Vybrané problémy OTS

PA. Identifikace parametrů generativního systému z datového systému

PB. Zjednodušení modelu generativního systému

1. redukcí rozlišení
2. vyloučením proměnných

PC. Identifikace struktury generativního systému

PA. Zobecnění datového na generativní dynamický systém

Identifikace parametrů generativního dynamického systému z datového systému

Postup

1. Volba maximální délky historie

   (omezení množiny možných masek)

2. Odhad funkcí přípustnosti p_B a p_{GB} z dat.

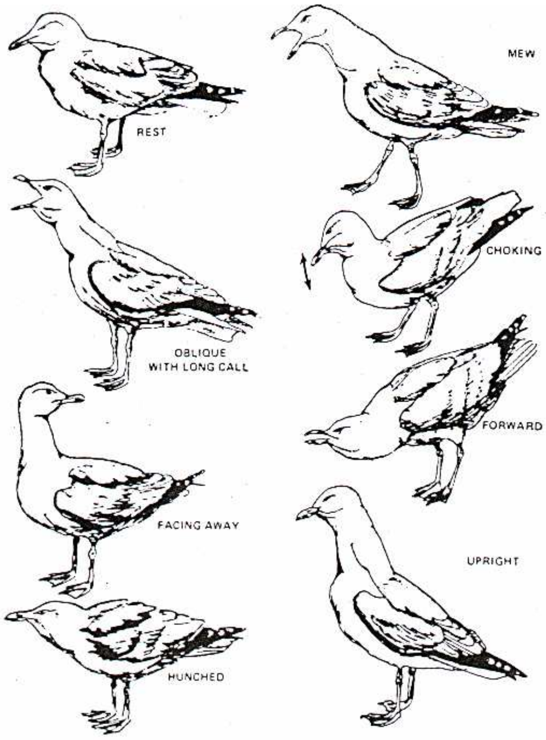
3. Vybere se generativní systém, který

   a. minimalizuje neurčitost při generování dat

   $$H(\mathbf{G} \mid \overline{\mathbf{G}}) = H(\mathbf{G}, \overline{\mathbf{G}}) - H(\overline{\mathbf{G}}),$$

   b. je nejjednodušší.

Výběrů může být více.



(a)

$s_1 = v_1(t-1)$

$s_2 = v_2(t-1)$

$s_3 = v_1(t)$

$s_4 = v_2(t)$

$\dot{V}_1 = \dot{V}_2$	$e_1 = e_2$	$V_1 = V_2$
ATTACK	↔	0
UPRIGHT	↔	1
GRASS PULLING	↔	2
CHOKING	↔	3
RETREAT	↔	4

(b)

$S = \{s_1, s_2, s_3, s_4\}$

s_1	s_3
s_2	s_4

○ - generující

$G = \{s_3, s_4\}$ $\bar{G} = \{s_1, s_2\}$

délka historie $\Delta H = 2$

t =	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
v_1	1	1	0	3	3	3	3	3	4	3	3	0	2	1	1	1	1	4	4	4
v_2	4	3	4	3	3	3	3	3	4	4	3	3	4	2	1	1	1	4	4	4

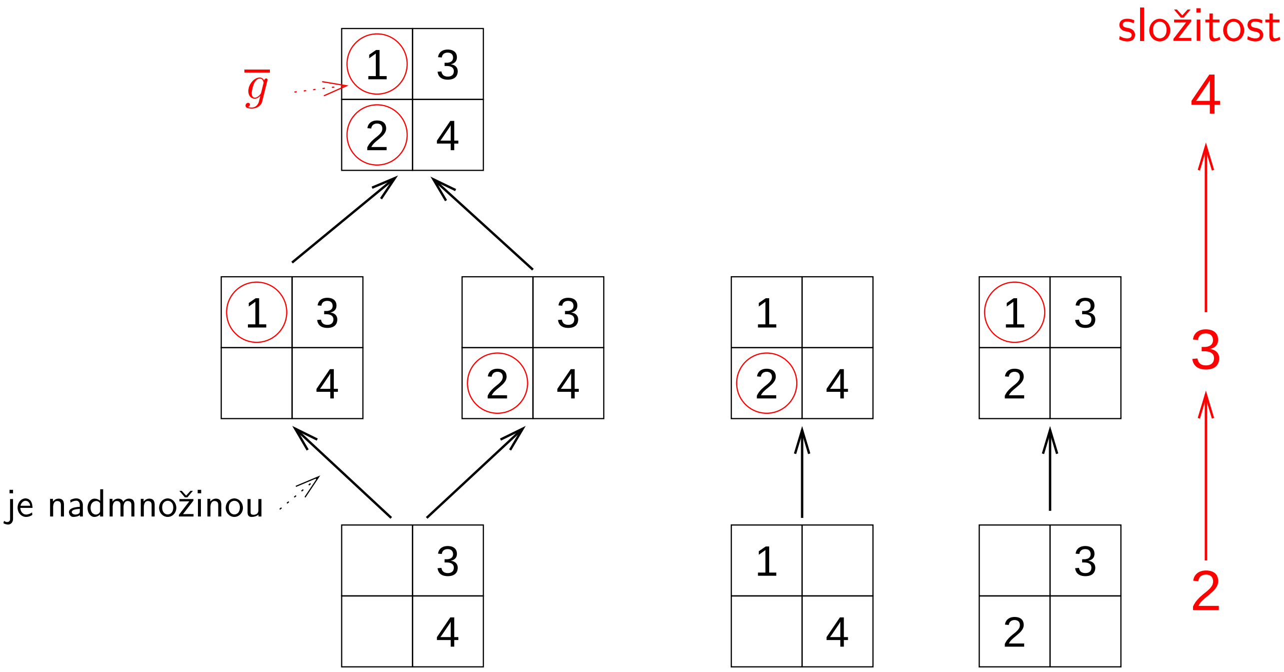
t =	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
v_1	4	4	0	2	2	4	4	4	4	4	4	4	2	2	2	2	2	0	2	1
v_2	4	3	1	4	2	2	4	4	4	4	4	4	4	4	1	1	1	3	3	1

