PC: Identifikace struktury zobecněného dynamického systému

Důležitý problém v obecné teorii systémů.

1. Podsystém a nadsystém.
2. Definice dekompozice systému.
3. Problém rekonstrukce systému:
   a. lokální a globální konzistence dynamických systémů,
   b. jednoduchá a iterativní spojovací procedura.
4. Problém identifikace struktury:
   a. generátor rekonstrukčních hypotéz,
   b. kvalita rekonstrukční hypotézy,
   c. identifikační procedura.
5. Příklad identifikace na skutečném systému.

Podsystém a nadsystém dynamického systému

Def: ${}^iF = ({}^iA, {}^iB;\ {}^iM, {}^ip_B)$ je podsystém systému $F = (A, B;\ M, p_B)$, když platí následující podmínky:

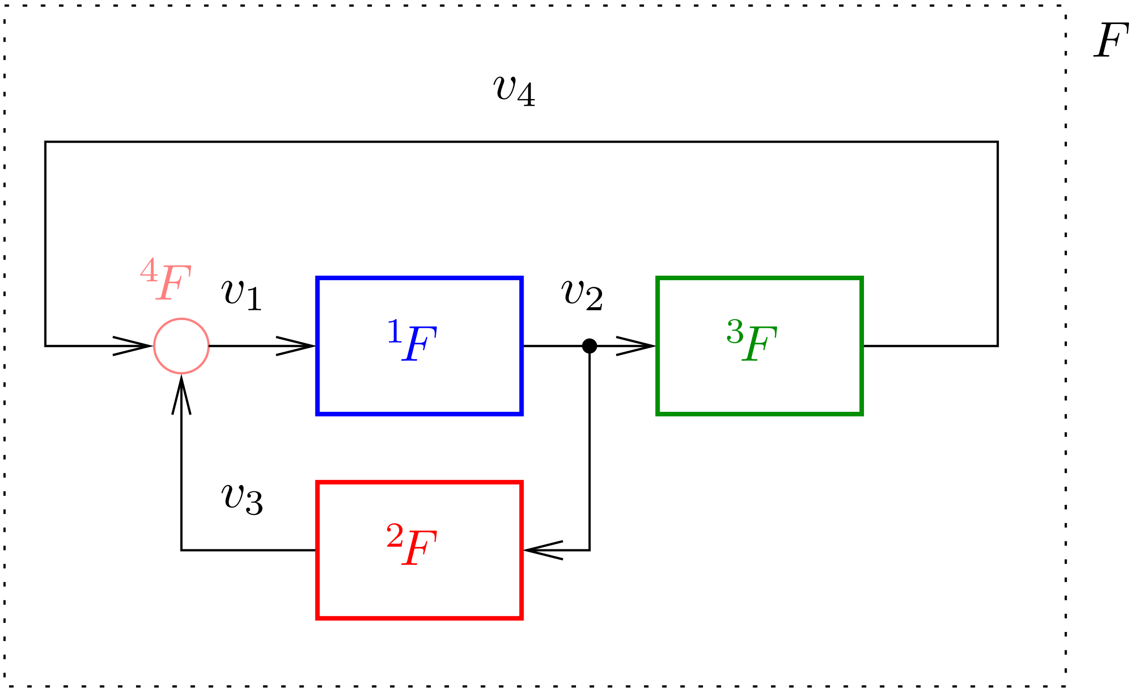
1. **kompatibilita s F** (ztotožnění atributů a parametrů)
   - má stejnou parametrickou množinu: ${}^iB = B$
   - obory hodnot základních proměnných V_j zachovány
2. **vnoření ${}^iF \sqsubseteq F$**
   a. množina vzorkovacích proměnných je vnořena: ${}^i\mathbf{S} \subseteq \mathbf{S}$
   b. (data pro proměnné v ${}^i\mathbf{S}$ jsou zachována)
   c. maska je vnořena ${}^iM \subseteq M$
   d. funkce přípustnosti ${}^ip_B$ je marginální k p_B

$\Rightarrow$ **Hierarchie podsystémů**

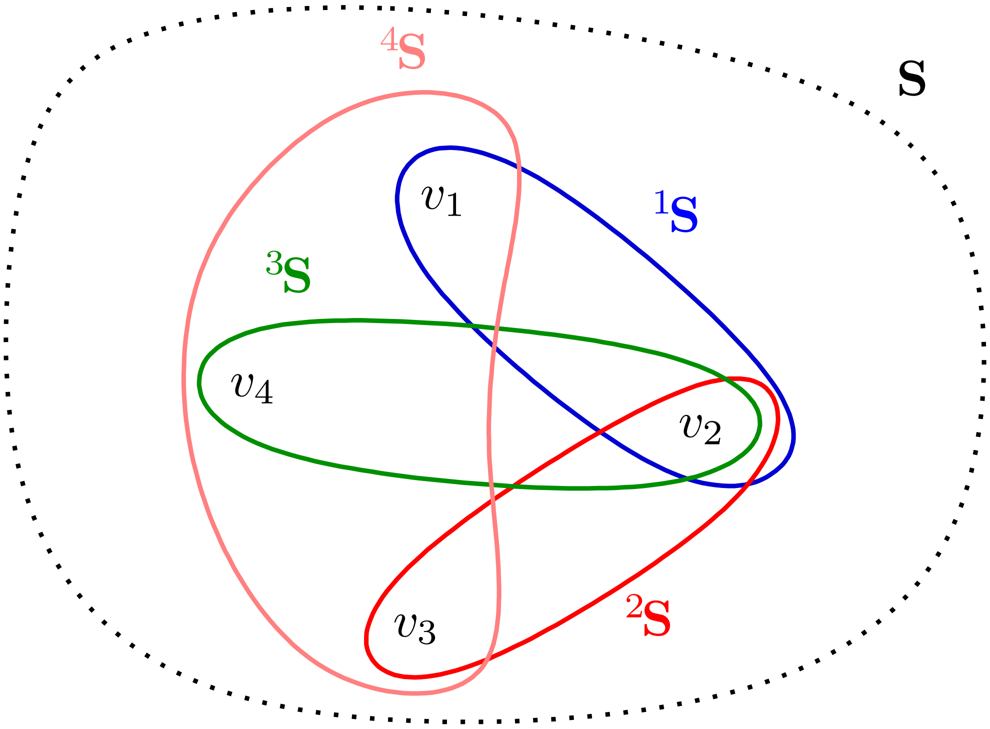
Konvence: $\mathbf{S}$ značí dále pouze množinu (vzorkovacích) proměnných dynamického systému F. Místo ${}^iF \sqsubseteq F$ budeme používat zkráceně ${}^i\mathbf{S} \sqsubseteq \mathbf{S}$.

Dekompozice systému

blokové vyjádření struktury



struktura jako rozklad množiny vzorkovacích proměnných



Dekompozice systému

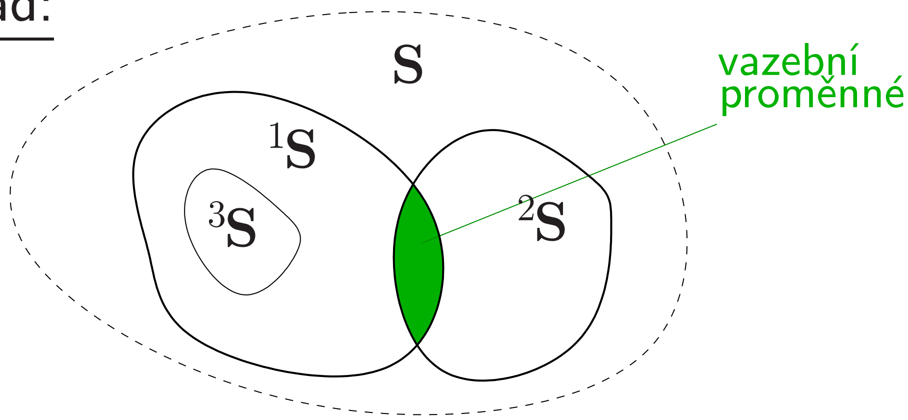
Celkový systém – obsahuje všechny proměnné.

Dekompozice: Množina podsystémů

$$G = \{{}^1\mathbf{S}, {}^2\mathbf{S}, \ldots, {}^q\mathbf{S}\}$$

celkového systému $\mathbf{S}$, taková, že žádné dva ${}^j\mathbf{S}$ a ${}^k\mathbf{S}$ nejsou navzájem podsystémy: ${}^j\mathbf{S} \not\sqsubseteq {}^k\mathbf{S}$

