COMP 424 - Artificial Intelligence
Lecture 13: Bayesian Networks

Instructor: Jackie CK Cheung (jcheung@cs.mcgill.ca)

Readings: R&N Ch 14

Quick Questions

1. What are some reasons for representing the world using probabilities, rather than by using a logical representation?
2. What are random variables?
3. What is Bayes Rule, and how do we use it to model beliefs?

Describing the World Probabilistically

- Recall these two opposite extremes:
  1. No independence assumptions: all random variables may depend on each other
     - E.g., $P(X, Y, Z) = P(X)P(Y|X)P(Z|X, Y)$
     - Highly complex model! Too many parameters to estimate.
  2. Full independence assumptions: all random variables are independent of each other
     - E.g., $P(X, Y, Z) = P(X)P(Y)P(Z)$
     - Model is too simple! Cannot capture interactions between variables.
- We need something in between!
  - Conditional independence to model *some* of the dependencies
  - Need systematic way to represent our independence assumptions

Review: Naïve Bayes model

Without any independence assumptions:

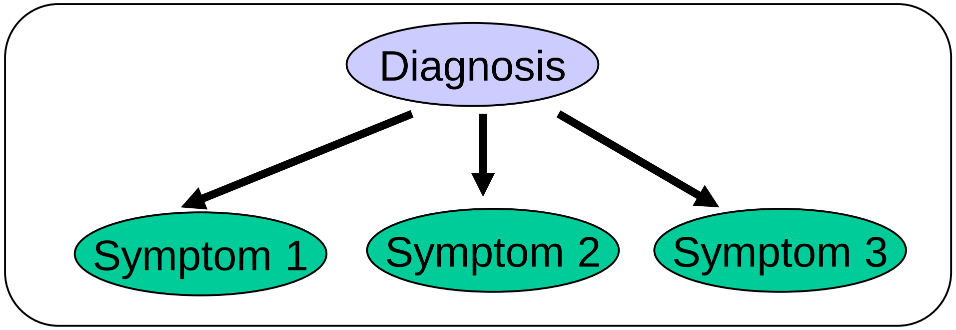
$$P(D, s_1, \ldots, s_N) = P(D)P(s_1|D)P(s_2|s_1, D) \ldots P(s_N|s_1, \ldots, s_{N-1}, D)$$

- By the chain rule of probability

With the Naïve Bayes assumption:

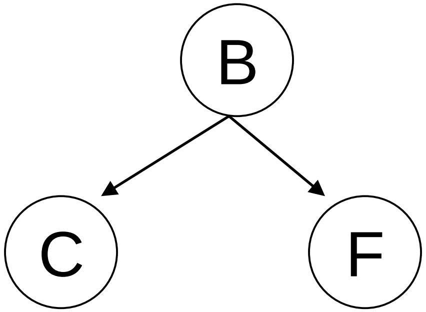
$$P(D, s_1, \ldots, s_N) = P(D)P(s_1|D)P(s_2|D) \ldots P(s_N|D)$$

We represented the model graphically like so:



Bayesian networks

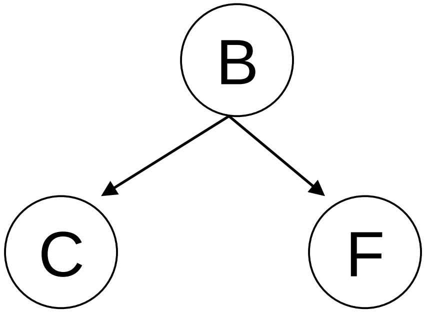
- **Bayesian networks** represent conditional independence relationships in a systematic way using a **graphical model**.
- Specify conditional independencies using graph structure.
- Graphical model = graph structure + parameters.



Bayesian networks - Basics

- **Nodes** are random variables
- **Edges** specify dependency between random variables
  - E.g., **B**: bronchitis, **C**: cough, **F**: fever (binary random variables)
  - Edges specify that **B** directly influences probability of **C**, **F**.
  - This results in **conditional probability distributions**:

P(B)	
***B* = 1**	***B* = 0**
0.18	0.82



P(C | *B)*		
	***C* = 1**	***C* = 0**
***B* = 0**	0.07	0.93
***B* = 1**	0.8	0.2

P(F | *B)*		
	***F* = 1**	***F* = 0**
***B* = 0**	0.05	0.95
***B* = 1**	0.9	0.10

Semantics of network structure

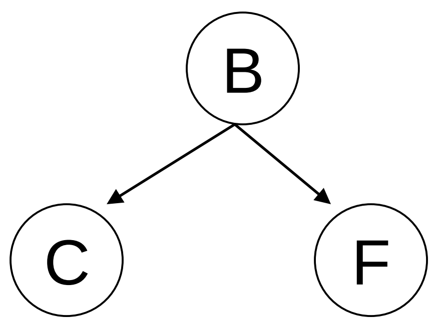
- In Bayesian networks, joint probability distribution is the product of these conditional probability distributions

$$P(x_1, x_2, \ldots, x_n) = \prod_{i=1}^{n} P(x_i | \text{parents}(X_i))$$

- e.g.,

$$P(B, C, F) = P(C|B)P(F|B)P(B)$$

P(B)	
***B* = 1**	***B* = 0**
0.18	0.82



P(C | B)		
	***C* = 1**	***C* = 0**
***B* = 0**	0.07	0.93
***B* = 1**	0.8	0.2

P(F | B)		
	***F* = 1**	***F* = 0**
***B* = 0**	0.05	0.95
***B* = 1**	0.9	0.10

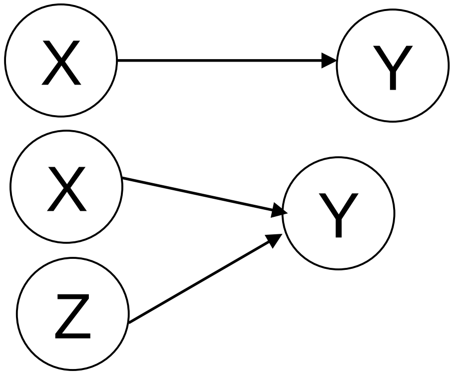
Bayesian networks, formally speaking

A Bayesian network is a **directed** graph, where:

- There is one node for each variable in the problem.
- Directed links (i.e. arcs) represent "direct influences".

How to interpret the arrows? What does it mean?