t =	41	42	43	44	45
v_1	0	2	1	1	4
v_2	3	3	3	3	4

(c)

Uspořádání všech masek dané hloubky historie

Pro délku historie $\Delta M = 2$ a $n = 2$ základní proměnné:



$M_2 \leq M_1$ je-li $g_1 = g_2$ a zároveň $\overline{g}_2 \subset \overline{g}_1$

$$N = (2^{\Delta M} - 1)^n - (2^{\Delta M - 1} - 1)^n$$

s_1	s_2	s_3	s_4	f_B	s_1	s_2	s_3	s_4	f_B
0	1	2	4	0.0227	2	3	1	1	0.0227
0	3	2	3	0.0455	2	3	1	3	0.0227
0	3	2	4	0.0227	2	4	1	2	0.0227
0	4	3	3	0.0227	2	4	2	1	0.0227
1	1	0	3	0.0227	2	4	2	2	0.0227
1	1	1	1	0.0455	2	4	2	4	0.0227
1	1	4	4	0.0227	3	3	0	3	0.0227
1	2	1	1	0.0227	3	3	3	3	0.0909
1	3	0	4	0.0227	3	3	4	4	0.0227
1	3	1	3	0.0227	3	4	3	3	0.0227
1	3	4	4	0.0227	4	2	4	4	0.0227
1	4	1	3	0.0227	4	3	0	1	0.0227
2	1	0	3	0.0227	4	4	2	4	0.0227
2	1	2	1	0.0455	4	4	3	4	0.0227
2	2	4	2	0.0227	4	4	4	3	0.0227
					4	4	4	4	0.1818

