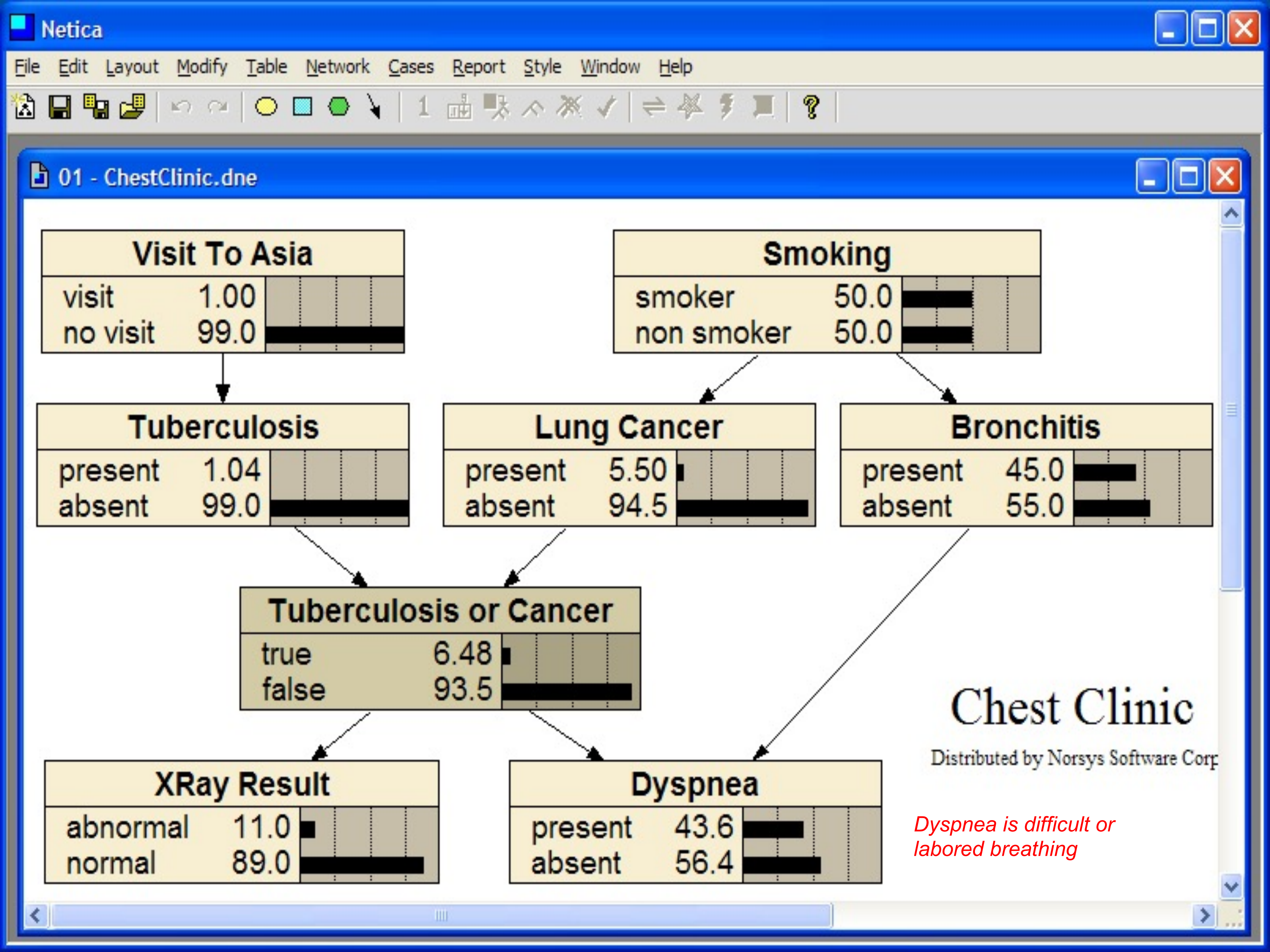


- If we see a lung tumor, the probability of heavy smoking and of exposure to toxics both go up