1. X is a parent of Y.
   -> *X has a direct influence on Y.*
2. X and Z are parents of Y.
   -> *X and Z have direct influence on Y.*

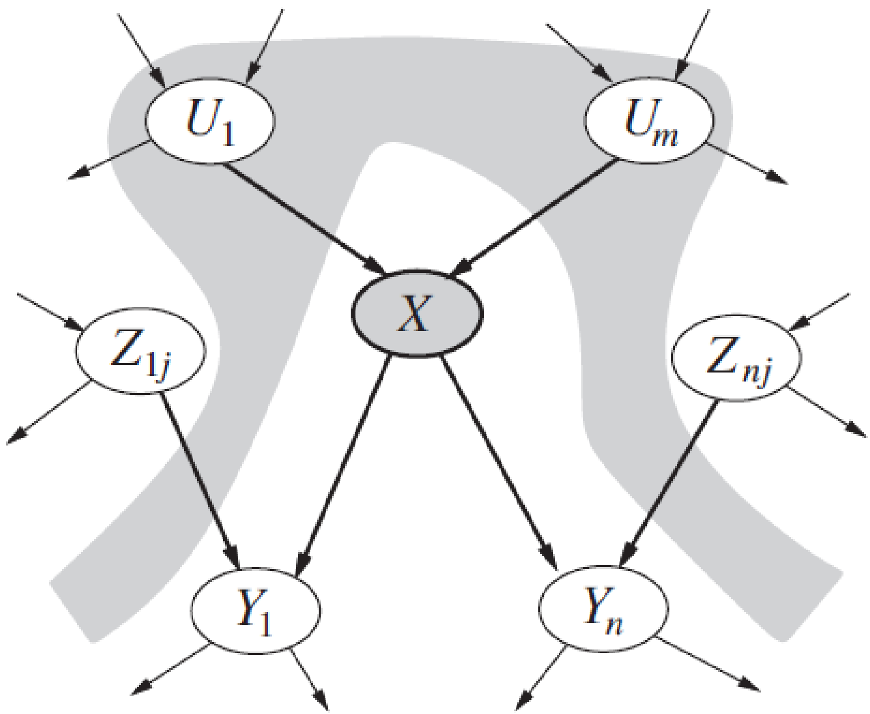


- **The graph cannot have directed cycles.**
- Each node X_i, has an associated conditional probability distribution, $P(X_i \mid \mathrm{parents}(X_i))$, that quantifies the effect of the parents on the node.

Network structure and conditional independence

1. A node is conditionally independent of its non-descendants, given its parents

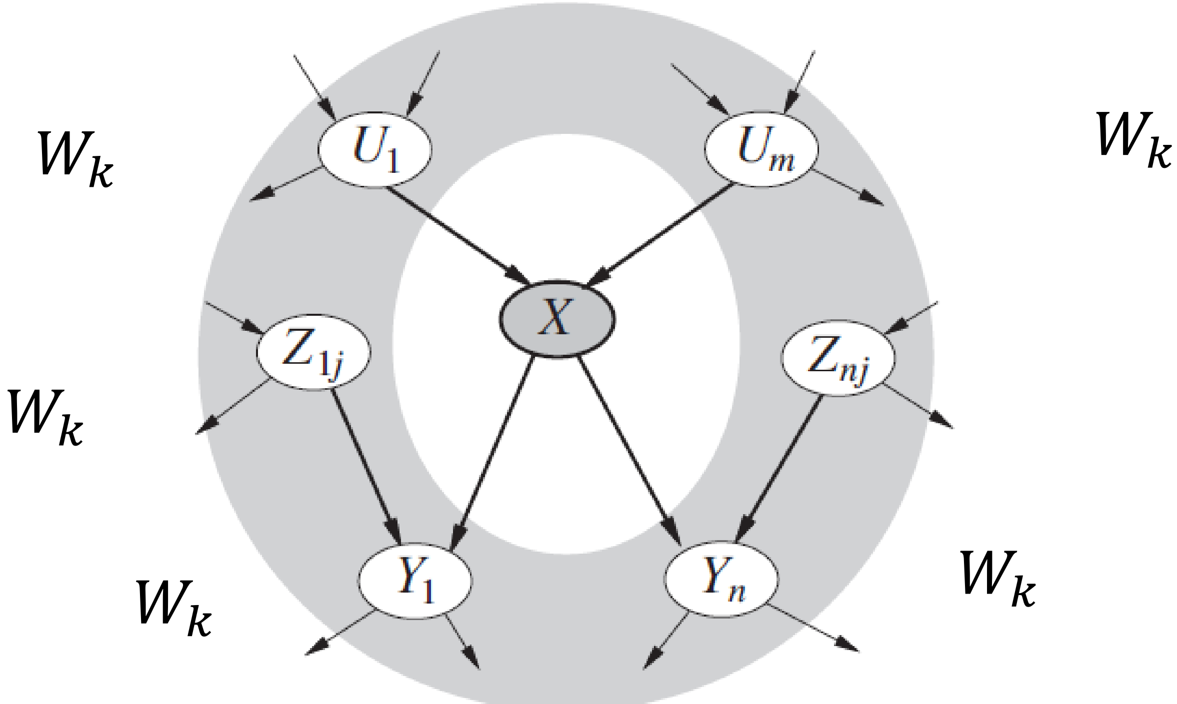
$$X \perp Z_{ij} \mid U_1, U_m \quad \forall i$$



Network structure and conditional independence

2. A node is conditionally independent of all other nodes, given its **Markov blanket** (parents, children, children's parents)

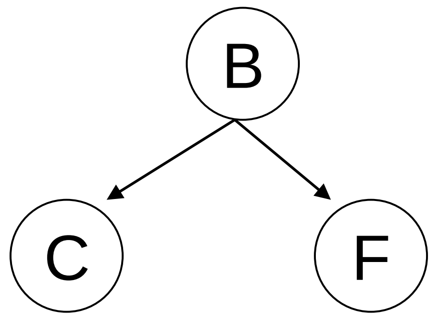
$$X \perp W_k \mid U_1, U_m, Y_1, Y_n, Z_{ij} \ \forall k$$



Example 1

B=patient has bronchitis, ***F***=patient has fever, ***C***=patient has cough

P(B)	
***B* = 1**	***B* = 0**
0.18	0.82



P(C | B)		
	***C* = 1**	***C* = 0**
***B* = 0**	0.07	0.93
***B* = 1**	0.8	0.2

P(F | B)		
	***F* = 1**	***F* = 0**
***B* = 0**	0.05	0.95
***B* = 1**	0.9	0.10

In above graph, $C \perp F \mid B$,
but not $C \perp F$ ("⊥" = "indep. of")

C is "conditionally independent" of F, given B.

Example 2 (from Poole and Mackworth)

- The agent receives a report that everyone is leaving a building and it must decide whether there is a fire in the building.
  - The **report** sensor is noisy (for eg. human error or mischief).
  - The fire **alarm** going off can cause everyone to **leave** but it's not always the case (for eg. everyone is in the middle of an exciting lecture).
  - The fire alarm usually goes off when there is a **fire** but the alarm could have been **tampered** with.
  - A fire also causes **smoke** to come out from the building.