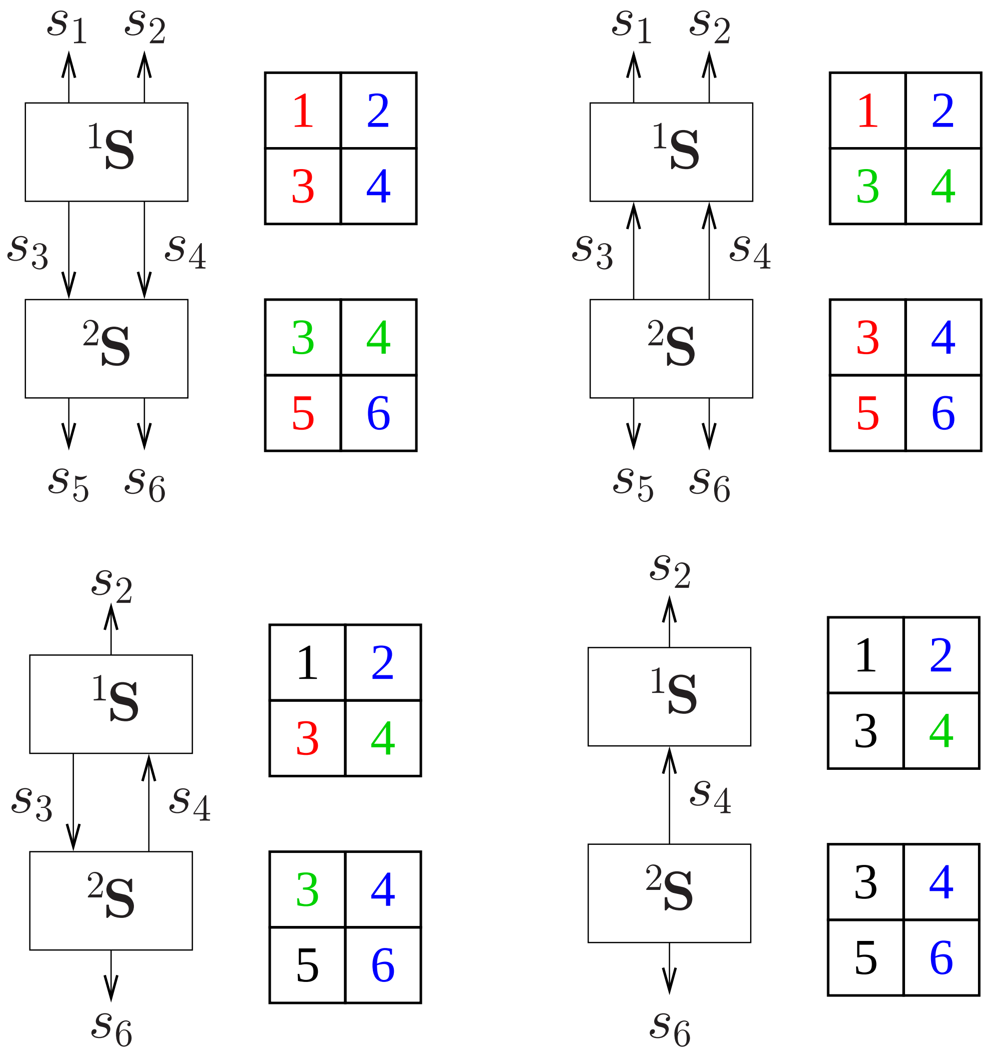
Protipříklad:



Podmínka iredundance: podsystém ${}^3\mathbf{S} \sqsubset {}^1\mathbf{S}$ nenese žádnou novou informaci o $\mathbf{S}$ a nepatří tedy do dekompozice systému $\mathbf{S}$.

- Vazební proměnné mezi podsystémy: $\mathbf{C}_{k,l} = {}^k\mathbf{S} \cap {}^l\mathbf{S}$
- Orientované vazby: rozklad proměnných na vstupní a výstupní.
- Proměnná může být deklarována jako výstupní jen v jednom podsystému (jednoznačnost řízení)

Rozklad proměnných na vazební vstupní, vazební výstupní generující, vazební výstupní generované a nevazební generující proměnné



- celkem 24 možností
- identifikace struktury systému není tímto rozkladem ovlivněna
- orientace vazby se pozná dle generativní neurčitosti příslušné proměnné vzhledem k 1. nebo 2. systému

Rekonstrukce a identifikace: úvod

Rekonstrukce systému

Konstrukce hypotézy o nejlepším celkovém systému $\mathbf{S}$ je-li dána jeho dekompozice $\{{}^1\mathbf{S}, {}^2\mathbf{S}, \ldots, {}^q\mathbf{S}\}$.

Aplikace:

1. inference celkového systému z dílčích
2. procedura nutná pro identifikaci

Identifikace struktury

Nejlepší dekompozice systému $\mathbf{S}$ na $\{{}^1\mathbf{S}, {}^2\mathbf{S}, \ldots, {}^q\mathbf{S}\}$.

Aplikace:

1. zjednodušení systému
   (např. rozpoznávání: jednodušší modely se odhadují lépe z dat)
2. nalezení struktury ve složitém systému
   (např. analýza kritických vazeb a závislostí)

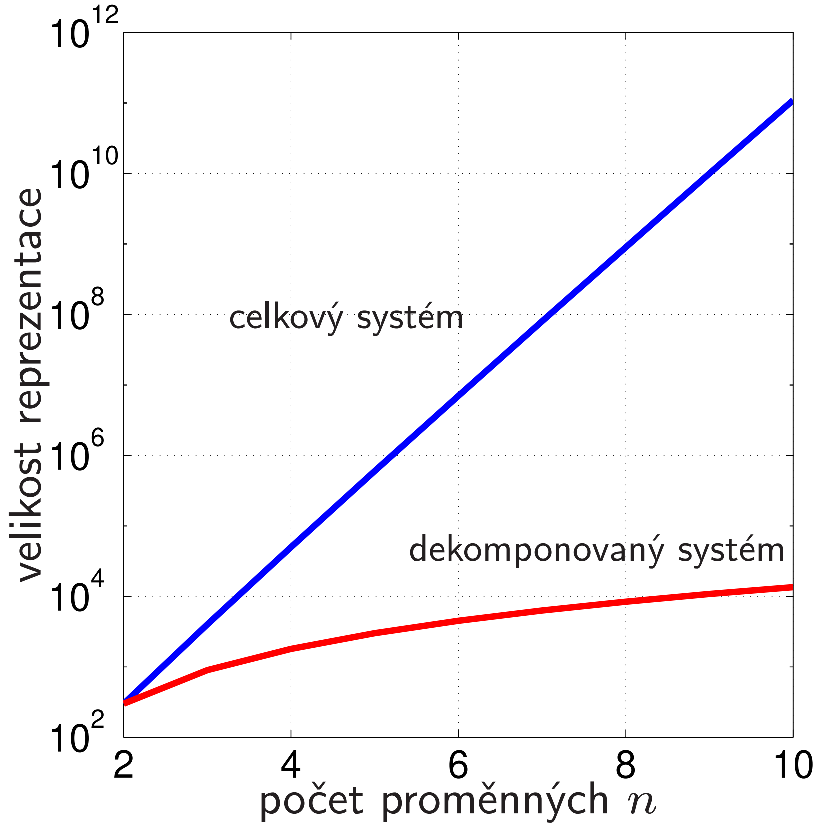
Velikost reprezentace celkového a dekomponovaného systému

k — počet stavů jedné proměnné

n — počet proměnných v systému

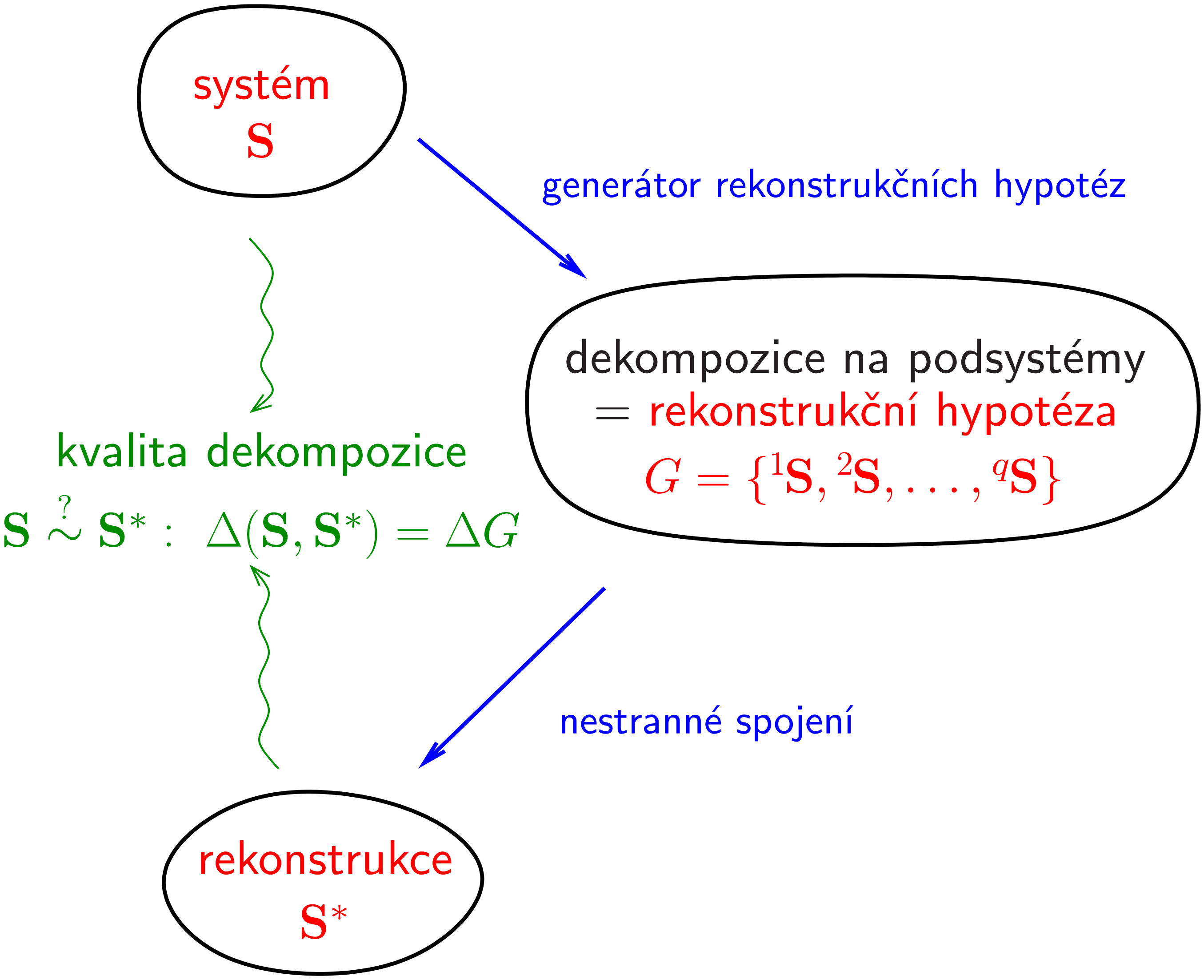
$(1+n)\,k^n$	velikost reprezentace celkového systému funkcí přípustnosti
$\frac{3}{2}\,k^2 n\,(n-1)$	velikost reprezentace dekompozice, kde každý podsystém má jen dvě proměnné

