Bayesian Networks and Decision Graphs

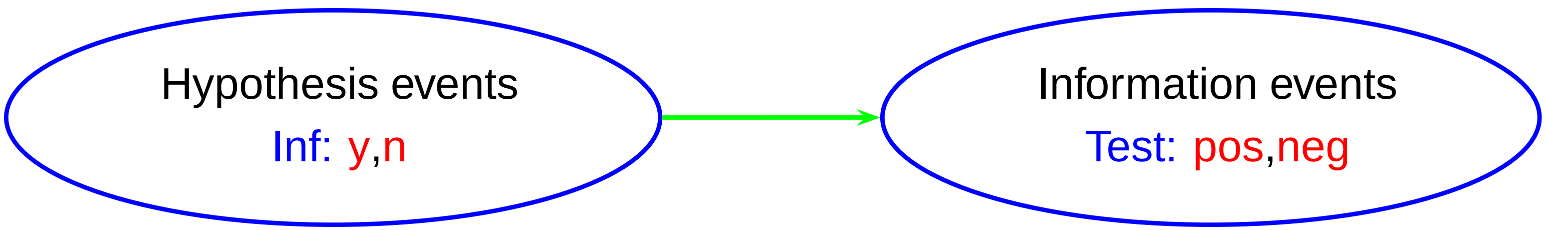
Chapter 3

Building models

Milk from a cow may be infected. To detect whether or not the milk is infected, you can apply a test which may either give a positive or a negative test result. The test is not perfect: It may give false positives as well as false negatives.

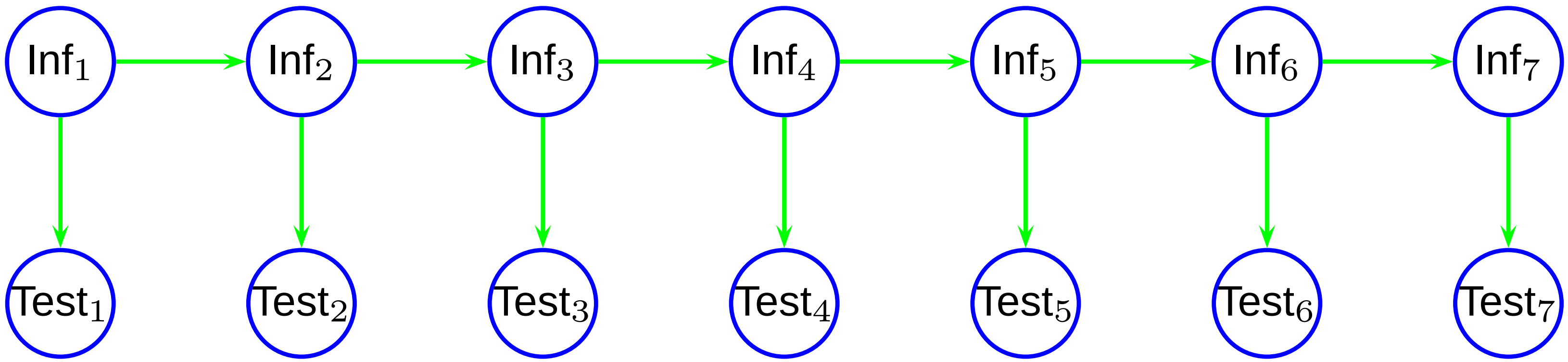
Building models

Milk from a cow may be infected. To detect whether or not the milk is infected, you can apply a test which may either give a positive or a negative test result. The test is not perfect: It may give false positives as well as false negatives.



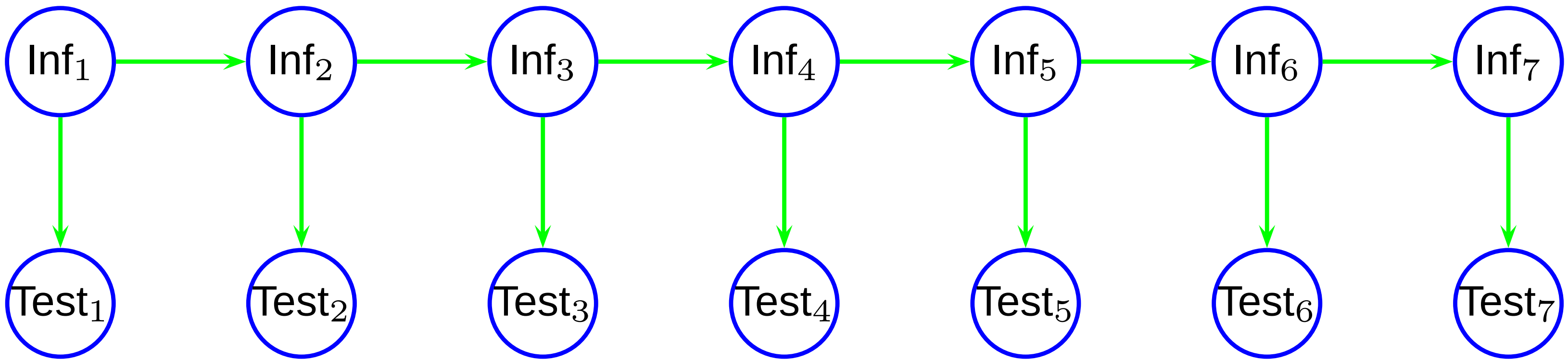
7-day model I

Infections develop over time:



7-day model I

Infections develop over time:



Assumption:

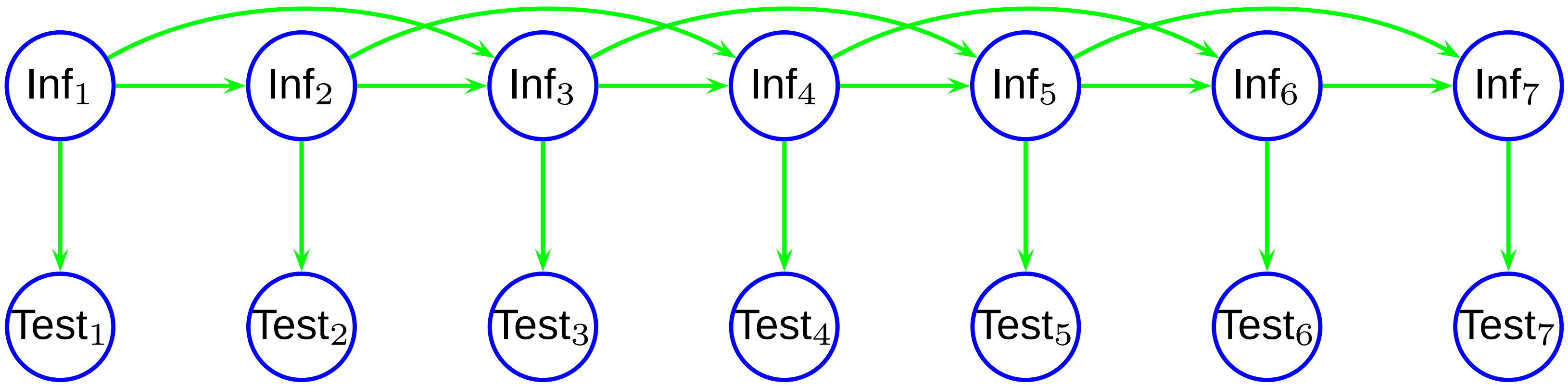
- The Markov property: If I know the present, then the past has no influence on the future, i.e.

  Inf$_{i-1}$ is d-separated from Inf$_{i+1}$ given Inf$_i$.

But what if yesterday's Inf-state has an impact on tomorrow's Inf-state?

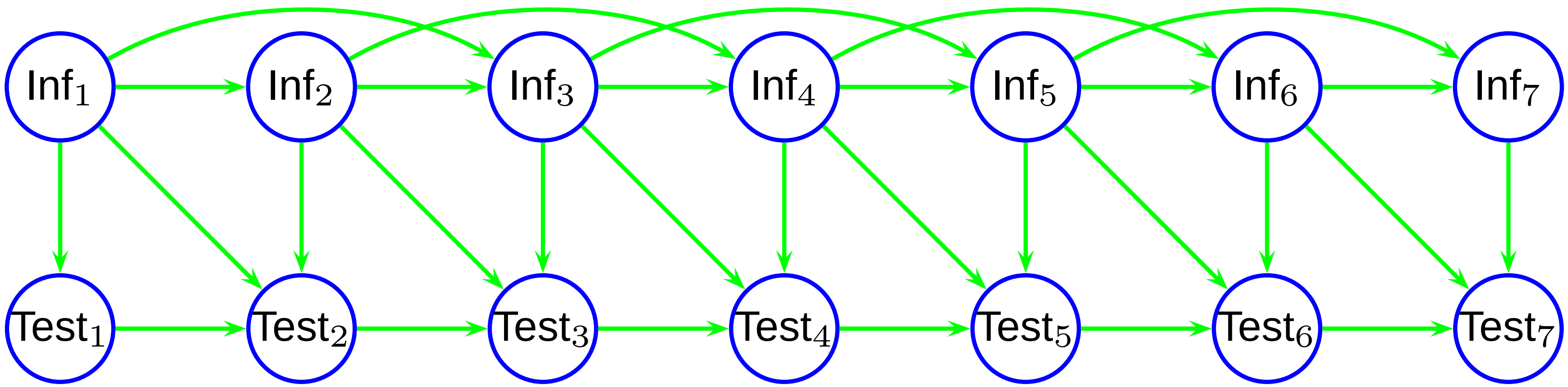
7-day model II

Yesterday's Inf-state has an impact on tomorrow's Inf-state:



7-day model III

The test-failure is dependent on whether or not the test failed yesterday:



Sore throat

I wake up one morning with a sore throat. It may be the beginning of a cold or I may suffer from angina. If it is a severe angina, then I will not go to work. To gain more insight, I can take my temperature and look down my throat for yellow spots.

Sore throat

I wake up one morning with a sore throat. It may be the beginning of a cold or I may suffer from angina. If it is a severe angina, then I will not go to work. To gain more insight, I can take my temperature and look down my throat for yellow spots.

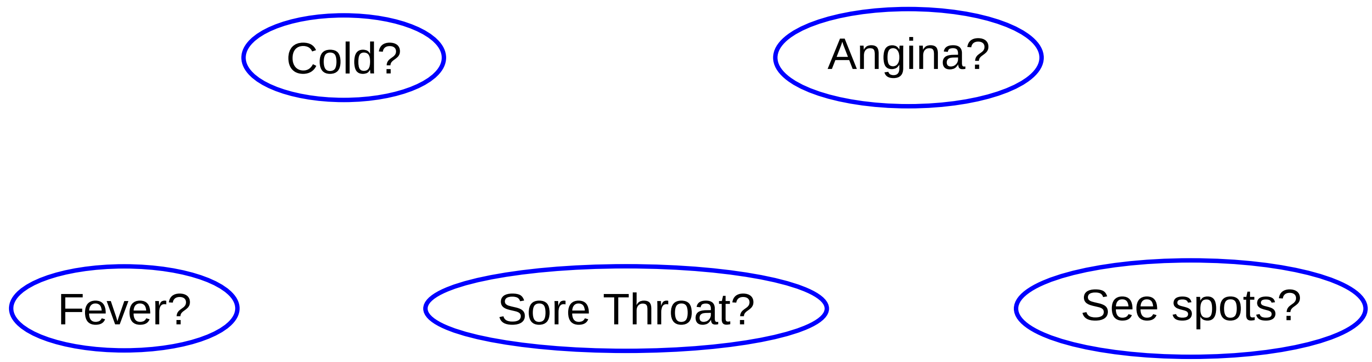
Hypothesis variables:

- Cold? - {n, y}
- Angina? - {no, mild, severe}

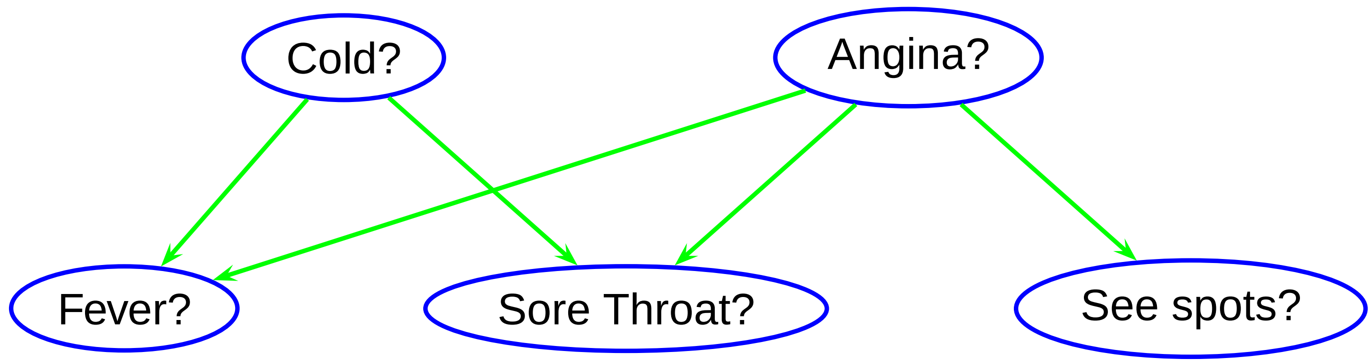
Information variables:

- Sore throat? - {n, y}
- See spots? - {n, y}
- Fever? - {no, low, high}

Model for sore throat

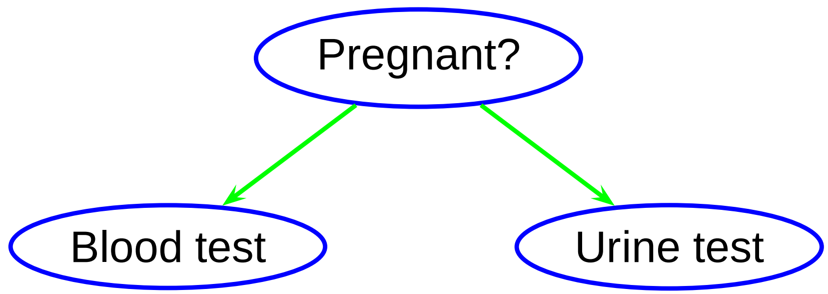


Model for sore throat



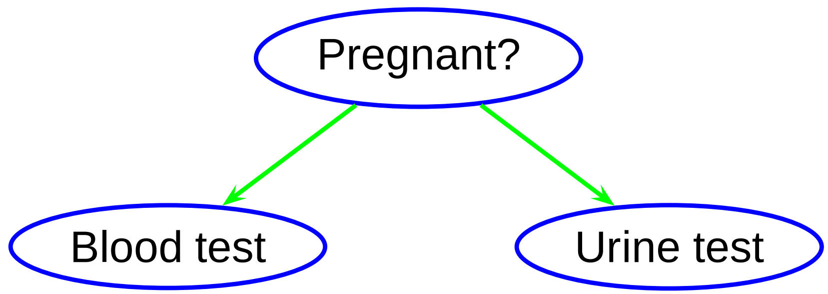
Insemination of a cow

Six weeks after the insemination of a cow, there are two tests: a Blood test and a Urine test.