Question: Is there a fire? Should the agent alert the fire brigade?

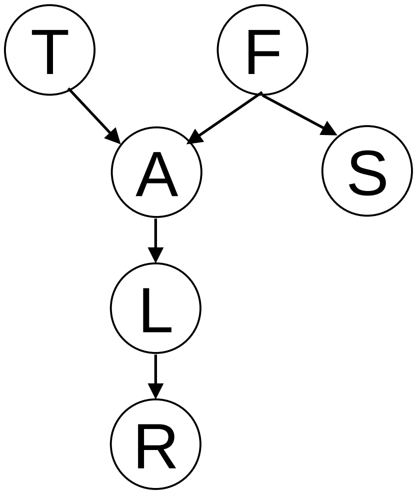
Constructing Belief Nets: Variables

Variables and domains:

- *Tampering* = *true* when there is tampering with the alarm.
- *Fire* = *true* when there is a fire.
- *Alarm* = *true* when the alarm sounds.
- *Smoke* = *true* when there is smoke.
- *Leaving* = *true* if there is an exodus of people.
- *Report* = *true* if there is a report given by someone of people leaving.

Are there any independence relationships between these variables?

Constructing Belief Nets: Structure



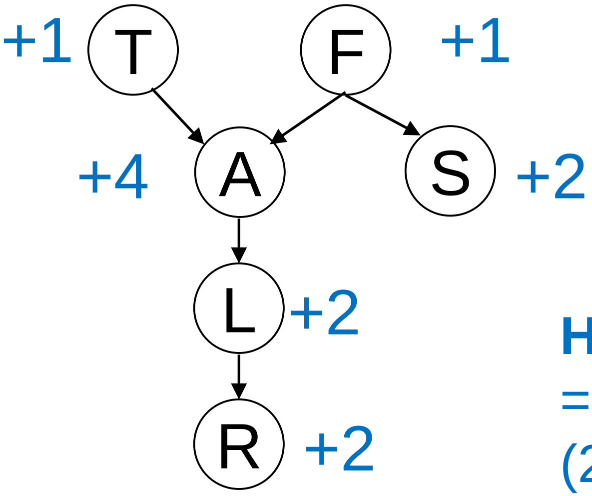
How many parameters?

Consider the variables in the order of causality:

- ***Fire*** is independent of ***Tampering***.
- ***Alarm*** depends on both ***Fire*** and ***Tampering***.
- ***Smoke*** depends only on ***Fire***. It is conditionally independent of ***Tampering*** and ***Alarm*** given whether there is a ***Fire***.
- ***Leaving*** only depends on ***Alarm*** and not directly on ***Fire*** or ***Tampering*** or ***Smoke***.
- ***Report*** depends directly only on ***Leaving***.

The network topology expresses the conditional independencies above.

Constructing Belief Nets: Structure



How many parameters?
= 12 are sufficient
(24 incl. complements.)

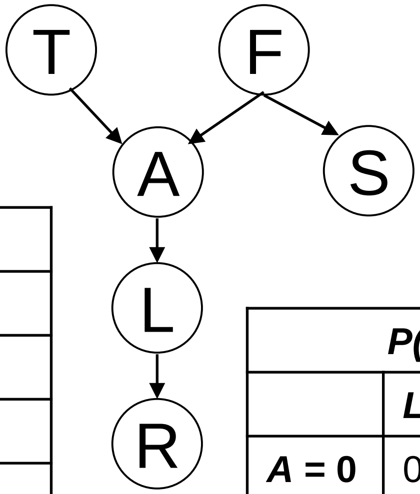
Consider the variables in the order of causality:

- ***Fire*** is independent of ***Tampering***.
- ***Alarm*** depends on both ***Fire*** and ***Tampering***.
- ***Smoke*** depends only on ***Fire***. It is conditionally independent of ***Tampering*** and ***Alarm*** given whether there is a ***Fire***.
- ***Leaving*** only depends on ***Alarm*** and not directly on ***Fire*** or ***Tampering*** or ***Smoke***.
- ***Report*** depends directly only on ***Leaving***.

The network topology expresses the conditional independencies above.

Constructing Belief Nets: CPDs

P(T)	
***T* = 1**	***T* = 0**
0.02	0.98



P(F)	
***F* = 1**	***F* = 0**
0.01	0.99

P(S | F)		
	***S* = 1**	***S* = 0**
***F* = 0**	0.01	0.99
***F* = 1**	0.90	0.10

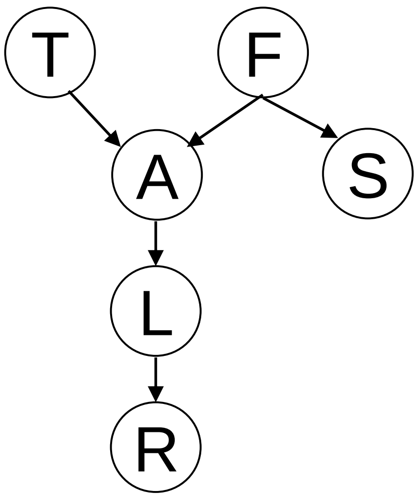
P(A| T, F)		
	***A* = 1**	***A* = 0**
***T*=0, *F*=0**	0.0001	0.9999
***T*=0, *F*=1**	0.99	0.01
***T*=1, *F*=0**	0.85	0.15
***T*=1, *F*=1**	0.5	0.5

P(L | A)		
	***L* = 1**	***L* = 0**
***A* = 0**	0.001	0.999
***A* = 1**	0.88	0.12

P(R | L)		
	***R* = 1**	***R* = 0**
***L* = 0**	0.01	0.99
***L* = 1**	0.75	0.25

Causality and Bayes Net Structure

- Directionality of edges *should* ideally specify causality, but this doesn't necessarily have to be the case.
  - E.g., fire and tampering cause alarm
  - Also we may not know direction of causality!



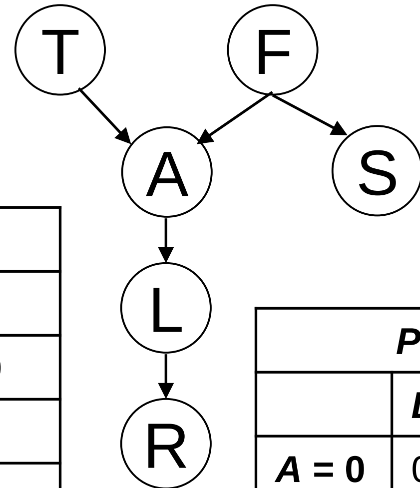
- Another graph structure (and corresponding CPTs) can produce the same joint probabilities!
- But not following causality usually results in more model parameters.