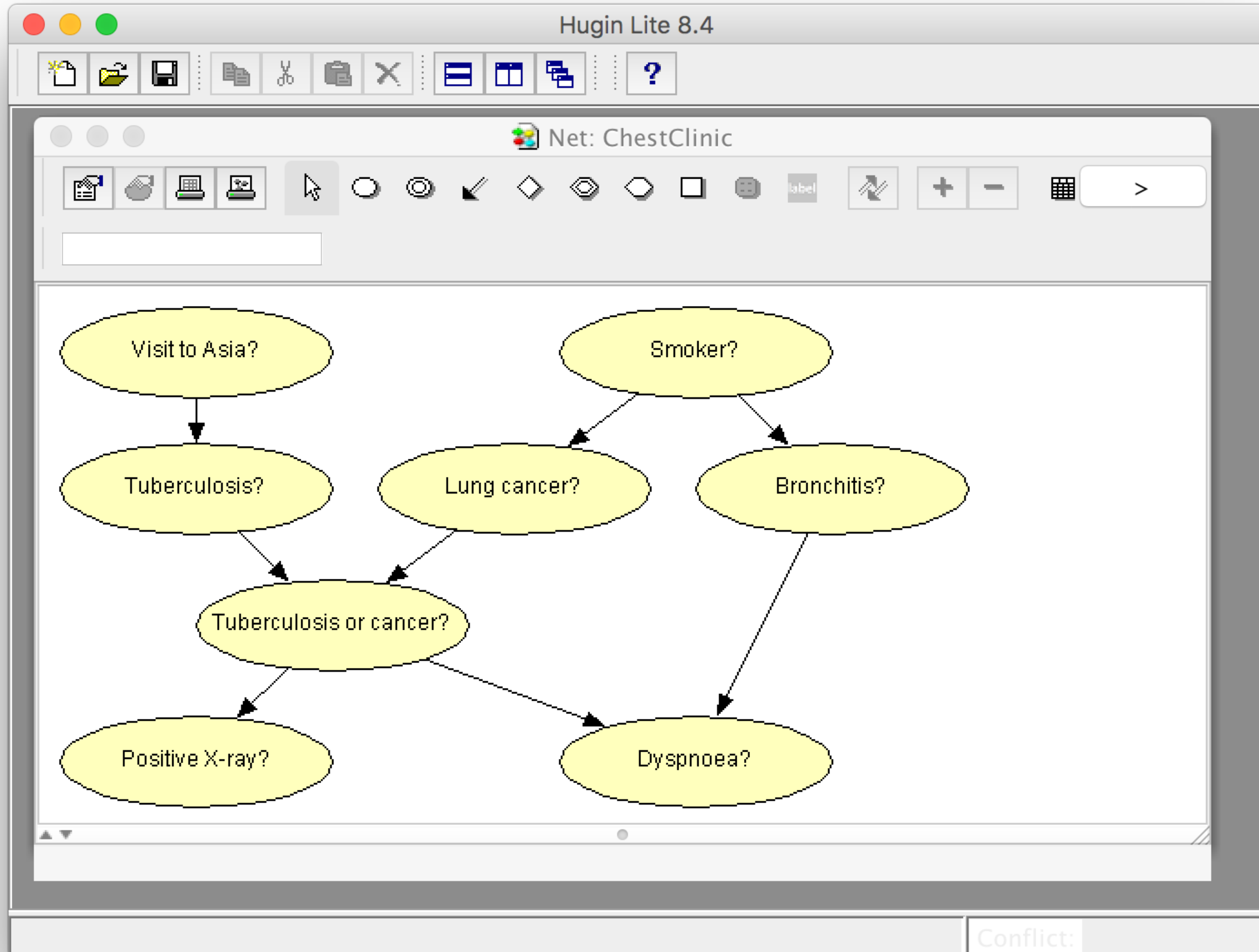
- If we then observe heavy smoking, the probability of exposure to toxics goes back down