Pro $k = 10$



- dekomponovaný syst.: méně proměnných $\Rightarrow$ lepší odhad z dat

Schéma identifikační procedury



- G a $\mathbf{S}$ nejsou porovnatelné   nelze srovnat kvalitu G a $\mathbf{S}$
- $\mathbf{S}^*$ a $\mathbf{S}$ jsou porovnatelné

Vzájemná konzistence dynamických systémů

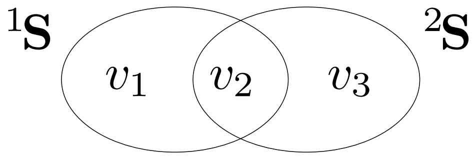
Lokální konzistence chování

Marginální funkce přípustnosti nad vazebními proměnnými musí být stejné

$$\mathbf{C}_{i,j} = {}^{i}\mathbf{S} \cap {}^{j}\mathbf{S}$$

$$[{}^{i}p_B \downarrow \mathbf{C}_{i,j}] = [{}^{j}p_B \downarrow \mathbf{C}_{i,j}]$$

Př: Lokálně *nekonzistentní* systémy:



$^{1}\mathbf{S}$

v_1	v_2	$^{1}p_B$
0	0	0.5
0	1	0.2
1	0	0.1
1	1	0.2

$^{2}\mathbf{S}$

v_2	v_3	$^{2}p_B$
0	0	0.4
0	1	0.25
1	0	0.15
1	1	0.2

v_2	$[{}^{1}p_B \downarrow \{v_2\}]$
0	0.6
1	0.4

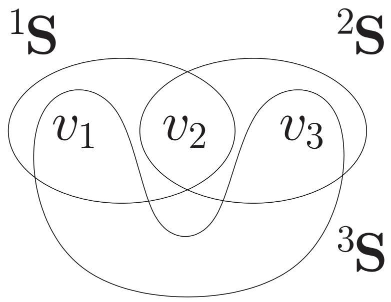
v_2	$[{}^{2}p_B \downarrow \{v_2\}]$
0	0.65
1	0.35

Pozn: podsystémy vzniklé rozkladem systému jsou lokálně konzistentní.

Globální konzistence chování

Sdruženou funkci přípustnosti p_B musí být možno zkonstruovat z marginálních ${}^i p_B,\ i \in N_q$.

Př: Globálně nekonzistentní systémy:



Lokálně konzistentní:

v_1	v_2	${}^1p_B$
0	0	$0\ (= a + x)$
0	1	0.7
1	0	0.3

v_2	v_3	${}^2p_B$
0	1	0.3
1	0	0.7
1	1	$0\ (= b + y)$

v_1	v_3	${}^3p_B$
0	0	0.4
0	1	$0.3\ (= a + b)$
1	0	0.3

v_1	v_2	v_3	p_B
0	0	1	a
0	1	1	b

$$a + x = 0 \Rightarrow a = x = 0 \qquad (a, x \geq 0)$$

$$b + y = 0 \Rightarrow b = y = 0 \qquad (y \geq 0)$$

$$a + b = 0.3$$

To je ve sporu.

Pozn: podsystémy vzniklé rozkladem systému jsou globálně konzistentní.

Rekonstrukce systému

Dáno: Dekompozice systému $G = \{{}^1\mathbf{S}, {}^2\mathbf{S}, \ldots, {}^q\mathbf{S}\}$.

Cíl: Nejlepší hypotéza o celkovém systému $\mathbf{S}$.

($s \in \mathcal{S}$ je stav a $|\mathcal{S}|$ je velikost stavového prostoru)

Postup:

1. **Určit množinu možných rekonstrukcí.**

$$
\begin{array}{lll}
{}^i p_B({}^i s) = \sum_{{}^i s \subset s} p_B(s) & \sum_{i=1}^{q} |{}^i\mathcal{S}| & \text{rovnic} \\
p_B(s) \geq 0 & |\mathcal{S}| & \text{nerovnic}
\end{array}
$$

2. **Vybrat nejlepší z nich.**

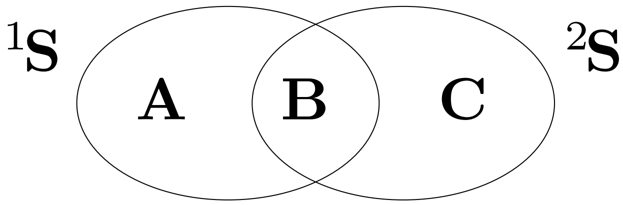
   Volba: (nestranná rekonstrukce)
   $\mathbf{S}$ neobsahuje jinou informaci než tu obsaženou v množině $\{{}^i\mathbf{S},\ i = 1, 2, \ldots, q\}$.

Implementace: Spojovací procedura.

Spojení dvou marginálních fcí přípustnosti

Marginální funkce přípustnosti:

rozklad množiny proměnných:



$$^1p_B\colon \mathbf{A} \times \mathbf{B} \longrightarrow \langle 0, 1\rangle$$
$$^2p_B\colon \mathbf{B} \times \mathbf{C} \longrightarrow \langle 0, 1\rangle$$

Nestranné spojení (o maximální entropii):

$$p_B^*(a,b,c) = (^1p_B * {}^2p_B)(a,b,c) = {}^1p_B(a,b) \cdot {}^2p_B(c \mid b)$$

Pozn: $^1p_B(b) = {}^2p_B(b)$ (kompatibilita)

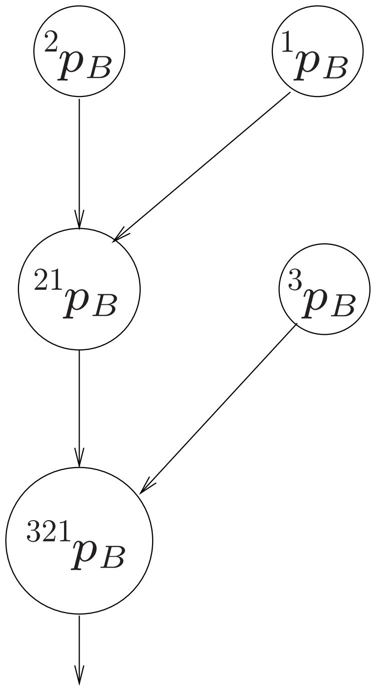
Speciální případy:

$$\mathbf{A} = \emptyset: \quad (^1p_B * {}^2p_B)(b,c) = {}^1p_B(b) \cdot {}^2p_B(c \mid b)$$
$$\mathbf{B} = \emptyset: \quad (^1p_B * {}^2p_B)(a,c) = {}^1p_B(a) \cdot {}^2p_B(c)$$

Spojovací procedura

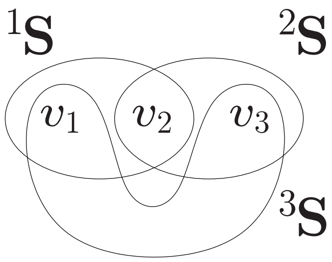
Jednoduchá procedura pro $\{{}^1\mathbf{S}, {}^2\mathbf{S}, \ldots, {}^q\mathbf{S}\}$:

$$p_B = {}^q p_B * \left({}^{q-1} p_B * \left(\cdots * \left({}^2 p_B * {}^1 p_B\right)\right)\right)$$



libovolné pořadí spojování

Problém: cyklické vazby mezi proměnnými



Je-li splněna podmínka $[p_B \downarrow {}^i\mathbf{S}] = {}^i p_B$, potom je spojení nestranné. Jinak nutno použít iterativní proceduru.

Iterativní spojovací procedura

Odstraní vliv cyklických vazeb.

Iterativní spojovací procedura $\mathrm{S}^*({}^1\mathbf{S}, {}^2\mathbf{S}, \dots, {}^q\mathbf{S}; \varepsilon)$

1. $k := 0,$ $\quad F_0 := {}^q p_B * {}^{q-1} p_B * \cdots * {}^1 p_B$
2. $k := k + 1,$ $\quad F_k := {}^q p_B * {}^{q-1} p_B * \cdots * \left({}^1 p_B * F_{k-1}\right)$
3. jestliže $\quad \max |F_k - F_{k-1}| < \varepsilon \quad$ stop, jinak opakuj 2.

Jestliže $\sum p_B(s) = 1$, pak řešení nalezeno s chybou $\pm\varepsilon$, jinak $\{{}^1\mathbf{S}, {}^2\mathbf{S}, \dots, {}^q\mathbf{S}\}$ není globálně konzistentní.

Výsledkem je nestranná rekonstrukce s chybou ε.

Nutná jen, pokud pro nějaké $i = 1, 2, \dots, q$ platí

$$[p_B \downarrow {}^i\mathbf{S}] \neq {}^i p_B$$

Rekonstrukční chyba

Dvě rozdělení $p^*(s)$, $p(s)$, $s \in \mathcal{S}$ nad stavovým prostorem $\mathcal{S}$.

