

729G78 Artificiell intelligens

Probabilistisk logik – Bayesianiska nätverk

Robin Keskisärkkä

HCS/IDA



LINKÖPINGS
UNIVERSITET

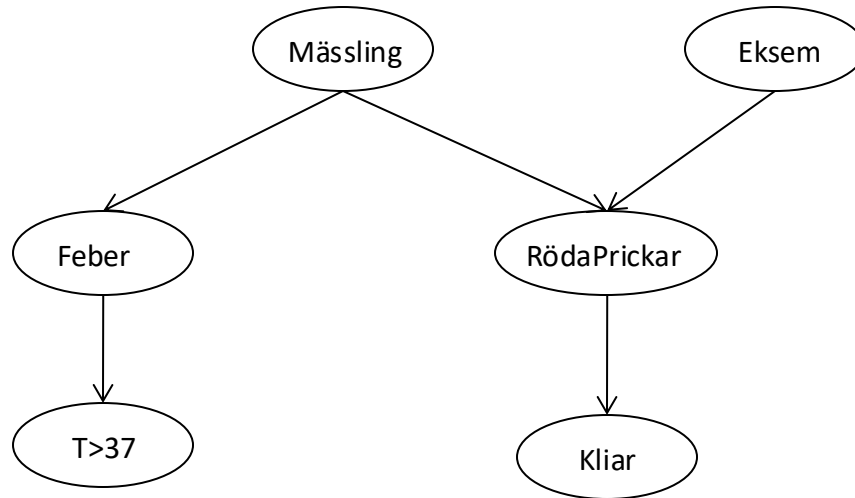
Bayesianska nätverk

- Stokastiska variabler representerade som noder i ett nätverk
- Riktade länkar mellan par av noder
- En tabell med övergångssannolikheter som anger effekten på en nod från föräldranoderna
- Grafen har inga riktade cykler (DAG)

Exempel

- Mässling ger feber och röda prickar. Eksem ger också röda prickar. Fever betyder att kroppstemperaturen överstiger 37 grader. Röda prickar kliar. Patienten kliar.
- Stokastiska variabler
 - *Mässling*
 - *Fever*
 - *RödaPrickar*
 - *Eksem*
 - $T > 37$
 - *Kliar*