$$H(s_1, s_2, s_3, s_4) - H(s_1, s_2) = 0.7535$$

$$H(s_1, s_3, s_4) - H(s_1) = 1.3149$$

$$H(s_2, s_3, s_4) - H(s_2) = 1.8024$$

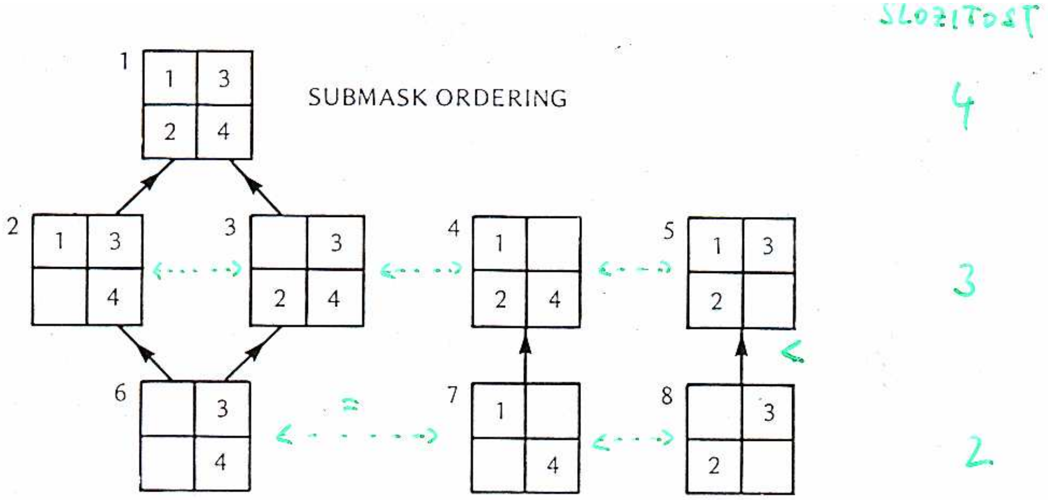
$$H(s_1, s_2, s_4) - H(s_2) = 1.6949$$

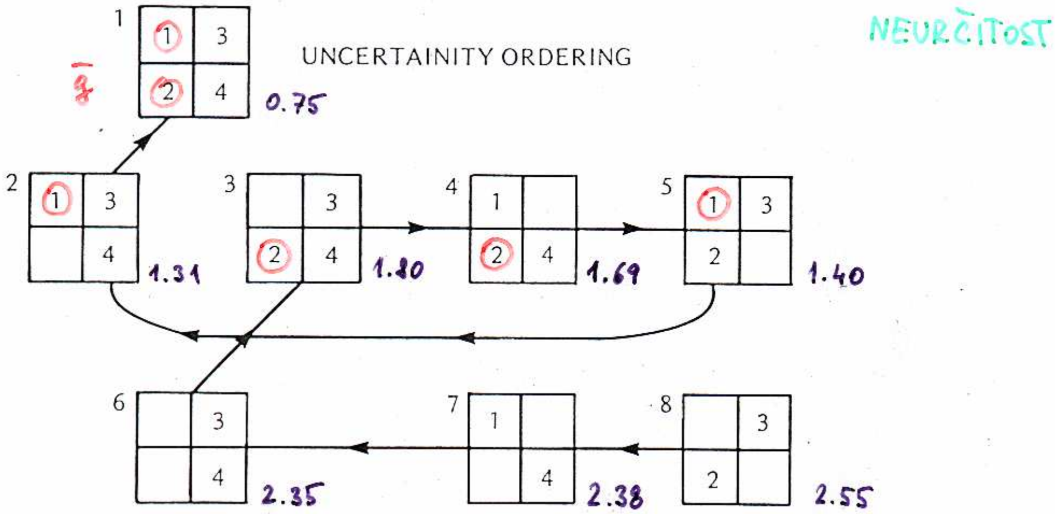
$$H(s_1, s_2, s_3) - H(s_1) = 1.4033$$

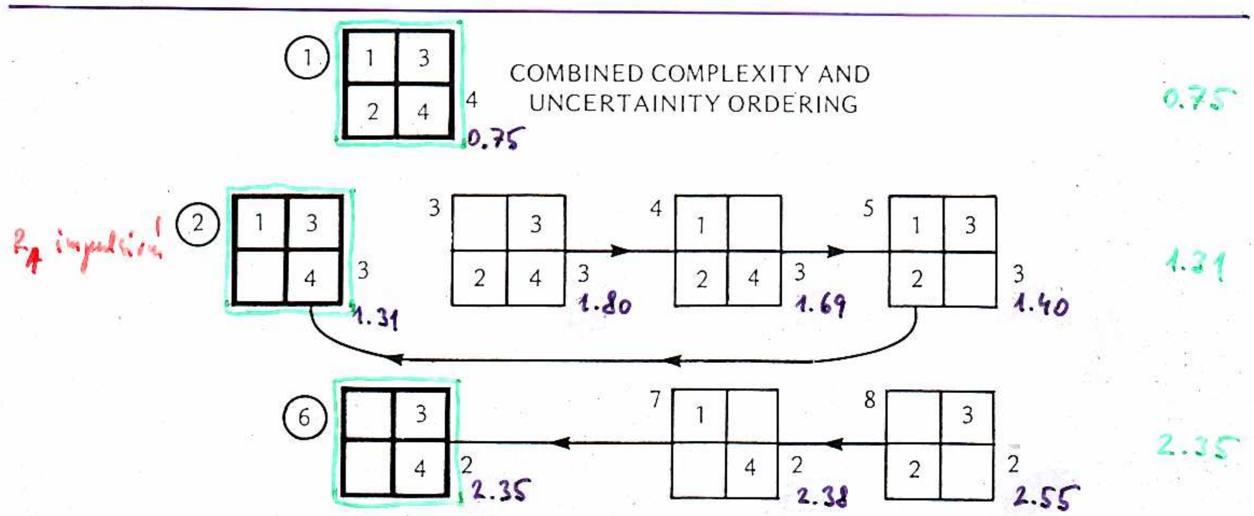
$$H(s_3, s_4) = 2.3539$$

$$H(s_1, s_4) = 2.3768$$

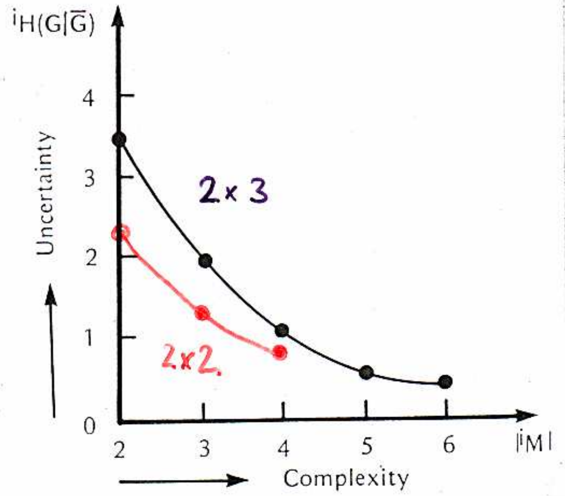
$$H(s_2, s_3) = 2.5506$$


SLOŽITOST
SUBMASK ORDERING
4
3
2
=


NEURČITOST
UNCERTAINITY ORDERING
0.75
1.31
1.80
1.69
1.40
2.35
2.38
2.55


COMBINED COMPLEXITY AND
UNCERTAINITY ORDERING
0.75
R4 impulsivní!
1.31
1.80
1.69
1.40
2.35
2.38
2.55

i	MASK	$\lvert{}^iM\rvert$	${}^iH(G\lvert\overline{G})$
1	1 3 5 / 2 4 6	6	0.41
2	_ 3 5 / 2 4 6	5	0.55
3	_ 3 5 / 2 4 _	4	1.07
4	_ 3 5 / _ _ 6	3	1.88
5	_ _ 5 / _ _ 6	2	3.38



R_1 dominuje (3) ale reaguje na R_2 (2)

R_2 reaguje

Zobecnění datového na dynamický systém: shrnutí

= **Identifikace parametrů**

Experimentální identifikace
vhodná pro zobecněné dynamické systémy bez zjevné struktury

Matematicko-fyzikální identifikace
vhodná pro lineární deterministické systémy jednoduché struktury

Pozn: Identifikace struktury

Rozklad systému na podsystémy.