Inference in Bayes Nets

- What's the point of all this? Answer questions about state of the world!
  - Find joint probability distribution
  - Answer questions using conditional probabilities
    - Determine causes
    - Find explanations
- Use probability rules to figure out answers!
- Key operations:
  - Rewrite joint probabilities as conditional probabilities
  - Marginalize out variables

Inference in BNs: Joint prob.

P(T)	
***T* = 1**	***T* = 0**
0.02	0.98



P(F)	
***F* = 1**	***F* = 0**
0.01	0.99

P(S | F)		
	***S* = 1**	***S* = 0**
***F* = 0**	0.01	0.99
***F* = 1**	0.90	0.10

P(A| T, F)		
	***A* = 1**	***A* = 0**
***T*=0, *F*=0**	0.0001	0.9999
***T*=0, *F*=1**	0.99	0.01
***T*=1, *F*=0**	0.85	0.15
***T*=1, *F*=1**	0.5	0.5

P(L | A)		
	***L* = 1**	***L* = 0**
***A* = 0**	0.001	0.999
***A* = 1**	0.88	0.12

P(R | L)		
	***R* = 1**	***R* = 0**
***L* = 0**	0.01	0.99
***L* = 1**	0.75	0.25

Full joint distribution: *P(Tampering, Fire, Alarm, Smoke, Leaving, Report) ??*

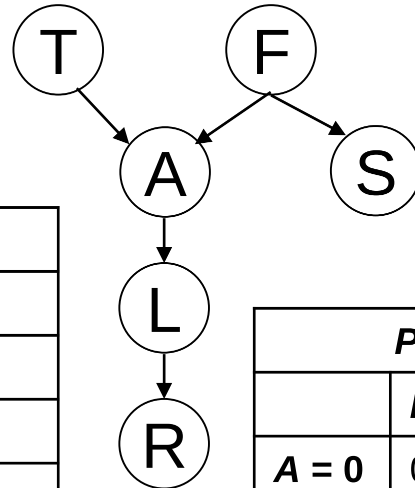
Use structure to solve this!

= *P(Tampering)* X *P(Fire)* X *P(Alarm | Tampering, Fire)* X *P(Smoke | Fire)*
  X *P(Leaving | Alarm)* X *P(Report | Leaving)*

= *0.02 x 0.01 x 0.5 x 0.9 x 0.88 x 0.75*

Inference in BNs: Joint prob.

P(T)	
***T* = 1**	***T* = 0**
0.02	0.98



P(F)	
***F* = 1**	***F* = 0**
0.01	0.99

P(S | F)		
	***S* = 1**	***S* = 0**
***F* = 0**	0.01	0.99
***F* = 1**	0.90	0.10

P(A| T, F)		
	***A* = 1**	***A* = 0**
***T*=0, *F*=0**	0.0001	0.9999
***T*=0, *F*=1**	0.99	0.01
***T*=1, *F*=0**	0.85	0.15
***T*=1, *F*=1**	0.5	0.5

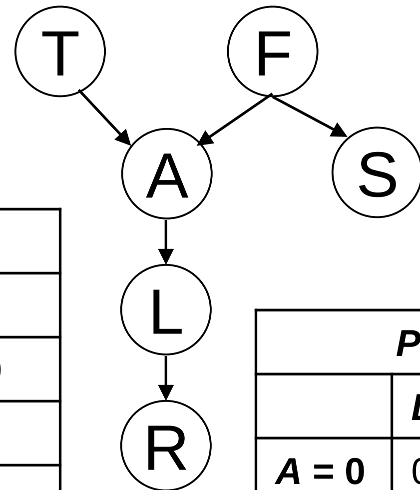
P(L | A)		
	***L* = 1**	***L* = 0**
***A* = 0**	0.001	0.999
***A* = 1**	0.88	0.12

P(R | L)		
	***R* = 1**	***R* = 0**
***L* = 0**	0.01	0.99
***L* = 1**	0.75	0.25

<u>Full joint distribution</u>: *P(~T, F, A, S, L, ~R) ??*

Inference in BNs: Joint prob.

P(T)	
***T* = 1**	***T* = 0**
0.02	0.98



P(F)	
***F* = 1**	***F* = 0**
0.01	0.99

P(S | F)		
	***S* = 1**	***S* = 0**
***F* = 0**	0.01	0.99
***F* = 1**	0.90	0.10

P(A| T, F)		
	***A* = 1**	***A* = 0**
***T*=0, *F*=0**	0.0001	0.9999
***T*=0, *F*=1**	0.99	0.01
***T*=1, *F*=0**	0.85	0.15
***T*=1, *F*=1**	0.5	0.5

P(L | A)		
	***L* = 1**	***L* = 0**
***A* = 0**	0.001	0.999
***A* = 1**	0.88	0.12

P(R | L)		
	***R* = 1**	***R* = 0**
***L* = 0**	0.01	0.99
***L* = 1**	0.75	0.25

Full joint distribution: *P(~T, F, A, S, L, ~R) ??*

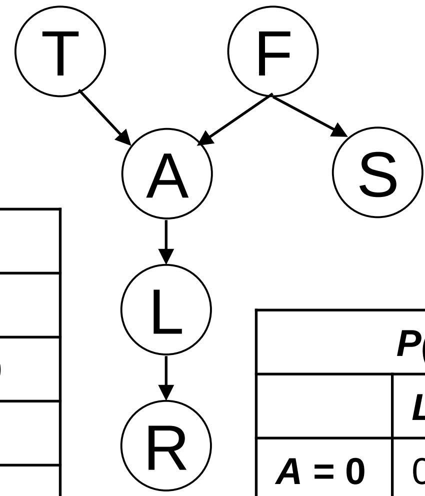
= *P(~T)* X *P(F)* X *P(A | ~T, F)* X *P(S | F)*

X *P(L | A)* X *P(~R | L)*

= *0.98 x 0.01 x 0.99 x 0.9 x 0.88 x 0.25*

Inference in BNs: Marginal prob.

