WARWICK
THE UNIVERSITY OF WARWICK

Bayesian Networks and decisions

June 2020

Contents

- Bayesian networks;
- Utility computation and policy comparison;
- Impact of evidence and queries.

Decision under uncertainty

- Policy-makers often need to make evidence-based decision about complex systems.
- This is even more challenging when multiple factors provide influence in combination.
- One challenge is how to combine the information provided by all factors and how use this information to evaluate candidate policies.

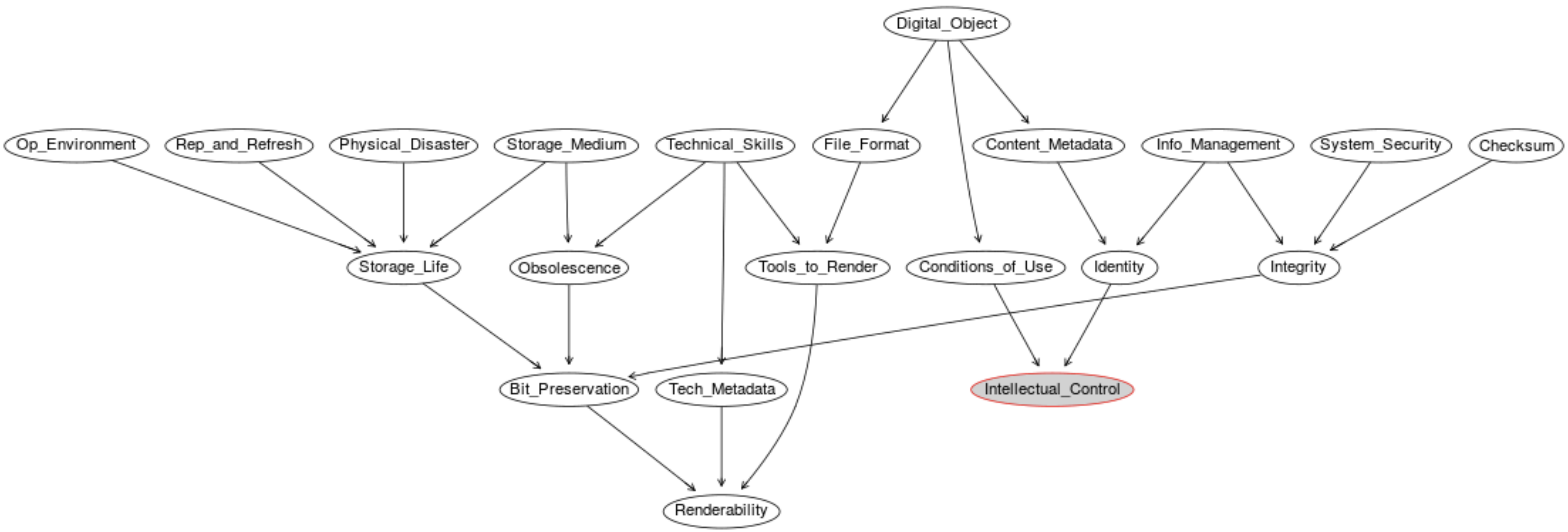
Decision under uncertainty

- To estimate the potential effectiveness of various policies, it is necessary to quantify the effect of various combinations of factors on the system of interest.
- For this purpose, Policy-makers rely on experimental data (when it is available) and expert knowledge.
- In this context, Bayesian Networks offer a useful approach designed to accommodate uncertainties and factors affecting the system into a joint probability model using both expert and data information.

Bayesian networks

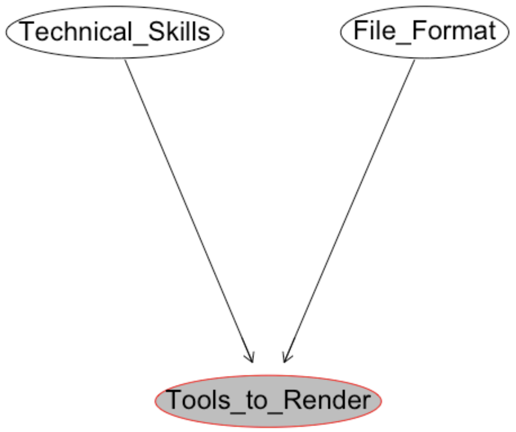
- Bayesian networks provide an intuitive framework to model the dependencies between the variables in a multivariate system.
- The main goal is to decompose a potentially complex system into simpler subsystems.
- Furthermore a causal network can be constructed based on expert knowledge about the problem;
- And structured expert judgement can be considered when data is missing or not available.

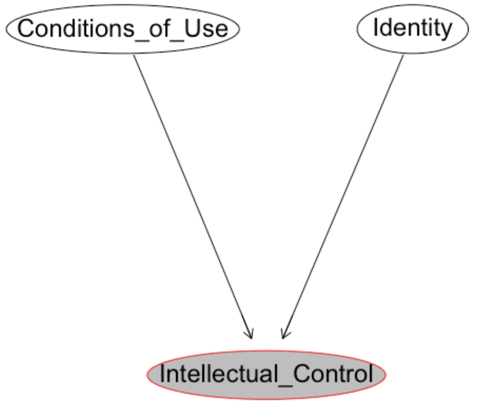
Digital preservation network



Why represent this problem as a Network?