Insemination of a cow

Six weeks after the insemination of a cow, there are two tests: a Blood test and a Urine test.

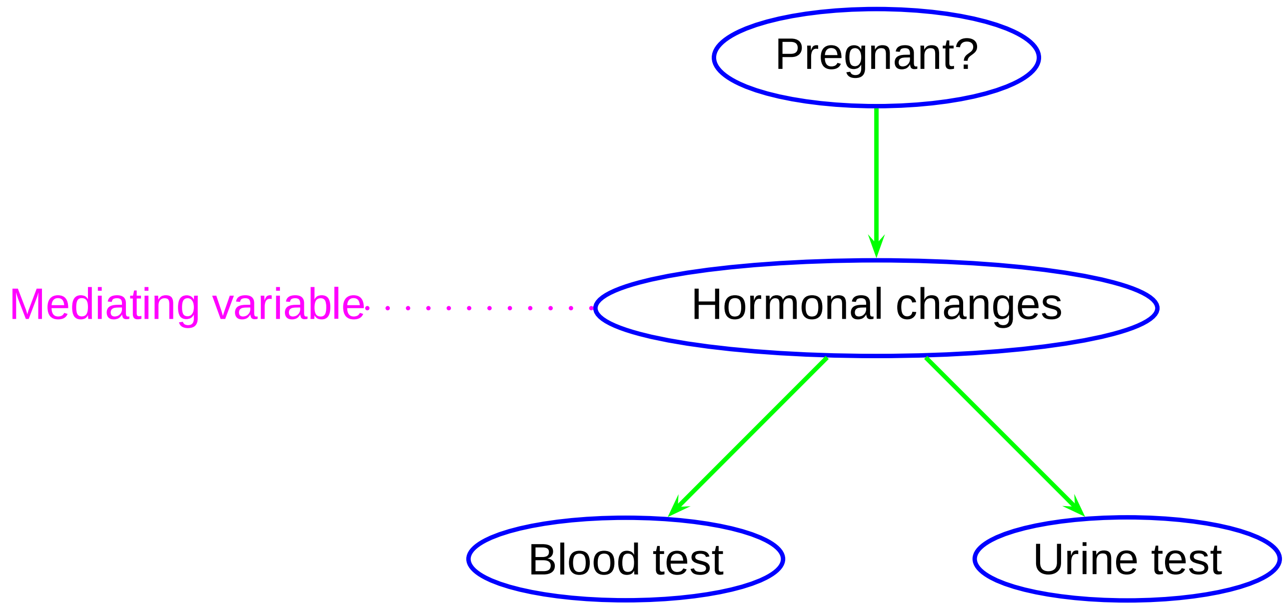


Check the conditional independences:

If we know that the cow is pregnant, will a negative blood test then change our expectation for the urine test?

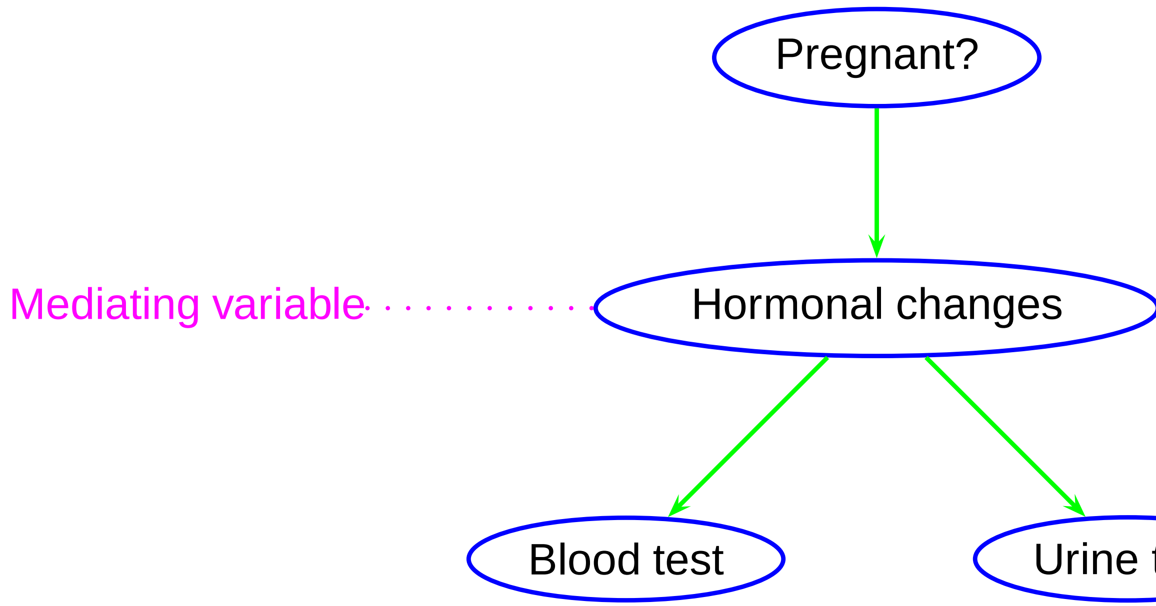
If it will, then the model does not reflect reality!

Insemination of a cow: A more correct model



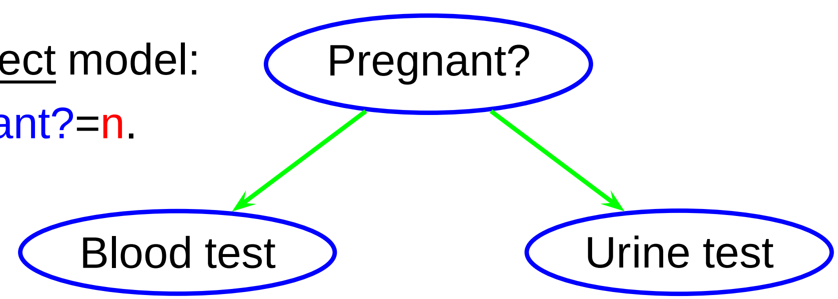
But does this actually make a difference?

Insemination of a cow: A more correct model



But does this actually make a difference?

Assume that both tests are negative in the <u>incorrect</u> model:
This will overestimate the probability for Pregnant?=n.

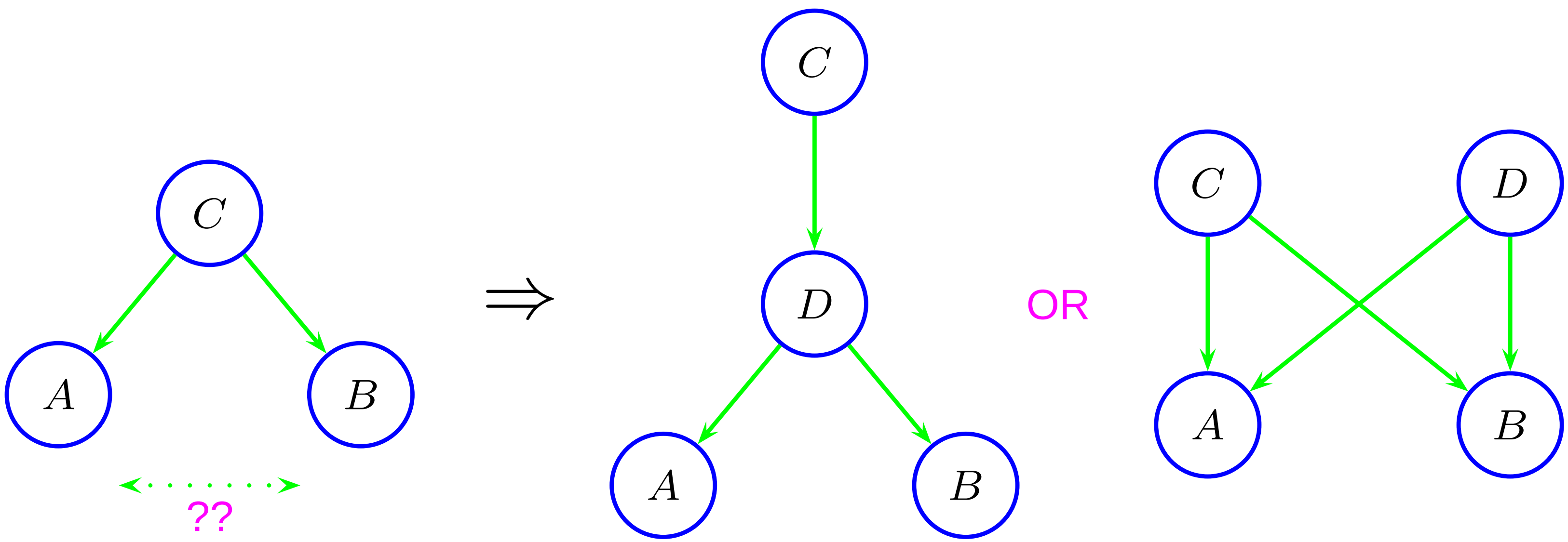


Why mediating variables?

Why do we introduce mediating variables:

- Necessary to catch the correct conditional independences.
- Can ease the specification of the probabilities in the model.

For example: If you find that there is a dependence between two variables A and B, but cannot determine a causal relation: Try with a mediating variable!



A simplified poker game

The game consists of:

- Two players.
- Three cards to each player.
- Two rounds of changing cards (max two cards in the second round)

What kind of hand does my opponent have?

A simplified poker game

The game consists of:

- Two players.
- Three cards to each player.
- Two rounds of changing cards (max two cards in the second round)

What kind of hand does my opponent have?

Hypothesis variable:

OH - {no, 1a, 2v, fl, st, 3v, sf}

Information variables:

FC - {0, 1, 2, 3} and SC - {0, 1, 2}

A simplified poker game

The game consists of:

- Two players.
- Three cards to each player.
- Two rounds of changing cards (max two cards in the second round)

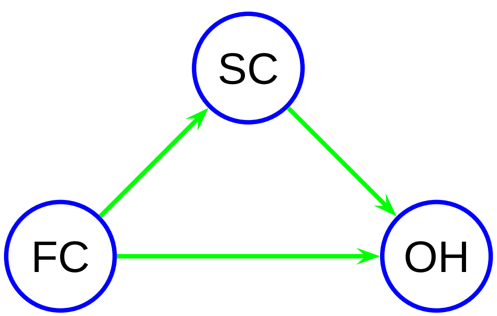
What kind of hand does my opponent have?

Hypothesis variable:

OH - {no, 1a, 2v, fl, st, 3v, sf}

Information variables:

FC - {0, 1, 2, 3} and SC - {0, 1, 2}



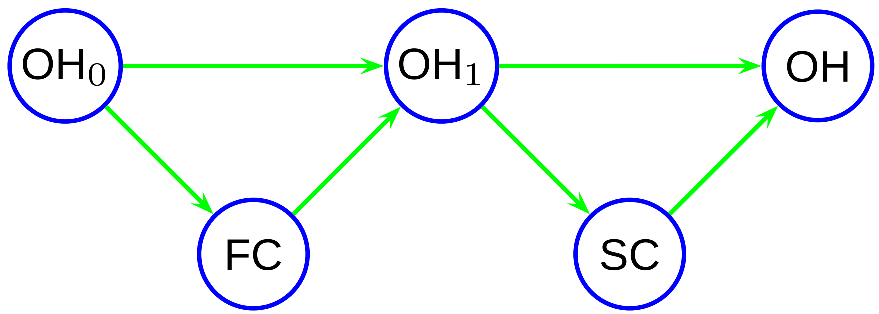
But how do we find:

$P(\text{FC})$, $P(\text{SC}|\text{FC})$ and $P(\text{OH}|\text{SC},\text{FC})$??

A simplified poker game: Mediating variables

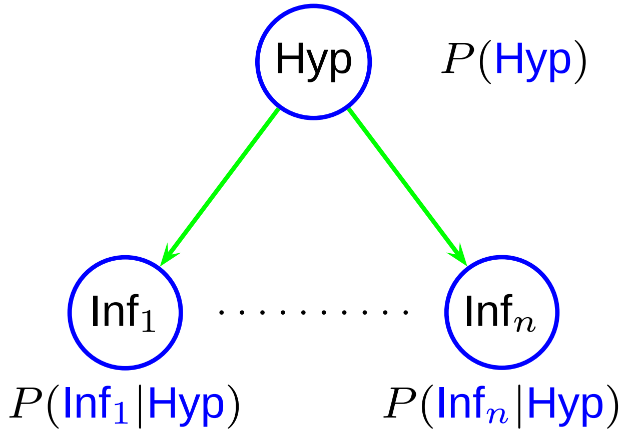
Introduce mediating variables:

- The opponent's initial hand, OH_0.
- The opponent's hand after the first change of cards, OH_1.