P(T)	
***T* = 1**	***T* = 0**
0.02	0.98



P(F)	
***F* = 1**	***F* = 0**
0.01	0.99

P(S | F)		
	***S* = 1**	***S* = 0**
***F* = 0**	0.01	0.99
***F* = 1**	0.90	0.10

P(A| T, F)		
	***A* = 1**	***A* = 0**
***T*=0, *F*=0**	0.0001	0.9999
***T*=0, *F*=1**	0.99	0.01
***T*=1, *F*=0**	0.85	0.15
***T*=1, *F*=1**	0.5	0.5

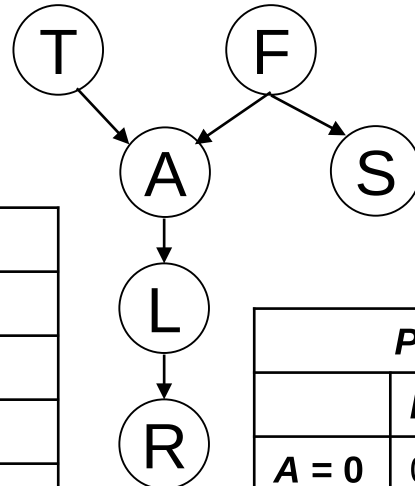
P(L | A)		
	***L* = 1**	***L* = 0**
***A* = 0**	0.001	0.999
***A* = 1**	0.88	0.12

P(R | L)		
	***R* = 1**	***R* = 0**
***L* = 0**	0.01	0.99
***L* = 1**	0.75	0.25

<u>Marginal probabilities</u>: Eg. Prob of getting a report ***P*(*R*)** ??

Inference in BNs: Marginal prob.

P(T)	
***T* = 1**	***T* = 0**
0.02	0.98



P(F)	
***F* = 1**	***F* = 0**
0.01	0.99

P(S | F)		
	***S* = 1**	***S* = 0**
***F* = 0**	0.01	0.99
***F* = 1**	0.90	0.10

P(A| T, F)		
	***A* = 1**	***A* = 0**
***T*=0, *F*=0**	0.0001	0.9999
***T*=0, *F*=1**	0.99	0.01
***T*=1, *F*=0**	0.85	0.15
***T*=1, *F*=1**	0.5	0.5

P(L | A)		
	***L* = 1**	***L* = 0**
***A* = 0**	0.001	0.999
***A* = 1**	0.88	0.12

P(R | L)		
	***R* = 1**	***R* = 0**
***L* = 0**	0.01	0.99
***L* = 1**	0.75	0.25

Marginal probabilities: Eg. Prob of getting a report ***P*(*R*)** ??

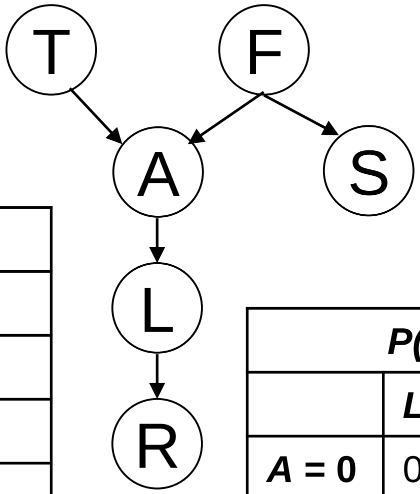
$= \quad P(R = 1) \qquad = \sum_{t,f,a,s,l} P(T=t, F=f, A=a, S=s, L=l, R = 1)$

$= \sum_{t,f,a,s,l} P(T)\, P(F)\, P(A|T,F)\, P(S|F)\, P(L|A)\, P(R=1|L)$

Sum over domain of marginalized vars: *T={0,1}, F={0,1}, A={0,1}, S={0,1}, L={0, 1}*

Inference in BNs: Causal reasoning

P(T)	
T = 1	*T* = 0
0.02	0.98



P(F)	
F = 1	*F* = 0
0.01	0.99

P(S | F)		
	S = 1	*S* = 0
F = 0	0.01	0.99
F = 1	0.90	0.10

P(A| T, F)		
	A = 1	*A* = 0
T=0, *F*=0	0.0001	0.9999
T=0, *F*=1	0.99	0.01
T=1, *F*=0	0.85	0.15
T=1, *F*=1	0.5	0.5

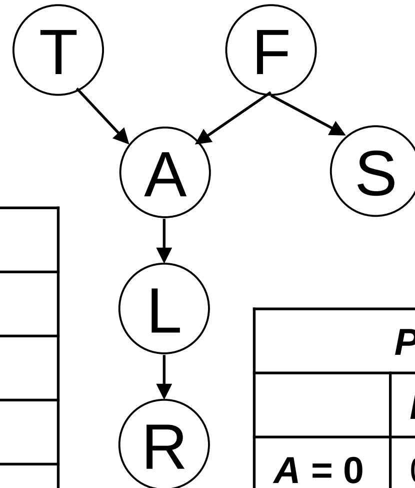
P(L | A)		
	L = 1	*L* = 0
A = 0	0.001	0.999
A = 1	0.88	0.12

P(R | L)		
	R = 1	*R* = 0
L = 0	0.01	0.99
L = 1	0.75	0.25

Causal reasoning: Eg. Prob of receiving a report in case of fire, *P(R | F)* ??

Inference in BNs: Causal reasoning

P(T)	
***T* = 1**	***T* = 0**
0.02	0.98



P(F)	
***F* = 1**	***F* = 0**
0.01	0.99

P(S | F)		
	***S* = 1**	***S* = 0**
***F* = 0**	0.01	0.99
***F* = 1**	0.90	0.10

P(A| T, F)		
	***A* = 1**	***A* = 0**
***T*=0, *F*=0**	0.0001	0.9999
***T*=0, *F*=1**	0.99	0.01
***T*=1, *F*=0**	0.85	0.15
***T*=1, *F*=1**	0.5	0.5

P(L | A)		
	***L* = 1**	***L* = 0**
***A* = 0**	0.001	0.999
***A* = 1**	0.88	0.12

P(R | L)		
	***R* = 1**	***R* = 0**
***L* = 0**	0.01	0.99
***L* = 1**	0.75	0.25

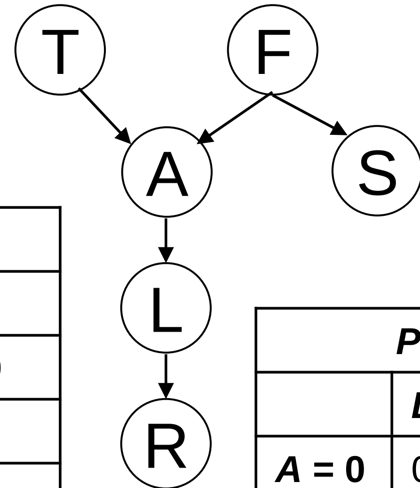
Causal reasoning: Eg. Prob of receiving a report in case of fire, $P(R \mid F)$??