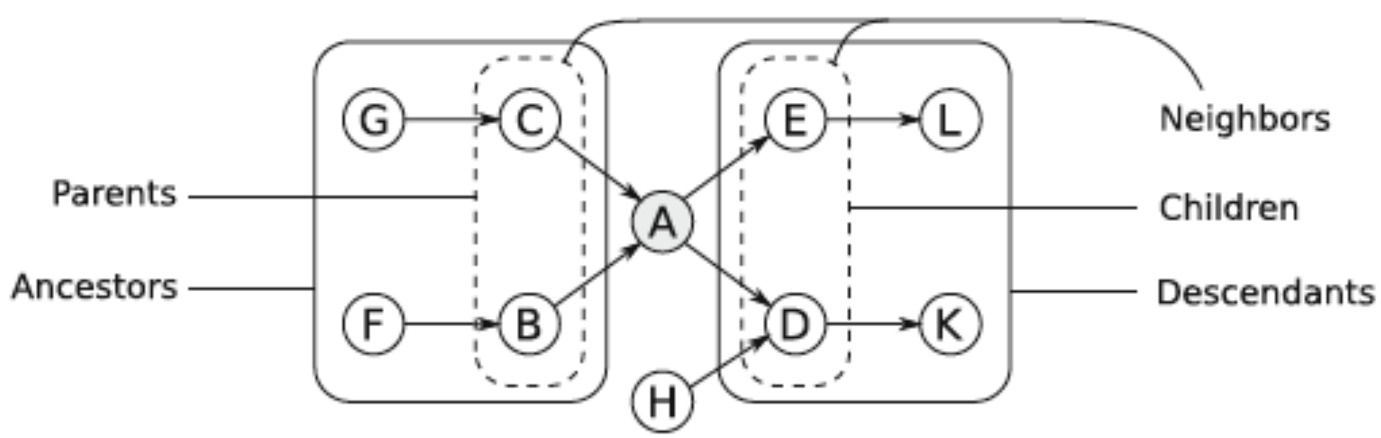
- A network is used here as a natural way to capture the notion of elements in a system (digital preservation) and their inter-relations.
- For instance, tools to render depend directly on technical skill and file format, but do not depend directly on Storage medium.
- Instead of assuming that Intellectual control depends on all the variables, we assume that it depends directly on conditions of use and identity.





Structure of a BN

- G is an ancestor of A and L is a descendant of A, as there is a path leading from G to A and from A to L.



- E is a child of A, as the path from A to E is composed by a single arc. And C is a parent of A.

Source: Bayesian Networks in R

Basic elements of a BN

- To account for the uncertain events in the system, we consider probability tables.
- Bayesian network (Pearl, 1988) represents these probability for all the variables jointly.
- It is defined by two basic elements:

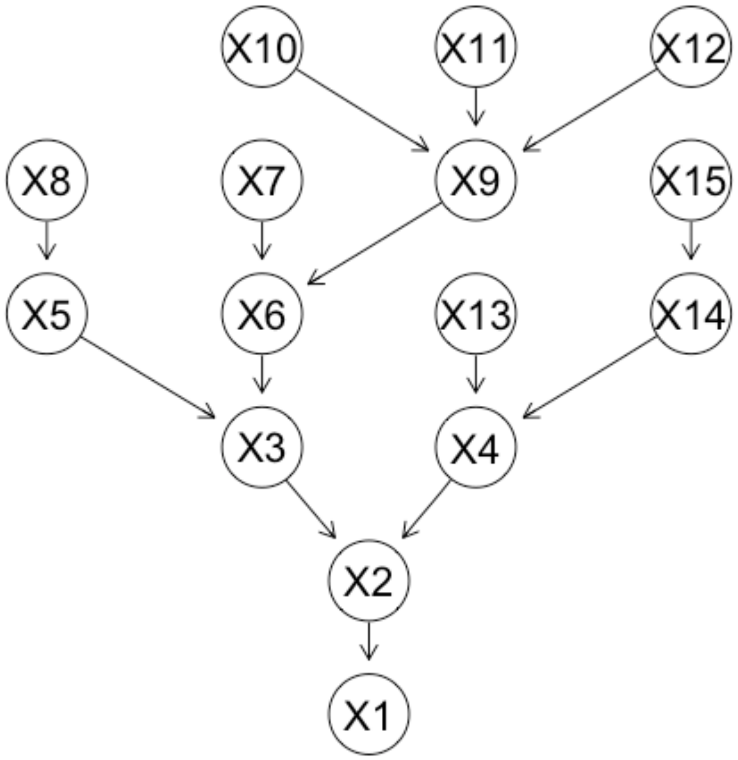
A graph G with each node corresponding to a variable

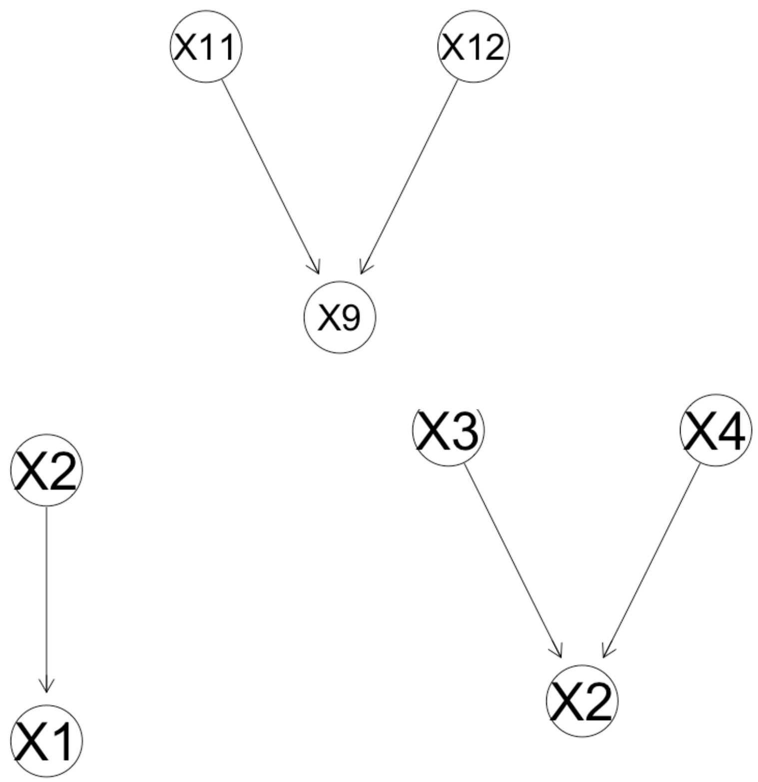
+

A set of local conditional distributions.

Divide to conquer: Local conditional distributions

- We consider the idea of conditional probabilities to divide a large multivariate problem in smaller ones based on conditional independence.





