BAYESIAN NETWORKS:

MODELING AND INFERENCE:

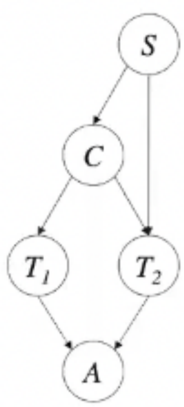
Inference

1. Queries

2. Algorithms

Modeling

Sensitivity Analysis



S	θ_s
male	.55
female	.45

S	C	$\theta_{c\|s}$
male	yes	.05
male	no	.95
female	yes	.01
female	no	.99

C	T_1	$\theta_{t_1\|c}$
yes	+ve	.80
yes	−ve	.20
no	+ve	.20
no	−ve	.80

S	C	T_2	$\theta_{t_2\|c,s}$
male	yes	+ve	.80
male	yes	−ve	.20
male	no	+ve	.20
male	no	−ve	.80
female	yes	+ve	.95
female	yes	−ve	.05
female	no	+ve	.05
female	no	−ve	.95

T_1	T_2	A	$\theta_{a\|t_1,t_2}$
+ve	+ve	yes	1
+ve	+ve	no	0
+ve	−ve	yes	0
+ve	−ve	no	1
−ve	+ve	yes	0
−ve	+ve	no	1
−ve	−ve	yes	1
−ve	−ve	no	0

C: Condition

T_1, T_2: Tests

S: Sex

32 worlds

QUERIES:

PRIOR MARGINALS:

Monitors

C = yes	3.2%
C = No	96.8%

T_1 = +ve	21.92%
T_1 = -ve	78.08%

Probabilities with no other evidence.

POSTERIOR MARGINALS:

We have evidence.

T_1 = +ve, T_2 = +ve

$Pr(\cdot)$

Prior

$Pr(\cdot \mid \beta)$

Posterior

β : T_1 = +ve, T_2 = +ve

C = yes	45.33%
C = No	54.67%

Male	83.02%
Female	16.98%

MPE: MOST PROBABLE EXPLANATION:

Evidence A = yes

$T_1 = T_2$ = +ve or

$T_1 = T_2$ = -ve.

32 instantiations

→ S, C, T_1, T_2 | A = yes

(16 instantiations)

Return the most probable instantiation of S, C, T_1, T_2

given evidence A = yes

→ C = NO, S = female, T_1 = -ve, T_2 = -ve } ~47.0%

MAP : MAXIMUM A POSTERIORI HYPOTHESIS:

Evidence:

A = yes

$$\left.\begin{array}{l} C = \text{no} \\ S = \text{male} \end{array}\right\} \sim 49.3\%$$

Pass evidence and set of parameters to maximize over.

MAP + all variables = MPE

COMPLEXITY OF INFERENCE:

- Variable elimination
- Conditioning

TOPOLOGY, TREEWIDTH

(measure of how connected)

n: # variables

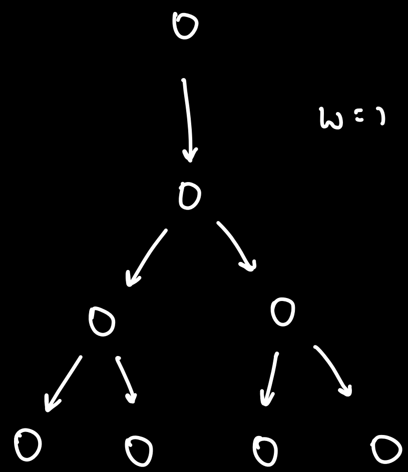
d: # values

w: treewidth

marginals

→ $O(n \cdot d^w)$

→ TREE



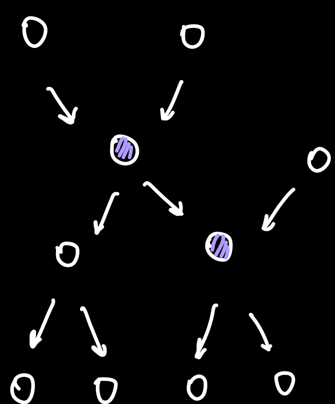
$w = 1$

→ POLYTREE [SINGLE - CONNECTED]

(Multiple parents)

$w = k$

where k : maximum number of parents a node can have.



$k = 2$

→ MULTIPLE - CONNECTED (DAG)

WEIGHTED MODEL COUNTING (WMC):

$\Delta : (A \vee B) \wedge \neg C$

A	B	C	weights
t	t	t	0.08
t	t	f	0.04
t	f	t	0.10
t	f	f	0.10
f	t	t	0.00
f	t	f	0.00
f	f	t	0.42
f	f	f	0.06

product of weights of literals.