$P(R = 1 \mid F = 1) = P(R = 1, F = 1) / P(F = 1)$

$= \sum_{t,a,s,l} P(T=t,F=1,A=a,S=s,L=l, R=1) / \sum_{t,a,s,l,r} P(T=t,F=1,A=a,S=s,L=l,R=r)$

Inference in BNs: Explanations

P(T)	
***T* = 1**	***T* = 0**
0.02	0.98



P(F)	
***F* = 1**	***F* = 0**
0.01	0.99

P(S | F)		
	***S* = 1**	***S* = 0**
***F* = 0**	0.01	0.99
***F* = 1**	0.90	0.10

P(A| T, F)		
	***A* = 1**	***A* = 0**
***T*=0, *F*=0**	0.0001	0.9999
***T*=0, *F*=1**	0.99	0.01
***T*=1, *F*=0**	0.85	0.15
***T*=1, *F*=1**	0.5	0.5

P(L | A)		
	***L* = 1**	***L* = 0**
***A* = 0**	0.001	0.999
***A* = 1**	0.88	0.12

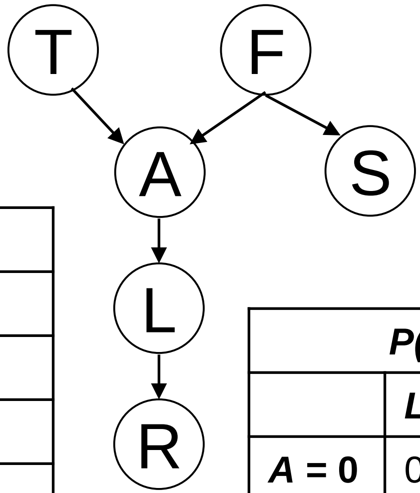
P(R | L)		
	***R* = 1**	***R* = 0**
***L* = 0**	0.01	0.99
***L* = 1**	0.75	0.25

Evidential reasoning or explanation.

- Suppose agent receives a report.
  - Prob that there is a fire?
  - Prob that there is tampering?

Inference in BNs: Explanations

P(T)	
***T* = 1**	***T* = 0**
0.02	0.98



P(F)	
***F* = 1**	***F* = 0**
0.01	0.99

P(S | F)		
	***S* = 1**	***S* = 0**
***F* = 0**	0.01	0.99
***F* = 1**	0.90	0.10

P(A| T, F)		
	***A* = 1**	***A* = 0**
***T*=0, *F*=0**	0.0001	0.9999
***T*=0, *F*=1**	0.99	0.01
***T*=1, *F*=0**	0.85	0.15
***T*=1, *F*=1**	0.5	0.5

P(L | A)		
	***L* = 1**	***L* = 0**
***A* = 0**	0.001	0.999
***A* = 1**	0.88	0.12

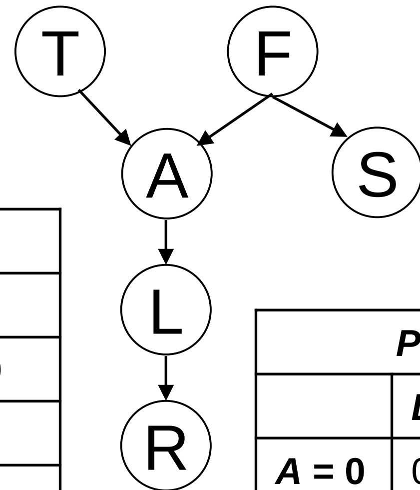
P(R | L)		
	***R* = 1**	***R* = 0**
***L* = 0**	0.01	0.99
***L* = 1**	0.75	0.25

Evidential reasoning or explanation.

- Suppose agent receives a report.
  - Prob that there is a fire? $P(F \mid R) = P(R, F) / P(R) = P(R \mid F)\, P(F) / P(R)$
  - Prob that there is tampering? $P(T \mid R) = P(T, R) / P(R)$
- Suppose agent sees smoke instead.
  - Prob that there is a fire?
  - Prob that there is tampering?

Inference in BNs: Explanations

P(T)	
***T* = 1**	***T* = 0**
0.02	0.98



P(F)	
***F* = 1**	***F* = 0**
0.01	0.99

P(S | F)		
	***S* = 1**	***S* = 0**
***F* = 0**	0.01	0.99
***F* = 1**	0.90	0.10

P(A| T, F)		
	***A* = 1**	***A* = 0**
***T*=0, *F*=0**	0.0001	0.9999
***T*=0, *F*=1**	0.99	0.01
***T*=1, *F*=0**	0.85	0.15
***T*=1, *F*=1**	0.5	0.5

P(L | A)		
	***L* = 1**	***L* = 0**
***A* = 0**	0.001	0.999
***A* = 1**	0.88	0.12

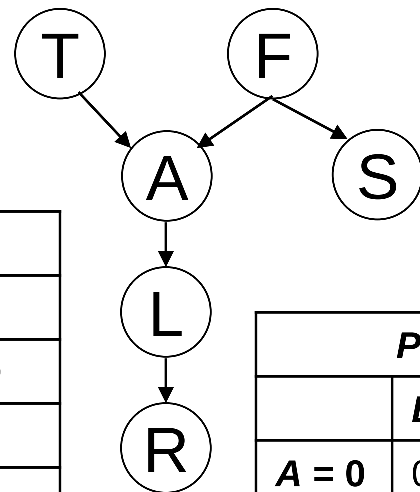
P(R | L)		
	***R* = 1**	***R* = 0**
***L* = 0**	0.01	0.99
***L* = 1**	0.75	0.25

Evidential reasoning or explanation.

- Suppose agent receives a report.
  - Prob that there is a fire? ***P(F | R) = P(R , F) / P(R) = P(R | F) P(F) / P(R)***
  - Prob that there is tampering? ***P(T | R) = P(T ,R) / P(R)***
- Suppose agent sees smoke instead.
  - Prob that there is a fire? ***P(F | S) = P(F , S) / P(S)***
  - Prob that there is tampering? ***P(T | S) = P(T , S) / P(S)***

Inference in BNs: Explanations

P(T)	
***T* = 1**	***T* = 0**
0.02	0.98



P(F)	
***F* = 1**	***F* = 0**
0.01	0.99

P(S | F)		
	***S* = 1**	***S* = 0**
***F* = 0**	0.01	0.99
***F* = 1**	0.90	0.10

P(A| T, F)		
	***A* = 1**	***A* = 0**
***T*=0, *F*=0**	0.0001	0.9999
***T*=0, *F*=1**	0.99	0.01
***T*=1, *F*=0**	0.85	0.15
***T*=1, *F*=1**	0.5	0.5