Steps to compare several policies

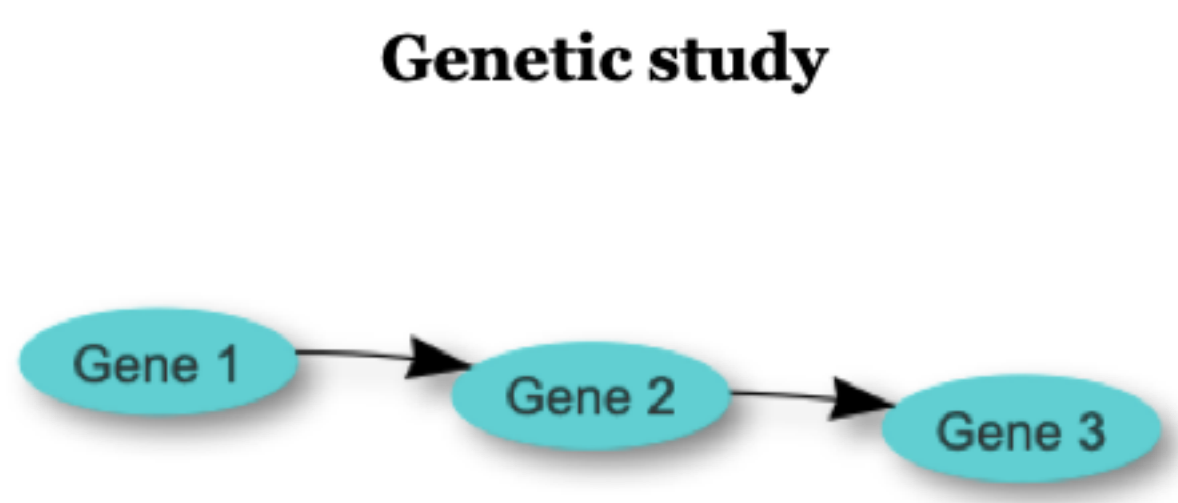
- Construct the network structure (define all variables and connections);
- Elicit or estimate the conditional probability tables;
- Estimate the marginal probabilities;
- Obtain the expected utility for each policy;
- Compare utilities which will be available to aid the decision maker.

Decision under uncertainty

- In the context of decision under uncertainty, we need a method to combine uncertain statements and knowledge about states of “the world”.
- Probability theory is the prevailing method for dealing with uncertainty, and it is the one in focus here.
- Having combined all the available information, expected utilities can be provided to support decision making under uncertainty.

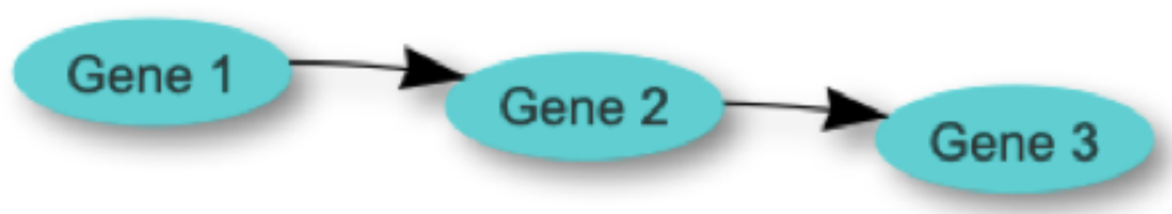
Example

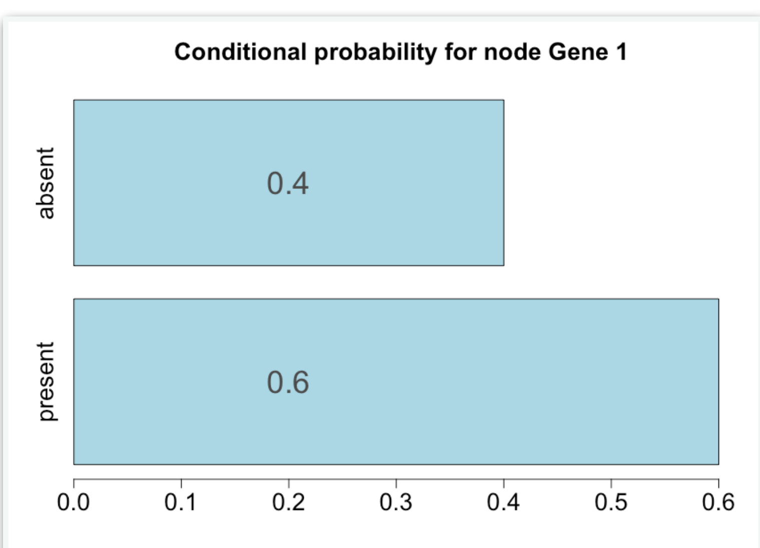
- Consider a genetic study such that 3 genes of interest are either present or absent.