Note: The states of OH_0 and OH_1 are different from OH.

Naïve Bayes models



We want the posterior probability of the hypothesis variable Hyp given the observations $\{\text{Inf}_1 = e_1, \ldots, \text{Inf}_n = e_n\}$:

$$P(\text{Hyp}|\text{Inf}_1 = e_1, \ldots, \text{Inf}_n = e_n) = \frac{P(\text{Inf}_1 = e_1, \ldots, \text{Inf}_n = e_n|\text{Hyp})P(\text{Hyp})}{P(\text{Inf}_1 = e_1, \ldots, \text{Inf}_n = e_n)}$$
$$= \mu \cdot P(\text{Inf}_1 = e_1|\text{Hyp}) \cdot \ldots \cdot P(\text{Inf}_n = e_n|\text{Hyp})P(\text{Hyp})$$

Note: The model assumes that the information variables are independent given the hypothesis variable.

Examples is a set of text documents along with their target values. V is the set of all possible target values. This function learns the probability terms $P(w_k|v_j)$, describing the probability that a randomly drawn word from a document in class v_j will be the English word w_k. It also learns the class prior probabilities $P(v_j)$.

1. *collect all words, punctuation, and other tokens that occur in Examples*
   - $Vocabulary \leftarrow$ the set of all distinct words and other tokens occurring in any text document from $Examples$
2. *calculate the required $P(v_j)$ and $P(w_k|v_j)$ probability terms*
   - For each target value v_j in V do
     - $docs_j \leftarrow$ the subset of documents from $Examples$ for which the target value is v_j
     - $P(v_j) \leftarrow \frac{|docs_j|}{|Examples|}$
     - $Text_j \leftarrow$ a single document created by concatenating all members of $docs_j$
     - $n \leftarrow$ total number of distinct word positions in $Text_j$
     - for each word w_k in $Vocabulary$
       - $n_k \leftarrow$ number of times word w_k occurs in $Text_j$
       - $P(w_k|v_j) \leftarrow \frac{n_k+1}{n+|Vocabulary|}$

CLASSIFY_NAIVE_BAYES_TEXT(*Doc*)

Return the estimated target value for the document Doc. a_i denotes the word found in the ith position within Doc.

- *positions* $\leftarrow$ all word positions in *Doc* that contain tokens found in *Vocabulary*
- Return v_{NB}, where

$$v_{NB} = \underset{v_j \in V}{\text{argmax}}\, P(v_j) \prod_{i \in positions} P(a_i|v_j)$$

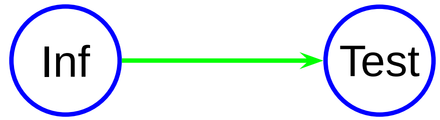
Summary: Catching the structure

1. Identify the relevant events and organize them in variables:
   - Hypothesis variables - Includes the events that are not directly observable.
   - Information variables - Information channels.
2. Determine causal relations between the variables.
3. Check conditional independences in the model.
4. Introduce mediating variables.

Where do the numbers come from?

- Theoretical insight.
- Statistics (large databases)
- Subjective estimates

Infected milk

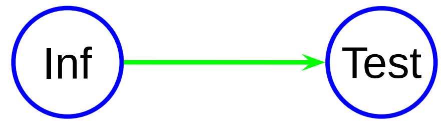


We need the probabilities:

- $P(\text{Test}|\text{Inf})$ - provided by the factory.
- $P(\text{Inf})$ - cow or farm specific.

Determining $P(\text{Inf})$: Assume that the farmer has 50 cows. The milk is poured into a container, and the dairy tests the milk with a very precise test. In average, the milk is infected once per month.

Infected milk



We need the probabilities:

- $P($Test$|$Inf$)$ - provided by the factory.
- $P($Inf$)$ - cow or farm specific.

Determining $P($Inf$)$: Assume that the farmer has 50 cows. The milk is poured into a container, and the dairy tests the milk with a very precise test. In average, the milk is infected once per month.

Calculations:

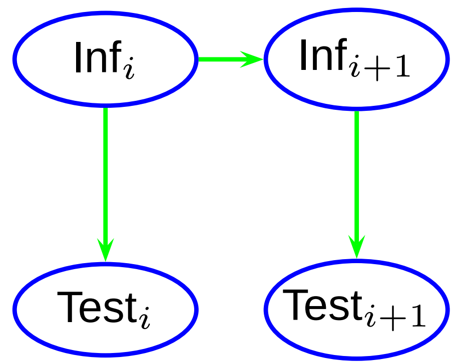
$$P(\text{\#Cows-infected} \geq 1) = \frac{1}{30} \quad \text{hence} \quad P(\text{\#Cows-infected} < 1) = 1 - \frac{1}{30} = \frac{29}{30}.$$

If $P(\text{Inf} = y) = x$, then $P(\text{Inf} = n) = (1 - x)$ and:

$$(1-x)^{50} = \frac{29}{30} \Leftrightarrow x = 1 - \left(\frac{29}{30}\right)^{\frac{1}{50}} \approx 0.00067$$

7-day model I

Infections develop over time:

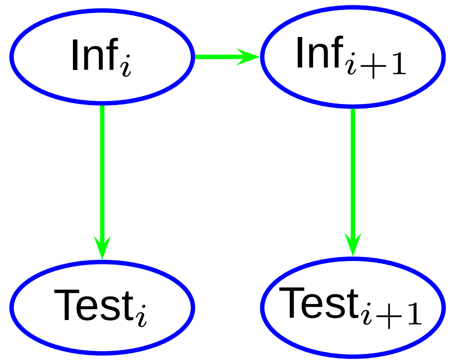


From experience we have:

- Risk of becoming infected? 0.0002
- Chance of getting cured from one day to another? 0.3

7-day model I

Infections develop over time:



From experience we have:

- Risk of becoming infected? 0.0002
- Chance of getting cured from one day to another? 0.3

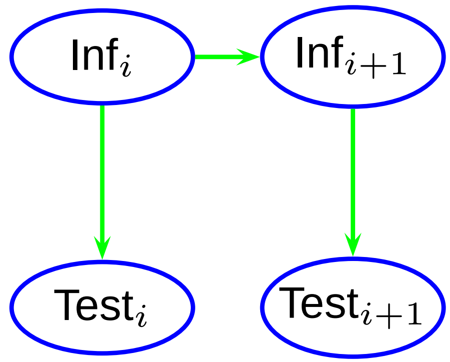
This gives us:

		Inf_i	
		y	n
Inf_{i+1}	y		
	n		

$P(\text{Inf}_{i+1} | \text{Inf}_i)$

7-day model I

Infections develop over time:



From experience we have:

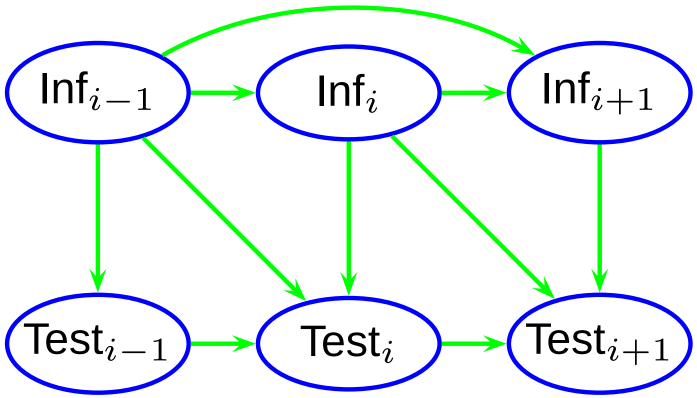
- Risk of becoming infected? 0.0002
- Chance of getting cured from one day to another? 0.3

This gives us:

		Inf_i y	Inf_i n
Inf_{i+1}	y	0.7	0.0002
	n	0.3	0.9998

$P(\text{Inf}_{i+1} \mid \text{Inf}_i)$

7-day model II



		Inf$_{i-1}$	
		y	n
Inf$_i$	y	0.6	1
	n	0.0002	0.0002

$P(\text{Inf}_{i+1} = y \mid \text{Inf}_{i-1}, \text{Inf}_i)$

That is:

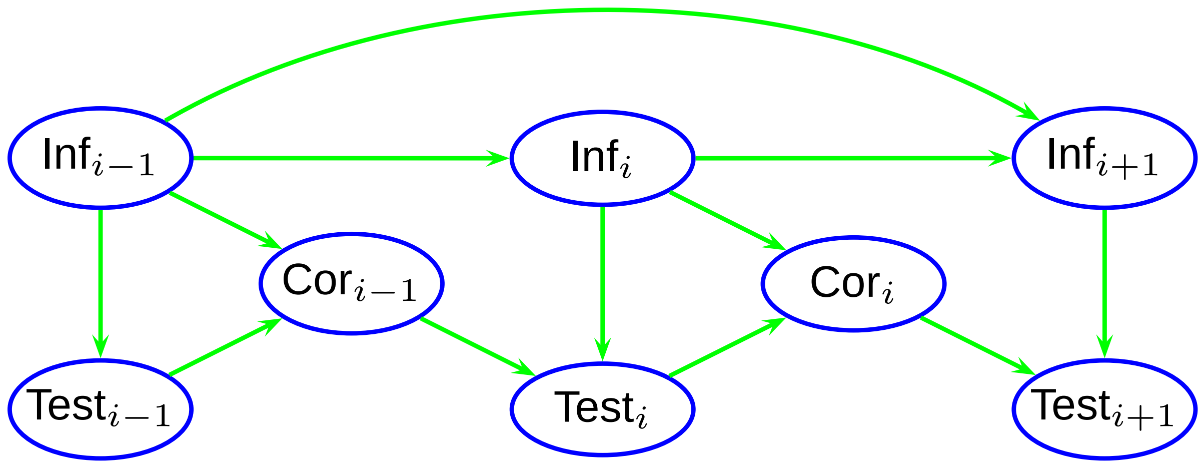
- An infection always lasts at least two days.
- After two days, the chance of being cured is 0.4.

However, we also need to specify $P(\text{Test}_{i+1} \mid \text{Inf}_{i+1}, \text{Test}_i, \text{Inf}_i)$:

- A correct test has a 99.9% of being correct the next time.
- An incorrect test has a 90% of being incorrect the next time.

This can be done much easier by introducing mediating variables!

7-day model III



We need the probabilities:

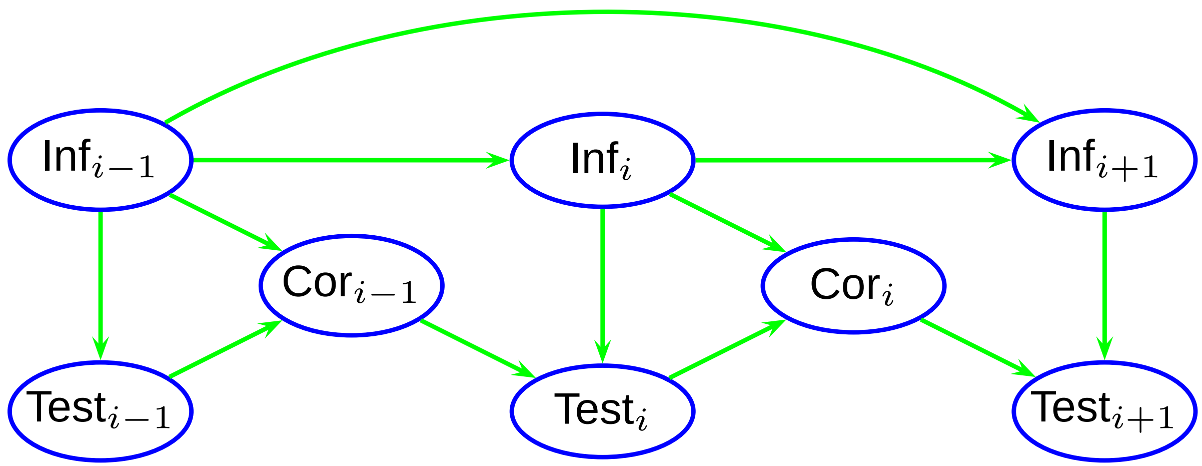
		Inf$_i$	
		y	n
Test$_i$	Pos		
	Neg		

$P(\text{Cor}_i = y | \text{Inf}_i, \text{Test}_i)$

		Inf$_i$	
		y	n
Cor$_{i-1}$	y		
	n		

$P(\text{Test}_i = Pos | \text{Inf}_i, \text{Cor}_{i-1})$

7-day model III



We need the probabilities:

		Inf$_i$ y	Inf$_i$ n
Test$_i$	Pos	1	0
	Neg	0	1

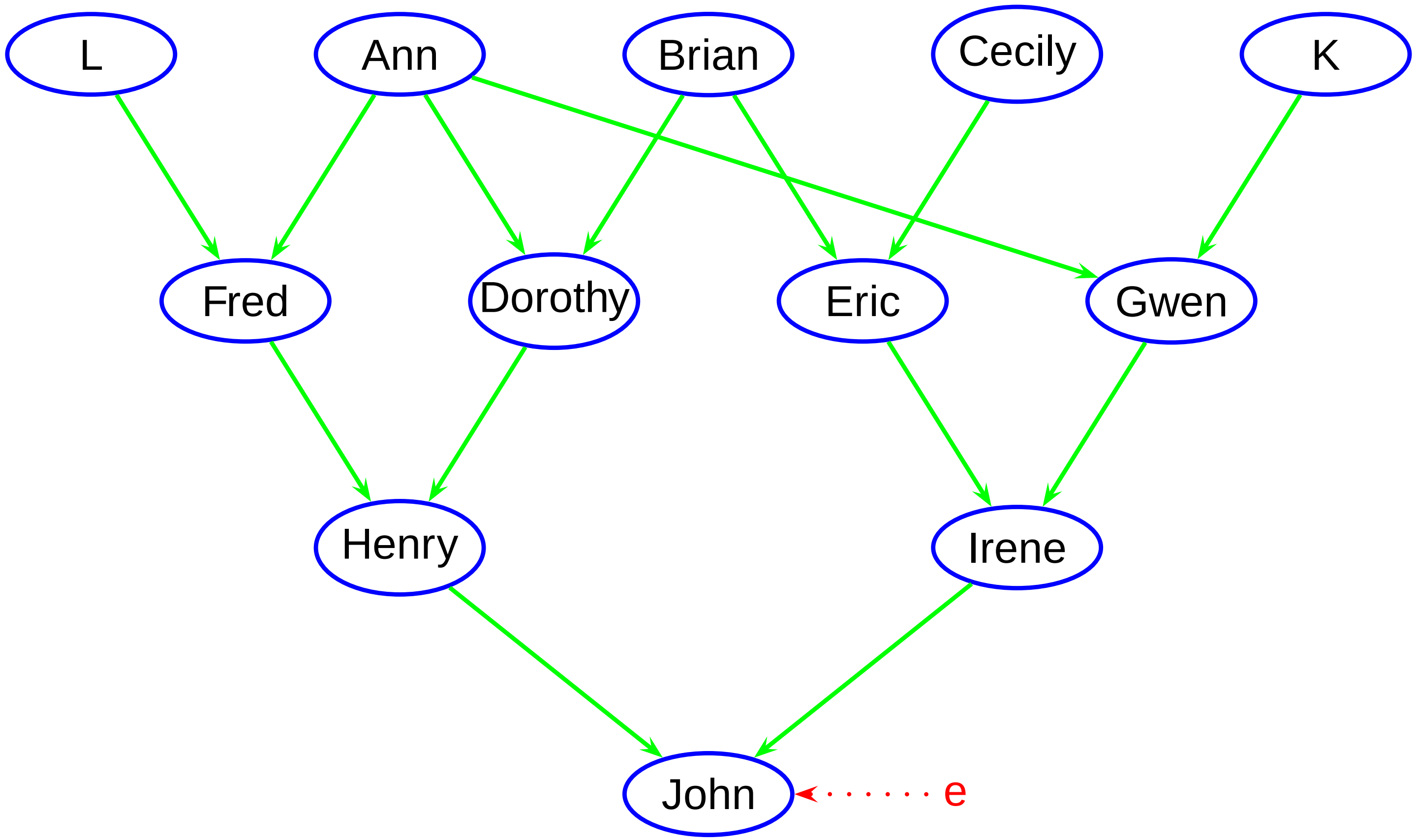
$P(\text{Cor}_i = y | \text{Inf}_i, \text{Test}_i)$

		Inf$_i$ y	Inf$_i$ n
Cor$_{i-1}$	y	0.999	0.001
	n	0.1	0.9

$P(\text{Test}_i = Pos | \text{Inf}_i, \text{Cor}_{i-1})$

Stud farm

Genealogical structure for the horses in a stud farm:



We get evidence e that John is sick.

Stud farm: Conditional probabilities I

The disease is carried by a recessive gene:

aa: sick,   aA: Carrier,   AA: Healthy

We should specify the probabilities:

		Mother		
		aa	aA	AA
	aa	(, ,)	(, ,)	(, ,)
Father	aA	(, ,)	(, ,)	(, ,)
	AA	(, ,)	(, ,)	(, ,)

$P(\text{Offspring}|\text{Father}, \text{Mother})$

Stud farm: Conditional probabilities I

The disease is carried by a recessive gene:

aa: sick, aA: Carrier, AA: Healthy

We should specify the probabilities:

		Mother		
		aa	aA	AA
	aa	(1, 0, 0)	(0.5, 0.5, 0)	(0, 1, 0)
Father	aA	(0.5, 0.5, 0)	(0.25, 0.5, 0.25)	(0, 0.5, 0.5)
	AA	(0, 1, 0)	(0, 0.5, 0.5)	(0, 0, 1)

$P(\text{Offspring}|\text{Father}, \text{Mother})$

Stud farm: Conditional probabilities II

But the other horses are not sick:

- John: aa, aA, AA.
- Other horses: aA, AA.

Prior probabilities:

$$P(aA) = 0.01 \quad \text{and} \quad P(AA) = 0.99.$$

Conditional probabilities:

		Irene	
		aA	AA
Henry	aA	(0.25,0.5,0.25)	(0,0.5,0.5)
	AA	(0,0.5,0.5)	(0,0,1)

$P(\text{John}|\text{Henry}, \text{Irene})$

		Mother	
		aA	AA
Father	aA	(,)	(,)
	AA	(,)	(,)

$P(\text{Offspring}|\text{Father}, \text{Mother})$

Stud farm: Conditional probabilities II

But the other horses are not sick:

- John: aa, aA, AA.
- Other horses: aA, AA.

Prior probabilities:

$$P(aA) = 0.01 \quad \text{and} \quad P(AA) = 0.99.$$

Conditional probabilities:

		Irene aA	Irene AA
Henry	aA	(0.25,0.5,0.25)	(0,0.5,0.5)
Henry	AA	(0,0.5,0.5)	(0,0,1)

$P(\text{John}|\text{Henry}, \text{Irene})$

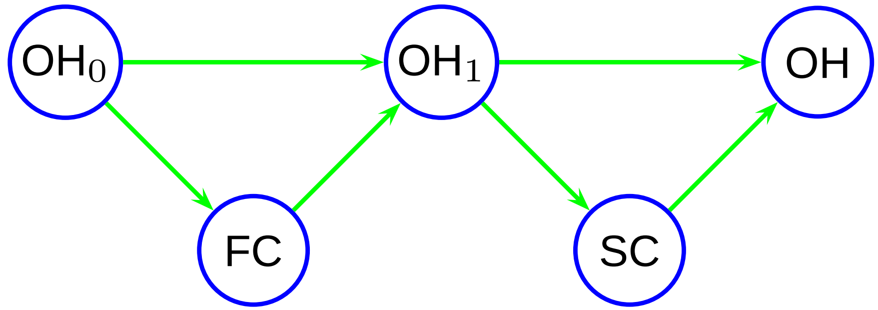
		Mother aA	Mother AA
Father	aA	(2/3,1/3)	(0.5,0.5)
Father	AA	(0.5,0.5)	(0,1)

$P(\text{Offspring}|\text{Father}, \text{Mother})$

Drop the first state and normalize:

$$(\boxed{0.25}, 0.5, 0.25) \Rightarrow (2/3, 1/3)$$

A simplified poker game I

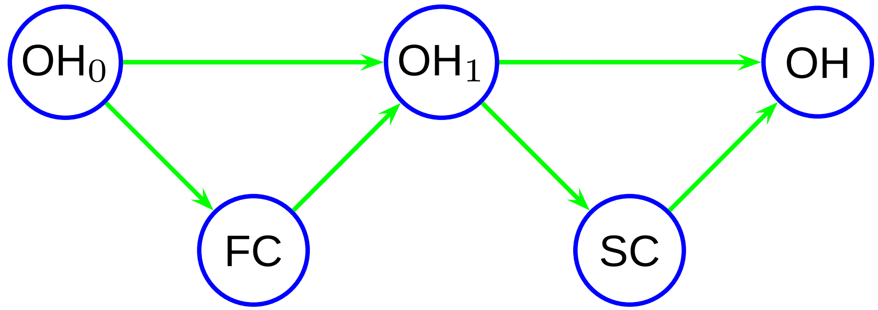


In order to find:

$$P(\text{OH}_0) = (\_\_\text{No}, \_\_\text{1a}, \_\_\text{2cons}, \_\_\text{2s}, \_\_\text{2v}, \_\_\text{fl}, \_\_\text{st}, \_\_\text{3v}, \_\_\text{sf})$$

we have to go into combinatorics: $\frac{\#\text{good}}{\binom{52}{3}}$.

A simplified poker game I



$$P(\text{OH}_0) \approx (0.167_{\text{No}}, 0.045_{\text{1a}}, 0.064_{\text{2cons}}, 0.466_{\text{2s}}, 0.169_{\text{2v}}, 0.049_{\text{fl}}, 0.035_{\text{st}}, 0.002_{\text{3v}}, 0.002_{\text{sf}})$$

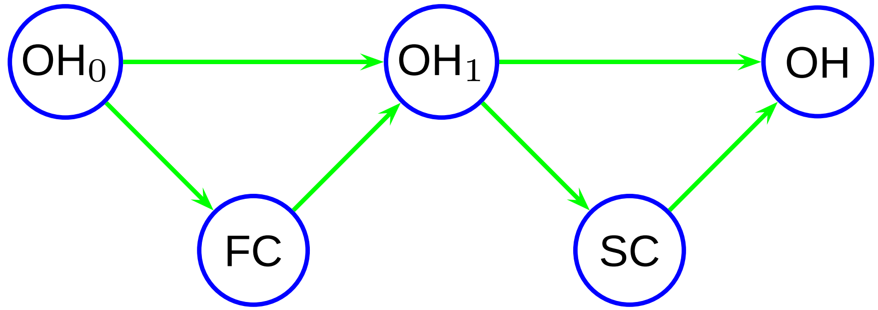
we have to go into combinatorics: $\frac{\#\text{good}}{\binom{52}{3}}$. For example,

$$P(\text{OH}_0 = \text{st}) = \frac{54 \cdot 4 \cdot 4 - 52}{\binom{52}{3}}.$$

Similar considerations apply to $P(\text{OH}_1 | \text{OH}_0, \text{FC})$. E.g.

$$P(\text{OH}_1 | \text{2cons}, 1) = (0_{\text{No}}, 0_{\text{1a}}, 0.374_{\text{2cons}}, 0.367_{\text{2s}}, 0.122_{\text{2v}}, 0_{\text{fl}}, 0.163_{\text{st}}, 0_{\text{3v}}, 0_{\text{sf}})$$

A simplified poker game II



Theoretical considerations are not enough:

$P(FC|OH_0)$ = What is my opponents strategy?

Assume the strategy:

$$\text{no} \rightarrow 3$$

$$\text{1a} \rightarrow 2$$

$$\text{2s} \vee \text{2cons} \vee \text{2v} \rightarrow 1$$

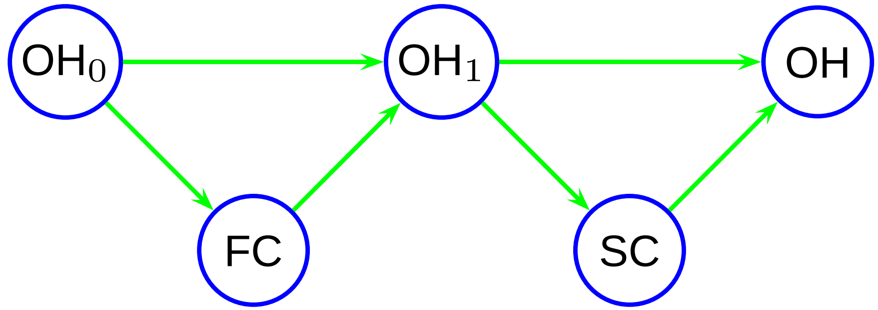
$$\text{2cons} \wedge \text{2s} \rightarrow 1 \qquad \text{(Keep 2s)}$$

$$\text{2cons} \wedge \text{2v} \vee \text{2s} \wedge \text{2v} \rightarrow 1 \qquad \text{(Keep 2v)}$$

$$\text{fl} \vee \text{st} \vee \text{3v} \vee \text{sf} \rightarrow 0$$

Note: the states 2cons $\wedge$ 2s, 2cons $\wedge$ 2v, 2s $\wedge$ 2v are redundant.

A simplified poker game II



Theoretical considerations are not enough:

P(FC|OH$_0$) = What is my opponents strategy?

Assume the strategy:

$$\text{no} \rightarrow 3$$

$$\text{1a} \rightarrow 2$$

$$\text{2s} \vee \text{2cons} \vee \text{2v} \rightarrow 1$$

$$\text{2cons} \wedge \text{2s} \rightarrow 1 \qquad \text{(Keep 2s)}$$

$$\text{2cons} \wedge \text{2v} \vee \text{2s} \wedge \text{2v} \rightarrow 1 \qquad \text{(Keep 2v)}$$

$$\text{fl} \vee \text{st} \vee \text{3v} \vee \text{sf} \rightarrow 0$$

Note: the states 2cons $\wedge$ 2s, 2cons $\wedge$ 2v, 2s $\wedge$ 2v are redundant.

However, knowing my system my opponent may “bluff”.

Transmission of symbol strings

A language L over $\{a, b\}$ is transmitted through a channel. Each word is surrounded by c. In the transmission some characters may be corrupted by noise and may be confused with others.

A five-letter word has been transmitted.