Kullbackova-Leiblerova vzdálenost

(informační vzdálenost, poměrná entropie)

$$L(p, p^*) = \sum_{s \in \mathcal{S}} p(s) \log \frac{p(s)}{p^*(s)}$$

Vlastnosti:

- $L(p, p^*) \geq 0$
- $L(p, p^*) \neq L(p^*, p)$ (není metrika)
- $L(p, p^*) = 0 \quad \Leftrightarrow \quad p \equiv p^*$
- je-li p^* nestranné spojení, pak
  1. $p^*(s) = 0 \Rightarrow p(s) = 0$
  2. $L(p, p^*) = H(S^*) - H(S)$ $L(p^*, p) \neq H(S) - H(S^*)$

Pro generativní systém

$$L(p, p^*) = \sum_{\overline{g} \in \overline{\mathbf{G}}} p(\overline{g}) \sum_{g \in \mathbf{G}} p(g \mid \overline{g}) \log \frac{p(g \mid \overline{g})}{p^*(g \mid \overline{g})}$$

$$= \cdots = H(\mathbf{G}^* \mid \overline{\mathbf{G}}^*) - H(\mathbf{G} \mid \overline{\mathbf{G}})$$

Rekonstrukční chyba: $\boxed{\Delta G = \Delta(\mathbf{S}, \mathbf{S}^*) = L(p, p^*)}$

Test statistické významnosti struktury systému

$$H_0 : p(a_i, b_j, c_k) = {}^1p(a_i, b_j) \cdot {}^2p(c_k \mid b_j)$$

skutečná četnost

$$n(a_i, b_j, c_k)$$

teoretická četnost

$$\begin{aligned} n^*(a_i, b_j, c_k) &= n \cdot {}^1p(a_i, b_j) \cdot {}^2p(c_k \mid b_j) \\ &\approx \frac{{}^1n(a_i, b_j) \cdot {}^2n(c_k \mid b_j)}{n} \end{aligned}$$

$$\chi^2 = \sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{k=1}^{K} \frac{(n(a_i, b_j, c_k) - n^*(a_i, b_j, c_k))^2}{n^*(a_i, b_j, c_k)}$$

- $I \cdot J \cdot K - I \cdot J - I \cdot K - J \cdot K - I - J - K - 1$ stupňů volnosti
- Lze jen u identifikace struktury systému

(máme n i n^*)

Identifikace struktury

Identifikace struktury

Nejlepší dekompozice systému $\mathbf{S}$ na množinu podsystémů $G = \{{}^1\mathbf{S}, {}^2\mathbf{S}, \ldots, {}^q\mathbf{S}\}$

Podproblémy

1. systematické generování rekonstrukčních hypotéz G
   - neredundance

$$\forall i, j \in \{1, 2, \ldots, q\} \colon {}^i\mathbf{S} \not\subseteq {}^j\mathbf{S}$$

   - pokrytí

$$\bigcup_i {}^i\mathbf{S} = \mathbf{S}$$

(**S**. . . množina proměnných systému)

2. Rekonstrukce celkového systému $\mathbf{S}^*$ z G.
3. Vyhodnocení rekonstrukční chyby $\Delta(G) = L(\mathbf{S}, \mathbf{S}^*)$

Bez podmínky nerendundance a pokrytí by byl počet hypotéz

$$2^{2^n - 1} - 1.$$

($n = 2 \Rightarrow$ 7 hypotéz, $n = 4 \Rightarrow$ 32767 hypotéz)

Zjemnění rekonstrukční hypotézy

Rekonstrukční hypotéza $G = \{{}^1\mathbf{S}, {}^2\mathbf{S}, \ldots, {}^q\mathbf{S}\}$

Zjemnění rekonstrukční hypotézy $\prec$

G_i je zjemněním G_j když pro každé ${}^x\mathbf{S} \in G_i$ existuje ${}^y\mathbf{S} \in G_j$ takové, že ${}^x\mathbf{S} \sqsubseteq {}^y\mathbf{S}$. Značíme

$$G_i \prec G_j$$

Bezprostřední zjemnění $\overset{*}{\prec}$

G_i je bezprostředním zjemněním G_j jestliže neexistuje G_k takové, že

$$G_i \prec G_k \quad \text{a} \quad G_k \prec G_j$$

Relace bezprostředního zjemnění tvoří svaz na množině všech rekonstrukčních hypotéz.

Pozn: nejde o zjemňování rozkladu, rekonstrukční hypotézy netvoří rozklad (proměnné se opakují).

Generátor rekonstrukčních hypotéz

Vstup: Rekonstrukční hypotéza $G = \{{}^1\mathbf{S}, {}^2\mathbf{S}, \dots, {}^q\mathbf{S}\}$

Výstup: Všechna bezprostřední zjemnění G

Procedura (prohledávání svazu $(G, \overset{*}{\prec})$)

1. Pro $i = 1, 2, \dots, q$ dělej kroky 2,3.
2. Jestliže $|{}^i\mathbf{S}| \geq 2$ potom nahraď ${}^i\mathbf{S}$ množinou všech podmnožin ${}^i\mathbf{S}$ o velikosti $|{}^i\mathbf{S}| - 1$.
3. Po odstranění redundantních podmnožin v každém rozkladu dostaneme seznam bezprostředních zjemnění G.

Pozn: může se opakovat na stejné úrovni zjemnění.

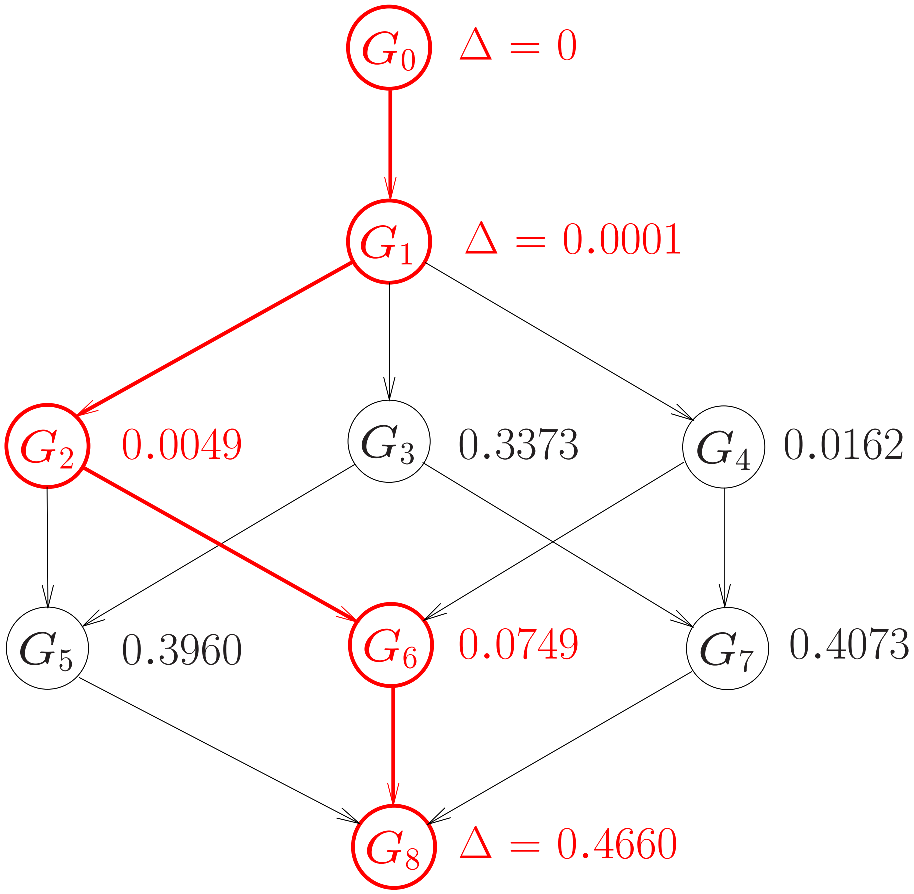
$$
\begin{aligned}
G &= \{\{1,2,3\},\{2,3,4\},\{1,4\}\} = \{{}^1\mathbf{S}, {}^2\mathbf{S}, {}^3\mathbf{S}\}\\
G_1 &= \{\{1,2\},\{1,3\},\{2,3,4\},\{1,4\}\}\\
G_2 &= \{\{1,2,3\},\{2,4\},\{3,4\},\{1,4\}\}\\
G_3 &= \{\{1,2,3\},\{2,3,4\}\}\\
G_4 &= \{\{1,2\},\{1,3\},\{2,3\},\{3,4\},\{2,4\},\{1,4\}\}\\
G_5 &= \{\{1,2\},\{1,3\},\{2,3,4\}\}\\
G_6 &= \{\{1,2,3\},\{2,4\},\{3,4\}\}
\end{aligned}
$$

Prohledávání svazu zjemnění

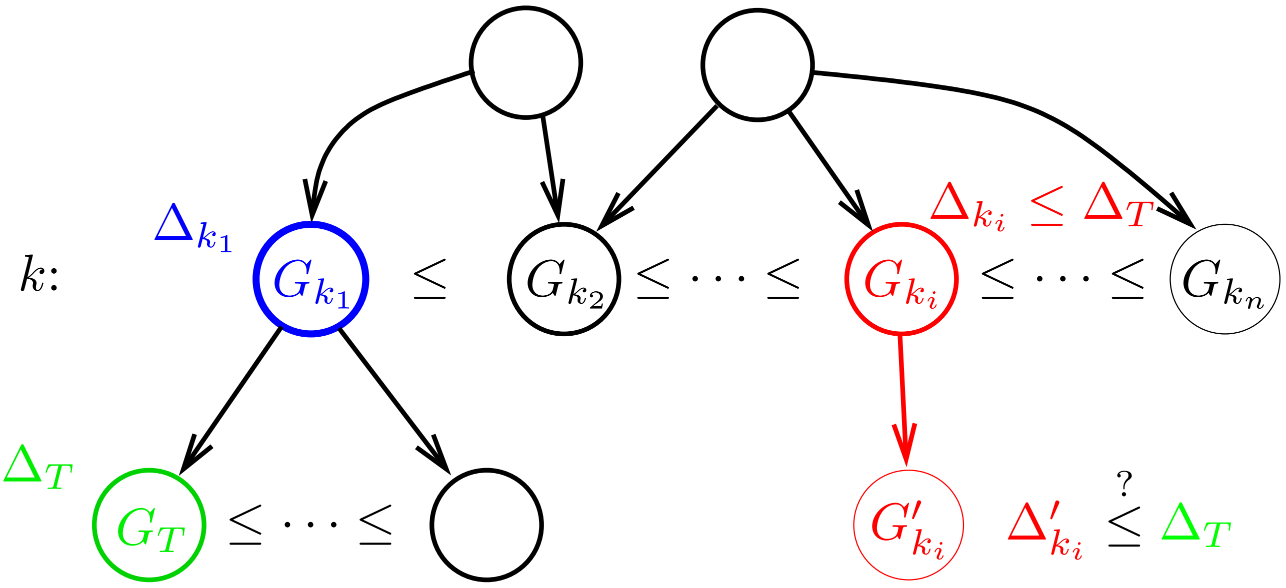
- Rekonstrukční chyba Δ monotonně roste podél každé větve svazu zjemnění: Je-li $G_i \succ G_k \succ G_j$ potom

$$\Delta(G_i) \leq \Delta(G_k) \leq \Delta(G_j)$$

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



Metoda větví a mezí na svazu zjemnění



1. vygeneruj úroveň k
2. nalezni uzel 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 $G_{k_2}, G_{k_3}, \ldots, G_{k_n}$ **jen pokud** $\Delta(G_{k_i}) \leq T$
6. (aktualizuj T, pokud nalezneš menší hodnotu prahu)
7. pokračuj stejně na úrovni $k+1$