Bayesianskt nätverk



Övergångssannolikheter

Ovillkorlig

<i>Mässling</i>	
$P(\text{mässling})$	$P(\neg \text{mässling})$
0,1	0,9

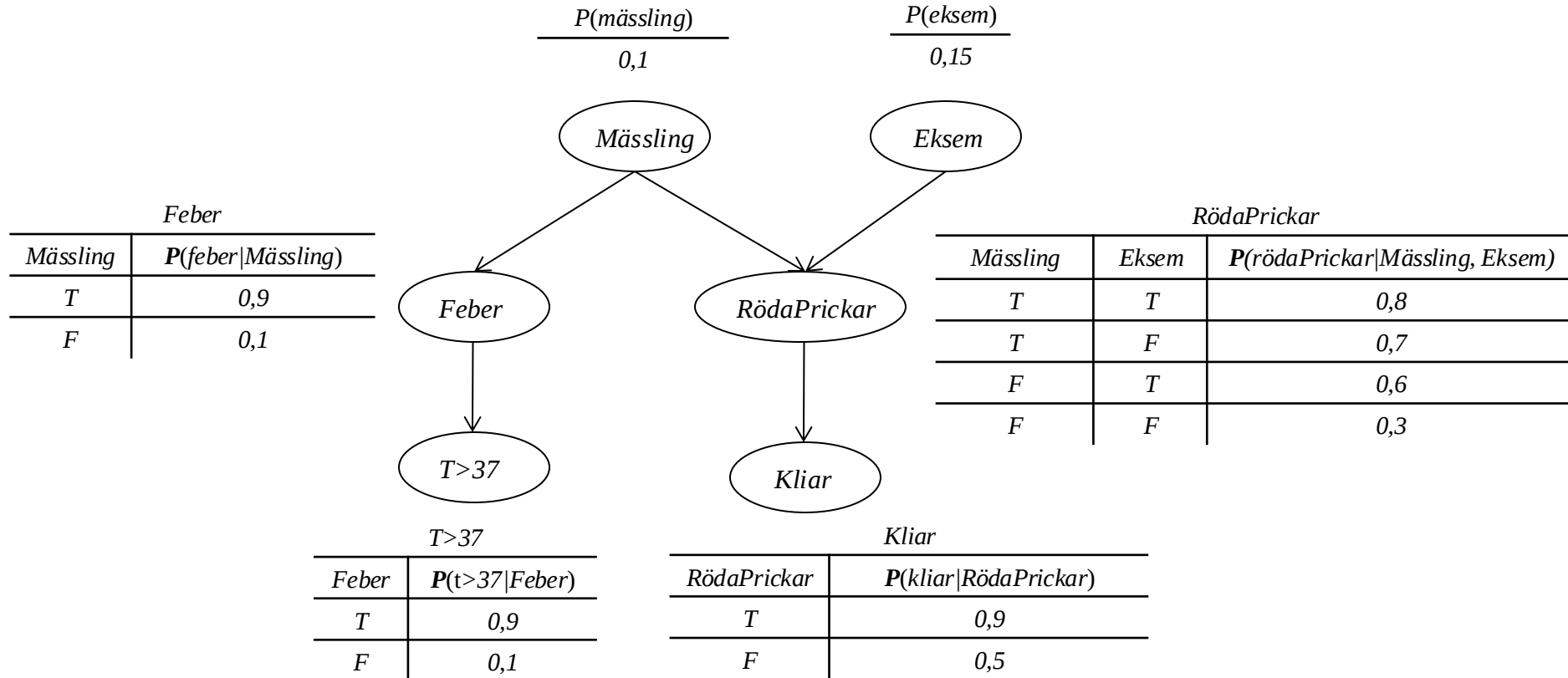
Ett beroende

$T > 37$		
<i>Feber</i>	$P(t > 37 \text{Feber})$	$P(\neg t > 37 \text{Feber})$
<i>T</i>	0,9	0,1
<i>F</i>	0,1	0,9

Beroende av två stokastiska variabler

<i>RödaPrickar</i>			
<i>Mässling</i>	<i>Eksem</i>	$P(\text{rödaPrickar} \text{Mässling}, \text{Eksem})$	$P(\neg \text{rödaPrickar} \text{Mässling}, \text{Eksem})$
<i>T</i>	<i>T</i>	0,8	0,2
<i>T</i>	<i>F</i>	0,7	0,3
<i>F</i>	<i>T</i>	0,6	0,4
<i>F</i>	<i>F</i>	0,3	0,7

Bayesianskt nätverk

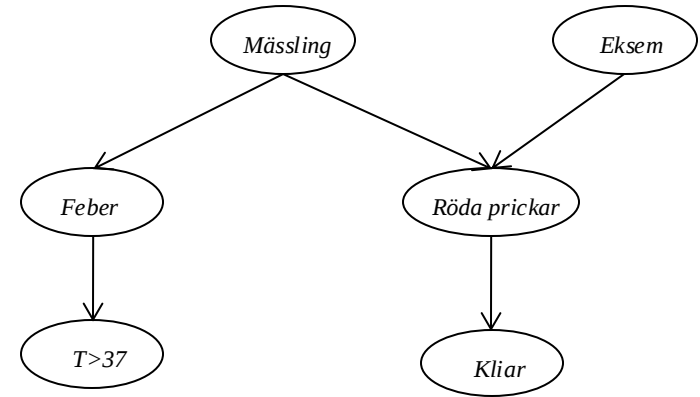


Inferens i bayesianska nät, 1

Det bayesianska nätverket beskriver domänen

$P(X_1=v_1, X_2=v_2, \dots, X_n=v_n)$ kan beräknas utifrån föräldrarnas påverkan

Ex. $P(\text{mässling, rödaPrickar, } \neg\text{kliar, eksem, } t>37, \text{ feber})?$