Jiný problém, probereme později.

PB. Zjednodušení dynamického systému

Dáno:

Zobecněný dynamický systém $F_B = (A, B;\, M_G, p_{GB})$

Cíl:

Zmenšení složitosti modelu systému při zachování generativní schopnosti

$$H(\mathbf{G} \mid \overline{\mathbf{G}}) = H(\mathbf{G}, \overline{\mathbf{G}}) - H(\overline{\mathbf{G}}) \to \text{min}$$

Zjednodušení

1. **Eliminací základních proměnných** $v_i \in A$ z obrazu systému (A, B).
2. **Redukcí rozlišení** základních proměnných.

Obojí definuje rozklad stavového prostoru $\mathbf{S}$ na třídy ekvivalence $\mathcal{R}$.

Agregační funkce na podmnožinách $R_i \in \mathcal{R}$ rozkladu:

$$p'_B(R_i) = \sum_{x \in R_i} p_B(x).$$

Problém výběru vhodného zjednodušení

Zjednodušení

$$F_B \to F'_{Bi}$$

Kritéria výběru

1. Složitost systému F'_{Bi}

$$|p'_{Bi}| = \Big| \{ c \in \mathbf{S}_i \mid p_{Bi}(c) > 0 \} \Big|$$

(počet stavů nenulové psti)

2. Generativní neurčitost F'_{Bi}

$$H(\mathbf{G}_i \mid \overline{\mathbf{G}}_i) = H(\mathbf{G}_i, \overline{\mathbf{G}}_i) - H(\overline{\mathbf{G}}_i)$$

- Sdružené kritérium definuje částečné uspořádání na rozkladech.
- Existuje několik nejlepších výběrů.

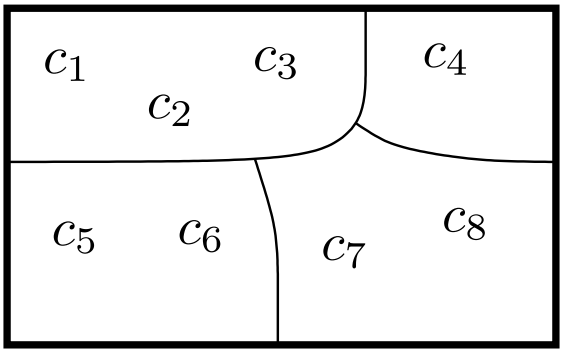
Zjednodušení dynamického systému vyloučením základních proměnných

$\mathbf{G} = \{s_1, s_3\}$, $\overline{\mathbf{G}} = \{\bar{s}_2, \bar{s}_4\}$
vyloučíme $v_2 \Rightarrow \{s_3, \bar{s}_4\}$

v_1		s_2	s_1
v_2	s_4		s_3

	s_1	$\bar{s}_2$	s_3	$\bar{s}_4$	$p(c)$
c_1	0	0	0	0	0.2
c_2	0	0	0	1	0.2
c_3	0	0	1	0	0.1
c_4	0	1	0	1	0.1
c_5	1	0	0	0	0.1
c_6	1	0	1	0	0.1
c_7	1	1	0	1	0.1
c_8	1	1	1	0	0.1

ekvivalentní stavy R_j



Po vyloučení v_2

	s_1	$\bar{s}_2$	$p' = \sum_{c_i \in R_j} p(c_i)$
R_1	0	0	0.5
R_2	0	1	0.1
R_3	1	0	0.2
R_4	1	1	0.2

vyl.	$\lvert p'_B \rvert$	$H(\mathbf{G} \mid \overline{\mathbf{G}})$
–	8	0.8047
v_1	3	0.4159
v_2	4	0.6097

ZJEDNODUŠENÍ GENERATIVNÍHO SYSTÉMU REDUKCÍ ROZLIŠENÍ

$$\left.\begin{array}{l} e\text{—elementary education;} \\ h\text{—completed high school;} \\ c\text{—college degree;} \\ g\text{—graduate degree.} \end{array}\right\} V_i$$

$$\big\downarrow \sigma$$

$$\left.\begin{array}{l} cg\text{—college or graduate degree;} \\ hc\text{—completed high school or college;} \\ eh\text{—no more than high school education;} \\ ehc\text{—any education except a graduate degree;} \\ hcg\text{—higher than elementary education.} \end{array}\right\} V_i'$$

