

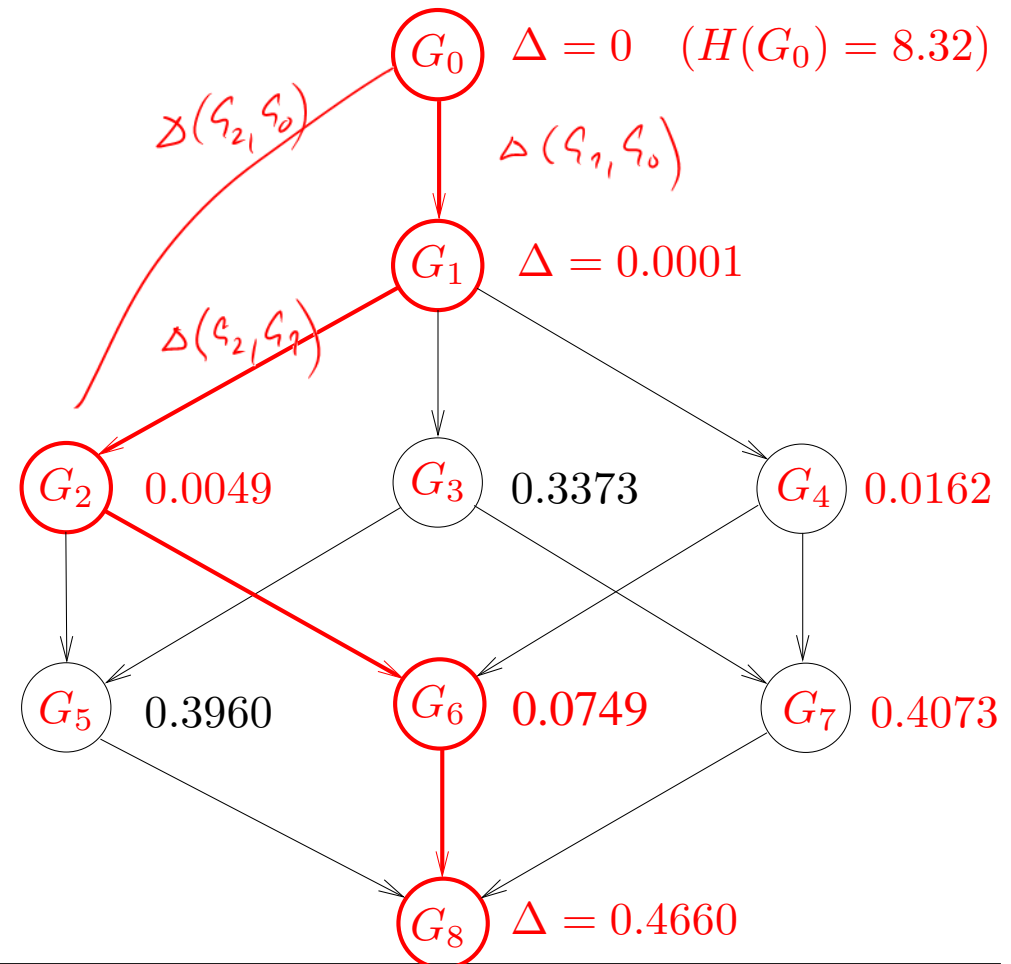
Prohledávání svazu zjemnění

- Rekonstrukční chyba Δ je monotonně neklesající podél každé cesty svazu zjemnění: Je-li $G_i \succ G_k \succ G_j$ potom $\Delta(G_i) \leq \Delta(G_k) \leq \Delta(G_j)$
- Rekonstrukční chyba Δ je aditivní podél každé cesty svazu zjemnění: je-li $G_i \succ G_k \succ G_j$, potom $\Delta(G_j, G_i) = \Delta(G_k, G_i) + \Delta(G_j, G_k)$

[Higashi 1983]

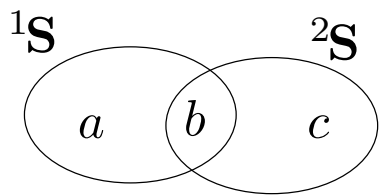
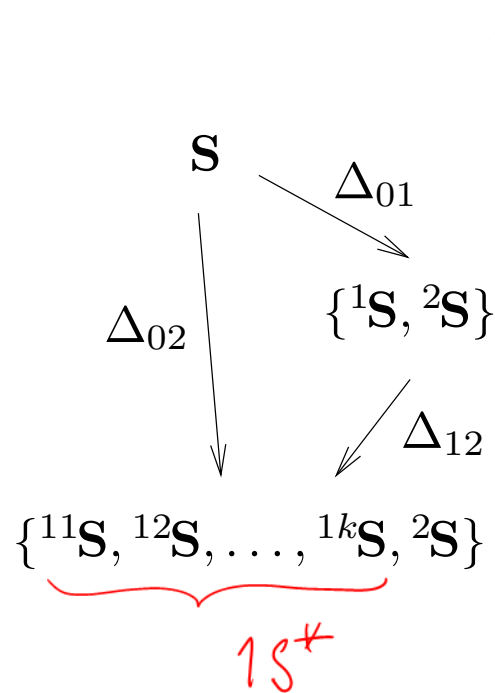
$\overline{G_0}$ $\overline{G_1}$ $\overline{G_2}$

- **Nejjednodušší postup:** Hledáme cestu, podle které Δ roste nejpomaleji
- **Složitější postup:** Na každé úrovni nás zajímá množina řešení s nejmenším Δ
 - Není nutno expandovat všechny větve: postačí metoda větví a mezí [Narendra&Fukunaga 1977]



Aditivita podél cesty ve svazu zjemnění rozkladu

- pro jednoduchost předpokládáme, že jednoduchá spojovací procedura stačí k nestrannému spojení