P(L | A)		
	***L* = 1**	***L* = 0**
***A* = 0**	0.001	0.999
***A* = 1**	0.88	0.12

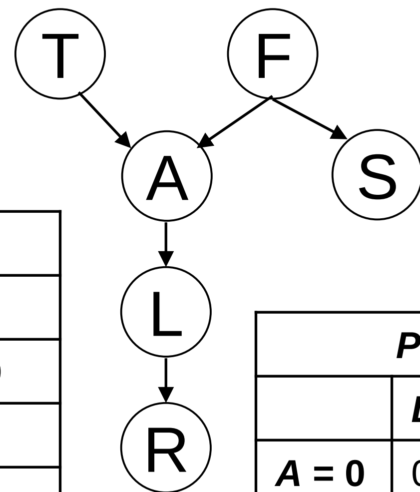
P(R | L)		
	***R* = 1**	***R* = 0**
***L* = 0**	0.01	0.99
***L* = 1**	0.75	0.25

Evidential reasoning or explanation. ***Compare posteriors with priors!***

- Suppose agent receives a report.
  - Prob that there is a fire? ***P(F | R)* = 0.2305**
  - Prob that there is tampering? ***P(T | R)* = 0.399**
- Suppose agent sees smoke instead.
  - Prob that there is a fire? ***P(F | S)* = 0.476**
  - Prob that there is tampering? ***P(T | S)* = 0.02**

Inference in BNs: Explanations

P(T)	
***T* = 1**	***T* = 0**
0.02	0.98



P(F)	
***F* = 1**	***F* = 0**
0.01	0.99

P(S | *F)*		
	***S* = 1**	***S* = 0**
***F* = 0**	0.01	0.99
***F* = 1**	0.90	0.10

P(A| T, F)		
	***A* = 1**	***A* = 0**
***T*=0, *F*=0**	0.0001	0.9999
***T*=0, *F*=1**	0.99	0.01
***T*=1, *F*=0**	0.85	0.15
***T*=1, *F*=1**	0.5	0.5

P(L | *A)*		
	***L* = 1**	***L* = 0**
***A* = 0**	0.001	0.999
***A* = 1**	0.88	0.12

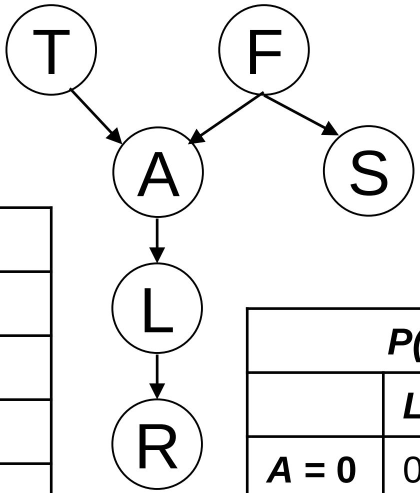
P(R | *L)*		
	***R* = 1**	***R* = 0**
***L* = 0**	0.01	0.99
***L* = 1**	0.75	0.25

Evidential reasoning or explanation.

- Suppose agent receives a report *and* sees smoke.

Inference in BNs: Explanations

P(T)	
***T* = 1**	***T* = 0**
0.02	0.98



P(F)	
***F* = 1**	***F* = 0**
0.01	0.99

P(S | F)		
	***S* = 1**	***S* = 0**
***F* = 0**	0.01	0.99
***F* = 1**	0.90	0.10

P(A| T, F)		
	***A* = 1**	***A* = 0**
***T*=0, *F*=0**	0.0001	0.9999
***T*=0, *F*=1**	0.99	0.01
***T*=1, *F*=0**	0.85	0.15
***T*=1, *F*=1**	0.5	0.5

P(L | A)		
	***L* = 1**	***L* = 0**
***A* = 0**	0.001	0.999
***A* = 1**	0.88	0.12

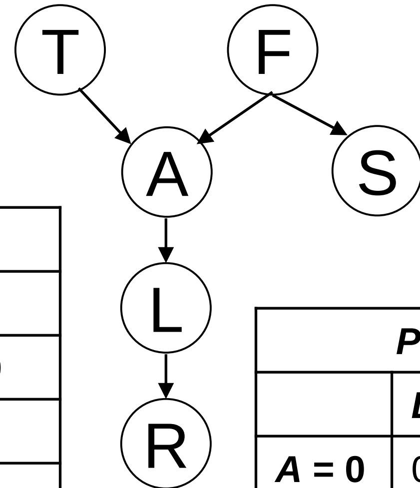
P(R | L)		
	***R* = 1**	***R* = 0**
***L* = 0**	0.01	0.99
***L* = 1**	0.75	0.25

Evidential reasoning or explanation.

- Suppose agent receives a report *and* sees smoke.
  - Prob that there is a fire? ***P(F | R, S) = P(F, R, S) / P(R , S) = 0.964***
  - Prob that there is tampering? ***P(T | R, S) = P(T ,R, S) / P(R, S) = 0.0286***

Inference in BNs: Explanations

P(T)	
***T* = 1**	***T* = 0**
0.02	0.98



P(F)	
***F* = 1**	***F* = 0**
0.01	0.99

P(S | F)		
	***S* = 1**	***S* = 0**
***F* = 0**	0.01	0.99
***F* = 1**	0.90	0.10

P(A| T, F)		
	***A* = 1**	***A* = 0**
***T*=0, *F*=0**	0.0001	0.9999
***T*=0, *F*=1**	0.99	0.01
***T*=1, *F*=0**	0.85	0.15
***T*=1, *F*=1**	0.5	0.5

P(L | A)		
	***L* = 1**	***L* = 0**
***A* = 0**	0.001	0.999
***A* = 1**	0.88	0.12

P(R | L)		
	***R* = 1**	***R* = 0**
***L* = 0**	0.01	0.99
***L* = 1**	0.75	0.25

Evidential reasoning or explanation.

- Suppose agent receives a report *and* sees smoke.
  - Prob that there is a fire? ***P(F | R, S) = P(F, R, S) / P(R , S)* = 0.964**
  - Prob that there is tampering? ***P(T | R, S) = P(T ,R, S) / P(R, S)* = 0.0286**

Compare to: ***P(F | R) =* 0.2305** and ***P(T | R) =* 0.399.** **This is called explaining away.**
***P(F | R, ~S) =* 0.0294** and ***P(T | R, ~S)* = 0.501.**

Types of queries for graphical models

1. **Unconditional probability query**
   - What is the prob of some value assignment for a subset of variables Y?

$$P(Y)$$

2. **Conditional probability query**
   - What is the prob of different value assignments for query variable Y, given evidence about variables Z?