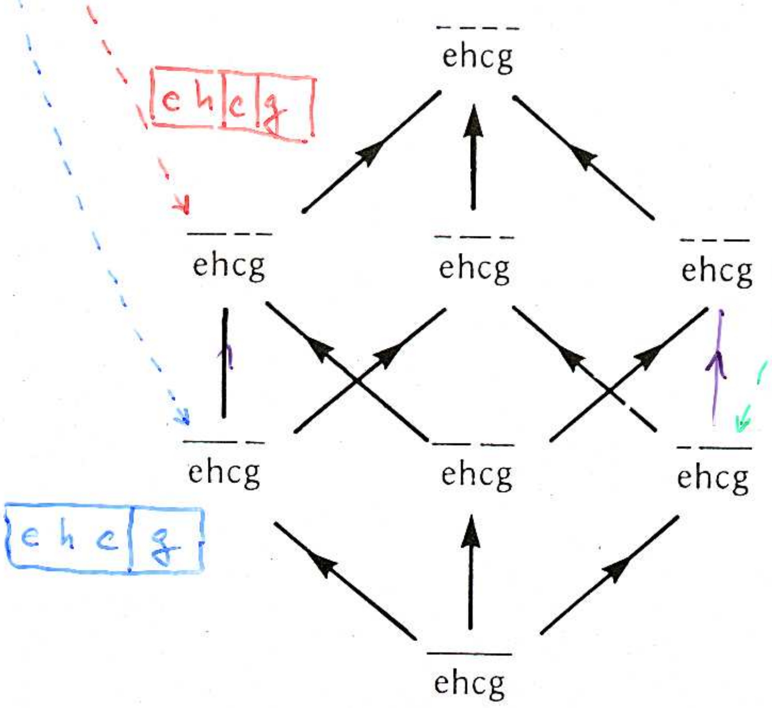


DIAGRAM ROZKLADŮ OBORU HODNOT PROMĚNNÉ v_i

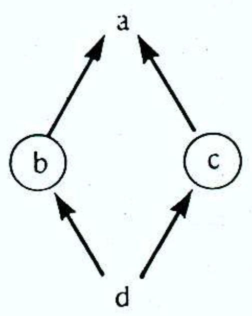
SVAZ

RELACE ZJEMNĚNÍ

$\mathcal{L} = (\{V_i/\sigma\}, <)$

JEDNA PROMĚNNÁ

DIAGRAM ROZKLADŮ PRO DVĚ PROMĚNNÉ : USPOŘÁDANÁ V_i



$a = \overline{0}\,\overline{1}\,\overline{2}$

$b = \overline{0\,1}\,\overline{2}$

$c = \overline{0}\,\overline{1\,2}$

$d = \overline{0\,1\,2}$

(a)

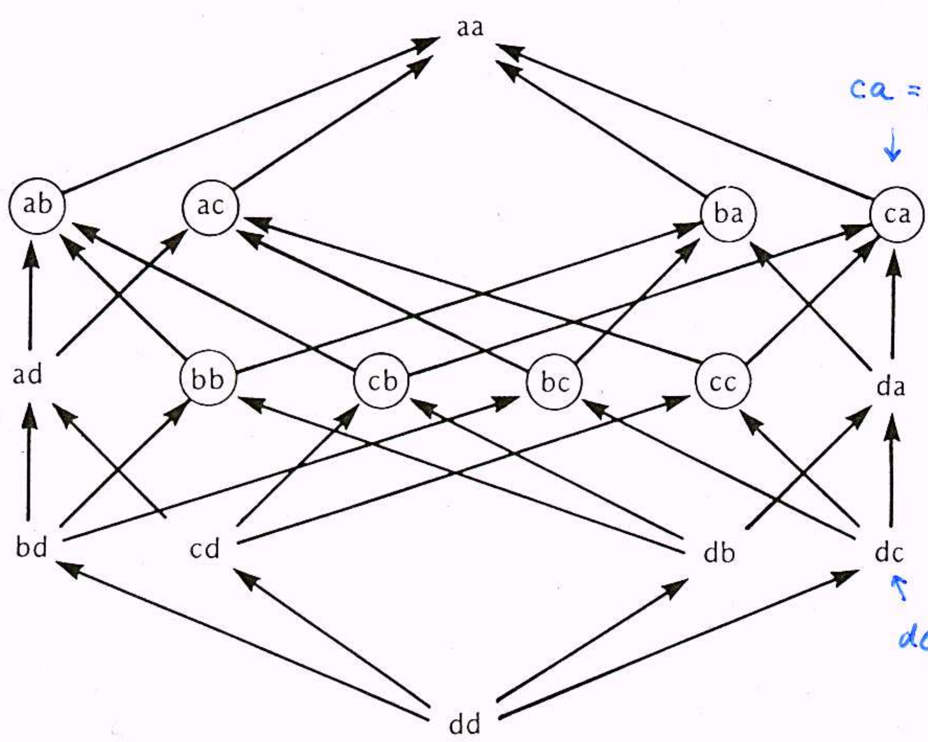
$V_i = \{0, 1, 2\}$

$X = \{a, b, c, d\}$

$\mathcal{L} = (X, \leq)$

$m = 3$
$n = 1$

$\Lambda_{m1} = 2^{m-1} = 4$



$ca = \overline{0}\,\overline{1\,2}, \overline{0}\,\overline{1}\,\overline{2}$

$\mathcal{L}^2 = (X^2, \leq)$

$dc = \overline{0\,1\,2}, \overline{0}\,\overline{1\,2}$

$m = 3$
$n = 2$

$\Lambda_{m2} = (2^{m-1})^n = 4^2 = 16$

$dc \equiv$ eliminace v_1 (nabývá jediné hodnoty)

