TSFS06 Diagnos och övervakning
Föreläsning 10 - Sannolikhetsbaserad diagnos och Bayesianska nätverk

Erik Frisk

Institutionen för systemteknik
Linköpings universitet
erik.frisk@liu.se

2020-05-12

Outline

Sneak-peak

Antag att residualerna r_1 och r_2 larmar i beslutsstrukturen

	F_1	F_2	F_3
r_1	0	X	X
r_2	X	0	X
r_3	X	X	0

En konsistensbaserad ansats skulle ge de minimala diagnoserna

$$\mathcal{D}_1 = OK(C_1) \land OK(C_2) \land \neg OK(C_3), \qquad \{F_3\}$$
$$\mathcal{D}_2 = \neg OK(C_1) \land \neg OK(C_2) \land OK(C_3), \qquad \{F_1, F_2\}$$

- kort introducerande demonstration i GeNIe

Outline

Konsistensbaserad diagnos

Definition (Diagnosis)

Givet en modell $\mathcal{M}$ och observationer $\mathcal{O}$ så är en modtilldelning $\mathcal{D}$ en diagnos omm mängden formler

$$\mathcal{M} \cup \mathcal{O} \cup \mathcal{D}$$

är konsistent.

Exempel, trippel redundans

$$\mathcal{M}: \begin{cases} y_1 = x + f_1 \\ y_2 = x + f_2 \\ y_3 = x + f_3 \\ OK(S_i) \rightarrow f_i = 0, \quad i = 1,2,3 \end{cases}, \qquad \mathcal{O} = \{y_1 = y_2 = 1, y_3 = 3\}$$

$$\mathcal{D}_1 = OK(S_1) \wedge OK(S_2) \wedge \neg OK(S_3)$$
$$\mathcal{D}_2 = OK(S_1) \wedge \neg OK(S_2) \wedge OK(S_3)$$
$$\mathcal{D}_3 = OK(S_1) \wedge \neg OK(S_2) \wedge \neg OK(S_3)$$

Konsistens hos $\mathcal{M} \cup \mathcal{O} \cup \mathcal{D}_1$ är ekvivalent med att det finns en lösning (x, f_3) till ekvationerna

$$1 = x,\ 1 = x,\ 3 = x + f_3$$

och motsvarande för $\mathcal{M} \cup \mathcal{O} \cup \mathcal{D}_2$ blir

$$1 = x,\ 1 = x + f_2,\ 3 = x$$

Osäkerheter och konsistensbaserad diagnos

Konsistens i de två fallen svarar mot att residualerna

$$r_1 = y_1 - y_2, \quad r_2 = y_1 - y_3$$

är exakt 0.

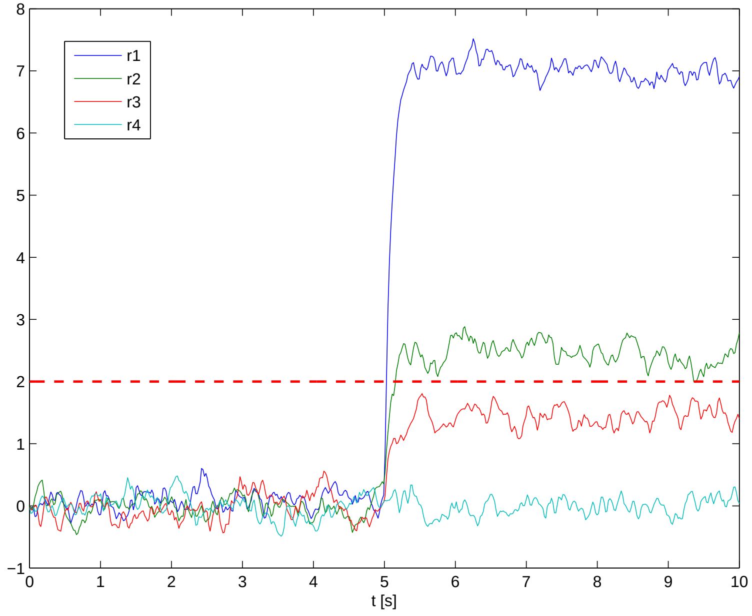
Problem

Med modellfel och mätbrus så är våra residualer aldrig exakt 0
$\Rightarrow$ vi tröskar våra residualer