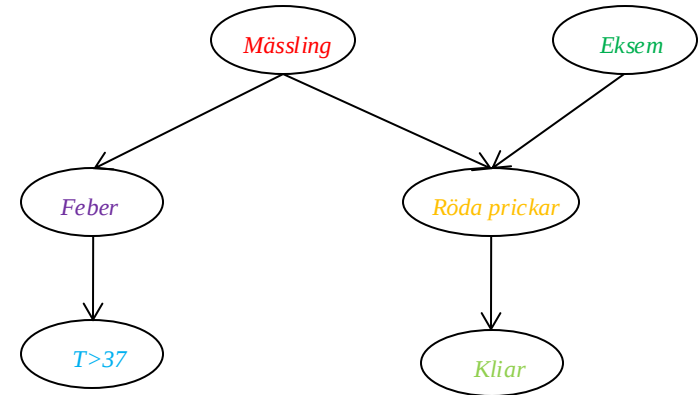
Inferens i bayesianska nät, 1

Vi kan skriva om uttrycket som produkten av varje variabel *givet* resterande variabler:

$$\begin{aligned} P(\text{mässling}, \text{rödaPrickar}, \neg\text{kliar}, \text{eksem}, t>37, \text{feber}) = \\ P(\text{mässling} | \text{rödaPrickar}, \neg\text{kliar}, \text{eksem}, t>37, \text{feber}) \cdot \\ P(\text{rödaPrickar} | \text{mässling}, \neg\text{kliar}, \text{eksem}, t>37, \text{feber}) \cdot \\ P(\neg\text{kliar} | \text{mässling}, \text{rödaPrickar}, \text{eksem}, t>37, \text{feber}) \cdot \\ P(\text{eksem} | \text{mässling}, \text{rödaPrickar}, \neg\text{kliar}, t>37, \text{feber}) \cdot \\ P(t>37 | \text{mässling}, \text{rödaPrickar}, \neg\text{kliar}, \text{eksem}, \text{feber}) \cdot \\ P(\text{feber} | \text{mässling}, \text{rödaPrickar}, \neg\text{kliar}, \text{eksem}, t>37) = \end{aligned}$$

Lite osnyggt eftersom många variabler inte påverkar den observerade variabeln. Men givet villkorligt oberoende kan vi stryka onödiga variabler. Ex:

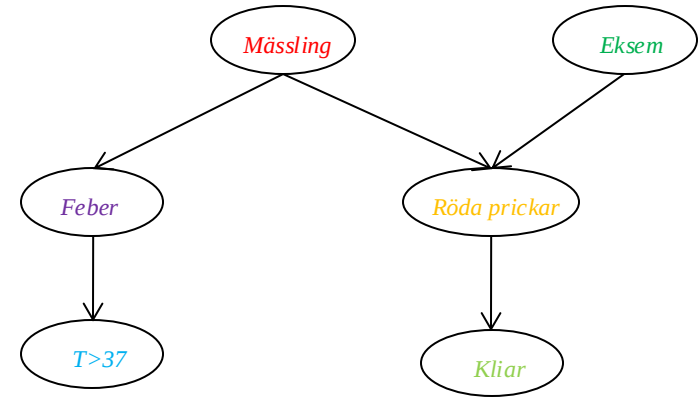
$$P(\text{RödaPrickar} | \text{Mässling}, \text{Kliar}, \text{Eksem}, T>37, \text{Feber}) = P(\text{RödaPrickar} | \text{Mässling}, \text{Eksem})$$