Some software tools

- Netica: Windows app for working with Bayes-ian belief networks and influence diagrams
  - A commercial product, free for small networks
  - Includes graphical editor, compiler, inference engine, etc.
  - To run in OS X or Linus you need Crossover
- Hugin: free demo versions for Linux, Mac, and Windows are available
- Various Python packages
- Aima-python code in probability4e.py


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
Dyspnea is difficult or labored breathing

Same BBN model in Hugin app

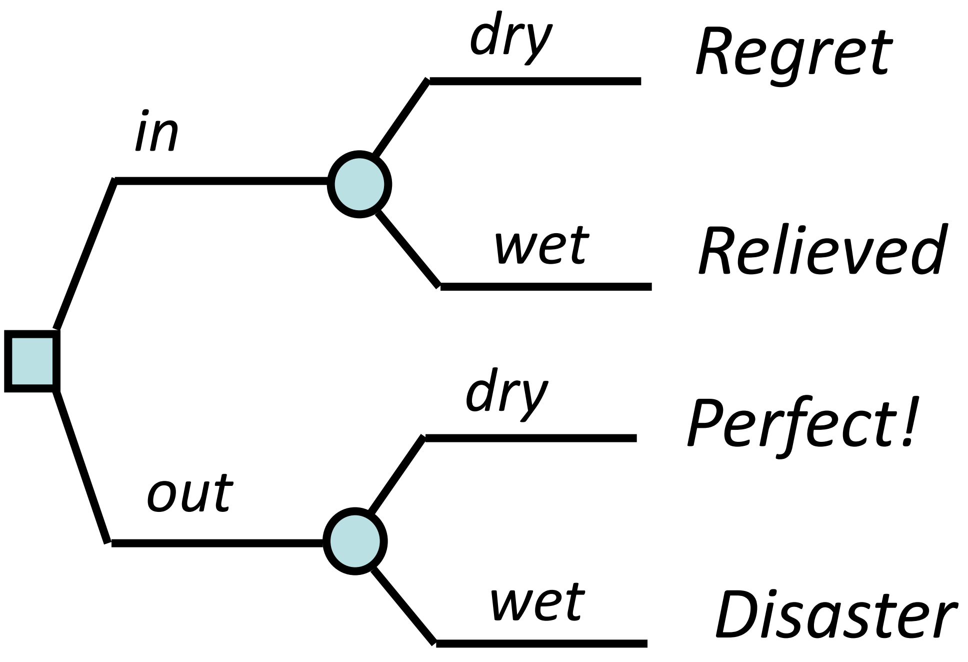


Decision making

- 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

Decision Problem

Should I have my party inside or outside?



Value Function

A numerical score over all possible states allows a BBN to be used to make decisions