Vi avgör om vi är tillräckligt nära konsistens genom att tröskla residualer/teststorheter

$$|r(t)| > J \Rightarrow \text{generera larm}$$

Trösklar och beslut



Beslut under osäker information

- felisoleringsalgoritmen tar ej i beaktande om residualer är långt över sina trösklar eller precis över
- kvantiseringseffekter som kanske inte är önskvärda, vill ha en mer mjuk övergång mellan besluten
- Beslut under osäker information är ett stort vetenskapligt fält
- Sannolikheter ett naturligt verktyg (men ej det enda)

Konsistensbaserad till sannolikhetsbaserad diagnos

Konsistens av

$$\mathcal{M} \cup \mathcal{O} \cup \mathcal{D}$$

ersätts av något i stil med

$$P(\mathcal{D}|\mathcal{M}, \mathcal{O}) \quad \text{eller} \quad P(F_i|\mathcal{M}, \mathcal{O})$$

oftast den senare pga. komplexitetsegenskaper.

Diagnoser och sannolikheter

Antag att residualerna r_1 och r_2 larmar i beslutsstrukturen

	F_1	F_2	F_3
r_1	0	X	X
r_2	X	0	X
r_3	X	X	0

En konsistensbaserad ansats skulle ge de minimala diagnoserna

$$\mathcal{D}_1 = OK(C_1) \wedge OK(C_2) \wedge \neg OK(C_3),$$
$$\mathcal{D}_2 = \neg OK(C_1) \wedge \neg OK(C_2) \wedge OK(C_3)$$

En sannolikhetsbaserad ansats (med enkelfelsantagande) skulle ge resultat i stil med

$$P(\mathcal{B}_i|\mathcal{O}, \mathcal{M}) = \begin{cases} 0.01 & \text{if } \mathcal{B}_i = NF \\ 0.85 & \text{if } \mathcal{B}_i = F_1 \\ 0.93 & \text{if } \mathcal{B}_i = F_2 \\ 0.22 & \text{if } \mathcal{B}_i = F_3 \end{cases}$$

Diagnoser och sannolikheter

$$P(\mathcal{B}_i|\mathcal{O}, \mathcal{M}) = \begin{cases} 0.01 & \text{if } \mathcal{B}_i = NF \\ 0.85 & \text{if } \mathcal{B}_i = F_1 \\ 0.93 & \text{if } \mathcal{B}_i = F_2 \\ 0.22 & \text{if } \mathcal{B}_i = F_3 \end{cases}$$

- felmoder vs diagnoser

$$P(F_i|\mathcal{M}, \mathcal{O}) \text{ vs. } P(D_i|\mathcal{M}, \mathcal{O})$$

- Önskvärt, men ofta ej möjligt, med explicita uttryck för sannolikhetsfördelningar, dynamiska och olinjära modeller

$$x_{t+1} = f(x_t) + \varepsilon_t$$

- stokastiska filter (E/U)-Kalman Filter, partikelfilter, ...

Sannolikhetsbaserad diagnos

Vi vill modellera processen så att vi på ett effektivt sätt kan räkna ut storheter i stil med

$$P(\mathcal{D}|\mathcal{M}, \mathcal{O}) \quad \text{eller} \quad P(F_i|\mathcal{M}, \mathcal{O})$$

oftast den senare pga. komplexitetsegenskaper.

Outline

- Introduktion, sneak-peak
- Sannolikhetsbaserad diagnos
- Introducerande exempel
- Notation och lite repetition
- Sannolikhetsbaserade modeller, inferens, och komplexitet
- Bayesianska nätverk
- Kanoniska modeller
- Sammanfattning

Introducerande exempel

Tänk ett fall med två fel, de möjliga felmoderna är då

$$FM \in \{NF, f_1, f_2, f_1\&f_2\}.$$

En residual har konstruerats för att, i huvudsak, detektera fel f_1 men som också är känslig för fel f_2

Sannolikhet	Formel	Värde
A priori-sannolikhet för fel i	$P(f_i),\ i = 1,2$	0.02
Falsklarm	$P(r > J \mid FM = NF)$	0.01
Känslighet för enkelfel f_1	$P(r > J \mid FM = f_1)$	0.99
Känslighet för enkelfel f_2	$P(r > J \mid FM = f_2)$	0.30
Känslighet för dubbelfel $f_1\&f_2$	$P(r > J \mid FM = f_1\&f_2)$	0.99

Residualen överträder sin tröskel, vad är slutsatsen

- deterministiskt
- med sannolikheterna

Introducerande exempel, forts.