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

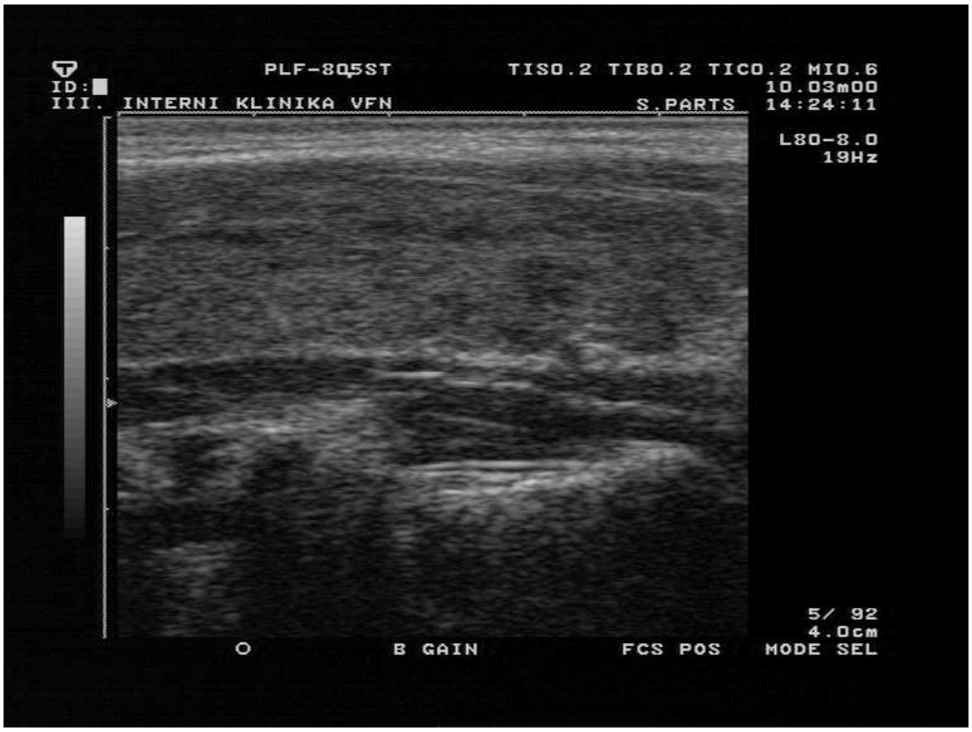
Identifikační procedura

1. Identifikuj zobecněný dynamický systém $\mathbf{S}$ z datového systému D.
2. Vytvoř rekonstrukční hypotézu $G_0 = \{\mathbf{S}\}$.
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

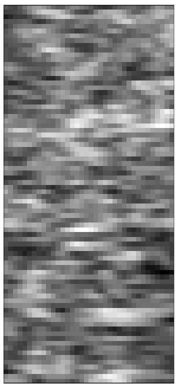
Zvolit co nejmenší model textury založený na funkci přípustnosti.

Důvody:

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

Rekonstrukční analýza

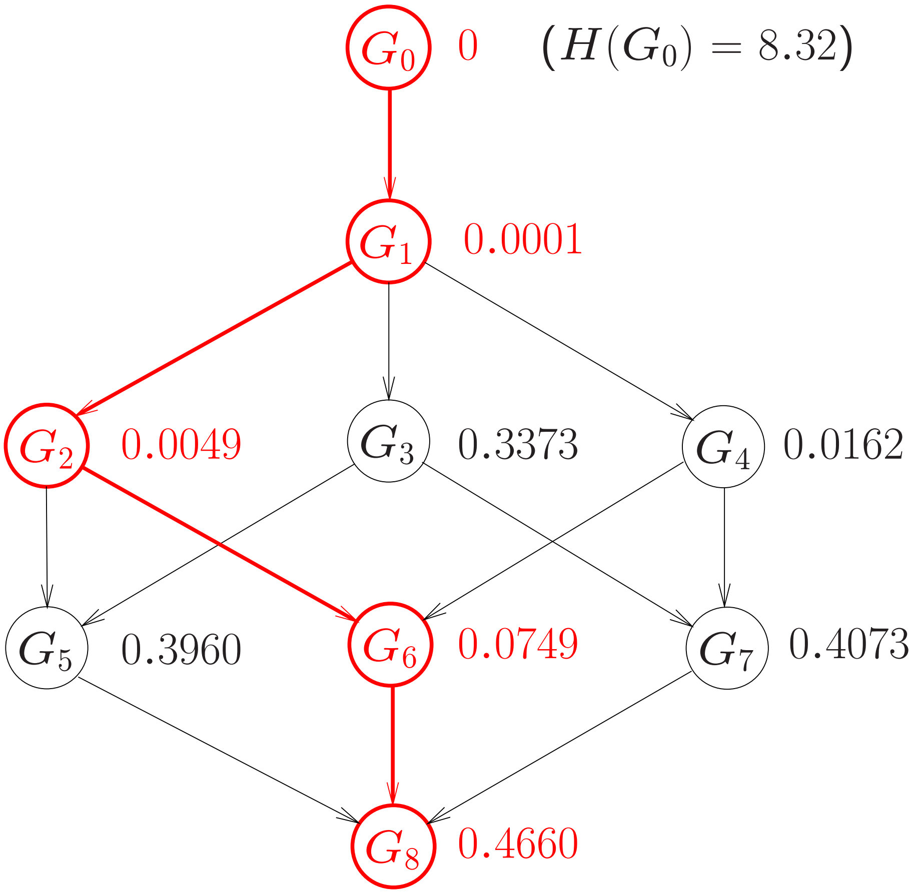
vzorek textury:



maska:

1	2
3	

svaz zjemnění



rekonstrukční hypotézy

$G_0 = 123$

$G_1 = 12|13|23$

$G_2 = 12|13$

$G_3 = 12|23$

$G_4 = 13|23$

$G_5 = 12|3$

$G_6 = 13|2$

$G_7 = 1|23$

$G_8 = 1|2|3$

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.

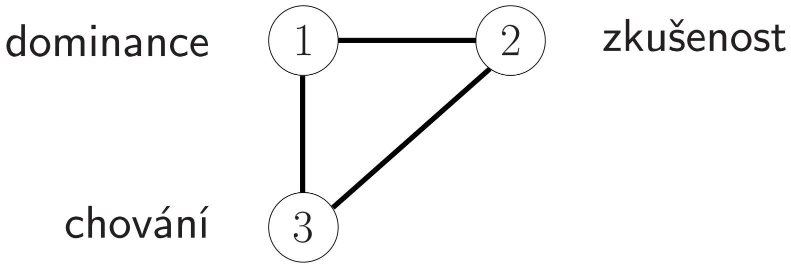
Studie: 114 jedinců 90-denních myší

Funkce přípustnosti

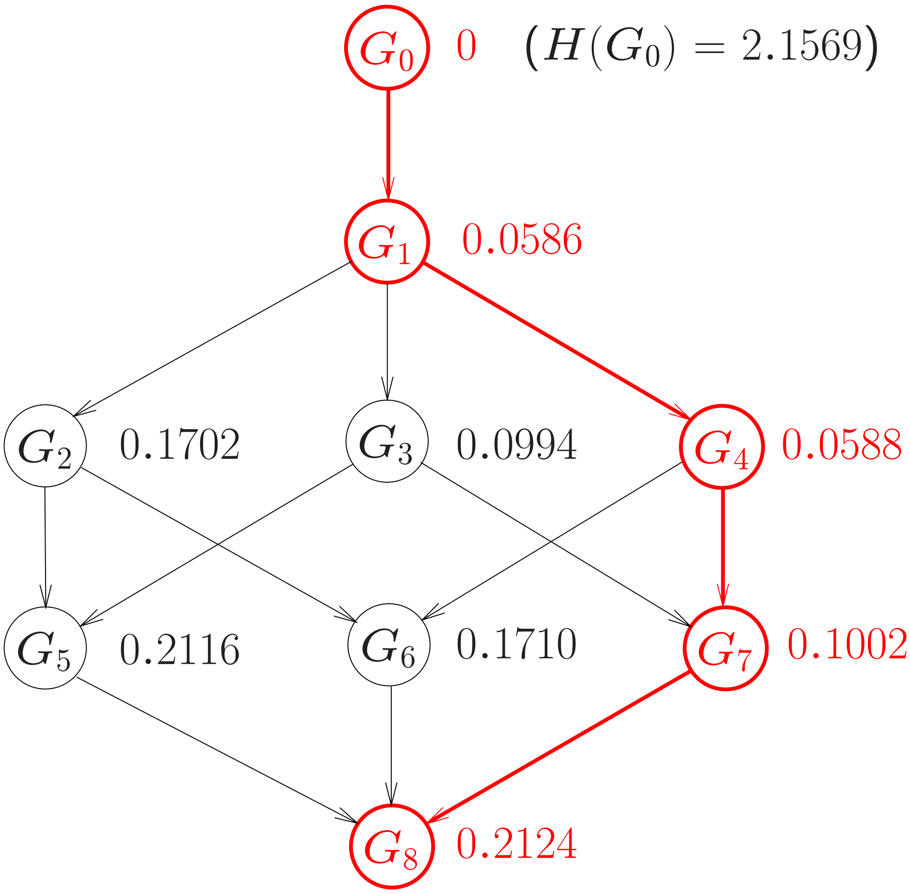
v_1	dominance	0	dominantní	
		1	podřízený	
v_2	sexuální zkušenost	0	nemá	
		1	má	
v_3	chování vůči potomstvu	0	infanticidní	
		1	rodičovské	
		2	nevšímavé	