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

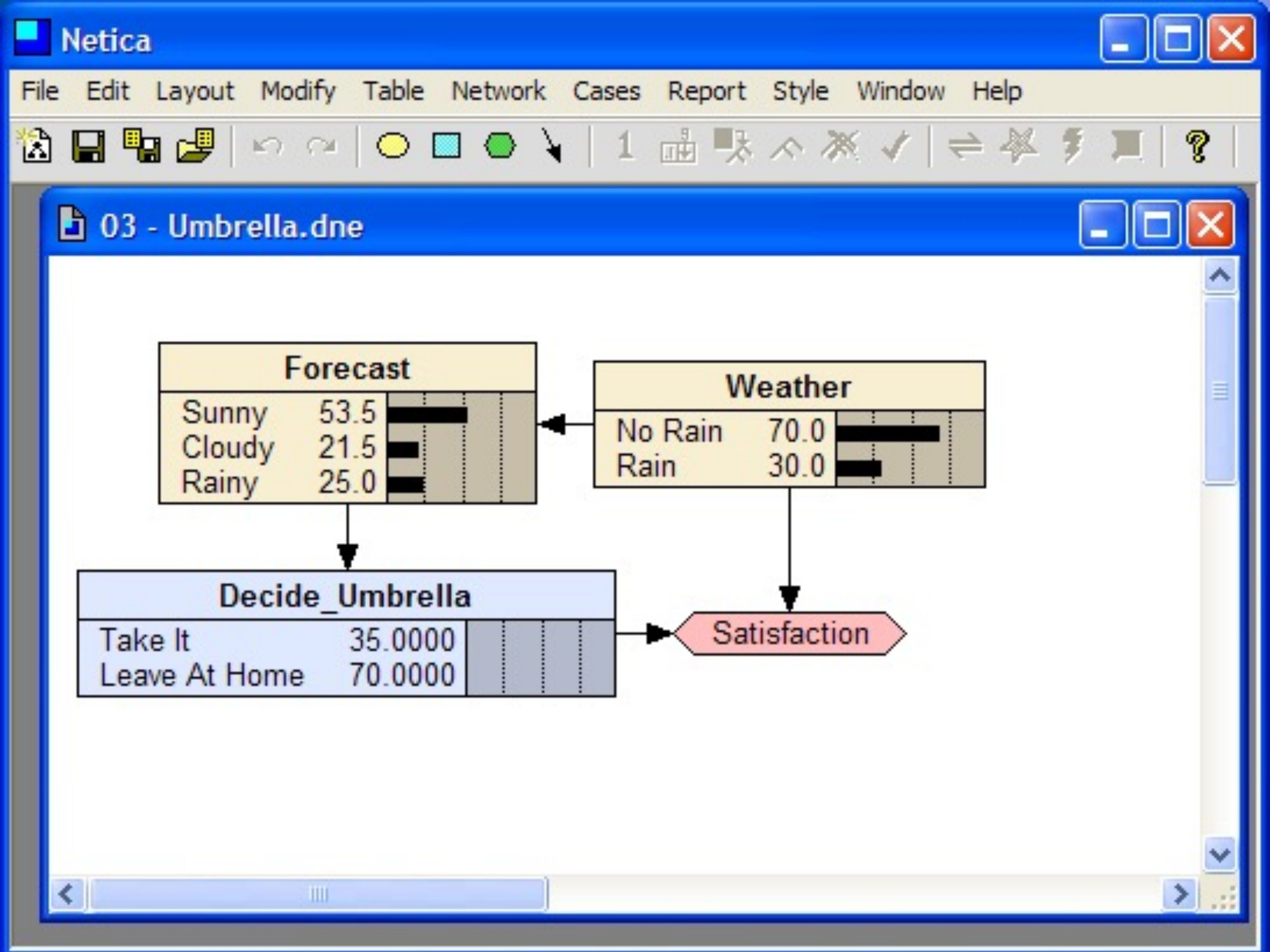
Using $ for the value helps our intuition