Oberoende antas mellan felen, dvs.

$$P(FM = f_1) = P(f_1, \neg f_2) = P(f_1)P(\neg f_2)$$

direkta räkningar ger då

$$P(FM = NF \mid r > J) = \frac{P(r > J \mid FM = NF)P(FM = NF)}{P(r > J)} = \frac{P(r > J \mid FM = NF)P(\neg f_1)P(\neg f_2)}{P(r > J)}$$

$$P(FM = f_1 \mid r > J) = \frac{P(r > J \mid FM = f_1)P(FM = f_1)}{P(r > J)} = \frac{P(r > J \mid FM = f_1)P(f_1)P(\neg f_2)}{P(r > J)}$$

$$P(FM = f_2 \mid r > J) = \frac{P(r > J \mid FM = f_2)P(FM = f_2)}{P(r > J)} = \frac{P(r > J \mid FM = f_2)P(\neg f_1)P(f_2)}{P(r > J)}$$

$$P(FM = f_1\&f_2 \mid r > J) = \frac{P(r > J \mid FM = f_1\&f_2)P(f_1)P(f_2)}{P(r > J)}$$

Introducerande exempel, forts.

Sätter man in värden fås att

$$P(FM = f \mid r > J) = \begin{cases} 27.2\% & \text{if } f = NF \\ 55.0\% & \text{if } f = f_1 \\ 16.7\% & \text{if } f = f_2 \\ 1.1\% & \text{if } f = f_1\&f_2 \end{cases}$$

- Behöver ej beräkna $P(r > J)$
- Krävdes en del handräkning, och exemplet var av väldigt enkel sort. Nu, hur generaliserar man detta till mer allmänna problem.

Outline

Förutsättningar, begränsningar, notation

- Endast diskreta stokastiska variabler
- Endast statiska modeller
- Går att generalisera båda dessa, men görs inte i den här kursen

Notation

Sannolikheten att en stokastisk variabel X (versal) har värdet x_i (gemen) skrivs med sannolikhetsfunktionen

$$P(X = x_i) \text{ eller kortare } P(x_i).$$

Om X endast kan ha värdena Sann eller Falsk skriver vi

$$P(x) \text{ och } P(\neg x)$$

för

$$P(X = \text{True}) \text{ och } P(X = \text{False}).$$

Vill vi beskriva sannolikheterna

$$P(FM = f|r > J) = \begin{cases} 27.2\% & \text{if } f = NF \\ 55.0\% & \text{if } f = f_1 \\ 16.7\% & \text{if } f = f_2 \\ 1.1\% & \text{if } f = f_1 \& f_2 \end{cases}$$

så kan en $\langle\rangle$-notation användas

$$P(FM|r > J) = \langle 0.27, 0.55, 0.17, 0.01 \rangle$$

Grundläggande samband/operationer

Marginalisering

$$P(y) = \sum_x P(x, y)$$

Kedjeregeln

$$P(x_1, \ldots, x_n) = \prod_i P(x_i|x_1, \ldots, x_{i-1})$$

exempelvis med $n = 3$

$$P(x_1, x_2, x_3) = P(x_1)\, P(x_2|x_1)\, P(x_3|x_1, x_2)$$

X och Y oberoende

$$P(x|y) = P(x)$$

Betingade sannolikheter

Viktig operation är att uppdatera sannolikheter (eng. belief) när ny data (evidence) inkommer.