①

$$\frac{{}^1p(a,b) \cdot {}^2p(c|b)}{{}^1p^*(a,b) \cdot {}^2p(c|b)} \Delta_{12} = \sum_{i,j,k} \frac{{}^1p(a_i, b_j) \cdot {}^2p(c_k | b_j) \cdot \log \frac{{}^1p(a_i, b_j) \cdot {}^2p(c_k | b_j)}{{}^1p^*(a_i, b_j) \cdot {}^2p(c_k | b_j)}}{{}^1p(a_i, b_j) \cdot {}^2p(c_k | b_j)} = L({}^1S, {}^1S^*)$$

$${}^1S^* = {}^{11}S * {}^{12}S * \dots * {}^{1k}S$$

$$\Delta_{02} \stackrel{?}{=} \Delta_{01} + \Delta_{12} \quad \text{②}$$

$$\sum_{i,j,k} p(a_i, b_j, c_k) \log \frac{p(a_i, b_j, c_k) \cdot \cancel{{}^1p(a_i, b_j)}}{\cancel{{}^1p(a_i, b_j)} \cdot \cancel{{}^2p(c_k | b_j)} \cdot \cancel{{}^1p^*(a_i, b_j)}} = \Delta_{02}$$

$$\Delta_{01} = L(S, {}^1S * {}^2S) = \sum_{i,j,k} p(a_i, b_j, c_k) \log \frac{p(a_i, b_j, c_k)}{p^*(a_i, b_j, c_k)} = {}^1p(a_i, b_j) \cdot {}^2p(c_k | b_j)$$

$$\Delta_{12} = L({}^1S * {}^2S, {}^1S^* * {}^2S) = L({}^1S, {}^1S^*) = \sum_{i,j} \frac{{}^1p(a_i, b_j)}{{}^1p^*(a_i, b_j)} \log \frac{{}^1p(a_i, b_j)}{{}^1p^*(a_i, b_j)}$$

$$\Delta_{02} = L(S, {}^1S^* * {}^2S)$$

$$\sum_k p(a_i, b_j, c_k)$$

Metoda větví a mezí na svazu zjemnění

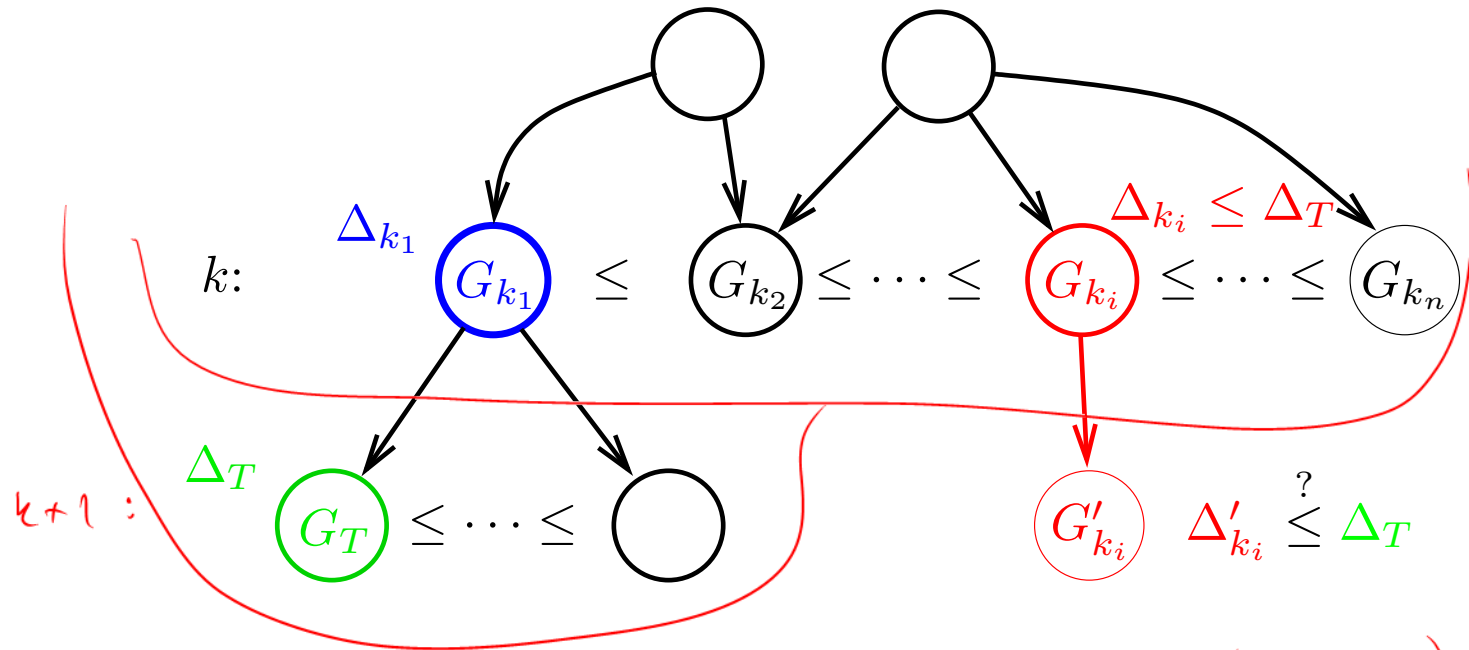
HACDA: N prvků

konstrukce $O(N)$

vložení $O(\log N)$

vyjmutí kořene $O(\log N)$

Sedgwick Algorithmus



1. vygeneruj úroveň k
2. nalezni uzal G_{k_1} s nejmenším Δ a expanduj ho
3. najdi G_T s nejmenším Δ mezi expandovanými uzly
4. zvol práh $T = \Delta(G_T)$
5. expanduj $G_{k_2}, G_{k_3}, \dots, G_{k_n}$ **jen pokud** $\Delta(G_{k_i}) \leq T$
6. (aktualizuj T , pokud nalezneš menší hodnotu prahu)
7. pokračuj stejně na úrovni $k + 1$

1. založ haldnu $H := (G_0, \Delta=0)$
2. $G_x :=$ vyjmi kořen z H
3. vlož všechny bezprostřední zjemnění G_x do H (test na duplikace)
4. opakuj 2-3 dokud $H \neq \emptyset$

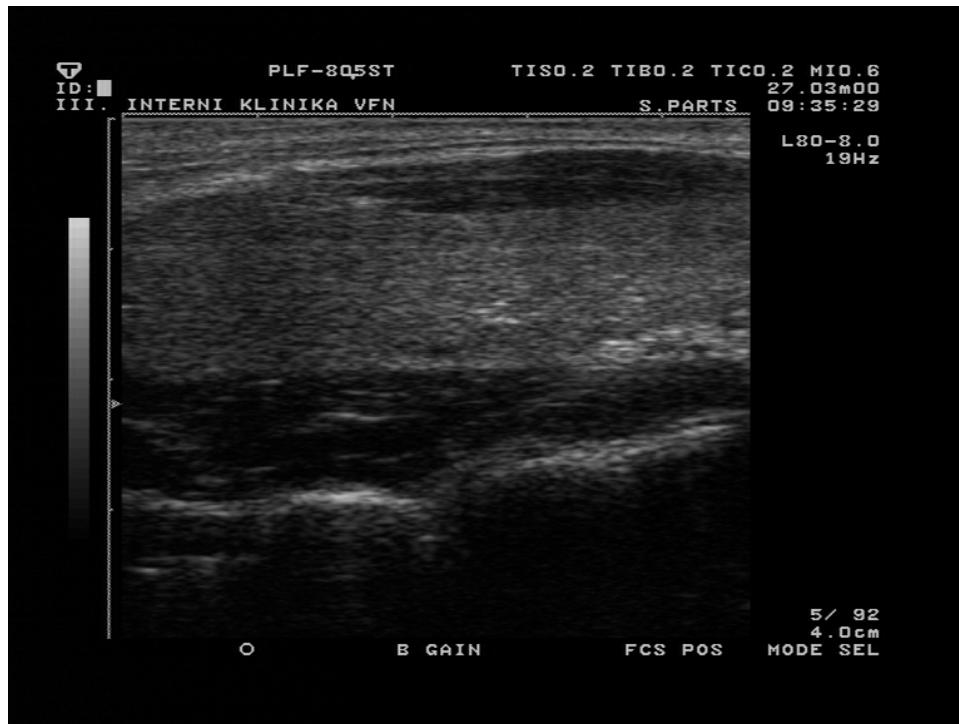
- Takto lze jen proto, že platí monotonicita Δ !
- Toto není obecná metoda větví a mezí

Identifikační procedura

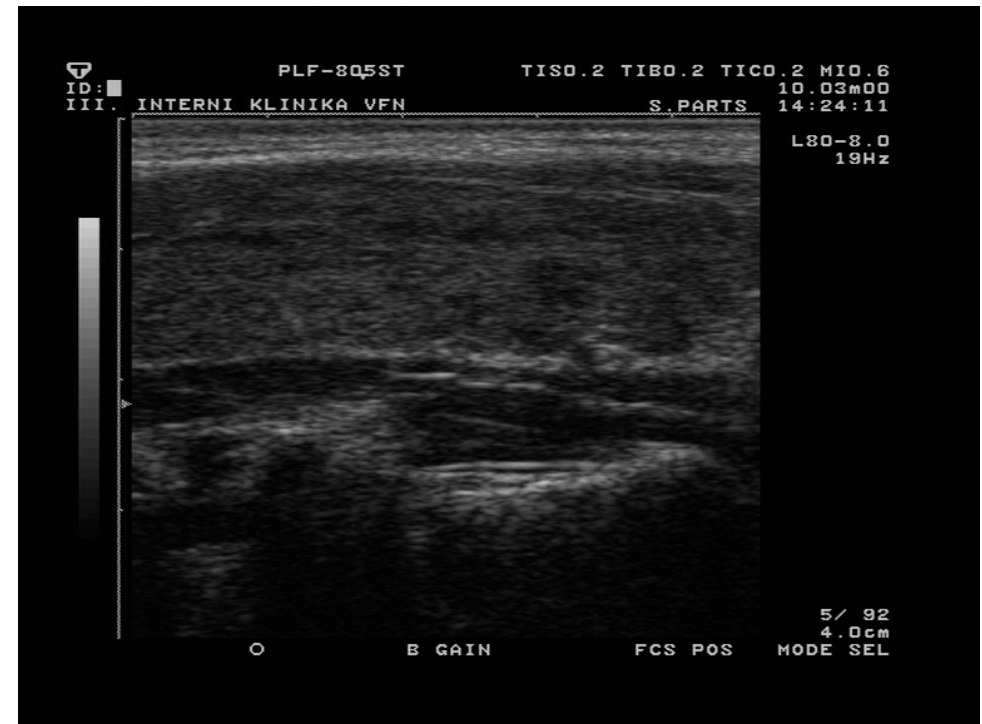
1. Identifikuj zobecněný dynamický systém S z datového systému D .
2. Vytvoř rekonstrukční hypotézu $G_0 = \{S\}$.
3. Prohledej svaz zjemnění hypotézy G_0 , na každé úrovni zjemnění zaznamenej hypotézu G_k o nejmenší hodnotě $\Delta(G_k)$.
4. Výsledkem je soubor nejlepších dekompozičních hypotéz klesajícího stupně strukturní složitosti.

Sonogram štítné žlázy v podélném řezu

zdravá



lymfocitická thyroitida



Příklad: Klasifikace textury

Aplikace

Diagnóza difúzních změn ve štítné žláze.

Pracovní hypotéza

Klasifikace je možná na základě textury.

Volba

Model založený na zobecněném dynamickém systému (kookurenční matici).

Základní problém

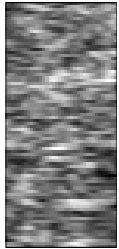
Zvolit co nejmenší model textury založený na funkci přípustnosti.

Důvody:

1. klasifikace je jednodušší
2. klasifikátor je mnohem snadnější naučit

Rekonstrukční analýza

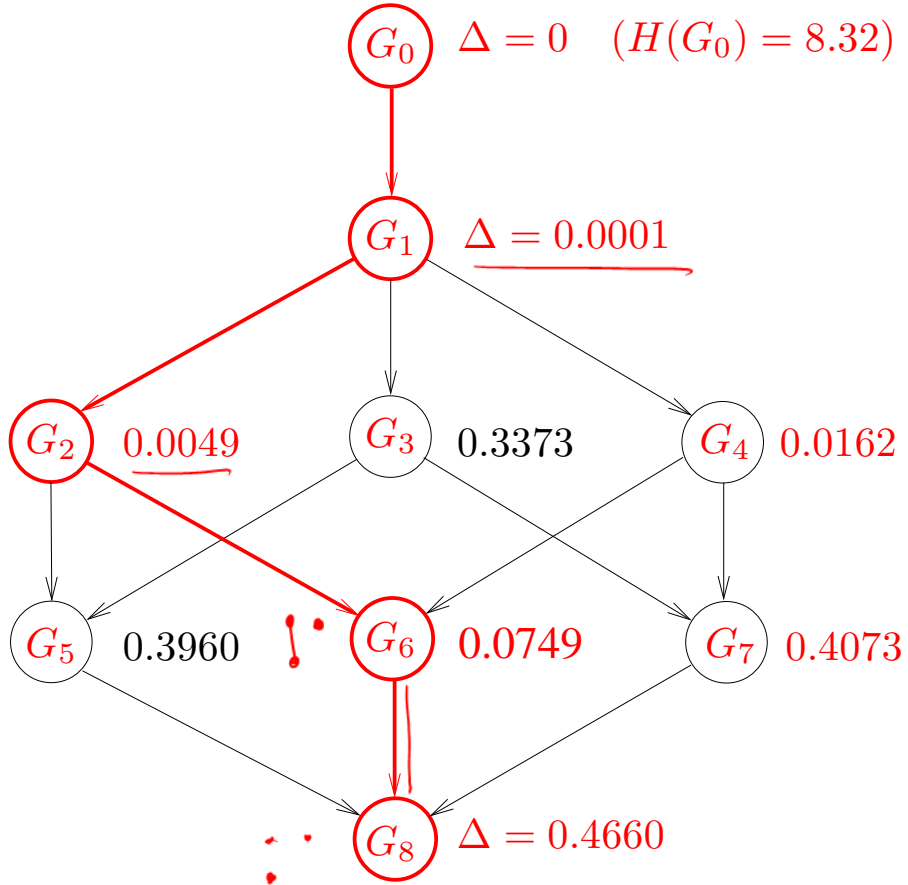
vzorek textury:



maska:

1	2
3	

svaz zjemnění



rekonstrukční hypotézy

- $G_0 = 123$ 4^3
- $G_1 = \underline{12}|\underline{13}|\underline{23}$ $3 \cdot 4^2$
- $G_2 = 12|13$
- $G_3 = 12|23$
- $G_4 = 13|23$
- $G_5 = 12|3$
- $G_6 = 13|2$
- $G_7 = 1|23$
- $G_8 = 1|2|3$

$p(a, b, c)$
 1 1 1
 2 2 3
 4^3
 4^2

Závěr: kookurenční matici má smysl počítat pouze ve svislém směru

Infanticida u myší

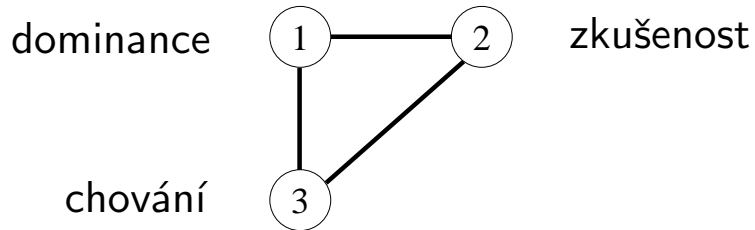
Test hypotézy sexuálního soutěžení: samec zvyšuje svůj sexuální úspěch na úkor konkurentů tím, že zabíjí jeho potomstvo a páří se s jejich matkou.

Studie: 114 jedinců 90-denních myší

Funkce přípustnosti stavu

	v_1	v_2	v_3	N
$\underline{v_1}$ dominance	0	<u>0</u>	<u>0</u>	28
	0	0	1	4
	0	0	2	2
	0	1	0	5
$\underline{v_2}$ sexuální zkušenost	0	1	1	25
	0	1	2	3
	1	0	0	5
$\underline{v_3}$ chování vůči potomstvu	1	0	1	9
	1	0	2	8
	1	1	0	7
	1	1	1	15
	1	1	2	3

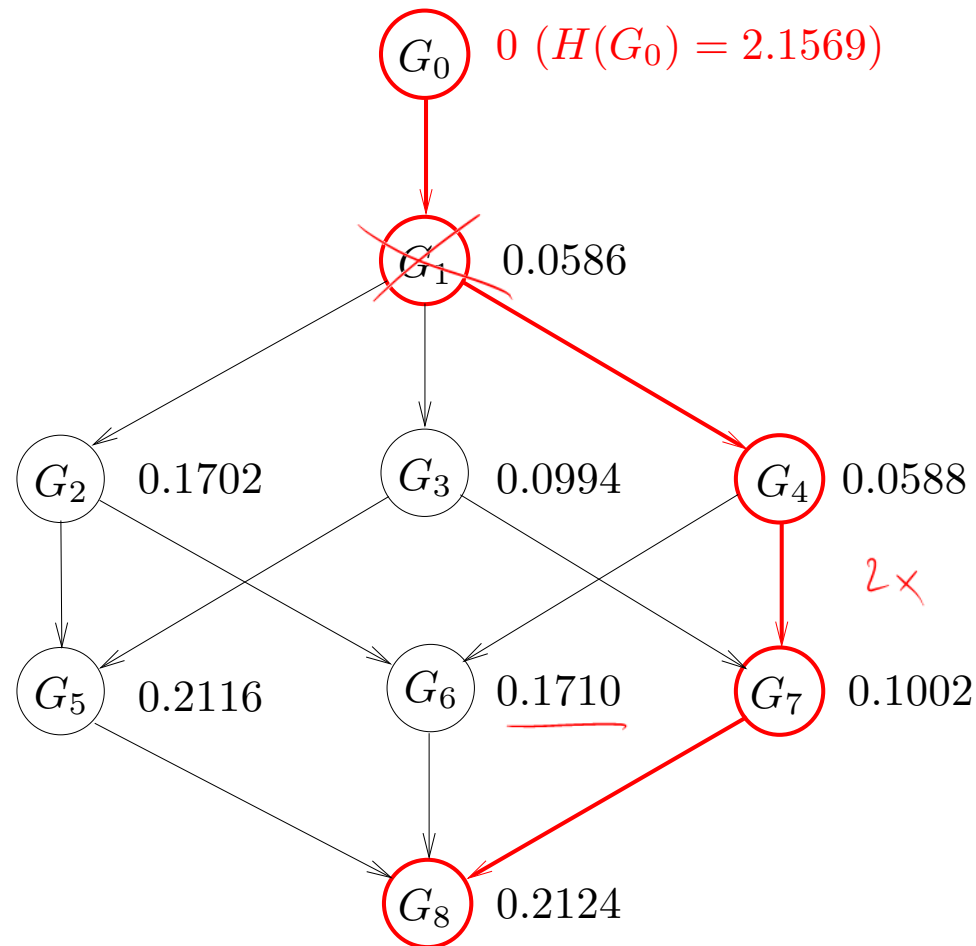
Infanticida u myší: rekonstrukční analýza



o

zkušenost

svaz zjemnění



rekonstrukční
hypotézy

$$G_0 = 123$$

$$\begin{array}{c} \bullet \\ \diagdown \\ \bullet \end{array} G_1 = 12|13|23$$

$$\begin{array}{c} \bullet \\ \text{---} \\ \bullet \end{array} G_2 = 12|13$$

$$\begin{array}{c} \bullet \\ \diagup \\ \bullet \end{array} G_3 = 12|23$$

$$\begin{array}{c} \bullet \\ \diagdown \\ \bullet \\ \diagup \\ \bullet \end{array} G_4 = 13|23$$

$$\begin{array}{c} \bullet \\ \text{---} \\ \bullet \\ \text{---} \\ \bullet \end{array} G_5 = 12|3$$

$$\begin{array}{c} \bullet \\ \text{---} \\ \bullet \\ \text{---} \\ \bullet \\ \text{---} \\ \bullet \end{array} G_6 = 13|2$$

$$\begin{array}{c} \bullet \\ \diagup \\ \bullet \\ \diagdown \\ \bullet \end{array} G_7 = 1|23$$

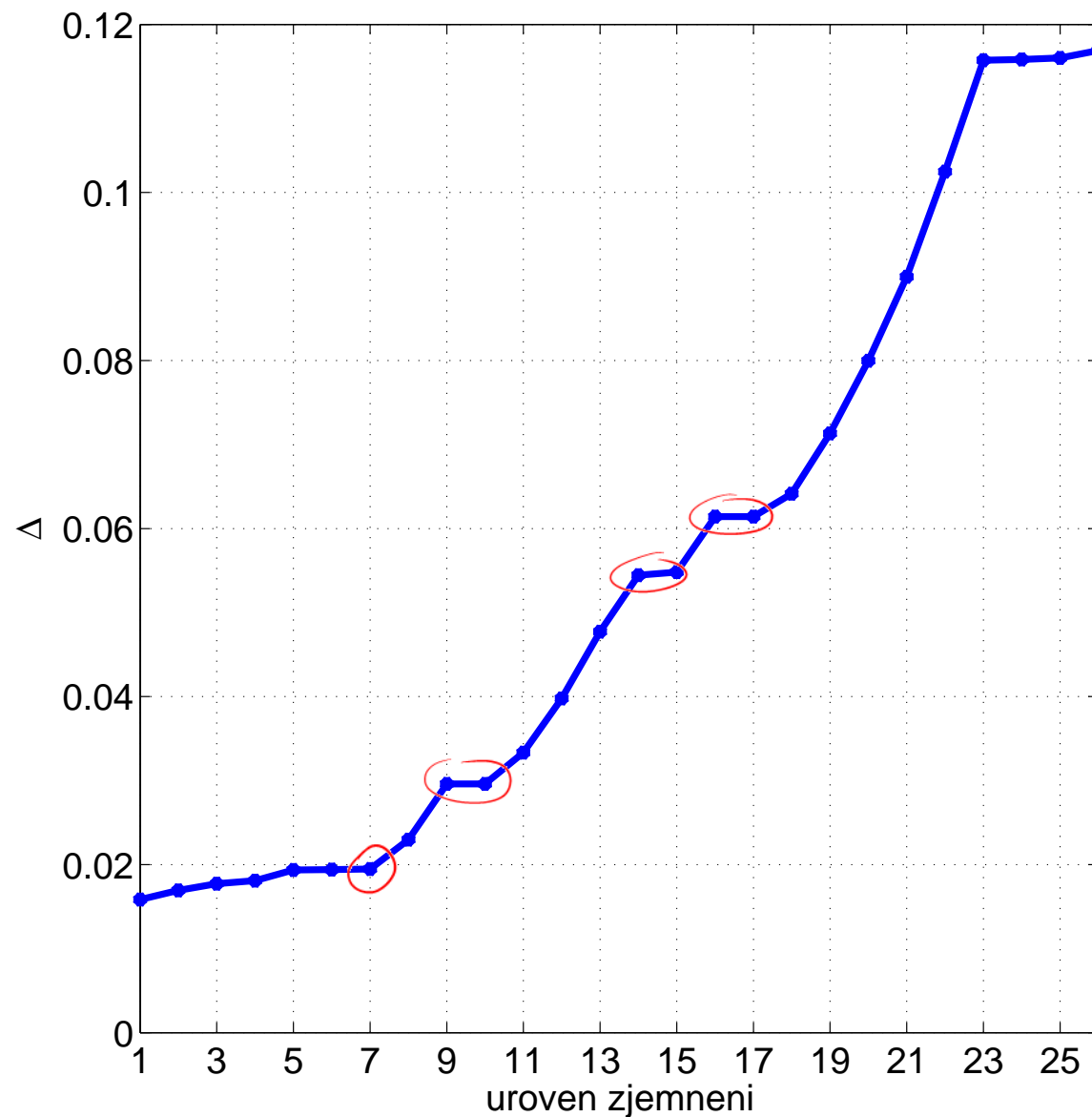
$$\begin{array}{c} \bullet \\ \text{---} \\ \bullet \\ \text{---} \\ \bullet \\ \text{---} \\ \bullet \\ \text{---} \\ \bullet \end{array} G_8 = 1|2|3$$

Závěry

- chování je ovlivněno oběma faktory, ale zkušenost je dvakrát významnější než dominance (ΔG_6 , ΔG_7)
- dominance a zkušenost jsou nezávislé ($\Delta G_3 \sim \Delta G_7$, první vazba, která se rozpadne)

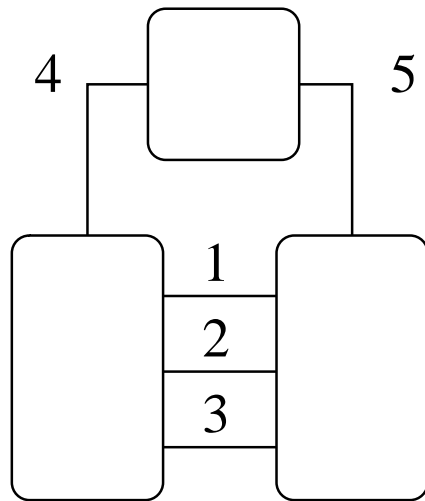
Struktura Parlamentu

k	G_k	$\Delta(G_k)$
1	1234 2345 3451 4512 5123	0.0158
2	1234 2345 3451 5123	0.0169
3	1234 2345 5123 451	0.0177
4	1234 2345 5123	0.0181
5	1234 5123 345 452	0.0194
6	1234 5123 345	0.0194
7	1234 5123 45	0.0195
8	5123 45 234 341 412	0.0230
9	45 234 341 412 512 123 235 351	0.0296
10	45 234 341 412 123 235 351	0.0296
11	45 234 341 412 235 351	0.0333
12	45 234 341 412 235 51	0.0398
13	45 234 341 235 51 12	0.0477
14	45 341 235 51 12 42	0.0544
15	45 341 235 51 12	0.0548
16	45 235 51 12 34 41 13	0.0614
17	45 235 51 12 34 13	0.0614
18	45 235 51 12 13	0.0641
19	45 235 51 12	0.0713
20	235 51 12 4	0.0800
21	235 12 4	0.0900
22	235 4 1	0.1025
23	4 1 23 35 52	0.1158
24	4 1 23 35	0.1158
25	4 1 35 2	0.1160
26	4 1 2 3 5	0.1169

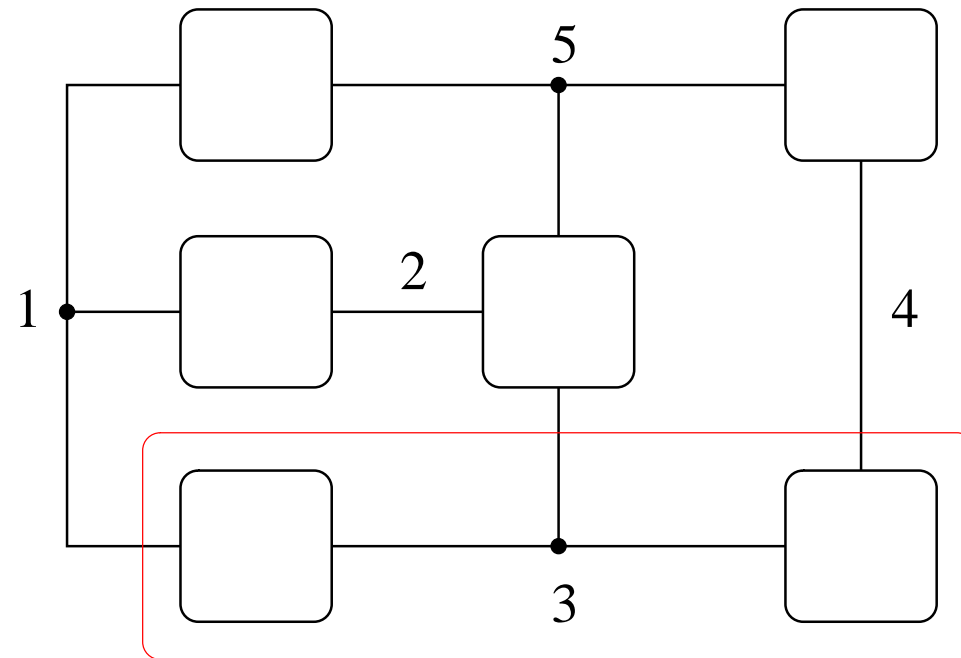


Parlament: interpretace úrovně 7 a 15+17

k	G_k	$\Delta(G_k)$
7	<u>1234</u> <u>5123</u> 45	0.0195



k	G_k	$\Delta(G_k)$
15	45 341 235 51 12	0.0548
17	45 235 51 12 34 13	0.0614



Posibilistická funkce přípustnosti

- posibilita π : stupeň možnosti
- $\pi = 1$. . událost je možná

x_i elementární jevy, $x_i \in S$

posibilita

pravděpodobnost

definiční vlastnosti

$$\pi: \exp S \rightarrow \langle 0, 1 \rangle$$

$$\pi(\{\}) = 0$$

$$\pi(S) = 1$$

$$\pi(\bigcup_i x_i) = \max_i \pi(x_i)$$

$$p: \exp S \rightarrow \langle 0, 1 \rangle$$

$$p(\{\}) = 0$$

$$p(S) = 1$$

$$p(\bigcup_i x_i) = \sum_i p(x_i)$$

odhad z dat

$$\pi(x_i) \approx \frac{N(x_i)}{\max_j N(x_j)}$$

$$p(x_i) \approx \frac{N(x_i)}{\sum_j N(x_j)}$$

Posibilistická míra neurčitosti

$$X_\alpha = \{x_i, \pi(x_i) \geq \alpha\} \quad \alpha\text{-řez}$$

$$L = \{\alpha, \exists x_i: \pi(x_i) = \alpha\} \cup \{0\} \quad \text{množina } \alpha\text{-hladin}$$

U-míra

$$U(\pi) = \sum_{j=2}^k (\alpha_j - \alpha_{j-1}) \log |X_{\alpha_j}| = \int_0^1 \log |X_\alpha| d\alpha$$

x_i	1	2	3	4	5	6
$\pi(x_i)$	0.1	0.0	0.3	1.0	0.3	0.2

příklad

j	α_j	X_{α_j}						$ X_{\alpha_j} $
1	0.0	×	×	×	×	×	×	6
2	0.1	×		×	×	×	×	5
→ 3	0.2			×	×	×	×	4 ←
4	0.3			×	×	×		3
5	1.0				×			1

$$U = \underbrace{0.1 \cdot \log 5}_{j=2} + \underbrace{0.1 \cdot \log 4}_{j=3} + \underbrace{0.1 \cdot \log 3}_{j=4} + \underbrace{0.7 \cdot \log 1}_{j=5}$$

Posibilistické spojení

Posibilistická marginální funkce

$$\pi(x) = \max_y \pi(x, y)$$

Posibilistická generativní neurčitost

$$U(\mathbf{G} \mid \bar{\mathbf{G}}) = U(\mathbf{G}, \bar{\mathbf{G}}) - U(\bar{\mathbf{G}})$$

Posibilistické spojení

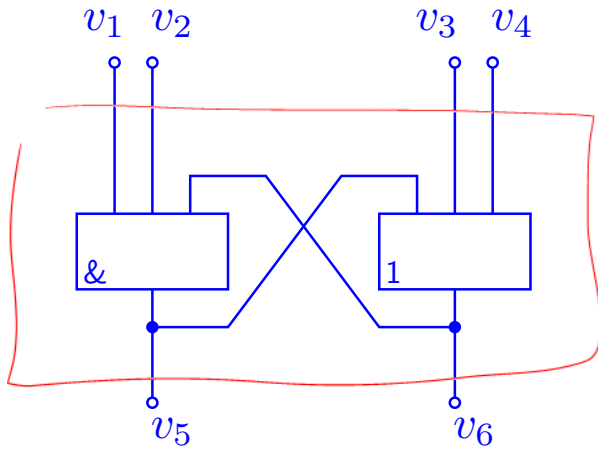
$$({}^1\pi * {}^2\pi)(a, b, c) = \min({}^1\pi(a, b), {}^2\pi(b, c))$$

není nutná iterativní spojovací procedura

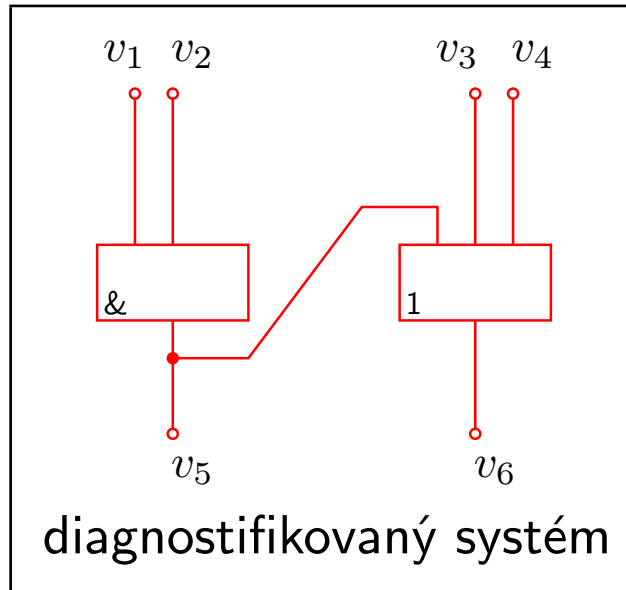
Rekonstrukční chyba

$$\Delta(G) = L(\pi, \pi^*) = \int_0^1 \log \frac{|X_\alpha(\pi^*)|}{|X_\alpha(\pi)|} d\alpha$$

Diagnostika obvodů



předpokládaný systém



diagnostifikovaný systém

není pozorován stav a

pozorované stavy

\bar{g}				g		\bar{V}
v_1	v_2	v_3	v_4	v_5	v_6	
0	0	0	0	0	0	1
0	0	0	1	0	1	1
0	0	1	0	0	1	1
0	0	1	1	0	1	
0	1	0	0	0	0	
0	1	0	1	0	1	
0	1	1	0	0	1	
0	1	1	1	0	1	
1	0	0	0	0	0	
1	0	0	1	0	1	
1	0	1	0	0	1	
1	0	1	1	0	1	
$a = 1$	1	0	0	0	0	
$b = 1$	1	0	0	1	1	
1	1	0	1	1	1	
1	1	1	0	1	1	
1	1	1	1	1	1	

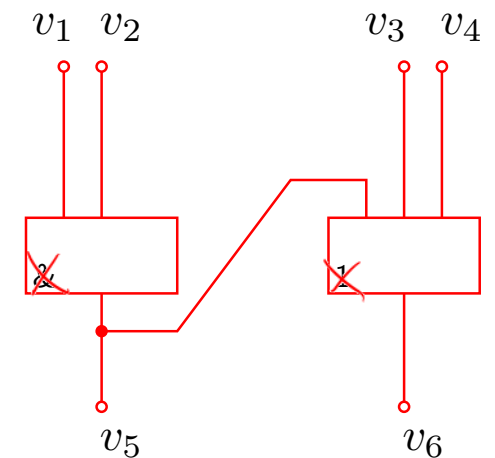
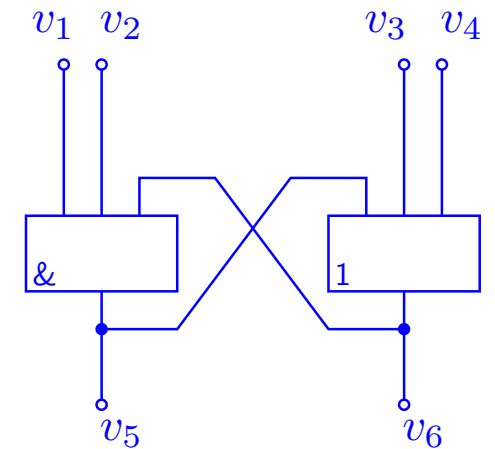
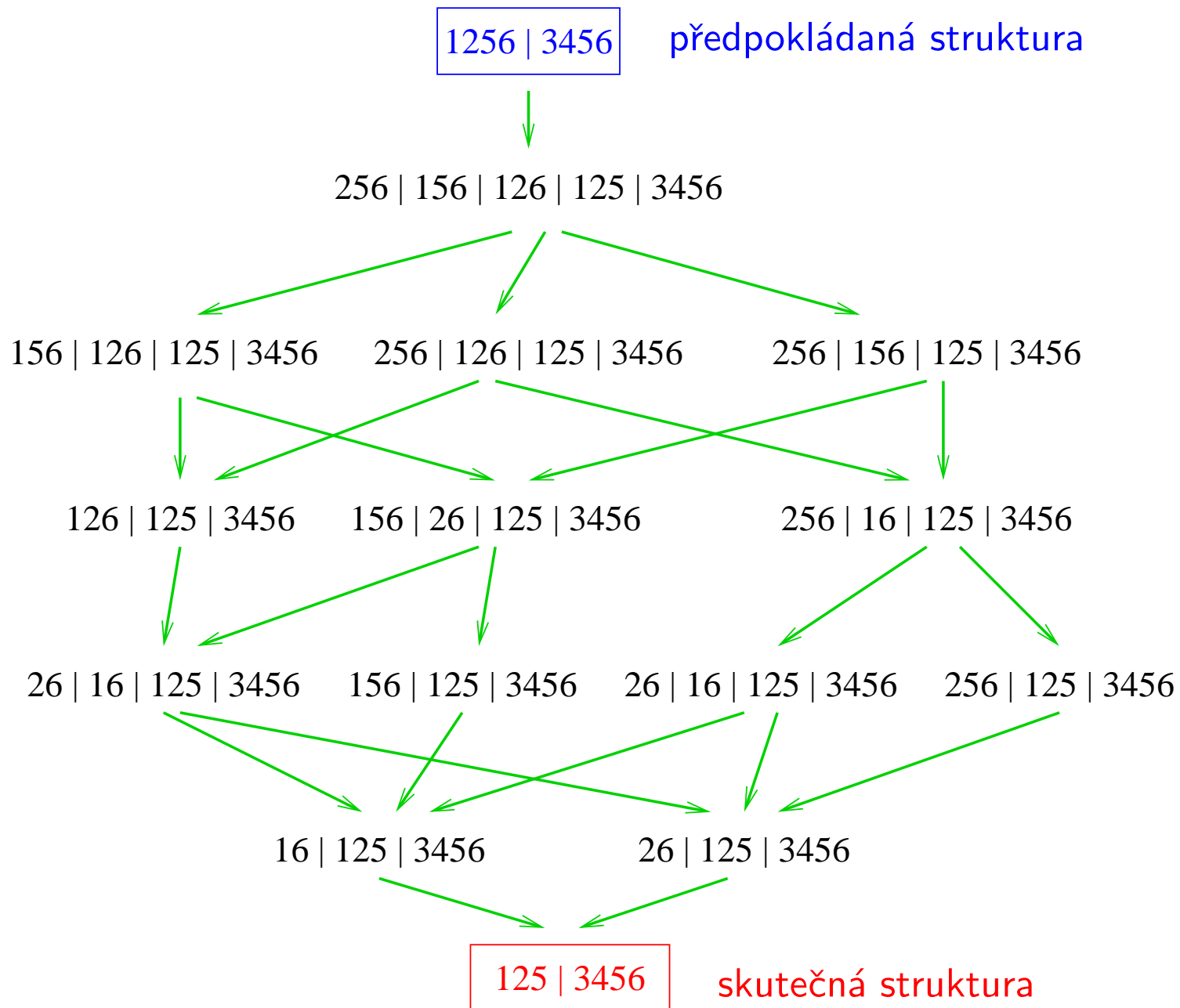
správná rekonstrukční hypotéza $\underline{1256} \mid \underline{3456} = G_A$

- $\Delta G_A > 0 \Rightarrow$ obvod je vadný
- $\Delta G_A = 0$ nedokazuje plnou funkcionalitu

nutná rekonstrukční analýza pro předchůdce G_A

nutná rekonstrukční analýza pro následníky G_A

Podgraf s nulovou rekonstrukční chybou



Pozn: $\Delta(1256 | 3456) = 3$