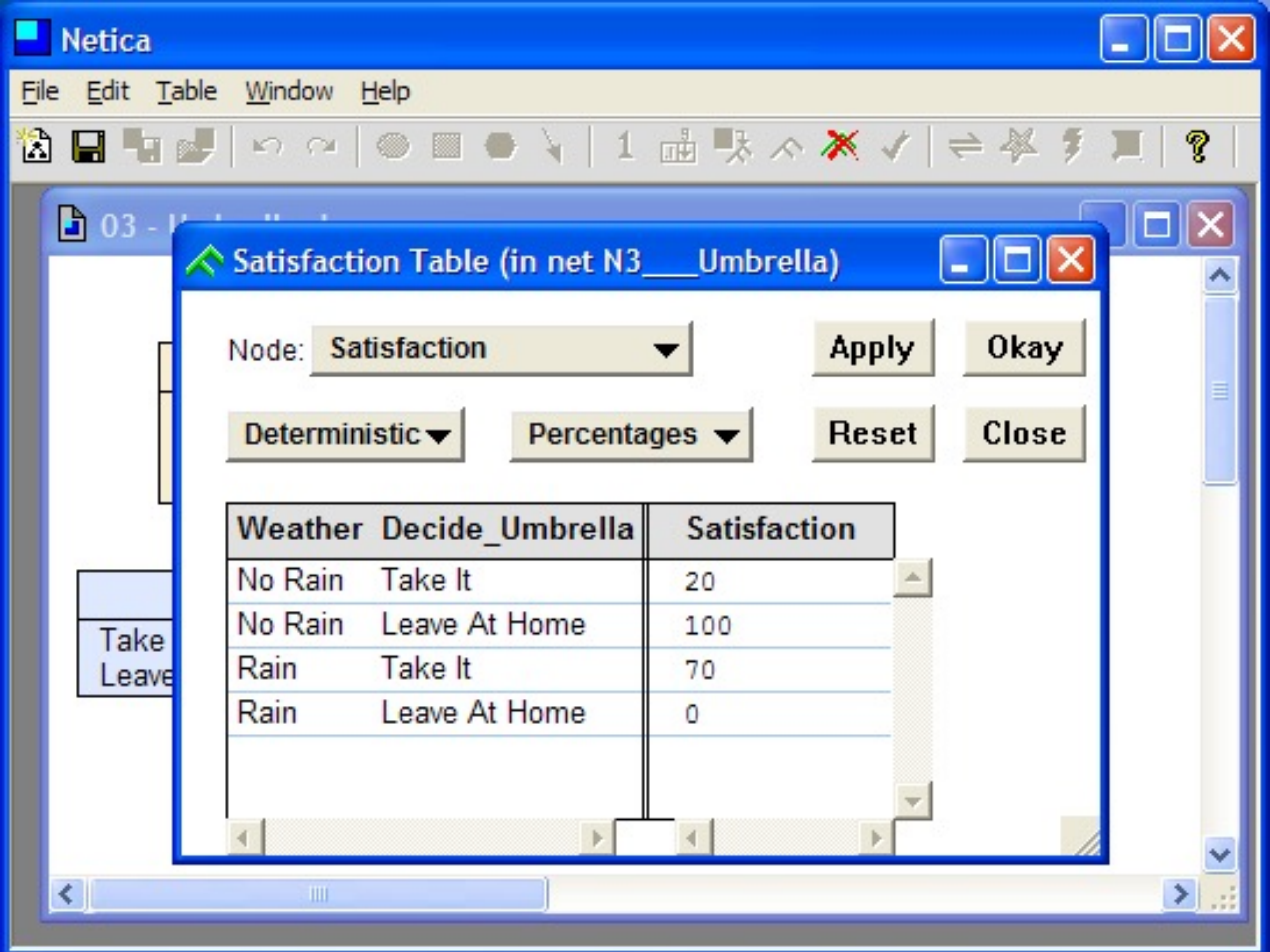
Decision Making with BBNs

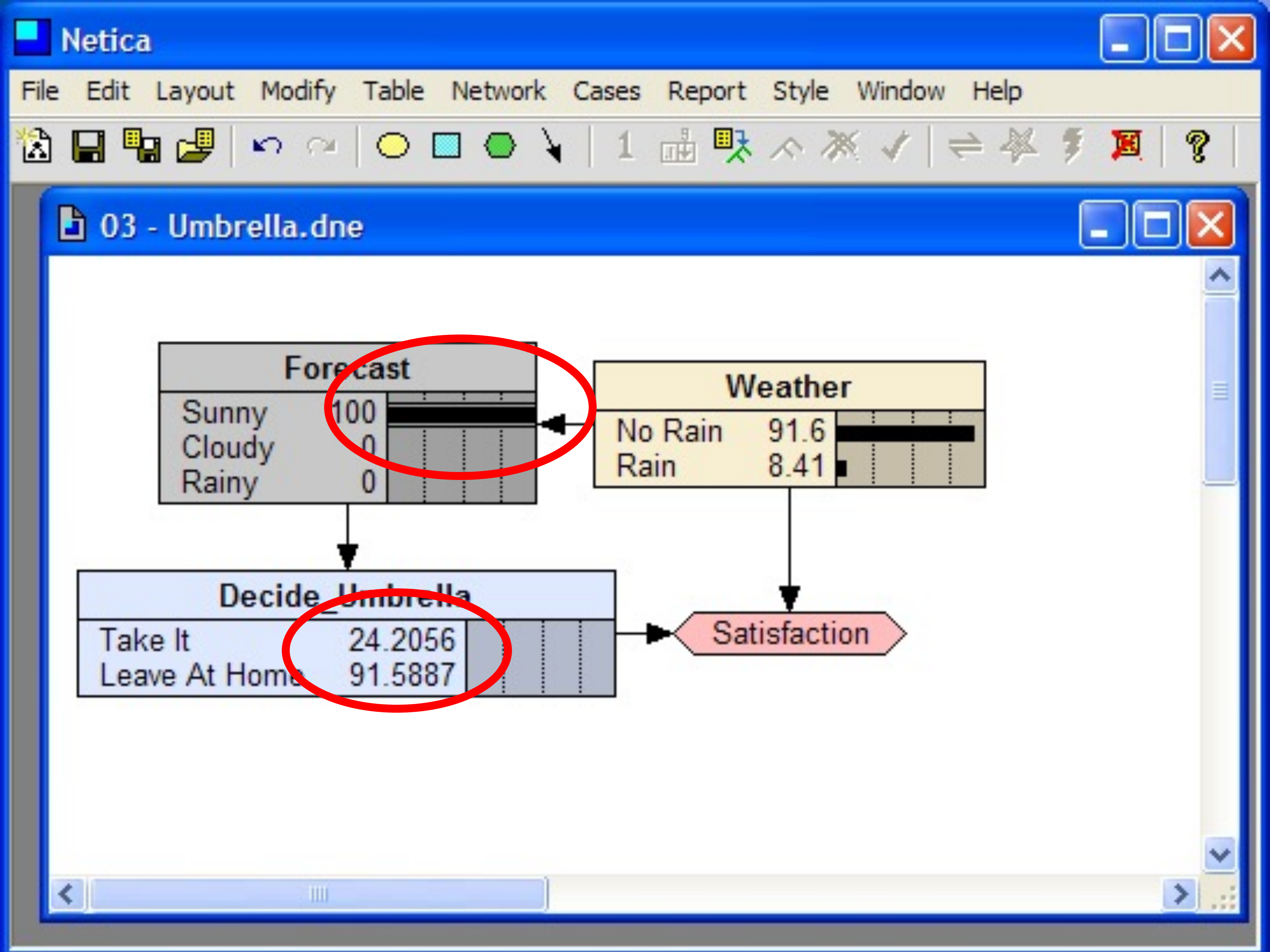
- Today's weather forecast might be either sunny, cloudy or rainy
- Should you take an umbrella when you leave?
- Your decision depends only on the forecast
  - Forecast "depends on" the actual weather
- Your satisfaction depends on your decision and the weather
  - Assign utility measure to each of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