DIAGRAM POSTIHUJE ZJEDNODUŠENÍ VYLUČOVÁNÍM PROMĚNNÝCH I REDUKCÍ ROZLIŠENÍ STAVOVÉHO PROSTORU

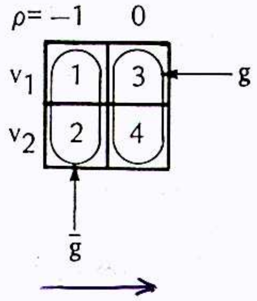
Postup při zjednodušení generativního systému

1. Vygeneruj všechna zjemnění, vypočti generativní neurčitost a spočti počet stavů nenulové pravděpodobnosti.
2. Zkonstruuj graf, jehož hrany směřují od uzlů s nižším počtem stavů k uzlům s vyšším nebo stejným počtem stavů a zároveň od uzlů s vyšší generativní neurčitostí k uzlům s nižší nebo stejnou generativní neurčitostí.
3. Uzly, které nemají následníka, reprezentují množinu řešení.

VOLBA ZJEDNODUŠENÍ GENERATIVNÍHO SYSTÉMU

	s_1	s_2	s_3	s_4	$f_B(\mathbf{c})$
$\mathbf{c} =$	0	0	0	1	0.2
	0	1	1	1	0.1
	0	1	1	2	0.1
	1	1	2	2	0.1
	1	2	1	2	0.2
	1	2	2	2	0.1
	2	2	0	0	0.2

. . .



$ac = \underbrace{\bar{0}\ \bar{1}\ \bar{2}}_{v_1}, \underbrace{\bar{0}\ \bar{1}\ \bar{2}}_{v_2}$

s_1	s_2	s_3	s_4	${}^{ac}f_B(\boldsymbol{\alpha})$
0	0	0	1	0.2
0	1	1	1	0.2
1	1	2	1	0.2
1	1	1	1	0.2
2	1	0	0	0.2

. . .

$cc = \underbrace{\bar{0}\ \overline{1\ 2}}_{v_1}, \underbrace{\bar{0}\ \overline{1\ 2}}_{v_2}$

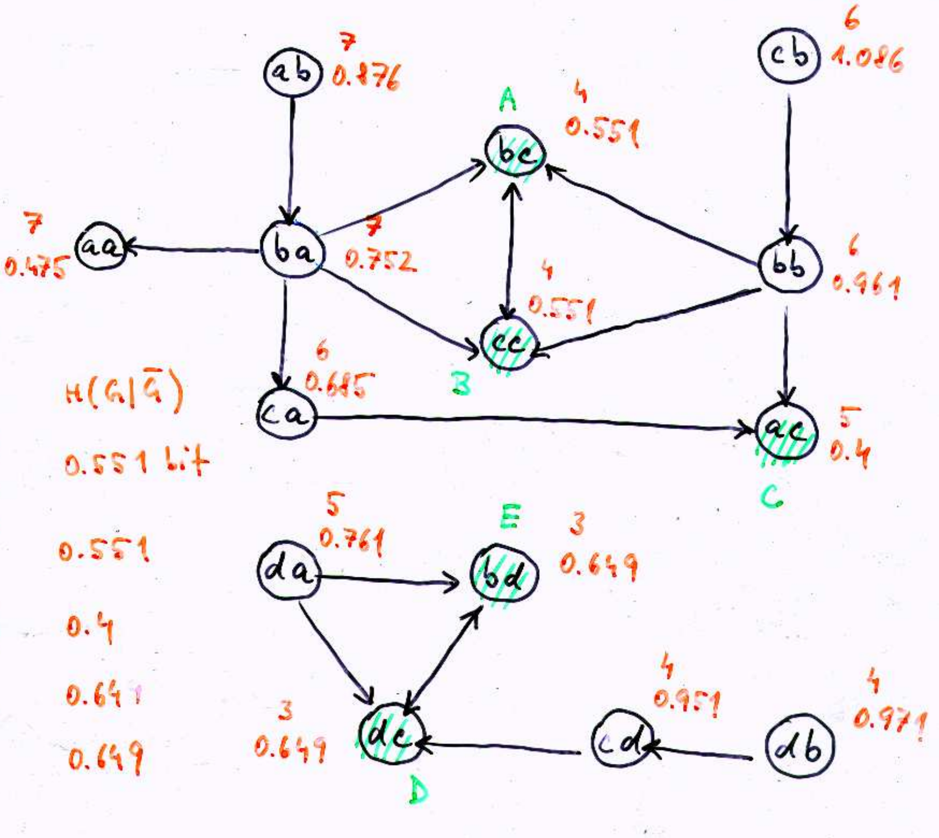
s_1	s_2	s_3	s_4	${}^{cc}f_B(\boldsymbol{\beta})$
0	0	0	1	0.2
0	1	1	1	0.2
1	1	1	1	0.4
1	1	0	0	0.2