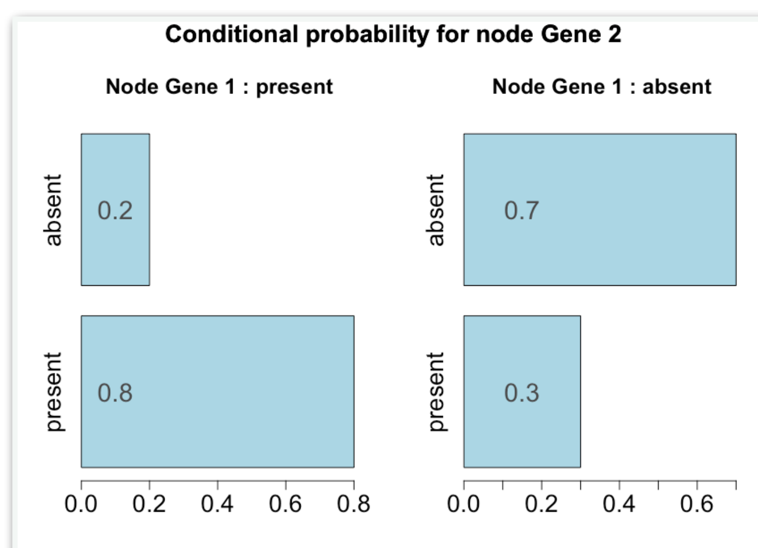


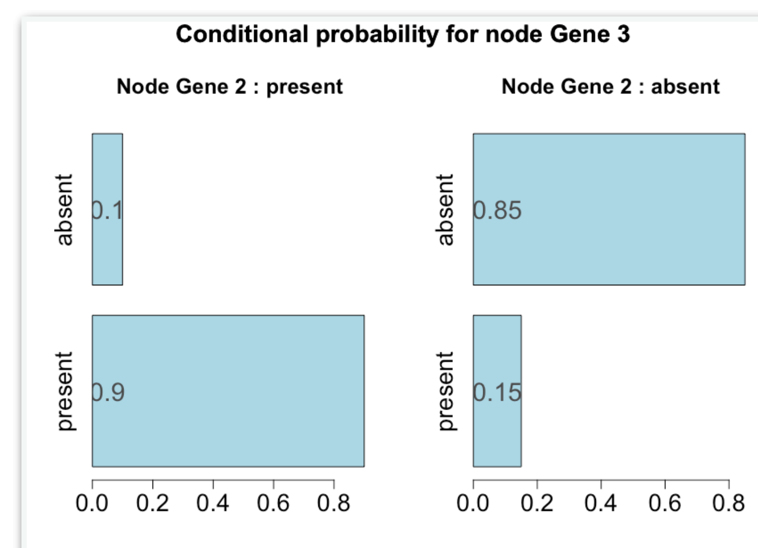
Local probabilities

Genetic study

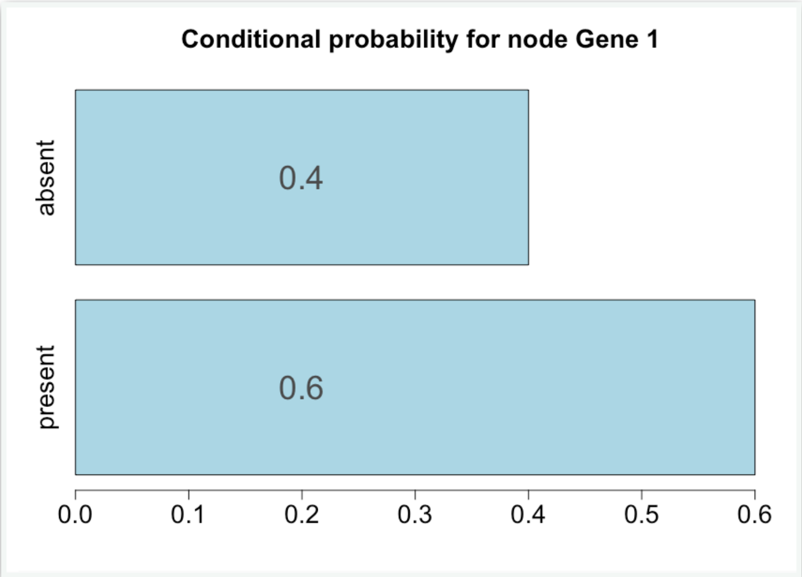


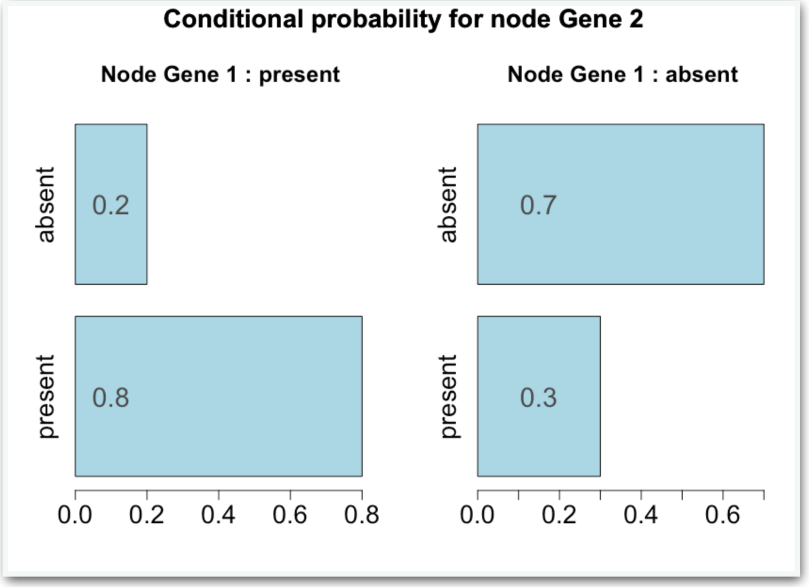


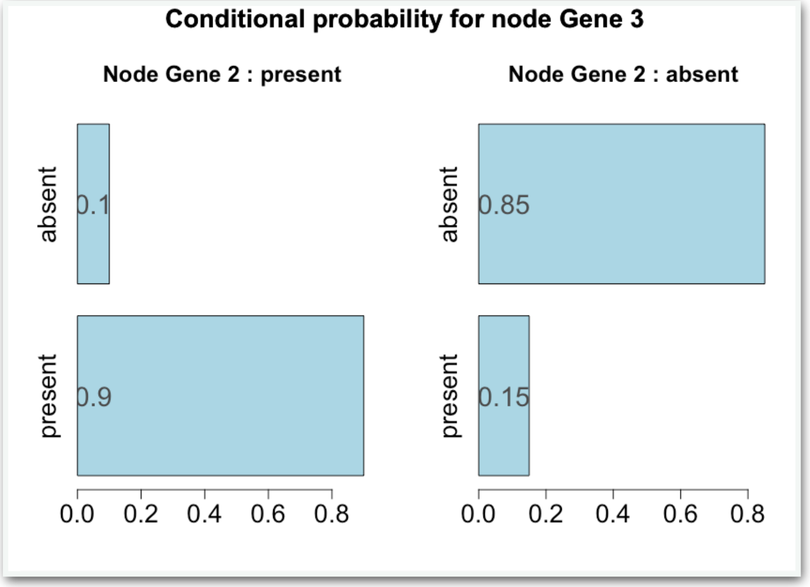




Computing probabilities







$$\begin{aligned} P(Gene\ 3 = Present) &= P(Gene\ 3 = Present \mid Gene\ 2 = Present)P(Gene\ 2 = Present) \\ &+ P(Gene\ 3 = Present \mid Gene\ 2 = Absent)P(Gene\ 2 = Absent) = 0.6 \end{aligned}$$