- ny data, kan vara när ett test larmar eller nya mätningar görs

Betingad sannolikhet

$$P(x|y) = \frac{P(x,y)}{P(y)} = \frac{P(y|x)P(x)}{P(y)}.$$

- $P(x)$ - prior
- $P(x|y)$ - posterior
- Tolkning: hur förändras kunskapen om X när vi får informationen att Y har värdet y

Outline

Sannolikhetsbaserade modeller

En sannolikhetsbaserad modell för de diskreta stokastiska variablerna $\mathcal{X} = \{X_1, \ldots, X_n\}$ är sannolikhetsfunktionen (joint probability mass function)

$$P(x_1, \ldots, x_n)$$

som ersätter ekvationerna som relaterar variablerna i en deterministisk modell.

I det introducerande exemplet, tre boolska variabler A, F_1, F_2 har en modell

$$P(a, f_1, f_2) = P(a|f_1, f_2)\, P(f_1|f_2)\, P(f_2) = P(a|f_1, f_2)\, P(f_1)\, P(f_2).$$

Sannolikhetsmodell för det introducerande exemplet

Sannolikhet	Formel	Värde
A priori-sannolikhet för fel i	$P(f_i),\ i = 1, 2$	0.02
Falsklarm	$P(r > J\|FM = NF)$	0.01
Känslighet för enkelfel f_1	$P(r > J\|FM = f_1)$	0.99
Känslighet för enkelfel f_2	$P(r > J\|FM = f_2)$	0.30
Känslighet för dubbelfel $f_1\&f_2$	$P(r > J\|FM = f_1\&f_2)$	0.99

$$P(a, f_1, f_2) = P(a|f_1, f_2)\, P(f_1|f_2)\, P(f_2) = P(a|f_1, f_2)\, P(f_1)\, P(f_2).$$

a	f_1	f_2	$P(a, f_1, f_2)$
False	False	False	0.9508
False	False	True	0.0137
False	True	False	0.0002
False	True	True	$4 \cdot 10^{-6}$
True	False	False	0.0096
True	False	True	0.0059
True	True	False	0.0194
True	True	True	0.0004

Inferens i sannolikhetsmodeller

Inferens (här)

Beräkna sannolikheter för vissa variabler givet värden på andra

$$P(F_1 | r_1 > J_1, r_3 > J_3)$$

Inferensuttryck

$$P(x|e) = \frac{P(x,e)}{P(e)} = \alpha\, P(x,e) = \alpha \sum_z P(x,e,z)$$

där normaliseringsfaktorn α bestäms ur

$$1 = \sum_x P(x|e) = \alpha \sum_x P(x,e)$$

Inferens i det introducerande exemplet

$$P(f_1|a) = \alpha\, P(f_1, a) = \alpha \sum_{f_2} P(f_1, f_2, a) =$$

$$= \alpha\,(P(f_1, \neg f_2, a) + P(f_1, f_2, a)) = \alpha(0.0194 + 0.0004) = \alpha \cdot 0.0198$$

Motsvarande för $P(\neg f_1|a)$ ger

$$P(\neg f_1|a) = \alpha \cdot 0.0155$$

och alltså

$$P(F_1|a) = \alpha\langle 0.0198, 0.0155\rangle = \langle 0.439, 0.561\rangle$$

Notera att, från de inledande räkningarna, så är

$$P(FM = f_1|a) = 0.55 \neq 0.561$$

Beror på att $FM = f_1$ var enkelfelsmoden bara medans f_1 är sann även i dubbelfelsmoden $FM = f_1 \& f_2$.

Modell och inferenskomplixitet

- inferens rättframt, utvärdera

$$P(x|e) = \alpha \sum_z P(x,e,z)$$

- n stycken (binära) variabler ger att $P(x_1, \ldots, x_n)$ har 2^n värden $\Rightarrow$ kombinatorisk explosion
- Nyckeln är att utnyttja oberoende mellan variabler, jmf sannolikhetsmodellen för det introducerande exemplet.
- Med n oberoende (binära) variabler blir det n parametrar.

$$P(x_1, \ldots, x_n) = \prod_{i=1}^{n} P(x_i)$$

dvs. exponentiellt antal parametrar har transformerats till linjär tillväxt.
- $P(x_1, \ldots, x_n)$ är gles

Här kommer Bayesianska nätverk in i bilden

Outline

Utnyttja oberoende

Oberoende

x_1, x_2 helt oberoende, x_3 beroende enbart av x_1 och x_2, och x_4 och x_5 är både beroende enbart av x_3

$$P(x_1, x_2, x_3, x_4, x_5) = P(x_1)P(x_2)P(x_3|x_1, x_2)P(x_4|x_3)P(x_5|x_3)$$

vilket ger 10 parametrar istället för $2^5 - 1 = 31$.