Inferens i bayesianska nät, 1

Vi kan skriva om uttrycket som produkten av varje variabel *givet* resterande variabler:

$$\begin{aligned} &P(\text{mässling}, \text{rödaPrickar}, \neg\text{kliar}, \text{eksem}, \text{t}>37, \text{feber}) = \\ &P(\text{mässling} \mid \text{rödaPrickar}, \neg\text{kliar}, \text{eksem}, \text{t}>37, \text{feber}) \cdot \\ &P(\text{rödaPrickar} \mid \text{mässling}, \neg\text{kliar}, \text{eksem}, \text{t}>37, \text{feber}) \cdot \\ &P(\neg\text{kliar} \mid \text{mässling}, \text{rödaPrickar}, \text{eksem}, \text{t}>37, \text{feber}) \cdot \\ &P(\text{eksem} \mid \text{mässling}, \text{rödaPrickar}, \neg\text{kliar}, \text{t}>37, \text{feber}) \cdot \\ &P(\text{t}>37 \mid \text{mässling}, \text{rödaPrickar}, \neg\text{kliar}, \text{eksem}, \text{feber}) \cdot \\ &P(\text{feber} \mid \text{mässling}, \text{rödaPrickar}, \neg\text{kliar}, \text{eksem}, \text{t}>37) = \end{aligned}$$



Inferens i bayesianska nät, 1

Vi kan skriva om uttrycket som produkten av varje variabel *givet* resterande variabler:

$$P(\text{mässling}, \text{rödaPrickar}, \neg\text{kliar}, \text{eksem}, \text{t}>37, \text{feber}) =$$

$$P(\text{mässling}) \cdot$$

$$P(\text{rödaPrickar} \mid \text{mässling}, \text{eksem}) \cdot$$

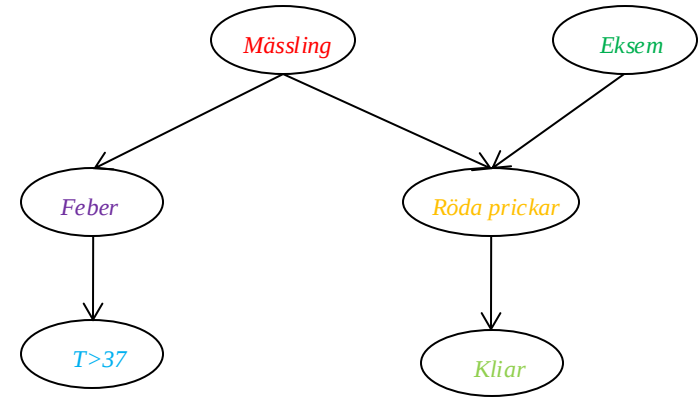
$$P(\neg\text{kliar} \mid \text{rödaPrickar}) \cdot$$

$$P(\text{eksem}) \cdot$$

$$P(\text{t}>37 \mid \text{feber}) \cdot$$

$$P(\text{feber} \mid \text{mässling}) =$$

Nu kan vi hitta alla värden i sannolikhetstabellerna!



Inferens i bayesianska nät, 1

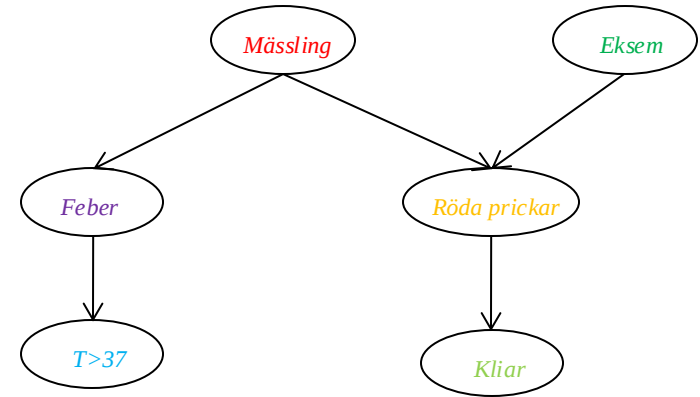
Kan skrivas om som produkten av varje variabel givet resterande:

$P(\text{mässling}, \text{rödaPrickar}, \neg\text{kliar}, \text{eksem}, t>37, \text{feber}) =$