Utility computation

- Suppose that there are two policies available:
  - Policy 1: surgery with utility 3 if gene 3 is present and utility 1 if gene 3 is absent;
  - Policy 2: take a medicine with utility 1 if gene 3 is present and utility 4 if gene 3 is absent;

	Gene 3 is present	Gene 3 is absent
Policy 1 (Surgery)	3	1
Policy 2 (Medicine)	0.5	4

If I knew that gene 3 was present, I would recommend surgery.

- What is the best decision?

Maximizing utilities

- The preferences that are relevant in a decision scenario must be expressed on a numerical scale through utilities of each decision.

$$P \succ Q <=> u(P) \geq u(Q).$$

- Recommendations are based on the principle of maximal expected utility.
- As the reasoning performed by a probabilistic network is normative, the outcome will provide a recommended course of action that maximizes the expected utility.

Note that a potentially complex decision is replaced by the comparison of real numbers.

Utility computation

- In our example:

	Gene 3 is present	Gene 3 is absent
Policy 1 (Surgery)	3	1
Policy 2 (Medicine)	0.5	4

- Then, P(gene3=present)=0.6 can be computed from the probability tables and <u>expected utilities</u> can be obtained.
  - E[U(surgery)] = (0.6 x 2) + (0.4 x 1) = 1.6
  - E[U(medicine)] = (0.6 x 0.5) + (0.4 x 4) = 1.9
- The best decision is to prescribe the medicine. Note that if P(gene3=present)=0.7 the best decision would be surgery.

Interventions

- In case we have found a causal network which represents well the system of interest then we can compute the effects of interventions.
- In this context, we will work with experimental settings, changing a configuration and seeing what happens (what if…).
- Evidence sensitivity analysis is the analysis of how sensitive the results of a belief update (propagation of evidence) is to variations in the set of evidence.

Interventions and decisions

- Investigating the impact of different subsets of the evidence on states of the hypothesis may help to determine subsets of the evidence acting in favor of or against each possible hypothesis.
- The sensitivity analysis could motivate changes in the system aiming its improvement.

Hard x Soft evidence

- Hard evidence: fixing the values of one or more variables in the network.
- Soft evidence: a new probability table for one or more variables in the network.