Betingat oberoende

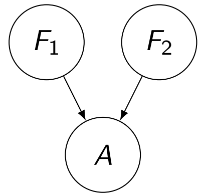
$$P(x_1, x_2, x_3) = P(x_1|x_2, x_3)P(x_2|x_3)P(x_3) = P(x_1|x_3)P(x_2|x_3)P(x_3)$$

Variablerna X_1 och X_2 ej oberoende, men

$$P(x_1, x_2|x_3) = \frac{P(x_1, x_2, x_3)}{P(x_3)} = \frac{P(x_1|x_3)P(x_2|x_3)P(x_3)}{P(x_3)} =$$
$$= P(x_1|x_3)P(x_2|x_3).$$

Bayesianska nätverk

I det introducerande exemplet med felen F_1, F_2, och alarm A så kan beroenden beskrivas med grafen



Kedjeregel och beroenden ger att

$$P(a, f_1, f_2) = P(a|f_1, f_2)P(f_1|f_2)P(f_2) = P(a|f_1, f_2)P(f_1)P(f_2)$$

dvs. en sannolikhetstabell för varje nod i grafen karakteriserar den totala sannolikhetsfunktionen

Kedjeregel för Bayesianska nätverk

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

Bayesianskt nätverk, definition

Definition (Bayesian network)

Let $\mathcal{X} = \{X_1, \ldots, X_n\}$ be a set of random variables with a finite set of values for each variable. A Bayesian network is then a pair $\mathcal{B} = \langle \mathcal{G}, \mathcal{P} \rangle$ where $\mathcal{G}$ is an acyclic directed graph, defined on the nodes $\mathcal{X}$, and $\mathcal{P}$ a set of conditional probability tables, one for each node in the graph, defined as

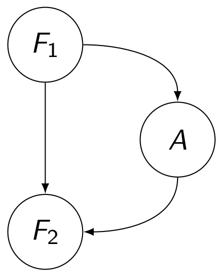
$$P(x_i|\text{parents}(x_i)).$$

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

Ett Bayesianskt nätverk är en representation av den totala sannolikhetsfunktionen (joint probability mass function) där beroendena mellan variablerna är explicit uttryckta i den acykliska grafen.

Bayesianskt nätverk

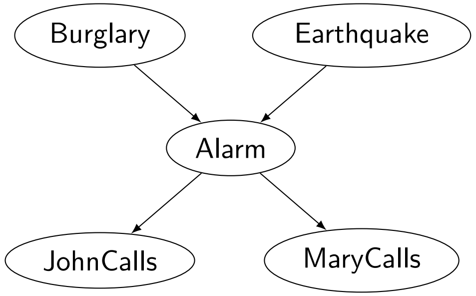
- Det finns ej ett unikt bayes-nät för en given sannolikhetsfunktion
- Varje variabelordning svarar mot en viss faktorisering av sannolikhetsfunktionen



$$P(a, f_1, f_2) = P(f_1)P(a|f_1)P(f_2|a, f_1)$$

- principiellt inga hinder mot cykler i beroendegrafen, men då tappar man möjligheter att göra effektiva inferensalgoritmer

Demonstrera (earthquake.xdsl)



B = Burglary
E = Earthquake
A = Alarm
J = JohnCalls
M = MaryCalls

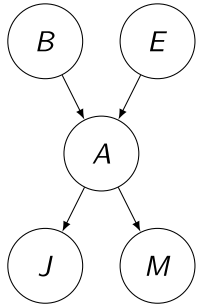
där $P(\text{Burglary}) = 0.001$ och $P(\text{Earthquake}) = 0.002$ samt