$P(\text{mässling}) \cdot P(\text{rödaPrickar} | \text{mässling}, \text{eksem}) \cdot P(\neg\text{kliar} | \text{rödaPrickar}) \cdot$

$P(\text{eksem}) \cdot P(t>37 | \text{feber}) \cdot P(\text{feber} | \text{mässling}) =$

$0,1 \cdot 0,8 \cdot 0,1 \cdot 0,15 \cdot 0,9 \cdot 0,9 = 0,000972 \text{ (ca 0,09\%)}$



Inferens i bayesianska nät, 2

- Mer generellt:

$$\mathbf{P}(X_1, X_2, \dots, X_n) = \prod_{i=1}^n \mathbf{P}(X_i \mid \text{Parents}(X_i))$$

om noden X_i bara beror av noder ovanför, dvs

$$\text{Parents}(X_i) \subseteq \{X_{i-1}, \dots, X_2, X_1\}$$

dvs villkorligt oberoende

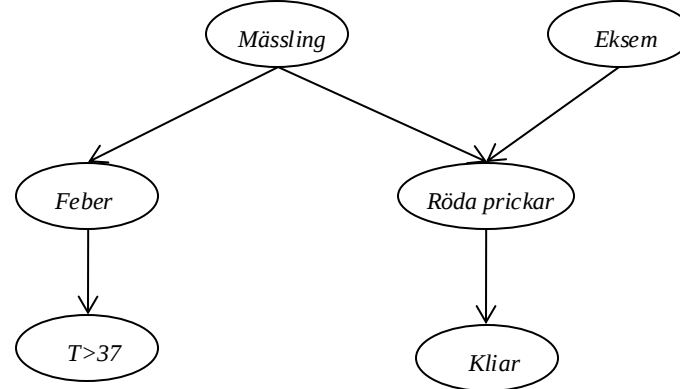
- I mässlingexemplet är t.ex. Kliar villkorligt oberoende allt givet Röda Prickar, dvs

$$\mathbf{P}(\text{Kliar} \mid T > 37, \text{Feber}, \text{Röda Prickar}, \text{Mässling}, \text{Eksem}) = \mathbf{P}(\text{Kliar} \mid \text{Röda Prickar})$$

$$P(a|b) = \frac{P(a \wedge b)}{P(b)} = \alpha \cdot P(a \wedge b)$$

Inferens i bayesianska nät, 3

- Betingade sannolikheter $P(a|b) = \alpha P(a \wedge b)$ eller mer generellt $\mathbf{P}(A|B) = \alpha \mathbf{P}(A, B)$ där $\mathbf{P}(A, B)$ kommer ur sambandet för föräldrars påverkan
- Exempel $P(kliar|eksem)$. Beror av såväl *Mässling* som *RödaPrickar* och räknas ut genom att summera över alla möjliga värden, dvs *sant* och *falskt*



Inferens i bayesianska nät, 4

- Diagnos
 - Från effekt till orsak
 $P(\neg mässling \mid kliar)$
- Kausalitet
 - Från orsak till effekt
 $P(kliar \mid \neg mässling)$
- Interkausalitet
 - Mellan orsaker till samma effekt
 $P(\neg mässling \mid eksem \wedge rödaPrickar)$
- Blandad inferens
 $P(eksem \mid kliar \wedge \neg mässling)$

Inferens i bayesianska nät, 5

- Beräkningstid och det generella fallet är NP-komplett
- Villkorliga oberoenden gör det ibland lättare
- Ofta approximativa lösningar