Decision Making with BBNs

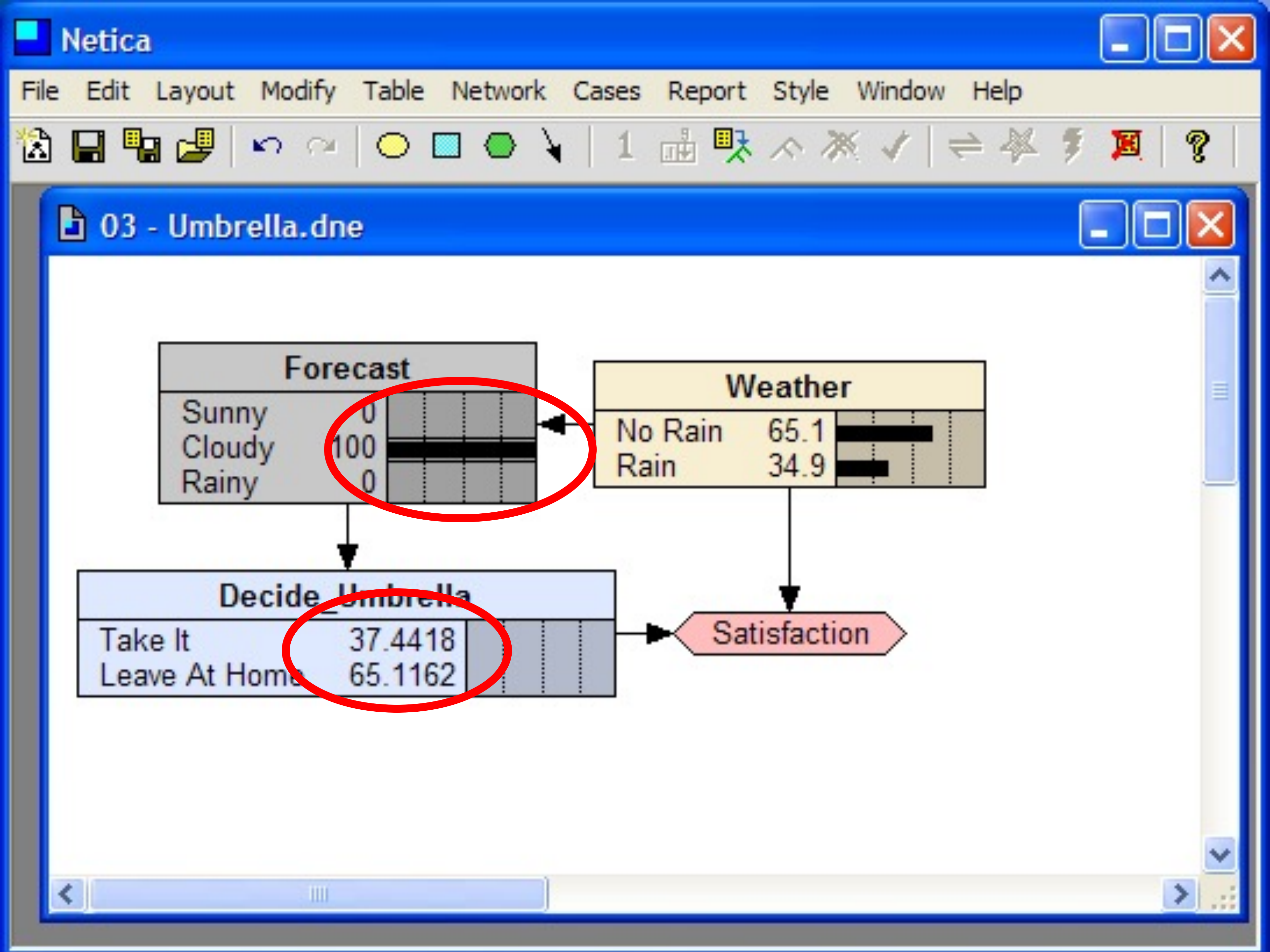
- Extend BBN framework to include two new kinds of nodes: **decision** and **utility**
- **Decision** node computes expected utility of a decision given its parent(s) (e.g., forecast) and a valuation
- **Utility** node computes utility value given its parents, e.g., a decision and weather
  - Assign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