B	E	$P(A\|B,E)$
falsk	falsk	0.001
falsk	sann	0.29
sann	falsk	0.94
sann	sann	0.95

A	$P(J\|A)$
falsk	0.05
sann	0.90

M	$P(M\|A)$
falsk	0.01
sann	0.70

Inferens i Bayesianska nätverk



B = Burglary
E = Earthquake
A = Alarm
J = JohnCalls
M = MaryCalls

$$P(b|j,m) = \alpha \sum_e \sum_a P(b,e,a,j,m) =$$
$$= \alpha \sum_e \sum_a P(b)P(e)P(a|b,e)P(j|a)P(m|a) =$$
$$\alpha\, P(b) \sum_e P(e) \sum_a P(a|b,e) \underbrace{P(j|a)P(m|a)}_{\text{beror ej på } e}$$

Algoritmer för inferens

- inte ämne för den här kursen
- kan delas in i två kategorier
  - exakt inferens
  - approximativ inferens
- exakt inferens NP-svårt
- variable elimination
- poly-tree, endast en väg mellan två noder
- poly-trees enkla, join-trees slår ihop noder för att få poly-trees

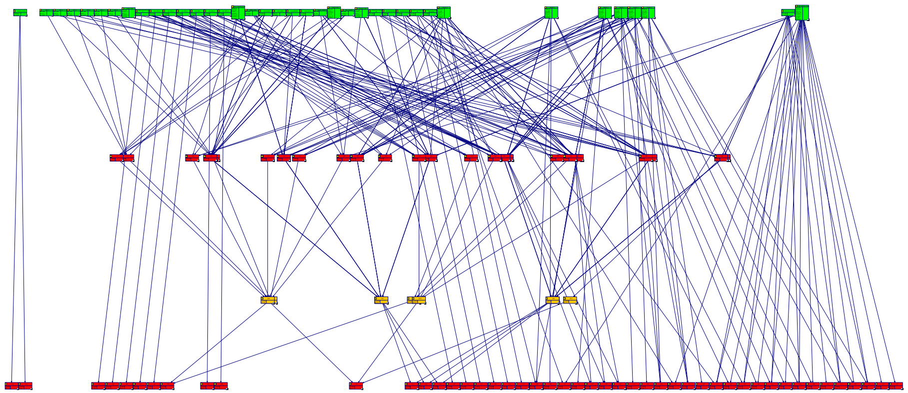
Demo, XPI fuel injection system, Scania (XPI.xdsl)

- 141 variabler
- 40 svarar mot komponenter som kan gå sönder
- Resten svarar mot observationer och diagnostester
- Många variabler har två värden, men vissa har upp till 8 möjliga
- En ventil har exempelvis möjligheterna

$$\{\text{Fuel leak}, \text{Electrical fault}, \text{Stuck or clogged}, \text{Wrong pressure}, \text{Emission fault}, \text{Corrosion or cavitation}, \text{Air leak}, \text{No Fault}\}$$

- $P(x_1, \ldots, x_{141})$ har i storleksordningen 10^{50} värden
- Utvecklat i examensarbetet "*Modeling of fuel injection system for troubleshooting*", Cyon. A, KTH, 2012.

XPI fuel injection system, Scania



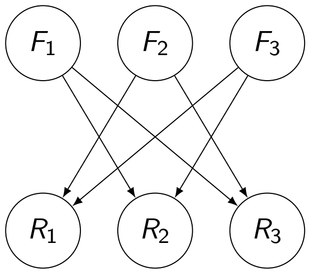
Outline

- Introduktion, sneak-peak
- Sannolikhetsbaserad diagnos
- Introducerande exempel
- Notation och lite repetition
- Sannolikhetsbaserade modeller, inferens, och komplexitet
- Bayesianska nätverk
- Kanoniska modeller
- Sammanfattning

Kanoniska modeller

- Grundproblem; ett stort antal parametrar/sannolikheter behöver bestämmas i en sannolikhetsmodell
- expertkunskap eller mycket data
- utnyttja strukturen i Bayes-nät, men problem dyker upp med nod med många (säg > 4) föräldrar
- Kanoniska modeller, parametriserade noder, mallar, ...
- XPI-modellen använder sig flitigt av kanoniska modeller, den vanligaste är *leaky or*-noder.