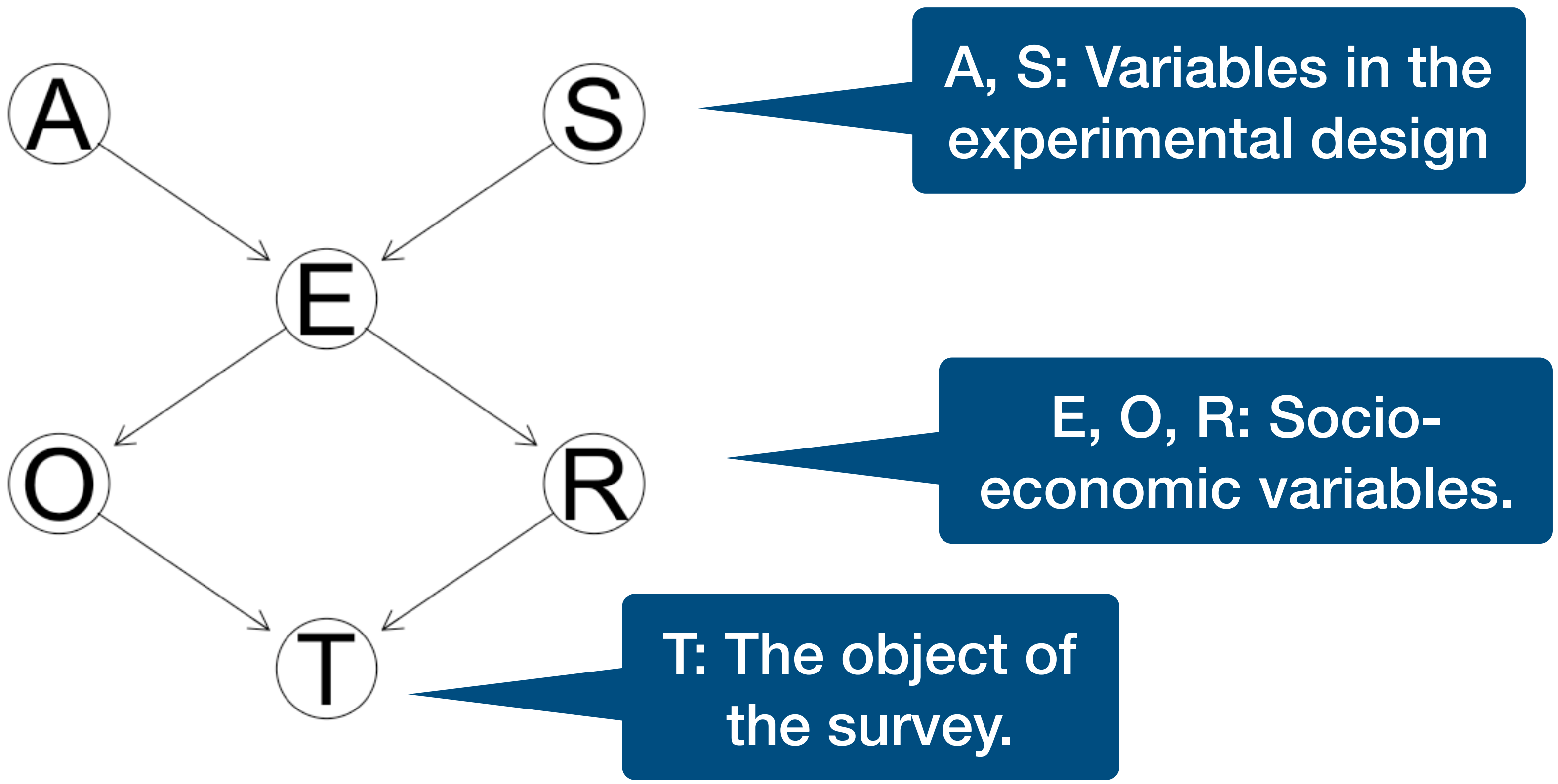
Transportation system example

- This data set that contains information about usage of different transportation systems (Travel) with a focus on cars and trains. It includes:
  - **Age** (A): It is recorded as young (<=30), adult (>30, <=60), and old (>60).
  - **Sex** (S): The biological sex (male - M or female - F).
  - **Education** (E): up to high school or university degree.
  - **Occupation** (O): Employee (emp) or a self employed (self) worker.
  - **Residence** (R): The size of the city the individual lives in (small or big).
  - **Travel** (T): The means of transport (car, train or other).

Probability distribution

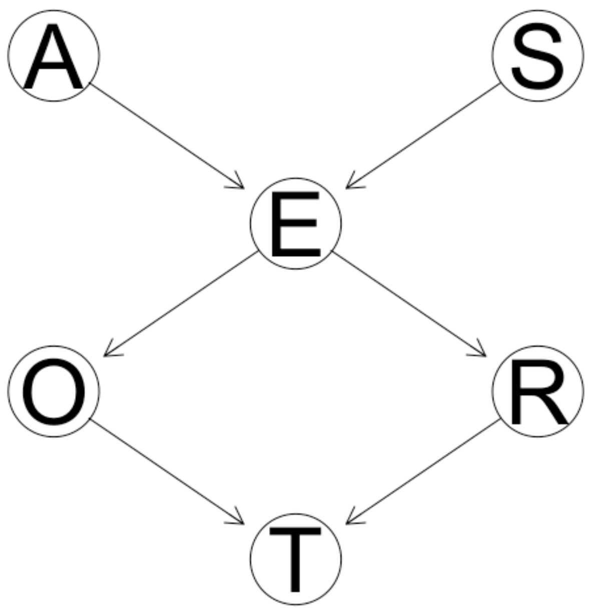
- Given the DAG, the joint probability distribution of the survey data variables factorises as:

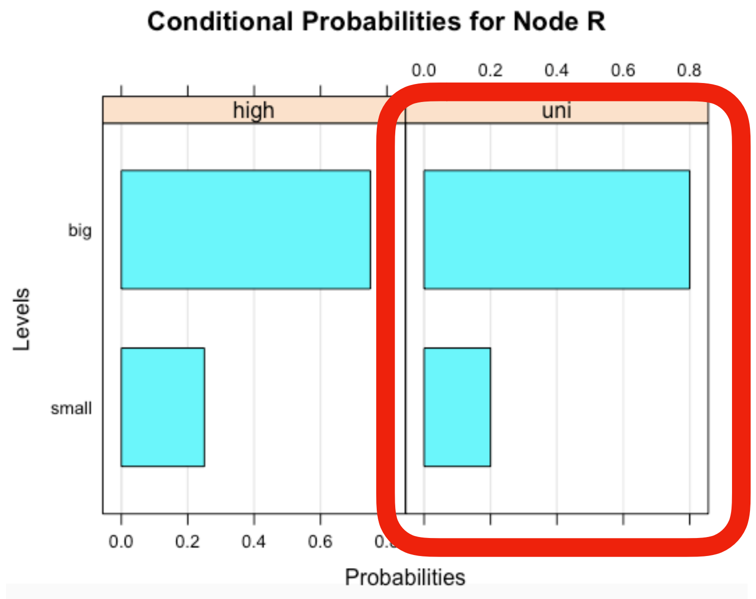
$$P(A, S, E, O, R, T) = P(T \mid O, R)\ P(R \mid E)\ P(O \mid E)\ P(E \mid A, S)\ P(S)\ P(A)$$

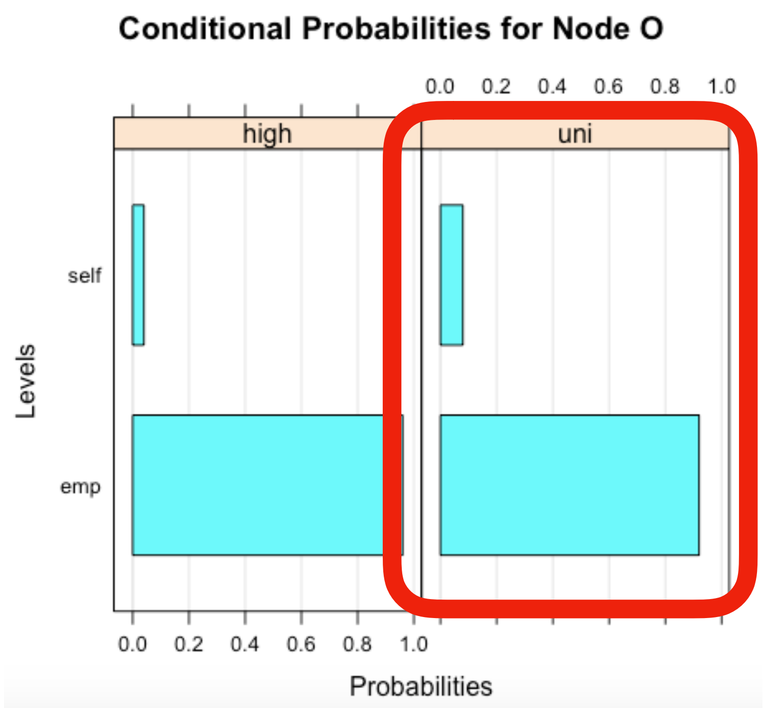


New evidence

- Suppose a new evidence is observed: **the educational level E is Uni**.