$a = \bar{0}\ \bar{1}\ \bar{2}$

$b = \overline{0\ 1}\ \bar{2}$

$c = \bar{0}\ \overline{1\ 2}$

$d = \overline{0\ 1\ 2}$

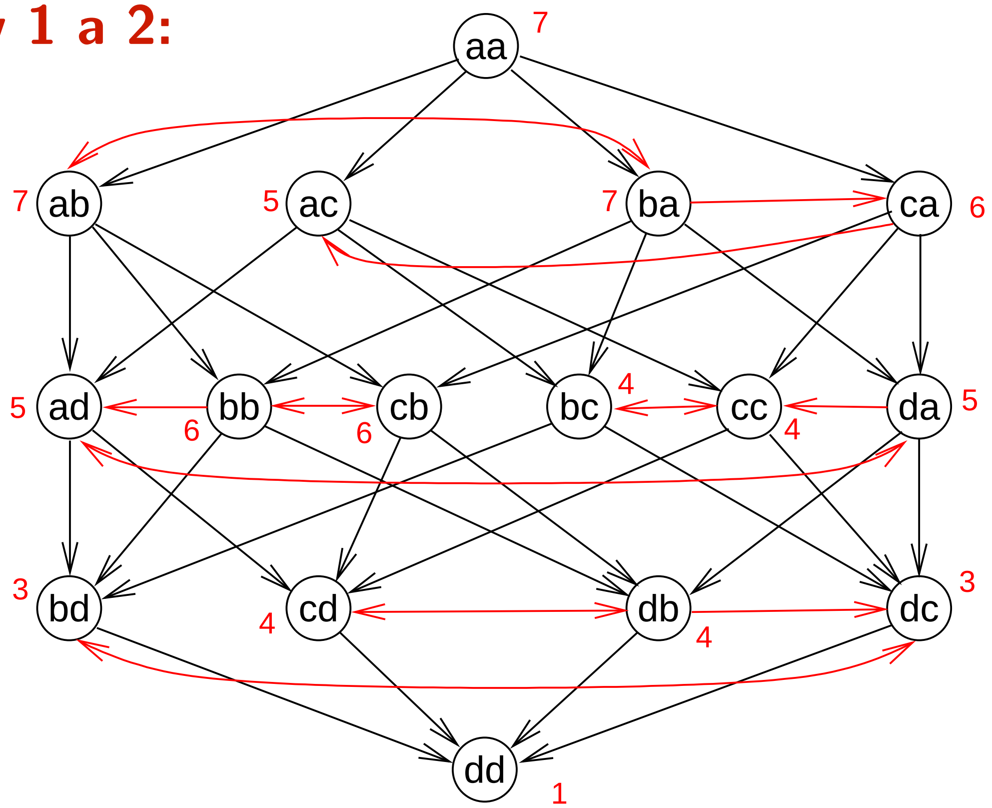


	$H(G\,|\,\bar{G})$
A: $\bar{0}\ \bar{1}\ \bar{2}, \bar{0}\ \overline{1\ 2}$	0.551 bit
B: $\bar{0}\ \overline{1\ 2}, \bar{0}\ \overline{1\ 2}$	0.551
C: $\bar{0}\ \bar{1}\ \bar{2}, \bar{0}\ \bar{1}\ \bar{2}$	0.4
D: $\bar{0}\ \bar{1}\ \bar{2}$	0.649
E: $\overline{0\ 1}\ \bar{2}$	0.649