Enkelt felisoleringsexempel



	F_1	F_2	F_3
r_1	0	X	X
r_2	X	0	X
r_3	X	X	0

- funktionen är or vid alarm-noderna
- vi vill kunna modellera falsklarm, missad detektion etc.
- deterministiska modeller räcker inte, det var ju osäkerheter som var den ursprungliga anledningen till att vi introducerade sannolikheter

Deterministisk funktion

För sambandet $y = f(x)$ så blir sannolikhetstabellen

$$P(y|x) = \begin{cases} 1 & y = f(x) \\ 0 & \text{otherwise} \end{cases}$$

Exempelvis för or-funktionen

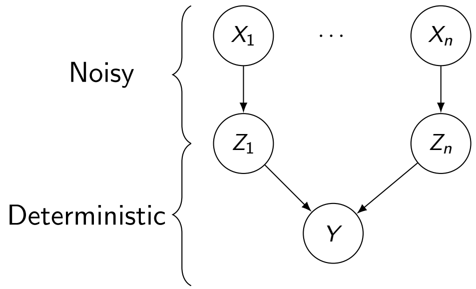
$$y = x_1 \vee x_2$$

så har vi

x_1	x_2	$P(y \mid x_1, x_2)$	$P(\neg y \mid x_1, x_2)$
false	false	0	1
false	true	1	0
true	false	1	0
true	true	1	0

Deterministiska modeller har 0 parametrar

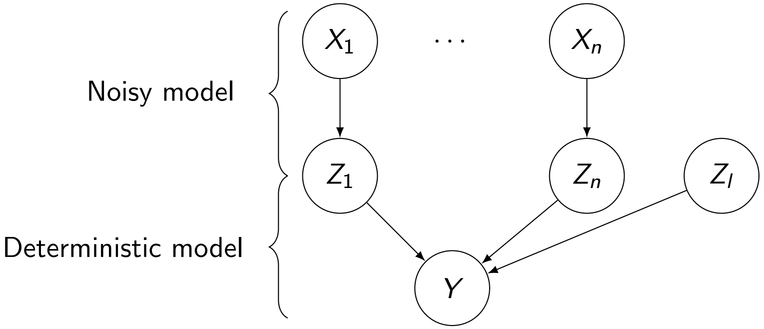
(binary) Noisy model



x_i	$P(z_i \mid x_i)$	$P(\neg z_i \mid x_i)$
False	0	1
True	c_1	$1 - c_1$

- 0 parametrar i den deterministiska delen
- 1 parameter (c_i) per variabel $\Rightarrow$ linjär tillväxt i parametrar
- $c_i = 1$ svarar mot deterministisk modell
- Att residual r_3 reagerar för fel f_1 respektive fel f_2 med sannolikhet 0.9 och 0.6 respektive kan då modelleras med noisy-or där $c_1 = 0.9$ och $c_2 = 0.6$.
- Inga falsklarm dock!

(binary) Noisy-leaky model



- Sannolikhetestabell för leak-node tillkommer
- Om sannolikheten för falsklarm är 0.01 så modelleras det i exemplet genom en noisy-leaky-or enligt tidigare och $P(Z_l) = \langle 0.99, 0.01 \rangle$.
- Noisy-leaky-or ofta bara noisy-or
- Noisy-or kan generaliseras till icke-binära variabler och kallas då noisy-max

Fortsättning på felisoleringsdemo

- Residualnoderna är leaky-or med falsklarmssannolikhet på 0.05 samt icke-ideala tester
- Illustrera hur BN kan användas vid feldetektion, bara felisolering, falsklarm

Outline

Sammanfattning

- Sannolikhetsbaserad diagnos
- Probabilistiska modeller
- Exakt inferens och komplexitetsproblem
- Bayesianska nätverk
- Kanoniska modeller

TSFS06 Diagnos och övervakning
Föreläsning 10 - Sannolikhetsbaserad diagnos och Bayesianska nätverk

Erik Frisk

Institutionen för systemteknik
Linköpings universitet
erik.frisk@liu.se

2020-05-12