SAT → yes, no

#SAT → 3 (models)

→ model counting (MC)

WMC:

$0.04 + 0.10 + 0.00 = 0.14$

$\Delta : (A \vee B) \wedge \neg C$

compile

NNF Circuit

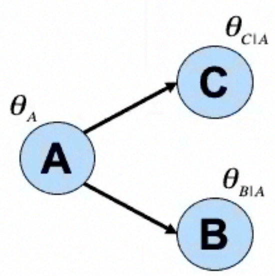
Smooth

Decomposable

Deterministic

} WMC in linear time.

REDUCING PROBABILISTIC INFERENCE TO WMC:



A	B	C	Pr(.)
T	T	T	$\theta_A \theta_{B\|A} \theta_{C\|A}$
T	T	F	$\theta_A \theta_{B\|A} \theta_{\neg C\|A}$
T	F	T	$\theta_A \theta_{\neg B\|A} \theta_{C\|A}$
T	F	F	$\theta_A \theta_{\neg B\|A} \theta_{\neg C\|A}$
F	T	T	$\theta_{\neg A} \theta_{B\|\neg A} \theta_{C\|\neg A}$
F	T	F	$\theta_{\neg A} \theta_{B\|\neg A} \theta_{\neg C\|\neg A}$
F	F	T	$\theta_{\neg A} \theta_{\neg B\|\neg A} \theta_{C\|\neg A}$
F	F	F	$\theta_{\neg A} \theta_{\neg B\|\neg A} \theta_{\neg C\|\neg A}$

8 worlds

Variables : A, B, C

Introduce a boolean variable for each parameter:

$P_1 : \theta_A$ $P_3 : \theta_{B|A}$ $P_5 : \theta_{B|\neg A}$ $P_7 : \theta_{C|A}$ $P_9 : \theta_{C|\neg A}$

$P_2 : \theta_{\neg A}$ $P_4 : \theta_{\neg B|A}$ $P_6 : \theta_{\neg B|\neg A}$ $P_8 : \theta_{\neg C|A}$ $P_{10} : \theta_{\neg C|\neg A}$

$$
\Delta = \begin{array}{lcl}
A & \Leftrightarrow & P_1 \\
\neg A & \Leftrightarrow & P_2 \\
\hline
A \wedge B & \Leftrightarrow & P_3 \\
A \wedge \neg B & \Leftrightarrow & P_4 \\
\neg A \wedge B & \Leftrightarrow & P_5 \\
\neg A \wedge \neg B & \Leftrightarrow & P_6 \\
\hline
A \wedge C & \Leftrightarrow & P_7 \\
A \wedge \neg C & \Leftrightarrow & P_8 \\
\neg A \wedge C & \Leftrightarrow & P_9 \\
\neg A \wedge \neg C & \Leftrightarrow & P_{10}
\end{array}
$$

Assign weights

$$WMC(\Delta \wedge \alpha) = Pr(\alpha)$$

$\Delta \rightarrow$ Boolean Circut

* One model of Δ:

$A \quad \neg B \quad C \quad P_1 \quad P_4 \quad P_7 \quad \neg P_2 \quad \neg P_3 \quad \neg P_5 \quad \neg P_6 \quad \neg P_8 \quad \neg P_9 \quad \neg P_{10}$

* Weights of literals:

- $w(A) = w(\neg A) = w(B) = w(\neg B) = w(C) = w(\neg C) = 1$

- $w(P_i) = \theta_i$

$w(P_1) = \theta_A$ $\qquad$ $w(P_7) = \theta_{C|A}$

$w(P_4) = \theta_{\neg B|A}$

- $w(\neg P_i) = 1$

$$w(\text{model}) = w(P_1)\, w(P_4)\, w(P_7) = \theta_A\, \theta_{\neg B|A}\, \theta_{c|A}$$

BAYESIAN:

n: variables

k: maximum # parents per node

d: maximum # values per variable

d^{k+1} } bound on size CPT

n CPTs.

Size of Bayesian Network: $O(n \cdot d^{k+1})$

Joint Probability d^n

Example

The flu is an acute disease characterized by fever, body aches and pains, and can be associated with chilling and a sore throat. The cold is a bodily disorder popularly associated with chilling and can cause a sore throat. Tonsillitis is inflammation of the tonsils which leads to a sore throat and can be associated with fever.