v_1	v_2	v_3	N
0	0	0	28
0	0	1	4
0	0	2	2
0	1	0	5
0	1	1	25
0	1	2	3
1	0	0	5
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



svaz zjemnění



rekonstrukční hypotézy

$G_0 = 123$

$G_1 = 12|13|23$

$G_2 = 12|13$

$G_3 = 12|23$

$G_4 = 13|23$

$G_5 = 12|3$

$G_6 = 13|2$

$G_7 = 1|23$

$G_8 = 1|2|3$

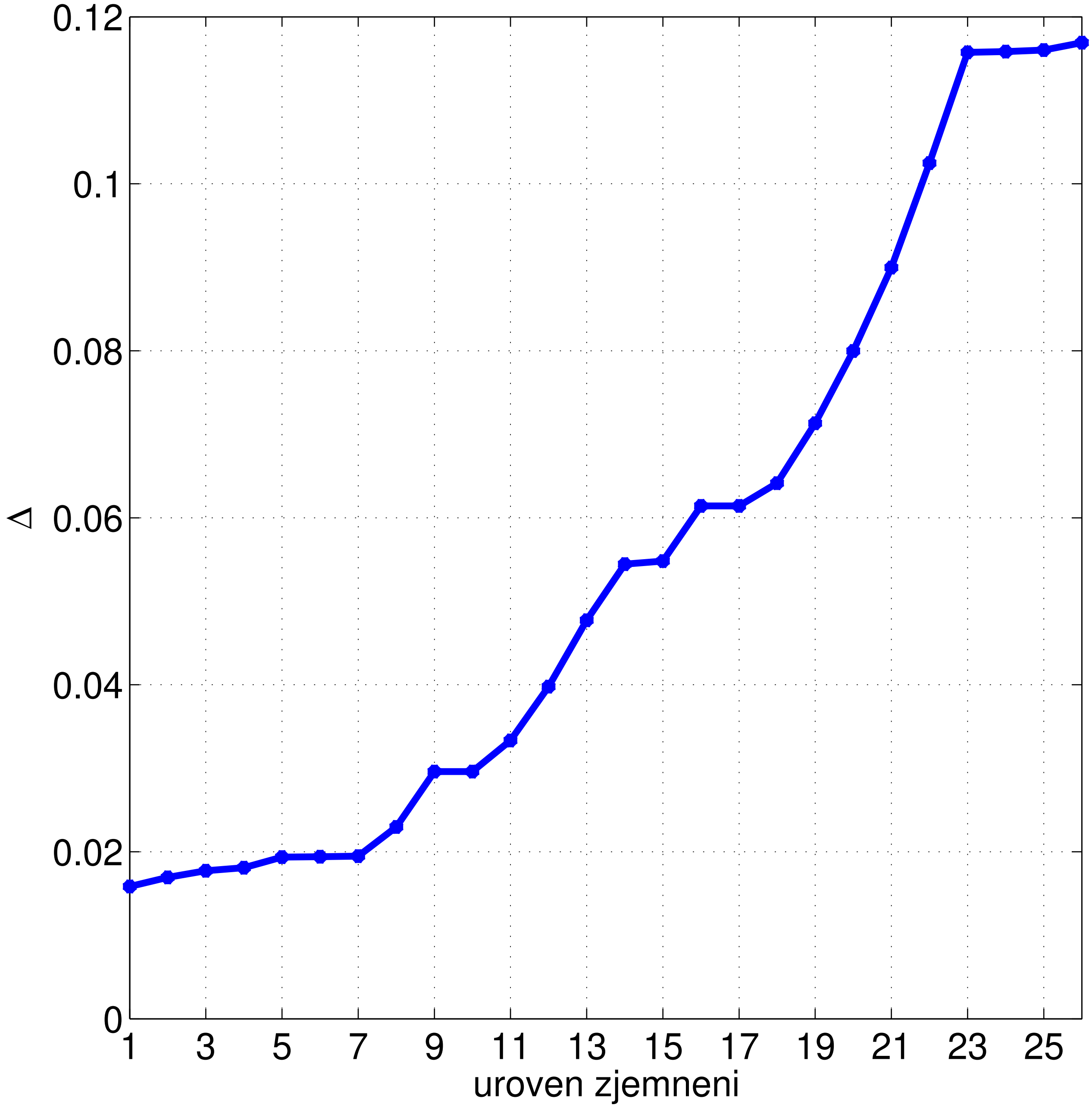
Závěry

- 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

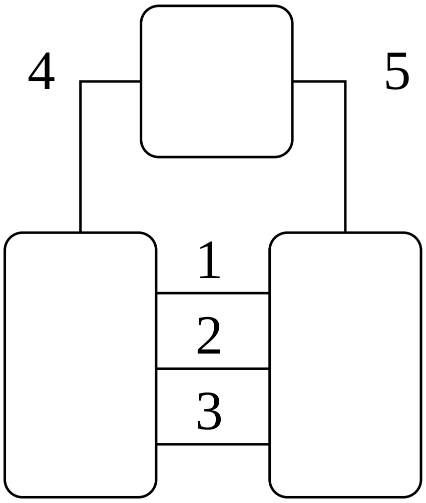
k	G_k	$\Delta(G_k)$
1	1234|2345|3451|4512|5123	0.0158
2	1234|2345|3451|5123	0.0169
3	1234|2345|5123|451	0.0177
4	1234|2345|5123	0.0181
5	1234|5123|345|452	0.0194
6	1234|5123|345	0.0194
7	1234|5123|45	0.0195
8	5123|45|234|341|412	0.0230
9	45|234|341|412|512|123|235|351	0.0296
10	45|234|341|412|123|235|351	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: rekonstrukční chyba na všech úrovních zjemnění

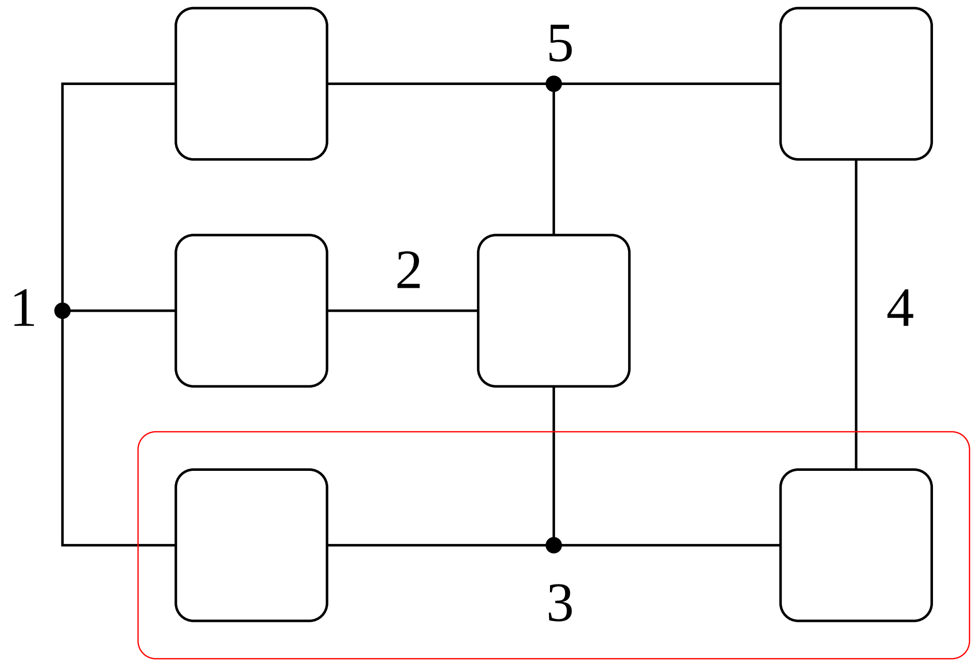


Parlament: interpretace úrovně 7 a 15+17

k	G_k	$\Delta(G_k)$
7	1234|5123|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
- $\pi = 1 \ldots$ událost je možná

x_i elementární jevy, $x_i \in S$

posibilita	pravděpodobnost
definiční vlastnosti	
$\pi\colon \exp S \to \langle 0, 1\rangle$	$p\colon \exp S \to \langle 0, 1\rangle$
$\pi(\{\}) = 0$	$p(\{\}) = 0$
$\pi(S) = 1$	$p(S) = 1$
$\pi\left(\bigcup_i x_i\right) = \max_i \pi(x_i)$	$p\left(\bigcup_i x_i\right) = \sum_i p(x_i)$
odhad z dat	
$\pi(x_i) \approx \dfrac{N(x_i)}{\max\limits_j N(x_j)}$	$p(x_i) \approx \dfrac{N(x_i)}{\sum\limits_j N(x_j)}$

Posibilistická míra neurčitosti

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

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

U-míra

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

---

příklad

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

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

$$U = \frac{1}{1.0} \left(\underbrace{0.1 \cdot \log 5}_{j=2} + \underbrace{0.1 \cdot \log 4}_{j=3} + \underbrace{0.1 \cdot \log 3}_{j=4} + \underbrace{0.7 \cdot \log 1}_{j=5} \right)$$

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

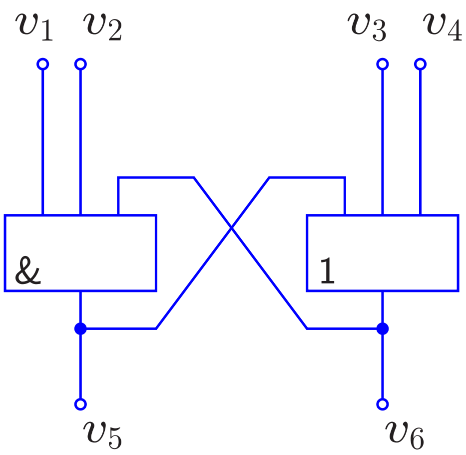
$$({}^1\pi * {}^2\pi)(a, b, c) = \min\left({}^1\pi(a, b),\ {}^2\pi(b, c)\right)$$

není nutná iterativní spojovací procedura

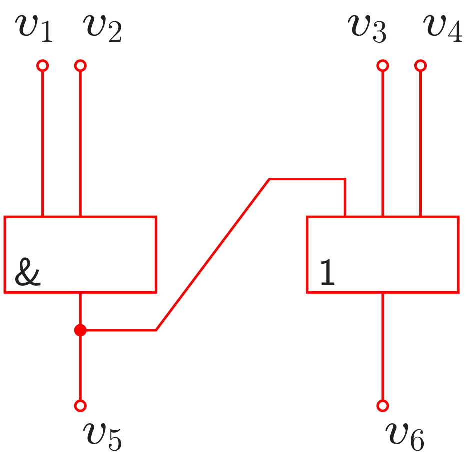
Rekonstrukční chyba

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

Diagnostika obvodů



předpokládaný systém



diagnostifikovaný systém

není pozorován stav a

pozorované stavy

	$\bar{g}$				g	
	v_1	v_2	v_3	v_4	v_5	v_6
	0	0	0	0	0	0
	0	0	0	1	0	1
	0	0	1	0	0	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