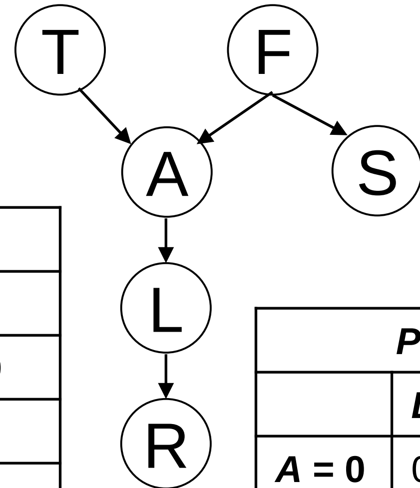
$$P(Y \mid Z = z)$$

3. **Maximum a posterior (MAP) queries**
   - Given evidence Z=z, what is the most likely assignment of values to variables Y?

$$MAP(Y \mid Z = z) \;=\; argmax_y P(Y = y \mid Z = z)$$

Inference in BNs: MAP queries

P(T)	
***T* = 1**	***T* = 0**
0.02	0.98



P(F)	
***F* = 1**	***F* = 0**
0.01	0.99

P(S | F)		
	***S* = 1**	***S* = 0**
***F* = 0**	0.01	0.99
***F* = 1**	0.90	0.10

P(A| T, F)		
	***A* = 1**	***A* = 0**
***T*=0, *F*=0**	0.0001	0.9999
***T*=0, *F*=1**	0.99	0.01
***T*=1, *F*=0**	0.85	0.15
***T*=1, *F*=1**	0.5	0.5

P(L | A)		
	***L* = 1**	***L* = 0**
***A* = 0**	0.001	0.999
***A* = 1**	0.88	0.12

P(R | L)		
	***R* = 1**	***R* = 0**
***L* = 0**	0.01	0.99
***L* = 1**	0.75	0.25

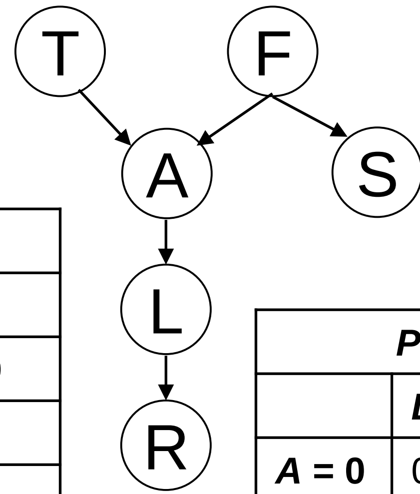
Calculating the MAP from the posteriors.

- Suppose agent receives a report.
  - Prob that there is a fire? ***P(F | R)* = 0.2305**.

What's the MAP?

Inference in BNs: MAP queries

P(T)	
***T* = 1**	***T* = 0**
0.02	0.98



P(F)	
***F* = 1**	***F* = 0**
0.01	0.99

P(S | F)		
	***S* = 1**	***S* = 0**
***F* = 0**	0.01	0.99
***F* = 1**	0.90	0.10

P(A| T, F)		
	***A* = 1**	***A* = 0**
***T*=0, *F*=0**	0.0001	0.9999
***T*=0, *F*=1**	0.99	0.01
***T*=1, *F*=0**	0.85	0.15
***T*=1, *F*=1**	0.5	0.5

P(L | A)		
	***L* = 1**	***L* = 0**
***A* = 0**	0.001	0.999
***A* = 1**	0.88	0.12

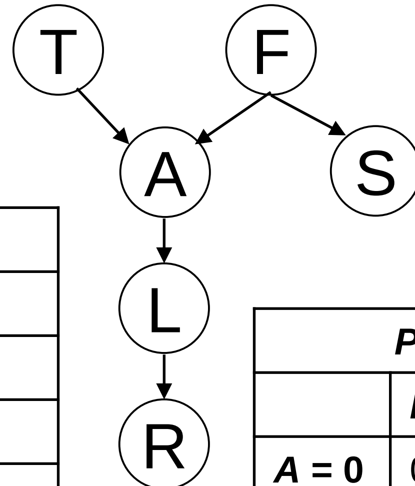
P(R | L)		
	***R* = 1**	***R* = 0**
***L* = 0**	0.01	0.99
***L* = 1**	0.75	0.25

Calculating the MAP from the posteriors.

- Suppose agent receives a report.
  - Prob that there is a fire? ***P(F | R)* = 0.2305**. ***F = 0***
  - Prob that there is tampering? ***P(T | R)* = 0.399** ***MAP?***

Inference in BNs: MAP queries

P(T)	
***T* = 1**	***T* = 0**
0.02	0.98



P(F)	
***F* = 1**	***F* = 0**
0.01	0.99

P(S | F)		
	***S* = 1**	***S* = 0**
***F* = 0**	0.01	0.99
***F* = 1**	0.90	0.10

P(A| T, F)		
	***A* = 1**	***A* = 0**
***T*=0, *F*=0**	0.0001	0.9999
***T*=0, *F*=1**	0.99	0.01
***T*=1, *F*=0**	0.85	0.15
***T*=1, *F*=1**	0.5	0.5

P(L | A)		
	***L* = 1**	***L* = 0**
***A* = 0**	0.001	0.999
***A* = 1**	0.88	0.12

P(R | L)		
	***R* = 1**	***R* = 0**
***L* = 0**	0.01	0.99
***L* = 1**	0.75	0.25

Calculating the MAP from the posteriors.

- Suppose agent receives a report.
  - Prob that there is a fire? ***P(F | R)* = 0.2305**. *F = 0*
  - Prob that there is tampering? ***P(T | R)* = 0.399** *T = 0*
- Suppose agent sees smoke AND receives a report.
  - Prob that there is a fire? ***P(F | R, S)*** *MAP?*

Other examples of MAP queries

- In speech recognition:
  - given a speech signal
  - determine sequence of words most likely to have generated signal.

- In medical diagnosis:
  - given a patient
  - determine the most probable diagnosis.

- In robotics:
  - given sensor readings
  - determine the most probably location of the robot.

Complexity of inference in Bayes Nets

- Given a Bayes net and a random variable X, deciding whether P(X=x) > 0 is NP-hard.

- <u>Bad news:</u>

  No general inference procedure that will work efficiently for all network configurations.

- <u>Good news:</u>

  For particular families of networks, inference can be done efficiently. E.g. Naïve Bayes Model!

Recall the Naïve Bayes Model

- A common assumption in early diagnosis is that the symptoms are independent of each other given the disease.
- Let $s_1, .., s_n$ be the symptoms exhibited by a patient (e.g. fever, headache, etc.). Let D be the patient's disease.
- Using the Naive Bayes assumption:

$$P(D, s_1, \ldots, s_n) = P(D)\, P(s_1 \mid D) \ldots P(s_n \mid D)$$

- Only $2n+1$ parameters!