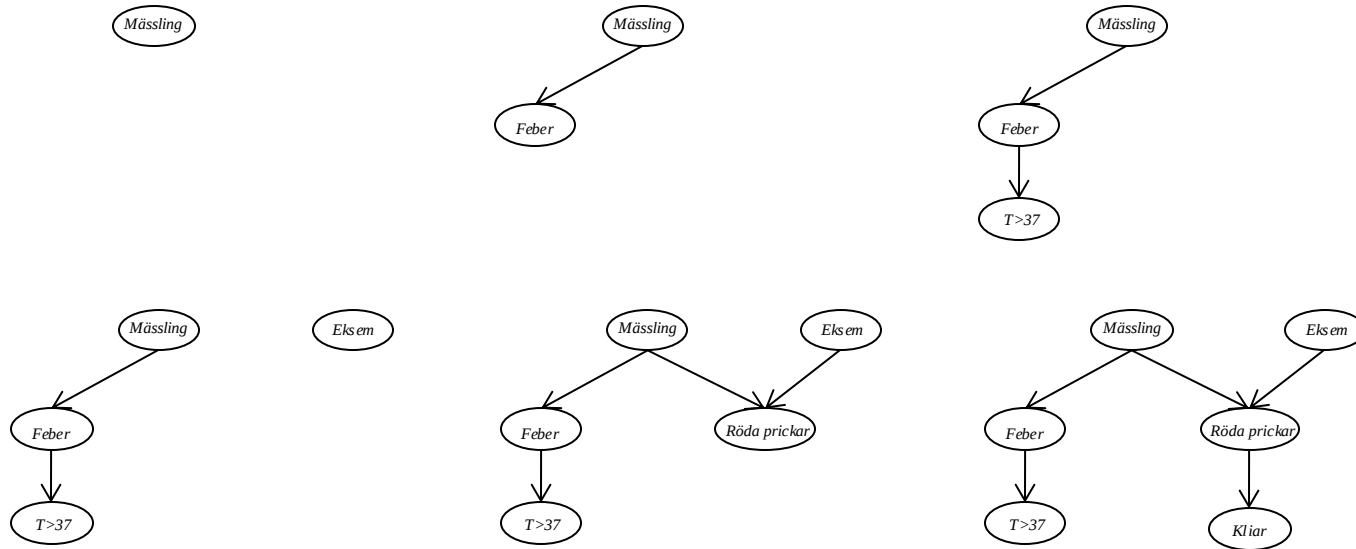
Konstruktion av Bayesianiska nätverk

1. Ordna de stokastiska variablerna, $X_1, \dots, X_n$
2. För varje nod, X_i , $i=1$ till n
 1. Stoppa in X_i i nätverket
 2. Lägg in bågar från de stokastiska variablerna $X_1, \dots, X_{i-1}$ så att
$$\mathbf{P}(X_i \mid X_1, \dots, X_{i-1}) = \mathbf{P}(X_i \mid \text{Parents}(X_i))$$
dvs. så att varje nod är villkorligt oberoende av de tidigare noderna i nätverket, givet sina föräldrar

Exempel

Noder: Feber, Mässling, Klier, RödaPrickar, Eksem, T>37

Sorterat: Mässling, Feber, T>37, Eksem, RödaPrickar, Klier



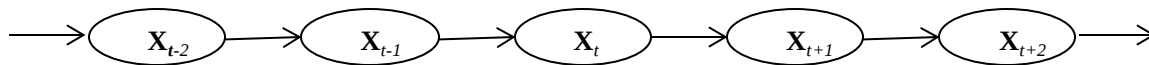
Variabelval

- Kausalsamband snarare än diagnossamband
- Från orsak till verkan
 - Grundorsak
 - Variabler som påverkas direkt av grundorsaken
 - Observerbara effekter

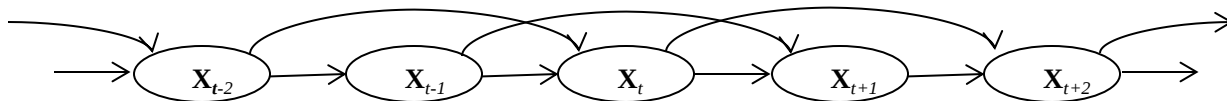
Probabilistiskt resonerande över tid

- Bayesianska nätverk beskriver en statisk situation
- Världen förändras med tiden, ex.:
 - Robot som går runt i världen
 - Medicinsk diagnos
 - Taligenkänning
- Tillstånd som förändras över tid, $\mathbf{X}_t$
- Observationer vid en viss tidpunkt, $\mathbf{E}_t$
- Diskreta observationstillfällen
- Sekvens av tillstånd, $\mathbf{X}_{a:b}$

Övergångssannolikheter

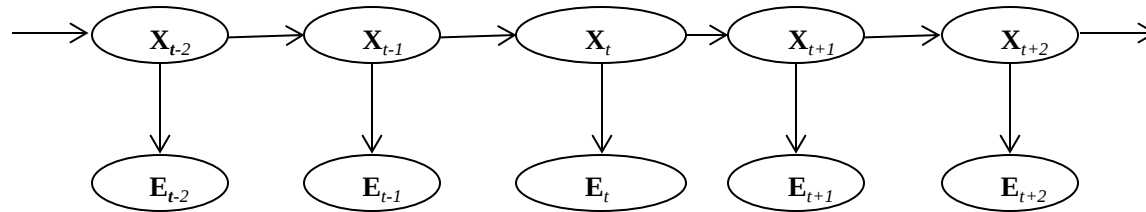


- $\mathbf{P}(X_t | X_{0:t-1}) = \mathbf{P}(X_t | X_0, X_1, X_2, \dots, X_{t-2}, X_{t-1})$
- Problem när t ökar
- Markovantagandet: *Nuvarande tillstånd beror bara av ett ändligt antal tidigare tillstånd*
- Första ordningens Markov process, beror bara av föregående tillstånd
- $\mathbf{P}(X_t | X_{0:t-1}) = \mathbf{P}(X_t | X_{t-1})$
- Andra ordningens markovprocess
- $\mathbf{P}(X_t | X_{0:t-1}) = \mathbf{P}(X_t | X_{t-2}, X_{t-1})$



Dold markovmodell (HMM)

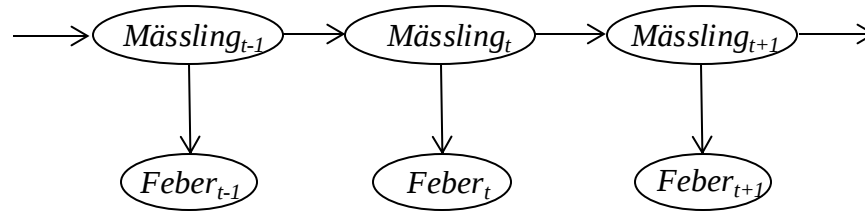
- Tillstånden i den vanliga markovmodellen observerbara
- I en dold markovmodell kan man bara se observationerna, $\mathbf{E}_t$



- Tillstånden är dolda, hidden
- Markovantagande för observationerna; de beror bara av nuvarande tillstånd
- $\mathbf{P}(\mathbf{E}_t \mid \mathbf{X}_{0:t-1}, \mathbf{E}_{0:t-1}) = \mathbf{P}(\mathbf{E}_t \mid \mathbf{X}_t)$

Exempel

- Sannolikheten för mässling beror av om man har feber eller inte



- Villkorliga sannolikhetstabeller

Övergångsmodell	
$Mässling_{t-1}$	$P(Mässling_t Mässling_{t-1})$
T	0,6
F	0,4

Observationsmodell	
$Mässling_t$	$P(Feber_t Mässling_t)$
T	0,9
F	0,3

Simultanfördelning

- Antag starttillståndet $\mathbf{P}(\mathbf{X}_0)$
- Övergångsmodellen $\mathbf{P}(\mathbf{X}_t | \mathbf{X}_{t-1})$
- Observationsmodellen $\mathbf{P}(\mathbf{E}_t | \mathbf{X}_t)$

$$\mathbf{P}(\mathbf{X}_{0:t}, \mathbf{E}_{1:t}) = \mathbf{P}(\mathbf{X}_0) \overset{\downarrow}{\prod}_{i=1} \mathbf{P}(\mathbf{X}_i | \mathbf{X}_{i-1}) \mathbf{P}(\mathbf{E}_i | \mathbf{X}_i)$$

Inferens i temporala modeller

- Filtrering, $\mathbf{P}(\mathbf{X}_t | \mathbf{e}_{1:t})$
 - Beräkna nuvarande sannolikhet givet alla bevis hittills
- Predicering, $\mathbf{P}(\mathbf{X}_{t+k} | \mathbf{e}_{1:t})$
 - Beräkna sannolikheten för ett framtida tillstånd
- "Smoothing", $\mathbf{P}(\mathbf{X}_k | \mathbf{e}_{1:t})$
 - Beräkna sannolikheten för ett tidigare tillstånd
- Mest sannolika förklaring, $\operatorname{argmax}_{\mathbf{x}_{1:t}} \mathbf{P}(\mathbf{x}_{1:t} | \mathbf{e}_{1:t})$
 - Beräkna den sekvens av tillstånd som genererat observationerna

Sammanfattning, 1

- Stokastisk variabel A med sannolikhet att vara sann $P(A=sant)$, skrivs också $P(a)$
 - T.ex. $P(a)=0,8$, eller $\mathbf{P}(A) = \langle 0,8 \ 0,2 \rangle$

- $\mathbf{P}(A,B)$ ger tabell med 4 (2^2) värden

	a	$\neg a$
b	0,2	0,3
$\neg b$	0,4	0,1

- n stokastiska variabler $\mathbf{P}(X_1, \dots X_n)$ ger 2^n värden
- Problem!
- Oberoende

$$P(\text{Tärning}=6 \mid \text{Tärning} = 1) = P(\text{Tärning}=6)$$

Sammanfattning, 2

- Bayes teorem, från kausalsamband till diagnossamband

$$P(a | b) = \frac{P(b | a)P(a)}{P(b)} = aP(b | a)P(a)$$

$$\mathbf{P}(A | B) = a\mathbf{P}(B | A)\mathbf{P}(A)$$

- Villkorligt oberoende

$$\mathbf{P}(Orsak | Obs_1...Obs_n) = \mathbf{P}(Orsak | Obs_1)...\mathbf{P}(Orsak | Obs_n)$$

- kombinerat med kedjeregeln

$$\mathbf{P}(Orsak, Obs_1...Obs_n) = \mathbf{P}(Orsak)\mathbf{P}(Orsak | Obs_1...Obs_n)$$

- ger Naive Bayes

$$\mathbf{P}(Orsak, Obs_1...Obs_n) = \mathbf{P}(Orsak) \prod_{i=1}^n \mathbf{P}(Orsak | Obs_i)$$

Sammanfattning, 3

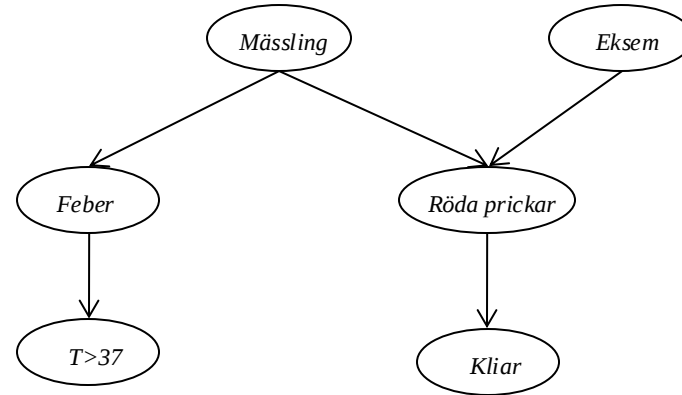
- Bayesianskt nätverk tar hänsyn till beroenden mellan variabler

Mässling och Eksem 1 värde vardera

Feber, $T > 37$ och Kliar 2 värden vardera

Röda Prickar 4 värden

Totalt 12 värden (jmf $2^6 = 64$ värden)



- HMM är en sekventiell variant av Bayesianska nätverk som fokuserar på resonemang över tid