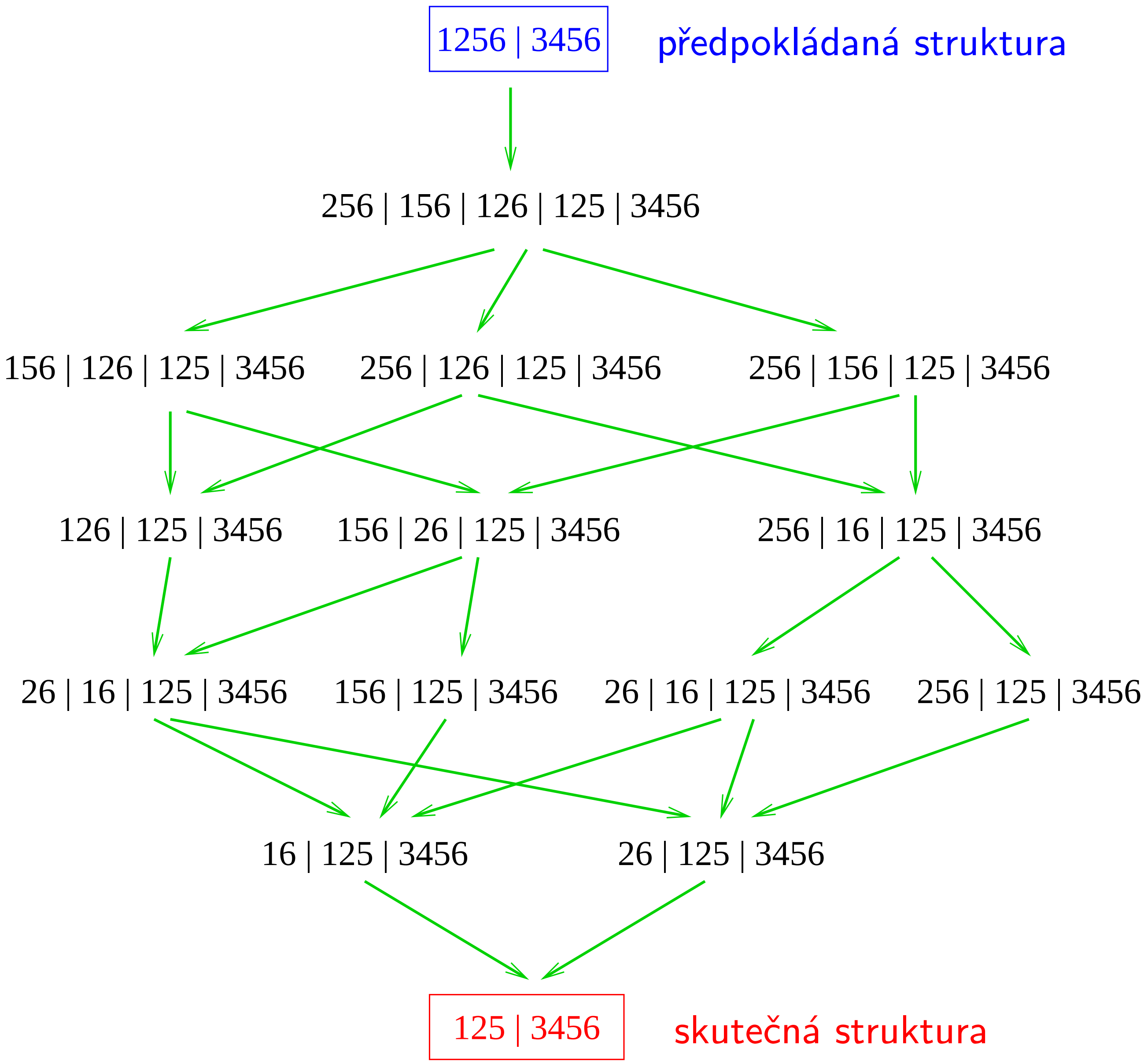
správná rekonstrukční hypotéza $1256 \mid 3456 = G_A$

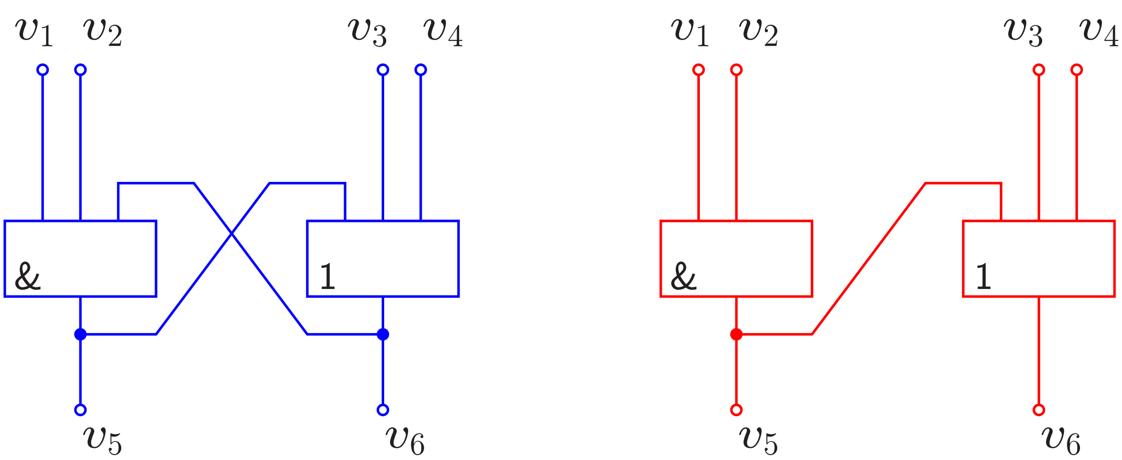
1. $\Delta G_A > 0 \Rightarrow$ obvod je vadný

   nutná rekonstrukční analýza pro předchůdce G_A

2. $\Delta G_A = 0$ nedokazuje plnou funkcionalitu

   rekonstrukční analýza pro zjemnění G_A





Pozn: $\Delta(1256 \mid 346) = 3$

Struktura svazu zjemnění

Problém: velikost svazu zjemnění

- r-ekvivalentní rekonstrukční hypotézy: stejný graf $\mathcal{R} = (V_n, E)$

  V_n – množina n proměnných systému,
  E – množina dvojic proměnných, které jsou obě v nějakém podsystému

- v každé třídě ekvivalence je právě jedna

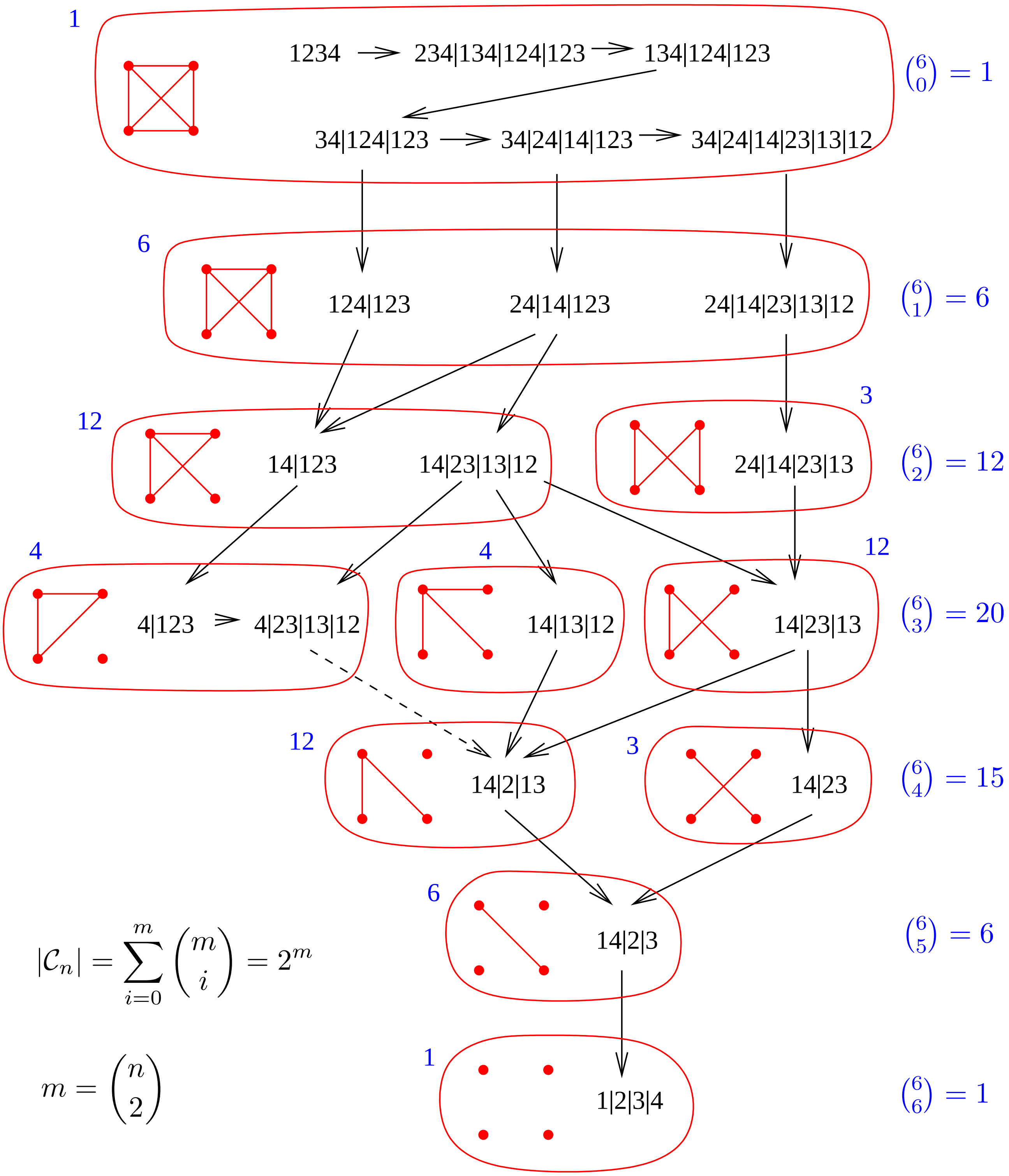
  $\mathcal{C}$-hypotéza: nejméně dekomponovaná ve třídě
  množina klik v grafu $\mathcal{R}$