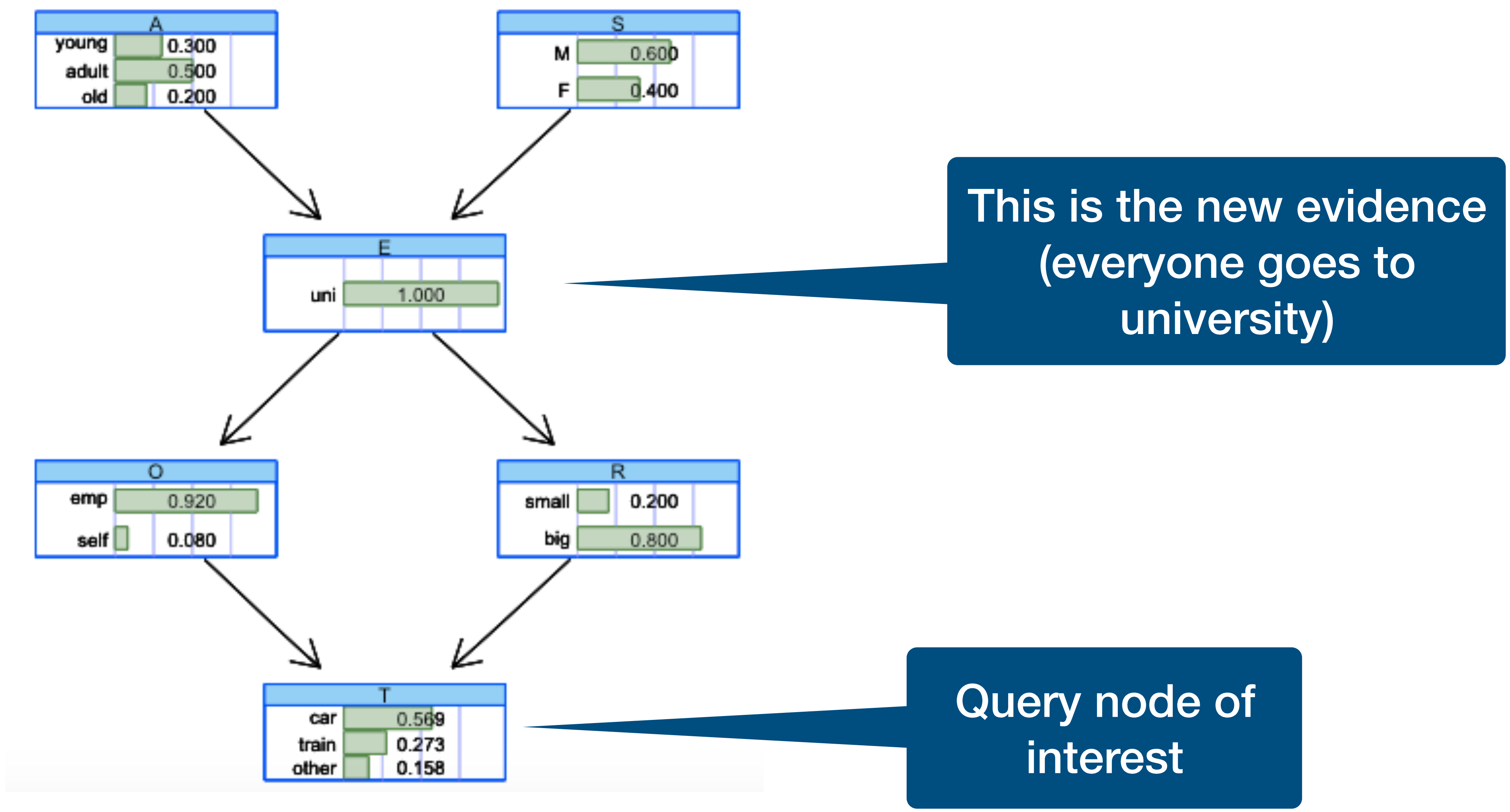






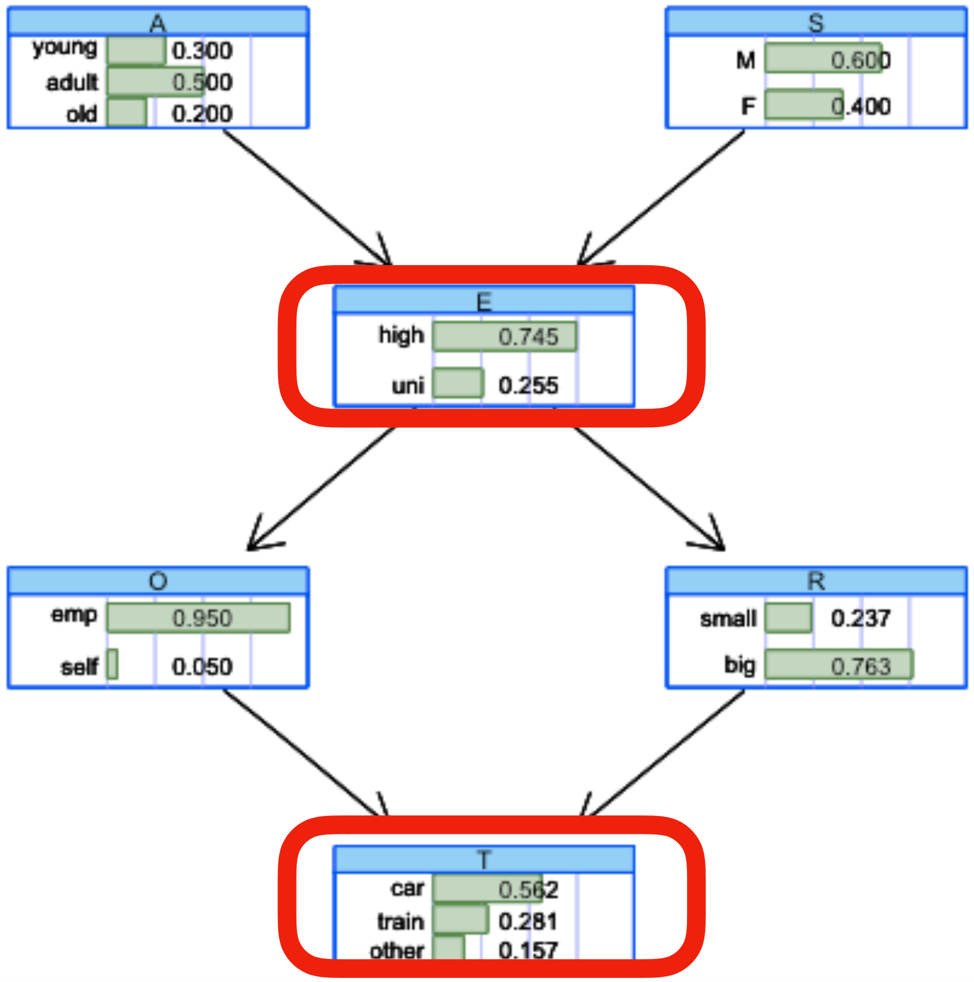
New evidence

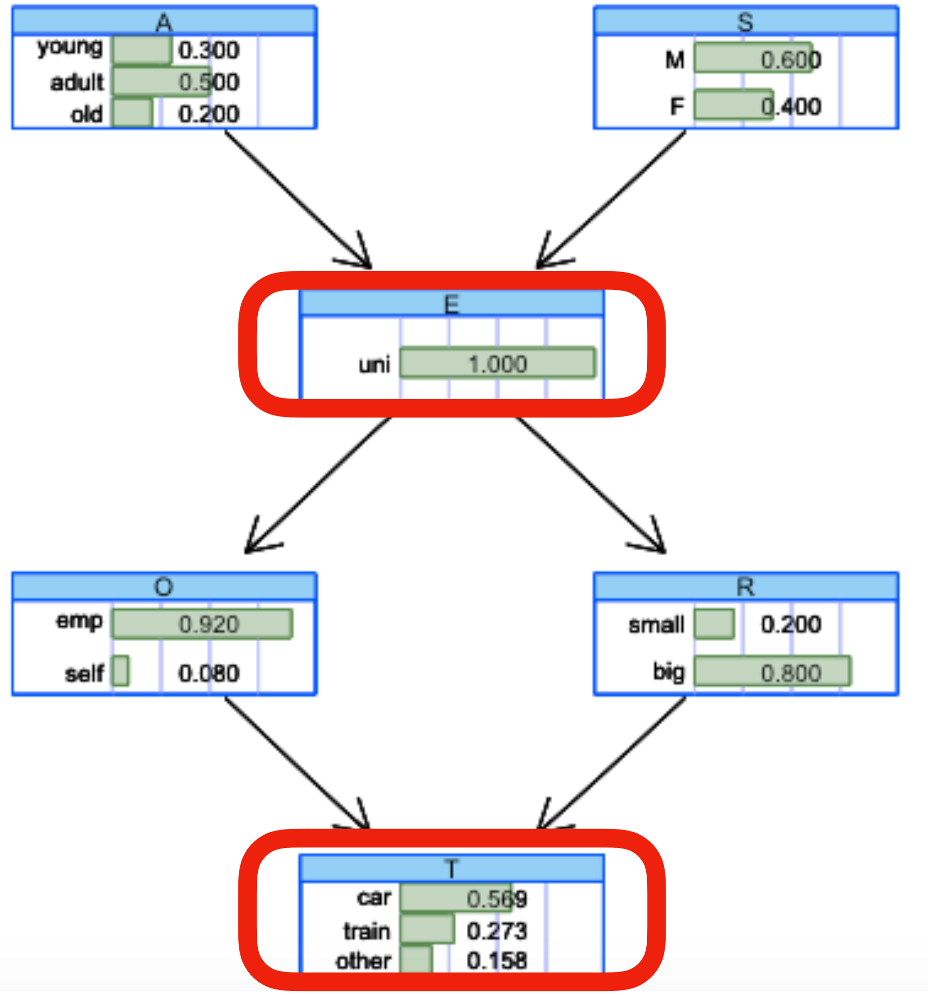
- We want to look into the probability of some event involving the other variables conditional on the evidence we have.



Impact of evidence

- Probabilities before evidence and after evidence that education level is university.

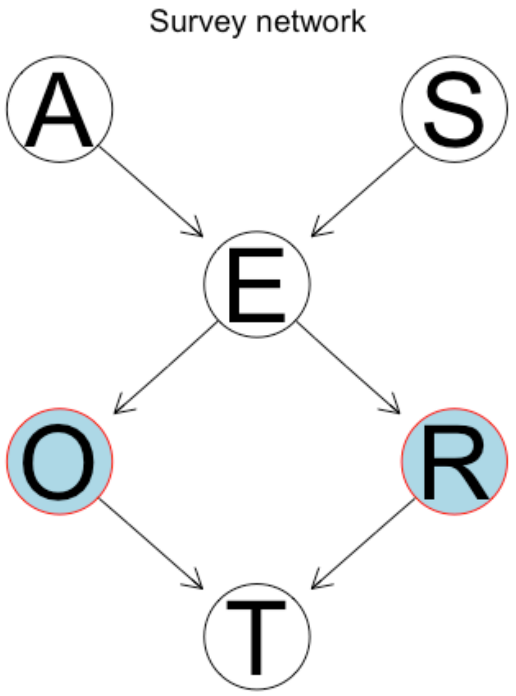




- The results indicate small impact of intervention (everyone goes to university).

Impact of a new policy

- Suppose we want to investigate the effect that residence (R) has on occupation (O).
- In particular, we want to investigate how occupation (O) changes for people living in big cities (R=big).



Impact of a new policy

- Suppose we have changed the structure of the BN such that now everybody lives and works on a big city.
- That is, the policy of interest is "to provide public housing in the city so that everybody can live and work there."
- What effect would this policy have on occupation?

Impact of a new policy

- The probability of O before intervention is

emp	self
0,9504	0,0496

- The probability of O after intervention is

emp	self
0,9462	0,0538

- Indicating a <u>very small effect</u> of the policy "providing public housing in the city" if the goal is to change occupation.