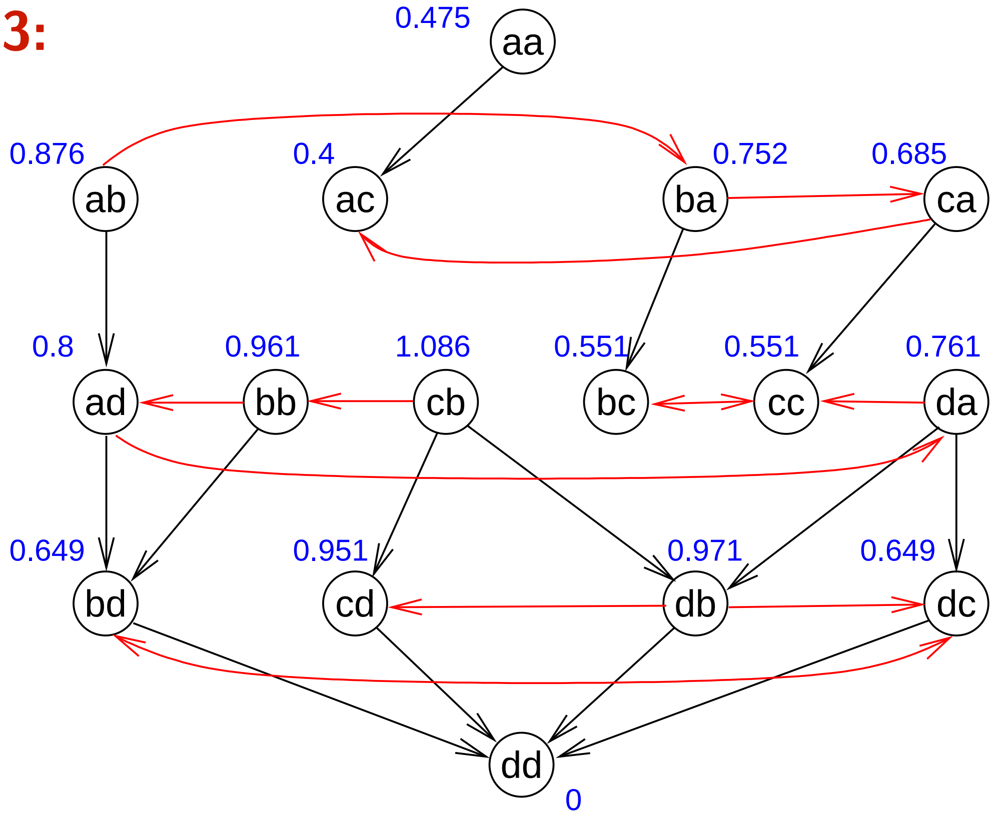
Postup na diagramu zjemnění rozlišení

1. Obrať orientaci všech hran.
2. Na jednotlivých úrovních diagramu doplň hrany, které směřují od vyššího k nižšímu nebo stejnému počtu stavů.
3. Odstraň všechny hrany, které směřují od nižší k vyšší generativní neurčitosti.
4. Odstraň tranzitivní hrany (není nutné).
5. Uzly, které nemají následníka, reprezentují množinu řešení.

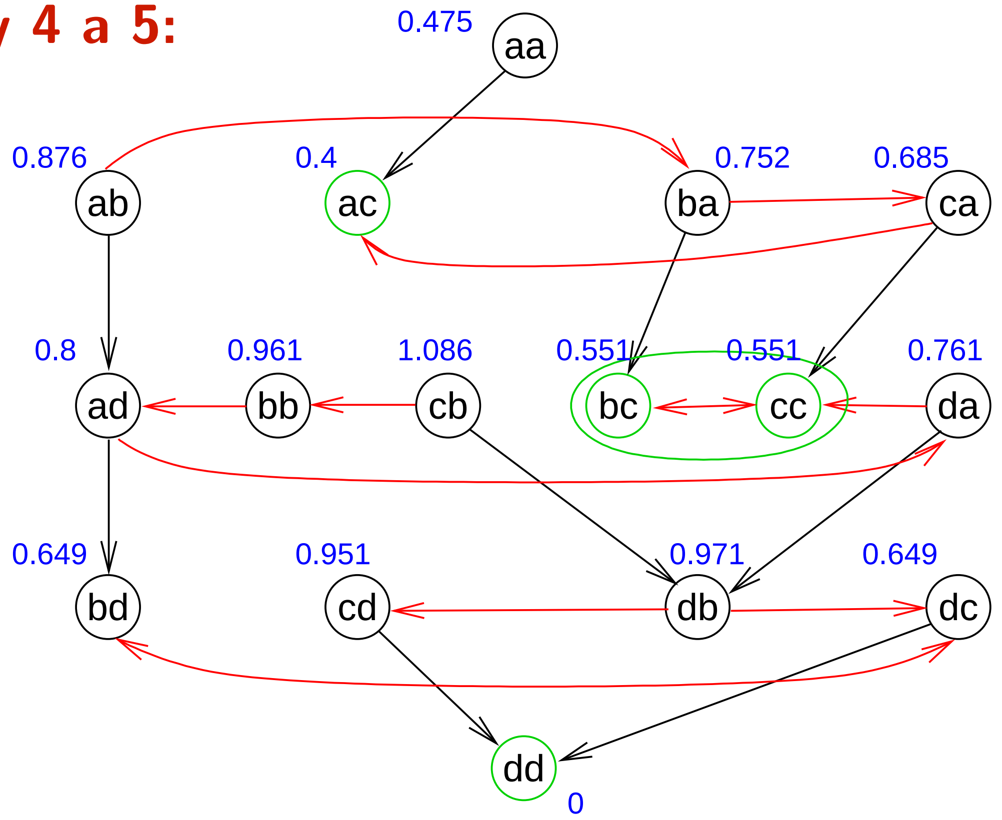
Kroky 1 a 2:



Krok 3:



Kroky 4 a 5:



Počet rozkladů oboru hodnot v_i s rozlišením na m úrovní

1. Obor hodnot V_i **není** úplně uspořádaný

$$\Lambda_m = \sum_{i=0}^{m-1} \binom{m-1}{i} \Lambda_i, \quad \Lambda_0 = 1$$

2. Obor hodnot V_i **je** úplně uspořádaný

$$\Lambda'_m = 2^{m-1}$$

m	2	3	4	5	6	7	8	9
Λ_m	2	5	15	52	203	877	4140	21147
Λ'_m	2	4	8	16	32	64	128	256

n proměnných v_i, $i = 1, 2, \ldots, n$ se stejným rozkladem

$$\Lambda_{mn} = (\Lambda_m)^n$$