  $\mathcal{P}$-hypotéza: nejvíce dekomponovaná ve třídě
  množina hran v grafu $\mathcal{R}$

- $\mathcal{P}$-hypotéza je zjemněním $\mathcal{C}$-hypotézy (existuje do ní cesta ve svazu zjemnění)

- Svaz zjemnění až na r-ekvivalenci a permutace indexů proměnných značíme $\mathcal{C}_n/i$ (n proměnných)

Příklad: Svaz $\mathcal{C}_4/i$



$$|\mathcal{C}_n| = \sum_{i=0}^{m} \binom{m}{i} = 2^m$$

$$m = \binom{n}{2}$$

Počet $\mathcal{G}$-hypotéz a $\mathcal{C}$-hypotéz

n	1	2	3	4	5	6	7
$\lvert\mathcal{G}_n\rvert$	1	2	9	114	6894	7785062	2414627396434
$\lvert\mathcal{C}_n\rvert$	1	2	8	64	1024	32768	2097152
$\lvert\mathcal{G}_n/i\rvert$	1	2	5	20	180	?	?
$\lvert\mathcal{C}_n/i\rvert$	1	2	4	11	34	156	1044

$\mathcal{G}_n$ — svaz zjemnění rekonstrukčních hypotéz

$\mathcal{C}_n$ — svaz zjemnění až na ekvivalenci r

$\mathcal{C}_n/i$ — svaz zjemnění až na permutace indexů

$\mathcal{C}_n/i$ — svaz zjemnění až na ekvivalenci r a permutace indexů

Hierarchizované prohledávání svazu zjemnění

▷ Generuj $\mathcal{C}$-hypotézy až na permutace indexů

  ▷ Generuj $\mathcal{G}$-hypotézy uvnitř třídy ekvivalence r

    ▷ Generuj permutace až na permutace na podmnožinách

- paralelizace prohledávání

Alternativní postup

▷ Generuj $\mathcal{C}$-hypotézy včetně permutací indexů ($\mathcal{C}_n$)

  ▷ Generuj $\mathcal{G}$-hypotézy uvnitř každé třídy ekvivalence

Prohledávání svazu zjemnění $\mathcal{C}$-hypotéz

Vstup: $\mathcal{C}$-hypotéza G_i, graf $r(G_i)$

Výstup: Všechna zjemnění $\mathcal{C}$-hypotézy

Opakuj 1–3 pro všechny hrany grafu $r(G_i)$:

1. Vyjmi hranu (a, b) z grafu.
2. Každý podsystém ${}^kS \in G_i$ obsahující a i b rozděl na dva podsystémy

$$ {}^kS \setminus \{a\} \quad \text{a} \quad {}^kS \setminus \{b\} $$

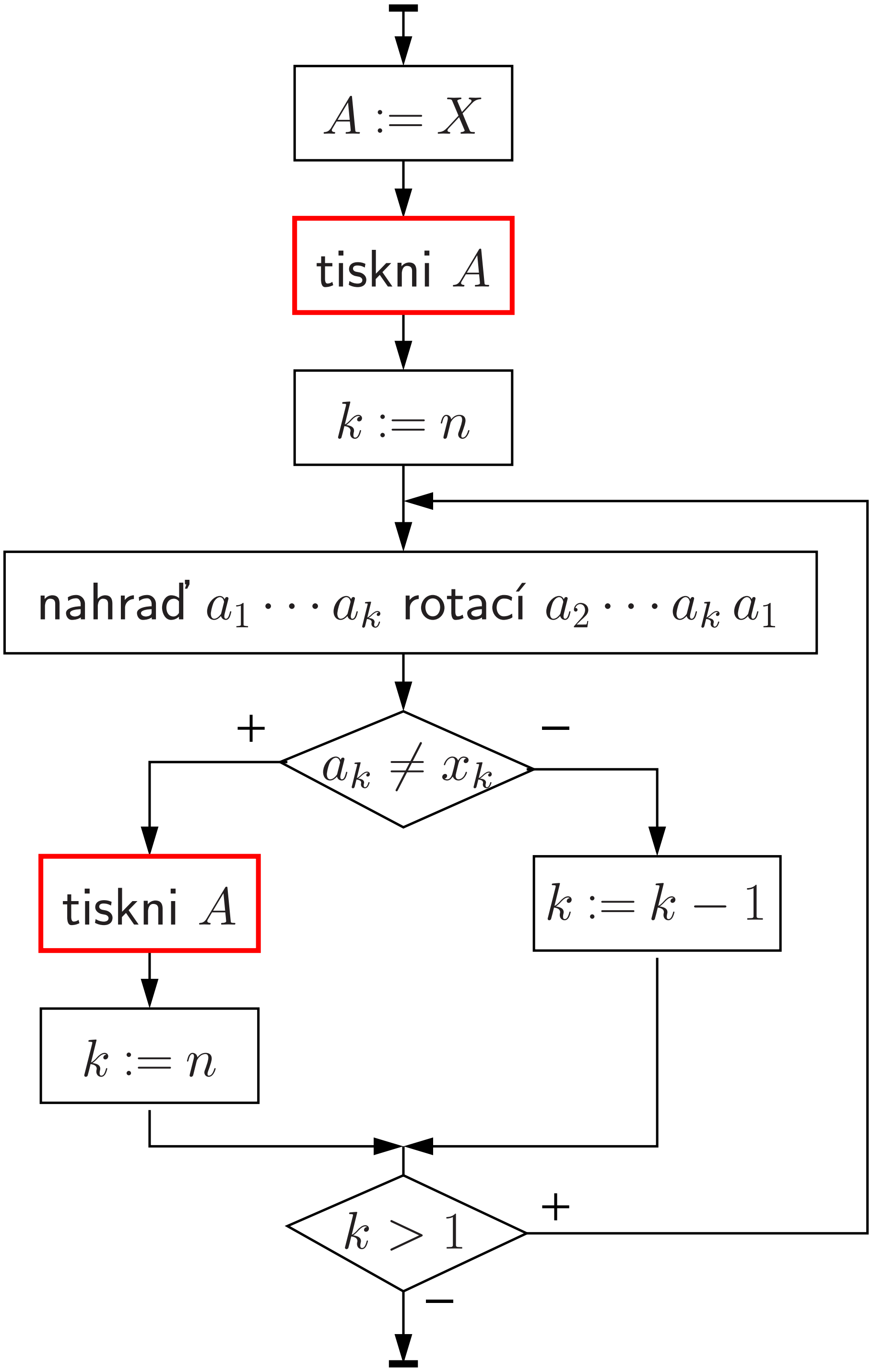
   a nahraď jimi ${}^kS$.
3. Odstraň z výsledku redundantní podsystémy

Pozn:

- Modifikace základní prohledávací procedury: podmínka $|{}^iS| \geq 2$ se nahradí podmínkou $|{}^iS| > 2$
  (zůstane v třídě $r(G)$)

Langdonův algoritmus (1967)

Vstup: množina $X = \{x_1, x_2, \cdots, x_n\}$ bez opakování
Výstup: všechny permutace $\pi_k(X)$



Shrnutí

PA Identifikace parametrů generativního systému z datového systému.

$\boxed{\bullet \text{ volbou masky}}$

kritéria

1. generativní neurčitost
   = jak dobře systém generuje generované proměnné

$$H(\mathbf{G} \mid \bar{\mathbf{G}}) = H(\mathbf{S}) - H(\bar{\mathbf{G}}), \quad H(\mathbf{S}) = H(\mathbf{G}, \bar{\mathbf{G}})$$

2. složitost = počet vzorkovacích proměnných

použití a vlastnosti

- nutná fáze modelování dynamických systémů
- existuje celá množina řešení

poznámky

- neparametrická metoda odhadu hustoty psti
- další možnosti: směs gausiánů, n nejbližších sousedů

Shrnutí

PB Zjednodušení modelu generativního systému.

- vylučováním proměnných
- redukcí rozlišení

kritéria

1. generativní neurčitost (jako v **PA**)
2. složitost
   $=$ počet stavů s nenulovou pravděpodobností

použití a vlastnosti

- redukce velikosti modelu před identifikací struktury
- existuje celá množina řešení

Shrnutí

PC Identifikace struktury systému.

- postupným zjemňováním dekompozice

kritéria

1. rekonstrukční neurčitost

$$\Delta(\mathbf{S}^*) = L(\mathbf{S}, \mathbf{S}^*)$$

$\mathbf{S}^*$ – rekonstrukce ze zvolených podsystémů
$\mathbf{S}$ – původní systém

2. míra dekompozice $=$ počet podsystémů

3. statistická významnost na hladině α

použití a vlastnosti

- redukce dimenze popisu
- analýza kritických vazeb a závislostí
- dolování znalostí z dat
- diagnostika systémů
- volba deskriptivních příznaků pro rozpoznávání
- existuje celá množina řešení