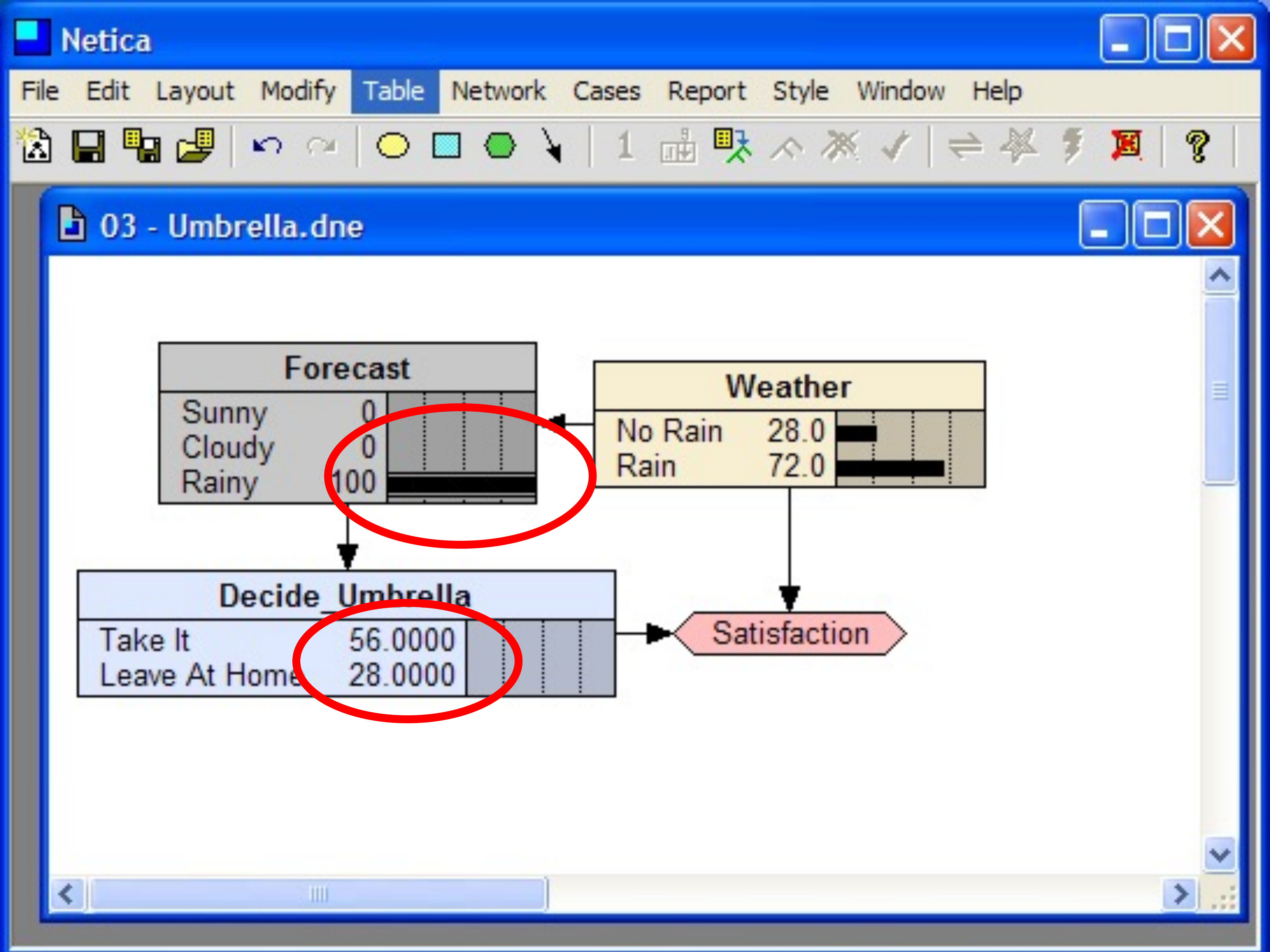

Netica
File Edit Layout Modify Table Network Cases Report Style Window Help
03 - Umbrella.dne
Forecast
Sunny 53.5
Cloudy 21.5
Rainy 25.0
Weather
No Rain 70.0
Rain 30.0
Decide_Umbrella
Take It 35.0000
Leave At Home 70.0000
Satisfaction


Netica
File
Edit
Table
Window
Help
03 -
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