Hypothesis variables:

Information variables:

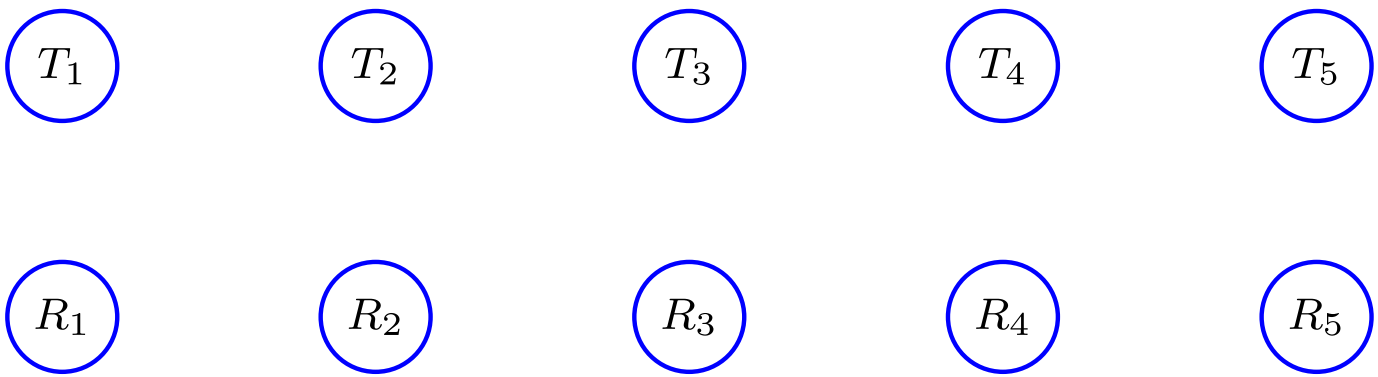
Transmission of symbol strings

A language L over $\{a, b\}$ is transmitted through a channel. Each word is surrounded by c. In the transmission some characters may be corrupted by noise and may be confused with others.

A five-letter word has been transmitted.

Hypothesis variables: T_1, T_2, T_3, T_4, T_5 (States: a, b)

Information variables: R_1, R_2, R_3, R_4, R_5 (States: a, b, c)



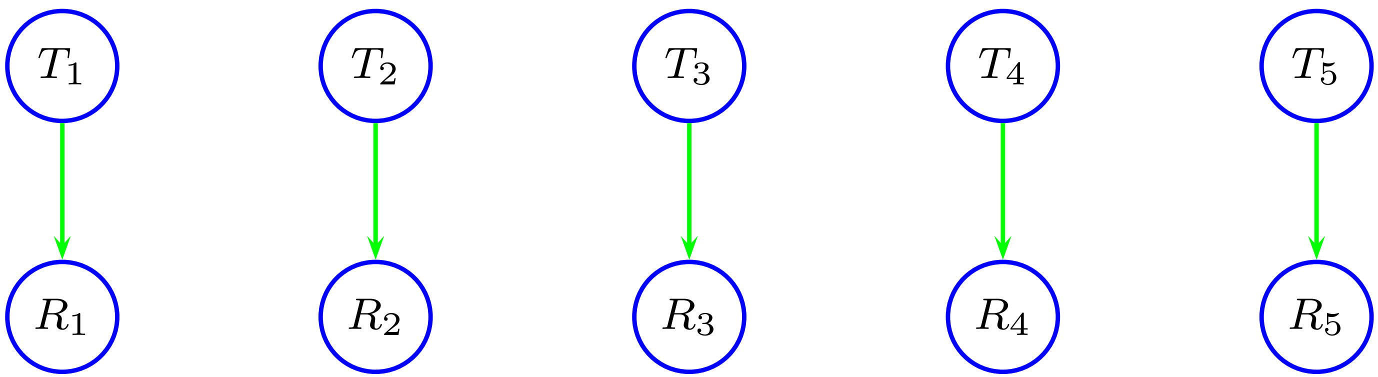
Transmission of symbol strings

A language L over $\{a, b\}$ is transmitted through a channel. Each word is surrounded by c. In the transmission some characters may be corrupted by noise and may be confused with others.

A five-letter word has been transmitted.

Hypothesis variables: T_1, T_2, T_3, T_4, T_5 (States: a, b)

Information variables: R_1, R_2, R_3, R_4, R_5 (States: a, b, c)

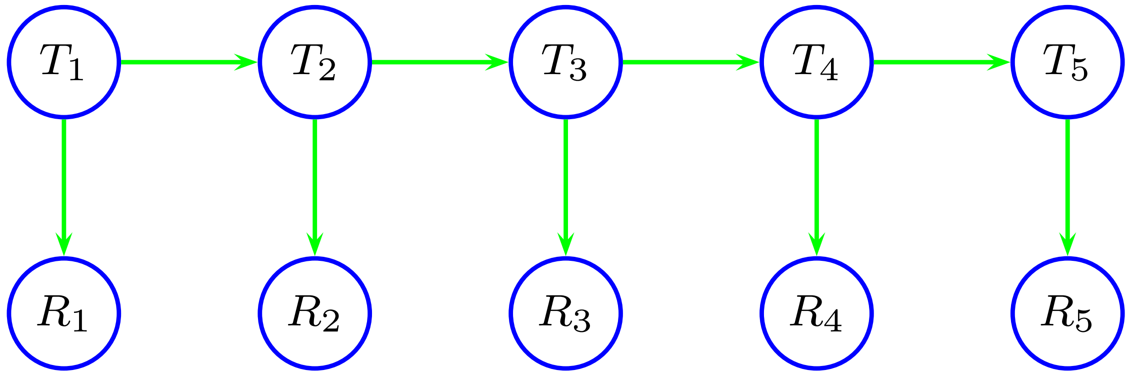


$P(R_i|T_i)$ can be determined through statistics:

		R_i		
		$R_i = a$	$R_i = b$	$R_i = c$
T_i	a	0.8	0.1	0.1
	b	0.15	0.8	0.05

Are the T_i's independent?

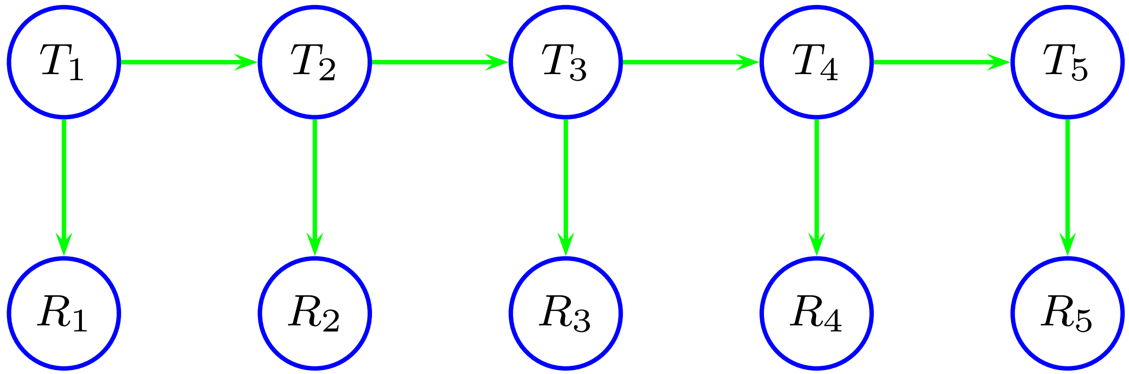
Transmission of symbol strings



To find $P(T_{i+1}|T_i)$: Look at the permitted words and their frequencies.

First 2 \ Last 3	aaa	aab	aba	abb	baa	bab	bba	bbb
aa	0.017	0.021	0.019	0.019	0.045	0.068	0.045	0.068
ab	0.033	0.040	0.037	0.038	0.011	0.016	0.010	0.015
ba	0.011	0.014	0.010	0.010	0.031	0.046	0.031	0.045
bb	0.050	0.060	0.057	0.057	0.016	0.023	0.015	0.023

Transmission of symbol strings

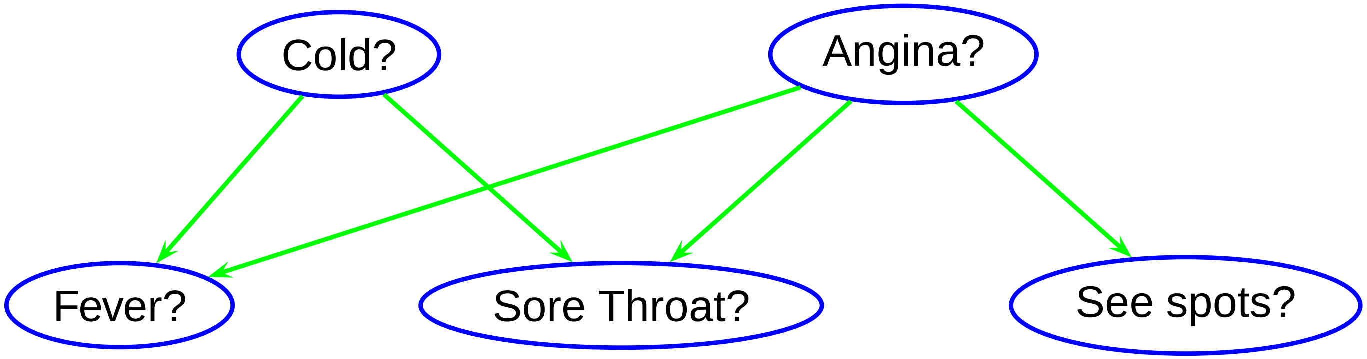


To find $P(T_{i+1}|T_i)$: Look at the permitted words and their frequencies.

		Last 3							
		aaa	aab	aba	abb	baa	bab	bba	bbb
First 2	aa	0.017	0.021	0.019	0.019	0.045	0.068	0.045	0.068
	ab	0.033	0.040	0.037	0.038	0.011	0.016	0.010	0.015
	ba	0.011	0.014	0.010	0.010	0.031	0.046	0.031	0.045
	bb	0.050	0.060	0.057	0.057	0.016	0.023	0.015	0.023

$$P(T_2 = a|T_1 = a) = \frac{P(T_2 = a, T_1 = a)}{P(T_1 = a)} = \frac{0.017 + 0.021 + \cdots + 0.068}{0.017 + \cdots + 0.068 + 0.033 + \cdots 0.015} = 0.6$$

Cold or angina? I



Subjective estimates:

$$P(\text{Cold?}) = (0.97, 0.03)$$

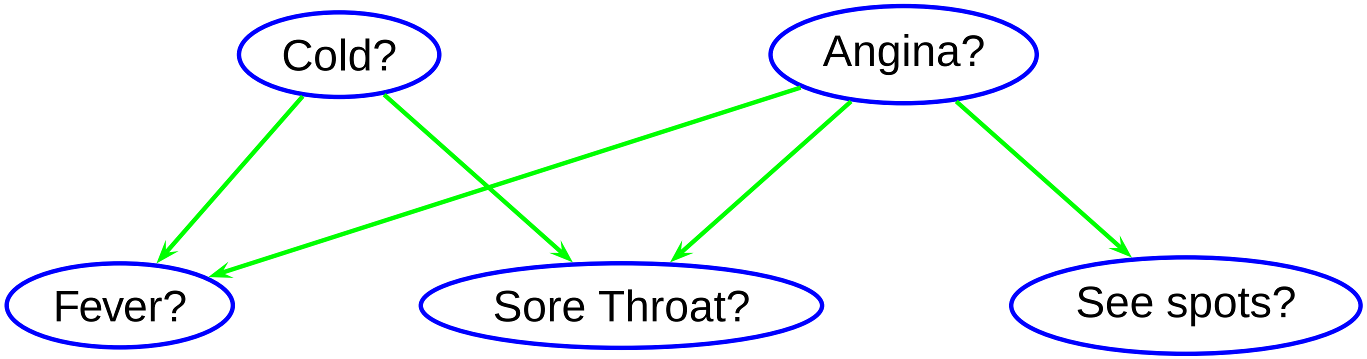
$$P(\text{Angina?}) = (0.993, 0.005, 0.002)$$

		Angina? no	Angina? mild	Angina? severe
See spots?	no	1	1	0.1
	yes	0	0	0.9

$P(\text{See spots?}|\text{Angina?})$

But how do we find e.g. $P(\text{Sore throat?}|\text{Angina?}, \text{Cold?})$?

Cold or angina? II



- If neither Cold? nor Angina?, then $P(\text{Sore throat?} = \text{y}) = 0.05$.
- If only Cold?, then $P(\text{Sore throat?} = \text{y}) = 0.4$.
- If only Angina? $=$ mild, then $P(\text{Sore throat?} = \text{y}) = 0.7$.
- If Angina? $=$ severe, then $P(\text{Sore throat?} = \text{y}) = 1$.

		Angina?		
		no	mild	severe
Cold?	no	0.05	0.7	1
	yes	0.4	??	1

$P(\text{Sore throat?} = \text{y} | \text{Cold?}, \text{Angina?})$

Cold or angina? III

We have the partial specification:

		Angina? no	Angina? mild	Angina? severe
Cold?	no	0.05	0.7	1
	yes	0.4	??	1

$P(\text{Sore throat?} = \text{yes} | \text{Cold?}, \text{Angina?})$

In order to find $P(\text{Sore throat} = \text{yes} | \text{Cold?} = \text{yes}, \text{Angina?} = \text{mild})$ assume that:

Out of 100 mornings, I have a "background" sore throat on 5 of them.

- 95 left: 40% "cold-sore" = 38
- 57 left: 70% "mild angina-sore" = 39.9

In total: $5 + 38 + 39.9 = 82.9 \rightarrow 85$.

Cold or angina? III

We have the partial specification:

		Angina? no	Angina? mild	Angina? severe
Cold?	no	0.05	0.7	1
	yes	0.4	??	1

$P(\text{Sore throat?} = \text{yes} \mid \text{Cold?}, \text{Angina?})$

In order to find $P(\text{Sore throat} = \text{yes} \mid \text{Cold?} = \text{yes}, \text{Angina?} = \text{mild})$ assume that:

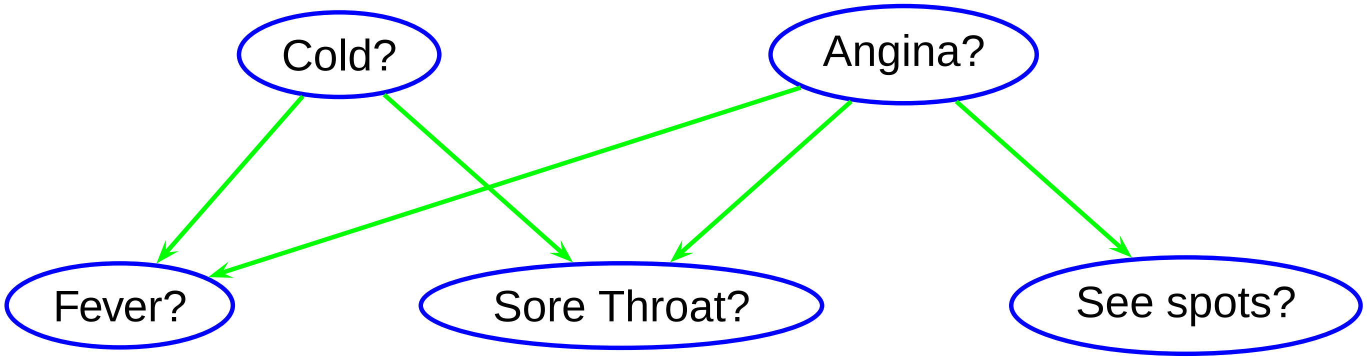
Out of 100 mornings, I have a "background" sore throat on 5 of them.

- 95 left: 40% "cold-sore" = 38
- 57 left: 70% "mild angina-sore" = 39.9

In total: $5 + 38 + 39.9 = 82.9 \rightarrow 85$.

		Angina? no	Angina? mild	Angina? severe
Cold?	no	0.05	0.7	1
	yes	0.4	0.85	1

Cold or angina? I



Subjective estimates:

$$P(\text{Cold?}) = (0.97, 0.03)$$

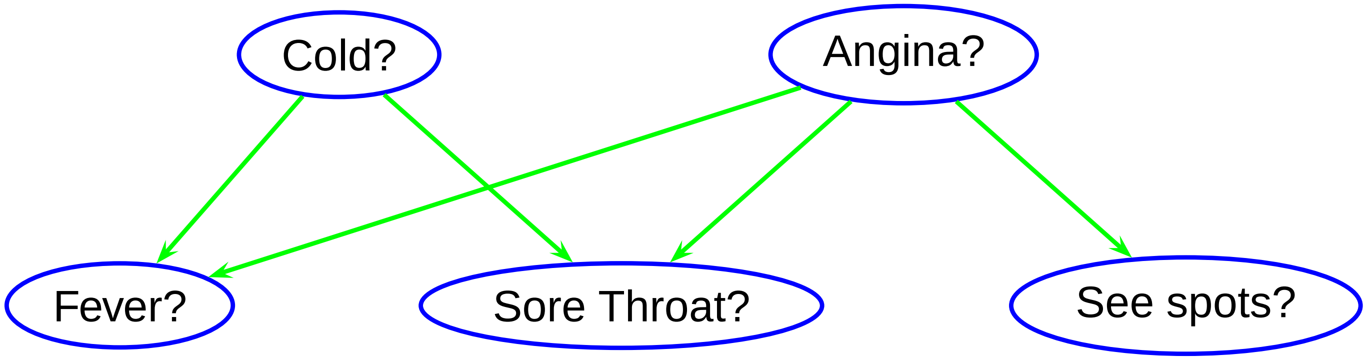
$$P(\text{Angina?}) = (0.993, 0.005, 0.002)$$

		Angina? no	Angina? mild	Angina? severe
See spots?	no	1	1	0.1
	yes	0	0	0.9

$P(\text{See spots?}|\text{Angina?})$

But how do we find e.g. $P(\text{Sore throat?}|\text{Angina?}, \text{Cold?})$?

Cold or angina? II



- If neither Cold? nor Angina?, then $P(\text{Sore throat?} = \text{y}) = 0.05$.
- If only Cold?, then $P(\text{Sore throat?} = \text{y}) = 0.4$.
- If only $\text{Angina?} = \text{mild}$, then $P(\text{Sore throat?} = \text{y}) = 0.7$.
- If $\text{Angina?} = \text{severe}$, then $P(\text{Sore throat?} = \text{y}) = 1$.

		Angina?		
		no	mild	severe
Cold?	no	0.05	0.7	1
	yes	0.4	??	1

$P(\text{Sore throat?} = \text{y} | \text{Cold?}, \text{Angina?})$

Cold or angina? III

We have the partial specification:

		Angina? no	Angina? mild	Angina? severe
Cold?	no	0.05	0.7	1
	yes	0.4	??	1

$P(\text{Sore throat?} = \text{yes} | \text{Cold?}, \text{Angina?})$

In order to find $P(\text{Sore throat} = \text{yes} | \text{Cold?} = \text{yes}, \text{Angina?} = \text{mild})$ assume that:

Out of 100 mornings, I have a “background” sore throat on 5 of them.

- 95 left: 40% “cold-sore” = 38
- 57 left: 70% “mild angina-sore” = 39.9

In total: $5 + 38 + 39.9 = 82.9 \rightarrow 85$.

Cold or angina? III

We have the partial specification:

		Angina? no	Angina? mild	Angina? severe
Cold?	no	0.05	0.7	1
	yes	0.4	??	1

$P(\text{Sore throat?} = \text{yes} | \text{Cold?}, \text{Angina?})$

In order to find $P(\text{Sore throat} = \text{yes} | \text{Cold?} = \text{yes}, \text{Angina?} = \text{mild})$ assume that:

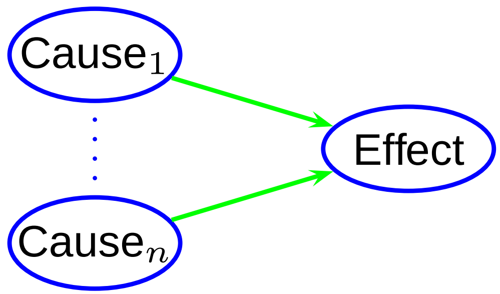
Out of 100 mornings, I have a “background” sore throat on 5 of them.

- 95 left: 40% “cold-sore” = 38
- 57 left: 70% “mild angina-sore” = 39.9

In total: $5 + 38 + 39.9 = 82.9 \rightarrow 85$.

		Angina? no	Angina? mild	Angina? severe
Cold?	no	0.05	0.7	1
	yes	0.4	0.85	1

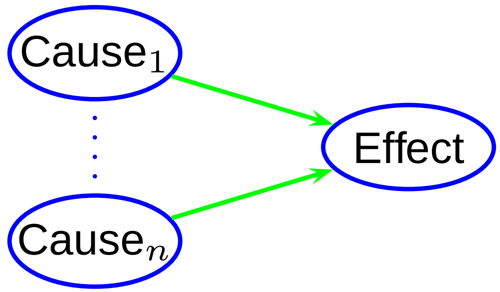
Several independent causes, in general



Cause$_i$ results in Effect with probability x_i.

$P($Effect $=$ yes$|$Combination of causes$)$ = ??

Several independent causes, in general



Cause$_i$ results in Effect with probability x_i.

$P(\text{Effect} = \text{yes} | \text{Combination of causes})$ = ??

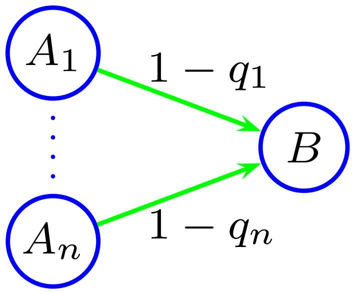
Way to look at it: Cause$_i$ results in Effect unless it is inhibited by "something".
The Inhibitor has probability $q_i = 1 - x_i$.

Assumption: The Inhibitors are independent.
That is, the probability of "Inh$_i$ and Inh$_j$" = $q_i q_j$.

Thus,

$$P(\text{Effect} = \text{yes} | \text{Cause}_i, \text{Cause}_j, \text{Cause}_k) = 1 - q_i q_j q_k$$

Noisy or



All nodes are binary.

All causes for B are listed explicitly.

In general:

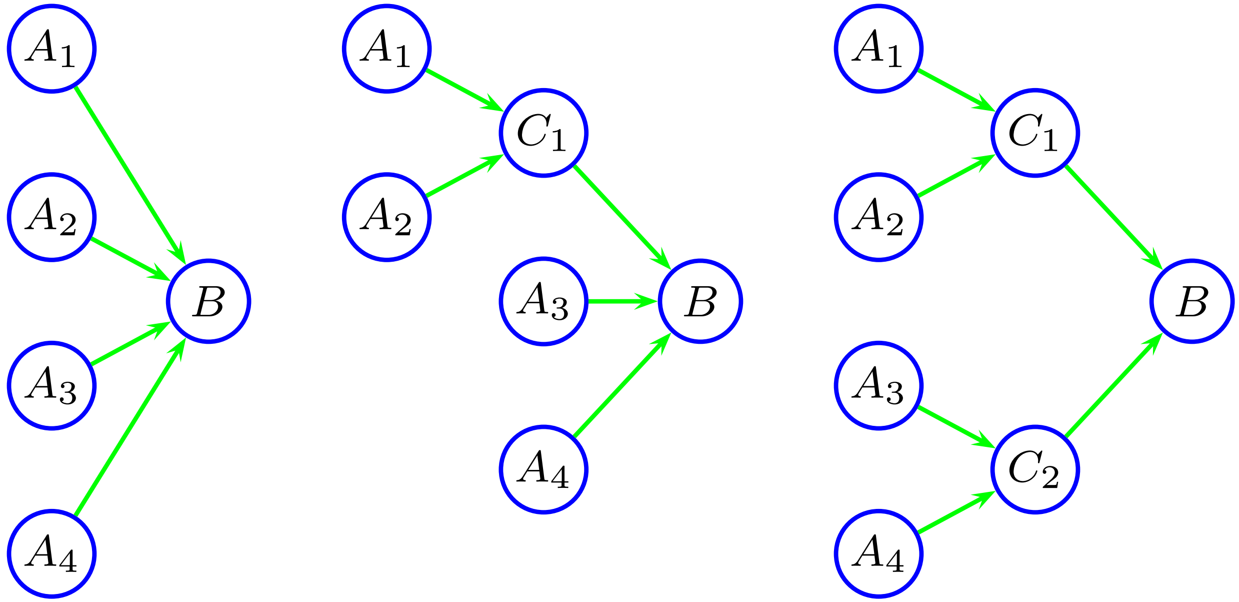
$$P(B = y \mid A_{i_1} = \cdots = A_{i_k} = y, \text{the rest} = n) = 1 - q_{i_1} \cdots q_{i_k}$$

If only A_1 and A_2 are on:

$$P(B = y \mid A_1 = y, A_2 = y, \text{the rest} = n) = 1 - q_1 q_2$$

Note: If $P(B = y \mid \text{All} = n) = x > 0$, then introduce a background cause C which is always on, and $q_c = 1 - x$.

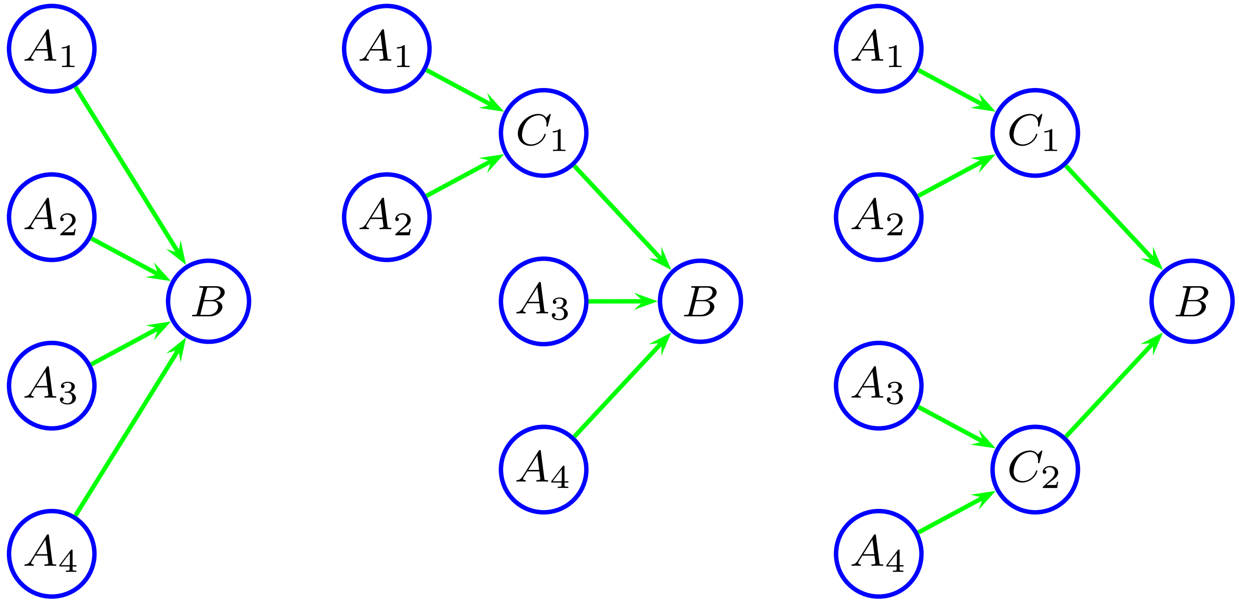
Divorcing



$3^5 = 243$ $\quad$ $3^4 + 3^3 = 108$ $\quad$ $3 \cdot 3^3 = 81$

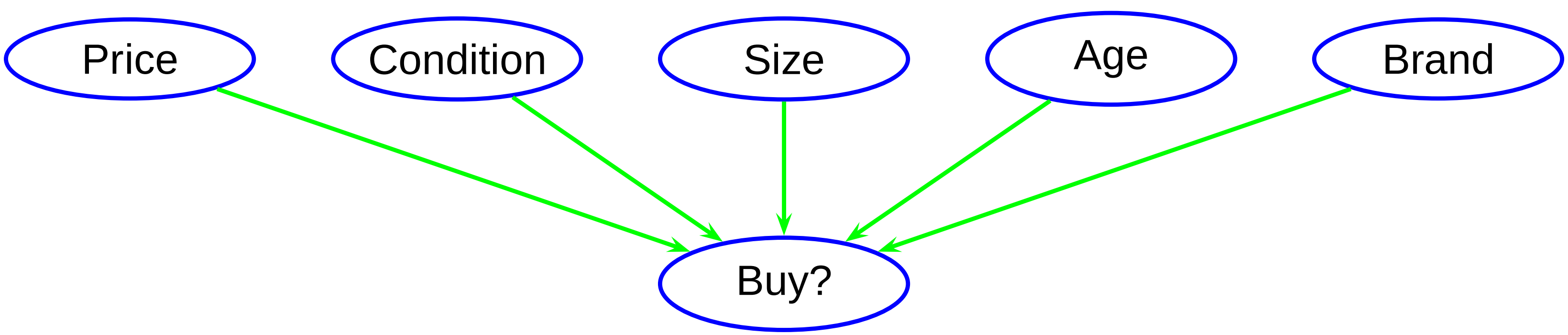
Can this always be done?

Divorcing

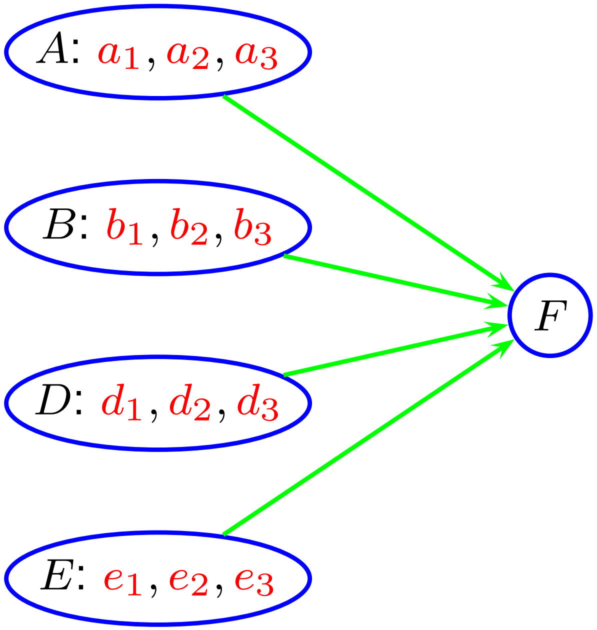


$3^5 = 243$ $\quad$ $3^4 + 3^3 = 108$ $\quad$ $3 \cdot 3^3 = 81$

Can this always be done?

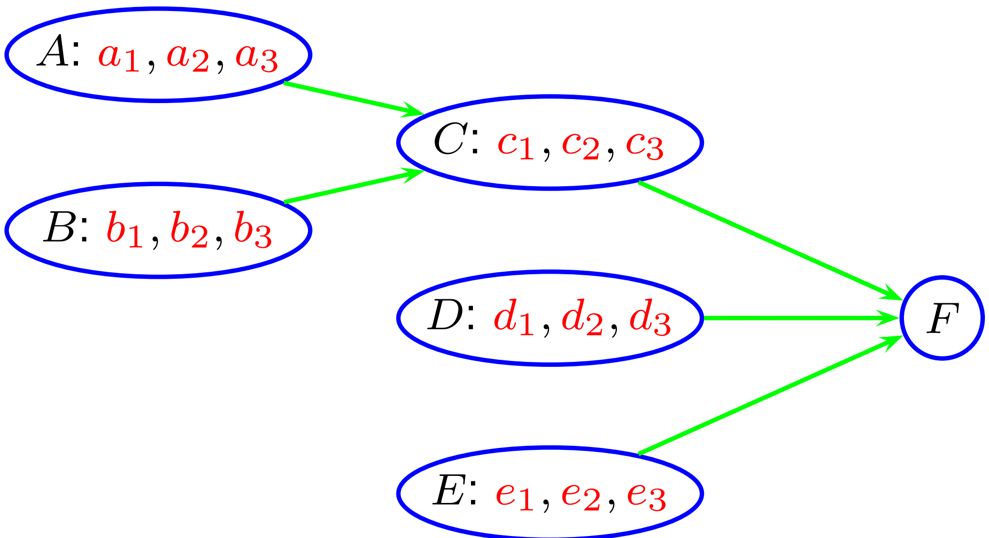


Divorcing: An example



$$P(F|a_1, b_2, D, E) = P(F|a_2, b_1, D, E)$$
$$P(F|a_1, b_1, D, E) = P(F|a_2, b_2, D, E)$$
$$P(F|a_3, b_i, D, E) = P(F|a_j, b_3, D, E)$$

Divorcing: An example

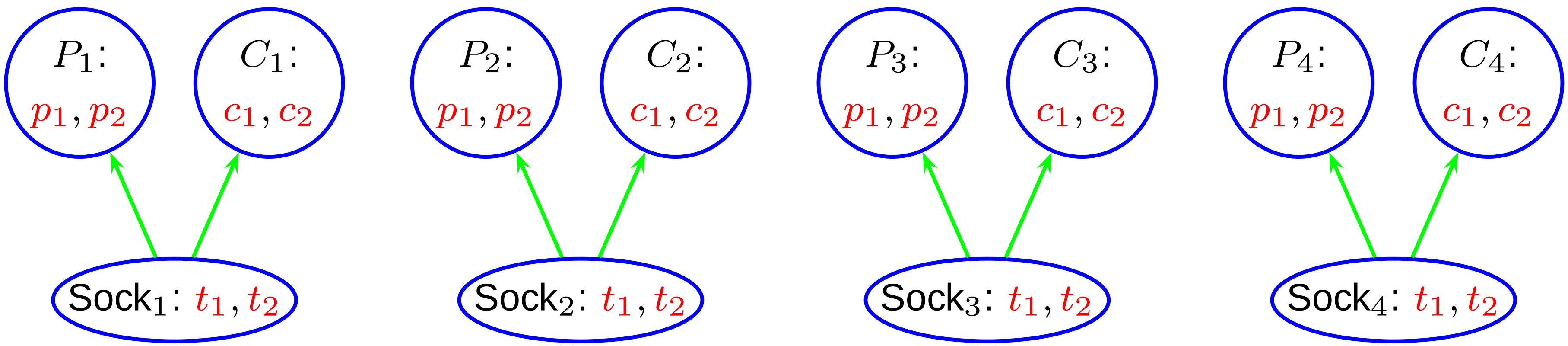


$$P(F|a_1, b_2, D, E) = P(F|a_2, b_1, D, E)$$
$$P(F|a_1, b_1, D, E) = P(F|a_2, b_2, D, E)$$
$$P(F|a_3, b_i, D, E) = P(F|a_j, b_3, D, E)$$

$$P(C|a_1, b_2) = P(C|a_2, b_1) = (1, 0, 0)$$
$$P(C|a_1, b_1) = P(C|a_2, b_2) = (0, 1, 0)$$
$$P(C|a_3, b_i) = P(C|a_j, b_3) = (0, 0, 1)$$
$$P(F|c_1, D, E) = P(F|a_1, b_2, D, E)$$
$$P(F|c_2, D, E) = P(F|a_1, b_1, D, E)$$
$$P(F|c_3, D, E) = P(F|a_3, b_i, D, E)$$

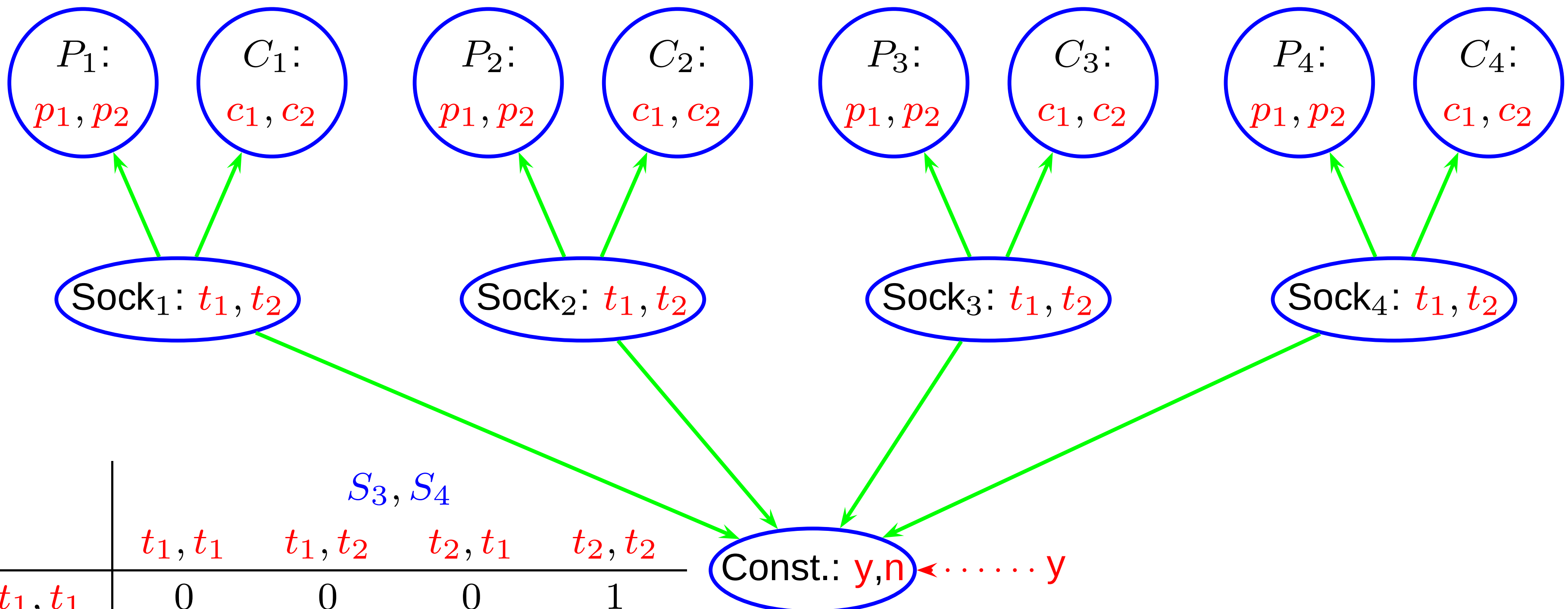
Logical constraints

I have washed two pairs of socks, and now it is hard to distinguish them. Still it is important for me to couple them correctly. The color and pattern give indications.



Logical constraints

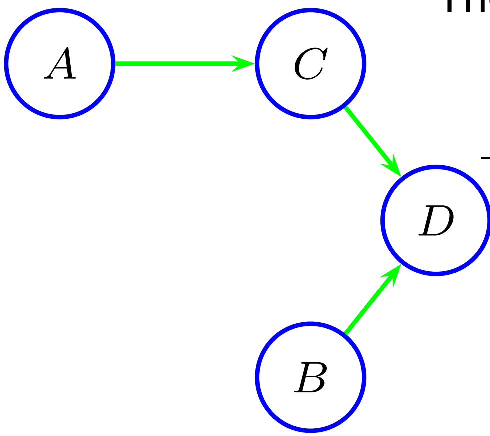
I have washed two pairs of socks, and now it is hard to distinguish them. Still it is important for me to couple them correctly. The color and pattern give indications.



S_1, S_2 \ S_3, S_4	t_1, t_1	t_1, t_2	t_2, t_1	t_2, t_2
t_1, t_1	0	0	0	1
t_1, t_2	0	1	1	0
t_2, t_1	0	1	1	0
t_2, t_2	1	0	0	0

$P(\text{Const.} = \text{y} | \text{Sock}_1, \ldots, \text{Sock}_4)$

Expert disagreement



There are three experts:

		B = y	B = n
C	y	0.4	0.7
	n	0.6	0.9

$P_1(D = y|B, C)$

		B = y	B = n
C	y	0.4	0.9
	n	0.4	0.7

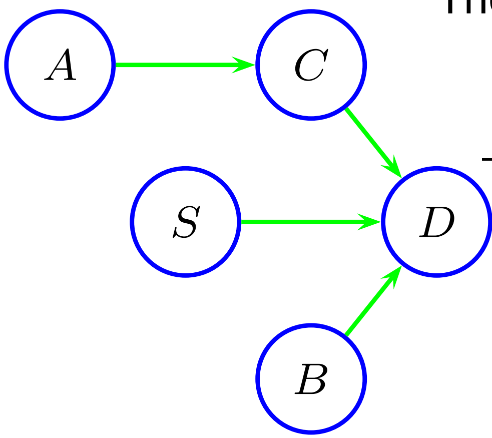
$P_2(D = y|B, C)$

		B = y	B = n
C	y	0.6	0.7
	n	0.5	0.9

$P_3(D = y|B, C)$

I believe twice as much in P_3 as I do in the others!

Expert disagreement



There are three experts:

		B = y	B = n
C	y	0.4	0.7
	n	0.6	0.9

$P_1(D = y \mid B, C)$

		B = y	B = n
C	y	0.4	0.9
	n	0.4	0.7

$P_2(D = y \mid B, C)$

		B = y	B = n
C	y	0.6	0.7
	n	0.5	0.9

$P_3(D = y \mid B, C)$

I believe twice as much in P_3 as I do in the others!

Encode the confidence in $P(S)$:

$$P(S) = (0.25, 0.25, 0.5)$$

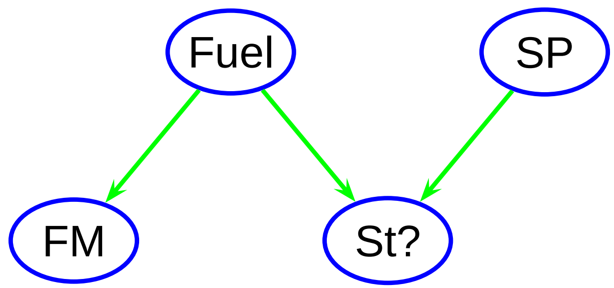
hence,

		B = y	B = n
C	y	$(0.4, 0.4, 0.6)$	$(0.7, 0.9, 0.7)$
	n	$(0.6, 0.4, 0.5)$	$(0.9, 0.7, 0.9)$

$P(D = y \mid B, C, S)$

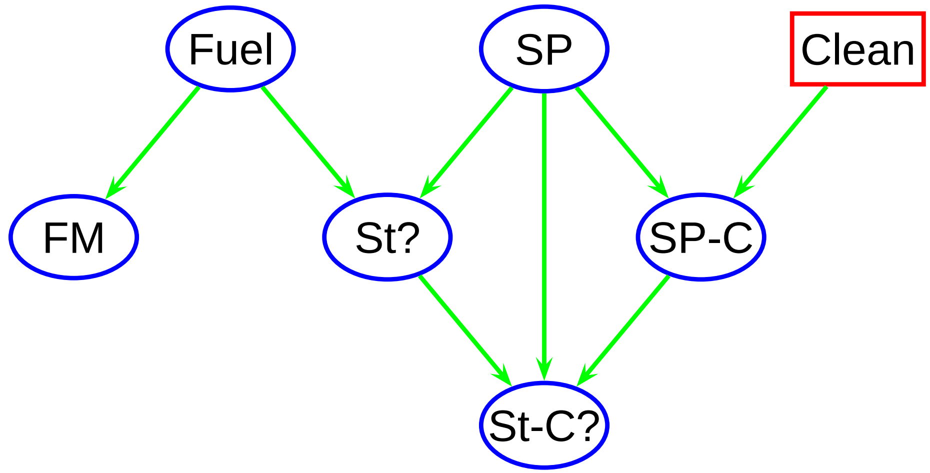
Interventions

Clean the spark plugs:

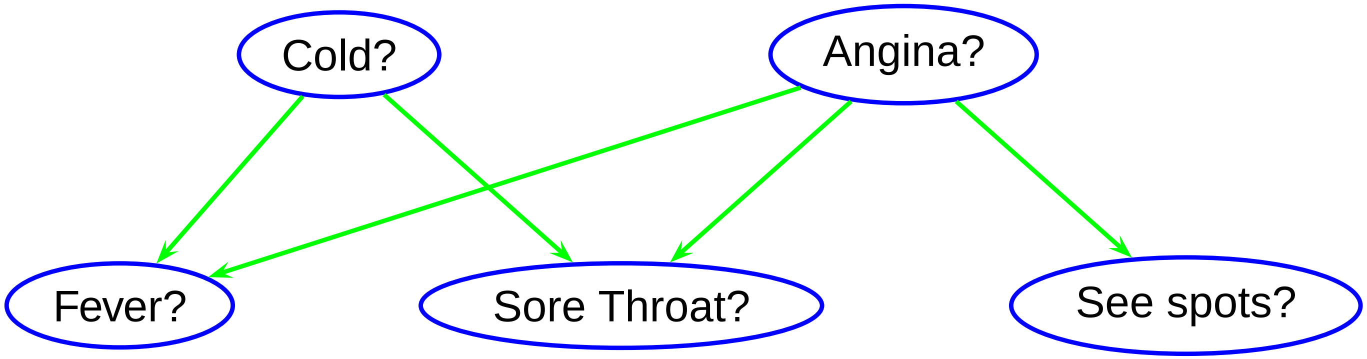


Interventions

Clean the spark plugs:



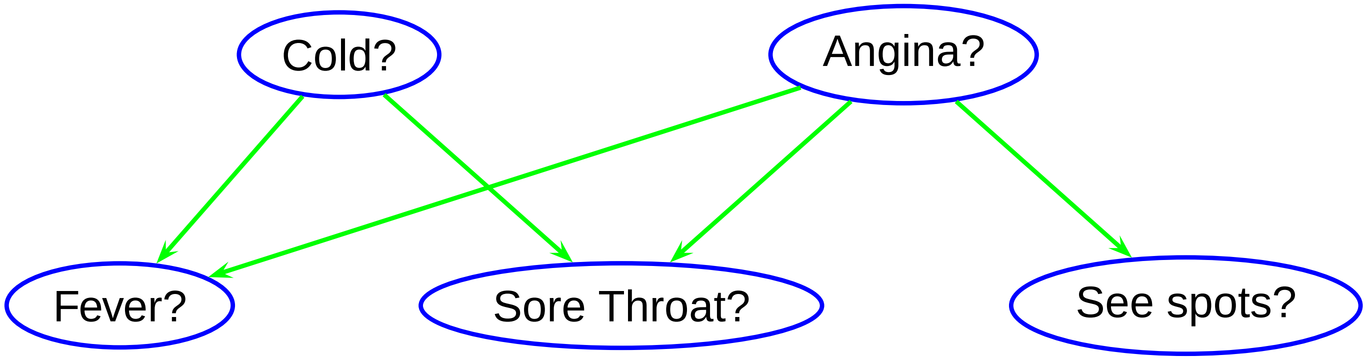
Joint probabilities I



It is not unusual to suffer from both cold and angina, so we look for the joint probability:

$$P(\text{Angina?}, \text{Cold?} \mid \bar{e})$$

Joint probabilities I



It is not unusual to suffer from both cold and angina, so we look for the joint probability:

$$P(\text{Angina?}, \text{Cold?}|\bar{e})$$

From the fundamental rule we have:

$$P(\text{Angina?}, \text{Cold?}|\bar{e}) = P(\text{Angina?}|\text{Cold?}, \bar{e})P(\text{Cold?}|\bar{e})$$

The probability:

- $P(\text{Cold?}|\bar{e})$ can be found by propagating $\bar{e}$.
- $P(\text{Angina?}|\text{Cold?}, \bar{e})$ can be found from $P(\text{Angina?}|\text{Cold?} = \text{yes}, \bar{e})$ and $P(\text{Angina?}|\text{Cold?} = \text{no}, \bar{e})$.

Joint probabilities II

From the fundamental rule we have:

$$P(\text{Angina?}, \text{Cold?}|\bar{e}) = P(\text{Angina?}|\text{Cold?}, \bar{e})P(\text{Cold?}|\bar{e})$$

Assume that:

$e = (\text{Fever?} = \text{no}, \text{SeeSpots?} = \text{yes}, \text{SoreThroat?} = \text{no})$.

We can calculate: $P(\text{Cold?}|\bar{e}) = (\quad , \quad)$

As well as:

- $P(\text{Angina?}|\text{Cold?} = \text{yes}, \bar{e}) = (\quad , \quad , \quad)$
- $P(\text{Angina?}|\text{Cold?} = \text{no}, \bar{e}) = (\quad , \quad , \quad)$

We can now calculate $P(\text{Angina?}, \text{Cold?}|\bar{e})$:

		Angina?		
		no	mild	severe
Cold?	no			
	yes			

Joint probabilities II

From the fundamental rule we have:

$$P(\text{Angina?}, \text{Cold?}|\bar{e}) = P(\text{Angina?}|\text{Cold?}, \bar{e})P(\text{Cold?}|\bar{e})$$

Assume that:

$e = (\text{Fever?} = \text{no}, \text{SeeSpots?} = \text{yes}, \text{SoreThroat?} = \text{no})$.

We can calculate: $P(\text{Cold?}|\bar{e}) = (0.997(n), 0.003(y))$

As well as:

- $P(\text{Angina?}|\text{Cold?} = \text{yes}, \bar{e}) = (0(n), 1(m), 0(s))$
- $P(\text{Angina?}|\text{Cold?} = \text{no}, \bar{e}) = (0(n), 0.971(m), 0.029(s))$

We can now calculate $P(\text{Angina?}, \text{Cold?}|\bar{e})$:

		Angina?		
		no	mild	severe
Cold?	no	0	0.968	0.029
	yes	0	0.003	0

Most probable explanation (MPE)

We can find the most probable configuration of Cold? and Angina? from:

$$P(\text{Angina?}, \text{Cold?} | \bar{e})$$

However, this can be achieved must faster:

- Use maximization instead of summation when marginalizing out a variable.

Most probable explanation (MPE)

We can find the most probable configuration of Cold? and Angina? from:

$$P(\text{Angina?}, \text{Cold?} | \bar{e})$$

However, this can be achieved must faster:

- Use maximization instead of summation when marginalizing out a variable.

This gives us MPE(Cold?)=no and MPE(Angina?)=mild.

Is the evidence reliable?

Since I see Fever? $=$ no and SoreThroat? $=$ no it seems questionable that I see spots!

- Can this warning be given by the system?
- Is the evidence coherent?

Is the evidence reliable?

Since I see Fever? $=$ no and SoreThroat? $=$ no it seems questionable that I see spots!

- Can this warning be given by the system?
- Is the evidence coherent?

For a coherent case covered by the model we expect the evidence to support each other:

$$P(e_1, e_2) > P(e_1)P(e_2)$$

We can measure this using:

$$\text{conf}(e_1, e_2) = \log_2 \frac{P(e_1)P(e_2)}{P(e_1, e_2)}$$

Thus, if $\text{conf}(e_1, e_2) > 0$ we take it as an indication that the evidence is conflicting.

Example

$$\text{conf}(\text{Fever?} = \text{no}, \text{SeeSpots?} = \text{yes}, \text{SoreThroat?} = \text{no})$$

$$= \log_2 \frac{P(\text{Fever?} = \text{no}) P(\text{SeeSpots?} = \text{yes}) P(\text{SoreThroat?} = \text{no})}{P(\text{Fever?} = \text{no}, \text{SeeSpots?} = \text{yes}, \text{SoreThroat?} = \text{no})}$$

$$= \log_2 \frac{0.960 \cdot 0.002 \cdot 0.978}{7.5131 \cdot 10^{-7}}$$

$$= \log_2(24993.47) = 11.32$$

Thus, we take it as an indication that the evidence is conflicting!

What are the crucial findings?

We would like to answer questions such as:

- What are the crucial findings?
- What if one of the findings were changed or removed?
- What set of findings would be sufficient for the conclusion?

Assume the conclusion that I suffer from mild angina:

What are the crucial findings?

We would like to answer questions such as:

- What are the crucial findings?
- What if one of the findings were changed or removed?
- What set of findings would be sufficient for the conclusion?

Assume the conclusion that I suffer from mild angina:

It is not enough with SeeSpots? $=$ yes:

- $P(\text{Angina?}|\text{SeeSpots?} = \text{yes}) = (0(n), 0.024(m), 0.976(s))$

However, SeeSpots? $=$ yes and SoreThroat? $=$ no is sufficient:

- $P(\text{Angina?}|\text{SeeSpots?} = \text{yes}, \text{SoreThroat?} = \text{no}) = (0(n), 0.884(m), 0.116(s))$

In this case findings on Fever? is irrelevant, e.g.:

- $P(\text{Angina?}|\text{SeeSpots?} = \text{yes}, \text{SoreThroat?} = \text{no}, \text{Fever?} = \text{high}) = (0(n), 0.683(m), 0.317(s))$

Sensitivity to variations in parameters

The initial tables:

Angina?	no		mild		severe	
Cold?	no	yes	no	yes	no	yes
no	0.995	0.6	0.3	0.15	0.001	0
yes	0.005	0.4	0.7	0.85	0.999	1

$P(\text{Sore throat?}|\text{Angina?}, \text{Cold?})$

Angina?	no	mild	severe
no	1	0.99	0
yes	0	0.01	1

$P(\text{See spots?}|\text{Angina?})$

Sensitivity to variations in parameters

The initial tables:

Angina?	no		mild		severe	
Cold?	no	yes	no	yes	no	yes
no	0.995	0.6	0.3	0.15	0.001	0
yes	0.005	0.4	0.7	0.85	0.999	1

$P($Sore throat?$|$Angina?, Cold?$)$

Angina?	no	mild	severe
no	1	0.99	0
yes	0	0.01	1

$P($See spots?$|$Angina?$)$

Assume that we have the parameters:

Angina?	no		mild		severe	
Cold?	no	yes	no	yes	no	yes
no	0.995	0.6	0.3	0.15	t	0
yes	0.005	0.4	0.7	0.85	0.999	1

$P($Sore throat?$|$Angina?, Cold?$)$

Angina?	no	mild	severe
no	1	0.99	0
yes	0	s	1

$P($See spots?$|$Angina?$)$

We want e.g.:

$$P(\text{Angina?} = \text{mild}|\bar{e})(t) \;;\; P(\text{Angina?} = \text{mild}|\bar{e})(s) \;;\; P(\text{Angina?} = \text{mild}|\bar{e})(s,t)$$

Sensitivity analysis

Theorem:

$$P(\bar{e})(t) = \alpha t + \beta = x(t)$$

Thus, we also have that $P(\text{Angina?} = \text{mild}, \bar{e})(t) = a \cdot t + b = y(t)$, and therefore:

$$P(\text{Angina?} = \text{mild} | \bar{e})(t) = \frac{y(t)}{x(t)}$$

Sensitivity analysis

Theorem:

$$P(\bar{e})(t) = \alpha t + \beta = x(t)$$

Thus, we also have that $P(\text{Angina?} = \text{mild}, \bar{e})(t) = a \cdot t + b = y(t)$, and therefore:

$$P(\text{Angina?} = \text{mild} | \bar{e})(t) = \frac{y(t)}{x(t)}$$

For $t = 0.001$ we have $x(t) = 7.513 \cdot 10^{-7}$ and $y(t) = 7.298 \cdot 10^{-7}$.

If we change t to 0.002 and propagate we get:

$$x(0.002) = 7.7286 \cdot 10^{-7} \qquad y(0.002) = 7.2975 \cdot 10^{-7}$$

We can now determine the coefficients α, β, a and b!