STEP 1:

Variable, Value

STEP 2:

Edge

STEP 3:

CPTs

STEP 1:

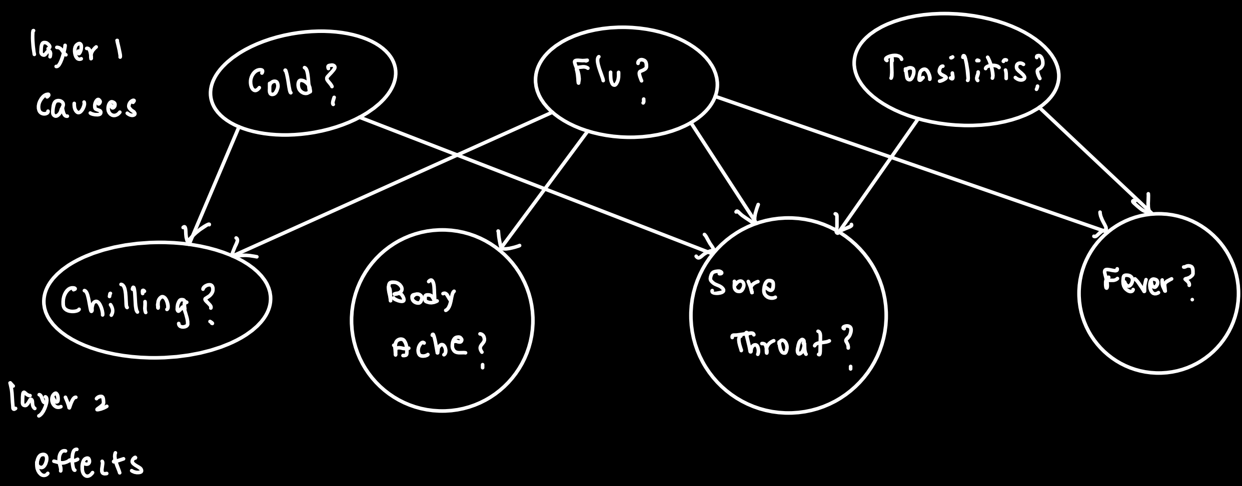
Variables:

Example

The flu is an acute disease characterized by fever, body aches and pains, and can be associated with chilling and a sore throat. The cold is a bodily disorder popularly associated with chilling and can cause a sore throat. Tonsillitis is inflammation of the tonsils which leads to a sore throat and can be associated with fever.

Values: T/F for each

STEP 2:



Query variable

Evidence variable

STEP 3:

3 Methods

1. Problem Statement
2. Subjective Beliefs
3. Learning from data.

Learning from data:

CPTs can also be estimated from medical records of previous patients

Case	Cold?	Flu?	Tonsillitis?	Chilling?	Bodyache?	Sorethroat?	Fever?
1	true	false	?	true	false	false	false
2	false	true	false	true	true	false	true
3	?	?	true	false	?	true	false
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

Data
→ Complete → Closed Form (Efficient)
→ Incomplete → EM: Expectation Maximization.

BN Structure + CPTs = Bayesian Netword

A. CPTs A = BN A
B. CPTs B = BN B

} Find the scores (probability) of each data.

Find maximum score

MAXIMUM LIKELIHOOD PRINCIPLE.

Pr(Data | BN)

Example

A few weeks after inseminating a cow, we have three possible tests to confirm pregnancy. The first is a scanning test which has a false positive of 1% and a false negative of 10%. The second is a blood test, which detects progesterone with a false positive of 10% and a false negative of 30%. The third test is a urine test, which also detects progesterone with a false positive of 10% and a false negative of 20%. The probability of a detectable progesterone level is 90% given pregnancy, and 1% given no pregnancy. The probability that insemination will impregnate a cow is 87%.

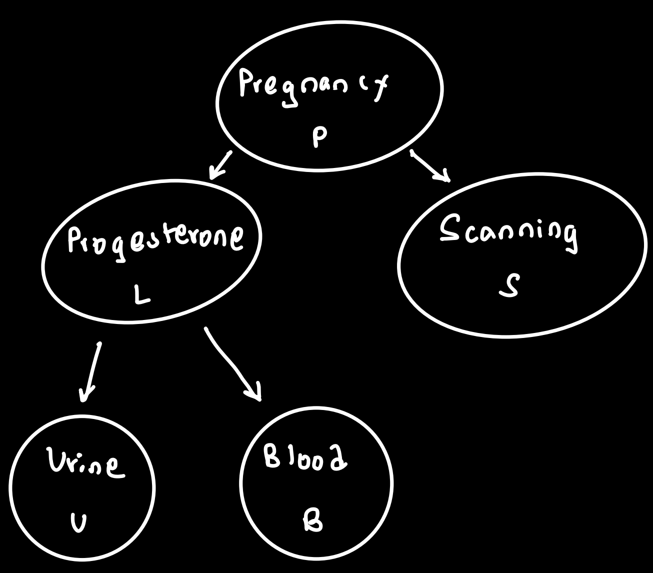
STEP 1:

Variables

Example

A few weeks after inseminating a cow, we have three possible tests to confirm pregnancy. The first is a scanning test which has a false positive of 1% and a false negative of 10%. The second is a blood test, which detects progesterone with a false positive of 10% and a false negative of 30%. The third test is a urine test, which also detects progesterone with a false positive of 10% and a false negative of 20%. The probability of a detectable progesterone level is 90% given pregnancy, and 1% given no pregnancy. The probability that insemination will impregnate a cow is 87%.

STEP 2:



STEP 3:

P	
yes	0.87
No	0.13

	P	S	
	yes	+ve	0.9
fn →	yes	-ve	0.1
fp →	no	+ve	0.01
	no	-ve	0.99

Evidence:

S = -ve, B = -ve, U = -ve

P = yes	10.21%
P = No	89.79%

Posterior

Marginal

if we want P = yes ≤ 5%

RECOMMENDATIONS:

1. fn for scanning } Something we can control.

   10% to 4.63%

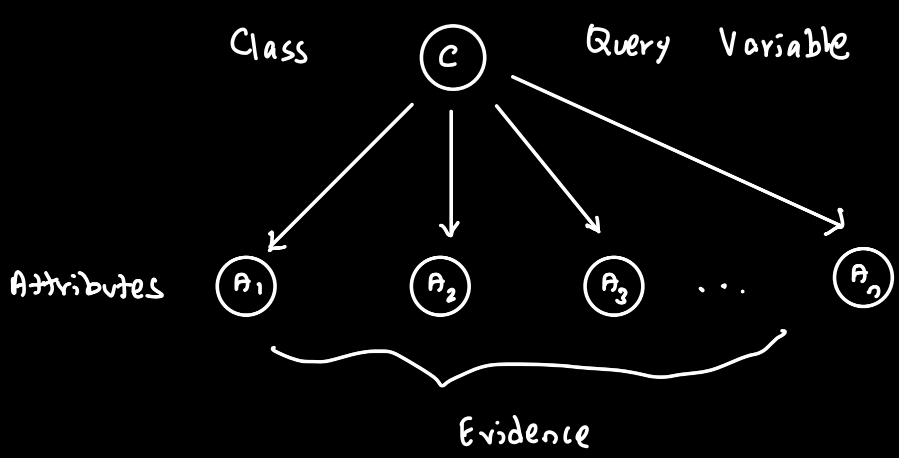
2. Success rate of procedure ×

   87% → 75.59%

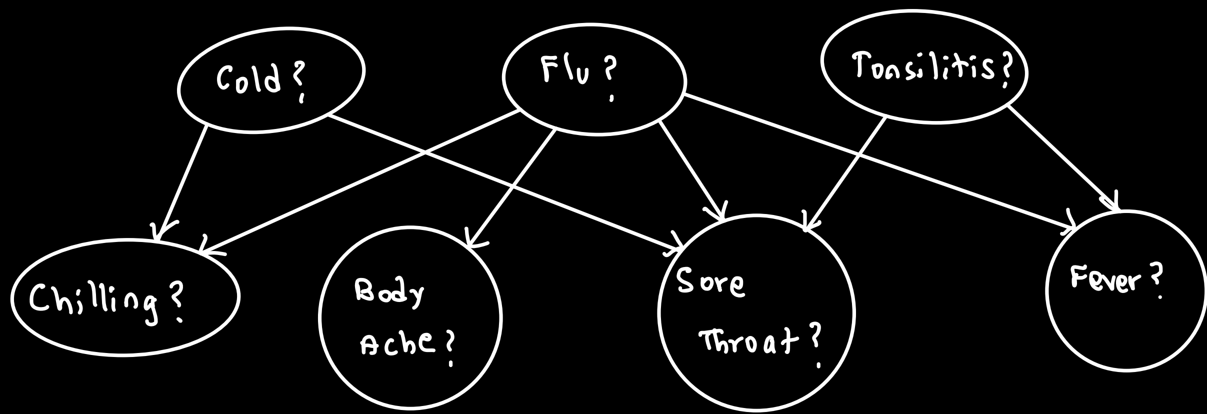
3. $P(L = yes \mid P = yes)$

   90% → 99.67% ×

NAIVE BAYES STRUCTURE:



TREE



SINGLE FAULT ASSUMPTION

If a person only has one of Cold/ Flu / Tonsilitis, then it can be represented as Naive Bayes: