

*PB051: Výpočetní metody v bioinformatice a
systémové biologii*

David Šafránek

11.5.2012

Tento projekt je spolufinancován Evropským sociálním fondem a státním rozpočtem České republiky.

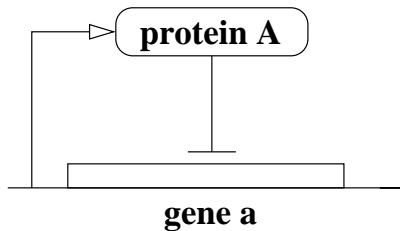


Obsah

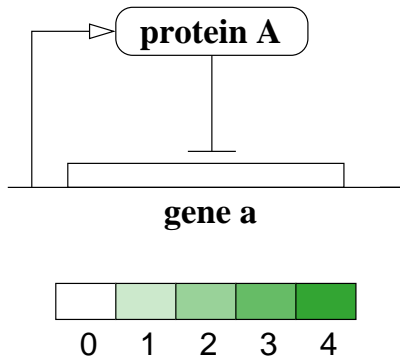
Boolovské síťe

Rekonstrukce boolovských sítí

Příklad modelu – autoregulace

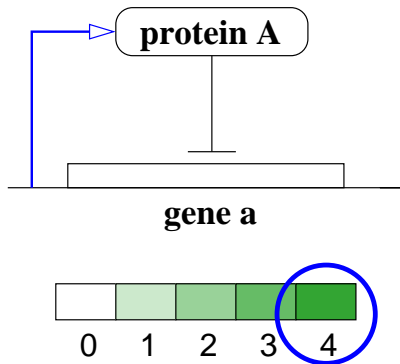


Příklad modelu – autoregulace



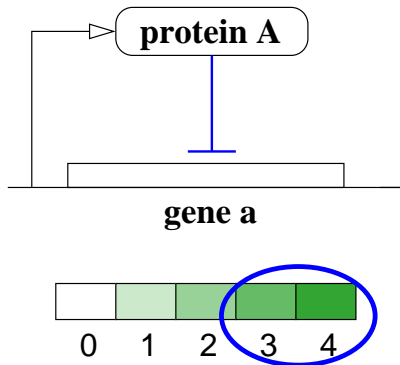
- identifikace diskretních úrovní exprese

Příklad modelu – autoregulace



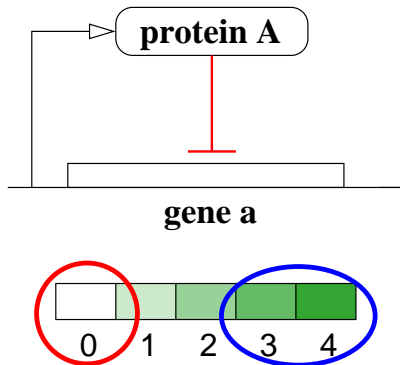
- spontánní (tzv. bázová) transkripce: $A \rightarrow 4$

Příklad modelu – autoregulace



- místo projevu regulace ($A \in \{3, 4\} \Rightarrow$ regulace aktivní)

Příklad modelu – autoregulace



- cílový bod regulace ($A \in \{3, 4\} \Rightarrow A \rightarrow 0$)

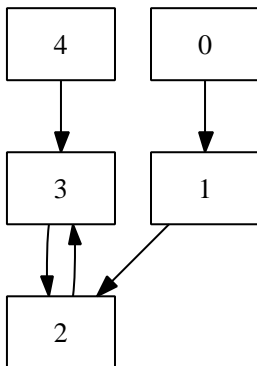
Stavový prostor – autoregulace

- přechodový systém $\langle S, T, S_0 \rangle$
 - S množina stavů, $S \equiv \{0, 1, 2, 3, 4\}$
 - $S_0 \subseteq S$ množina počátečních stavů
 - $T \subseteq S \times S$ přechodová relace:

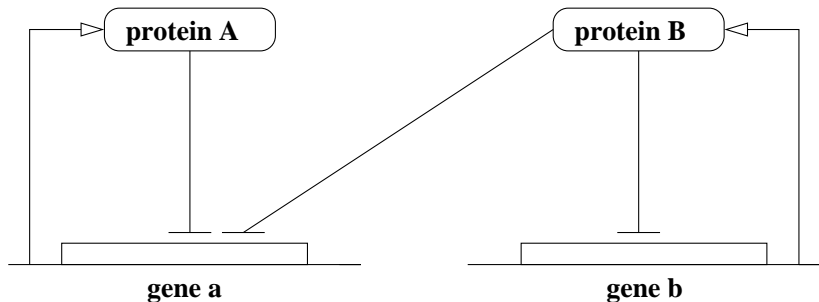
zdrojový stav	aktivní regulace	cílový stav
0	$\emptyset; [A \rightarrow 4]$	1
1	$\emptyset; [A \rightarrow 4]$	2
2	$\emptyset; [A \rightarrow 4]$	3
3	$A \rightarrow^- A; [A \rightarrow 0]$	2
4	$A \rightarrow^- A; [A \rightarrow 0]$	3

Stavový prostor – autoregulace

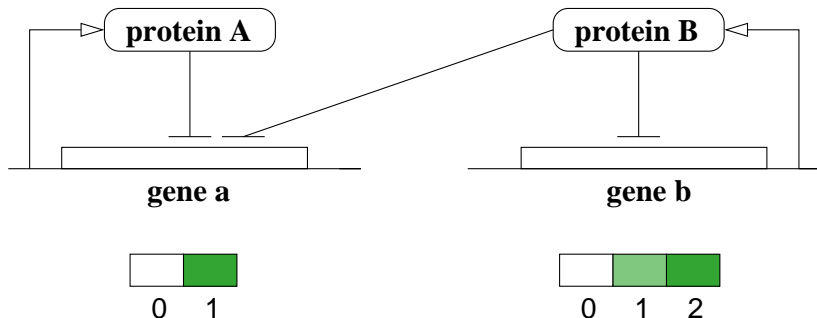
přechodový systém pro negativní autoregulaci $\langle S, T, S_0 = S \rangle$:



Příklad modelu složené regulace

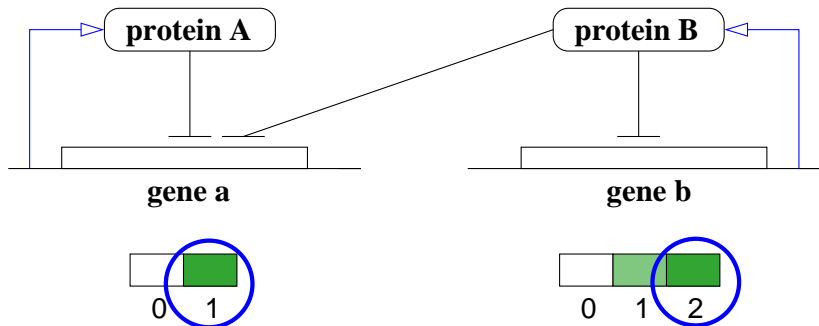


Diskrétní charakteristika dynamiky



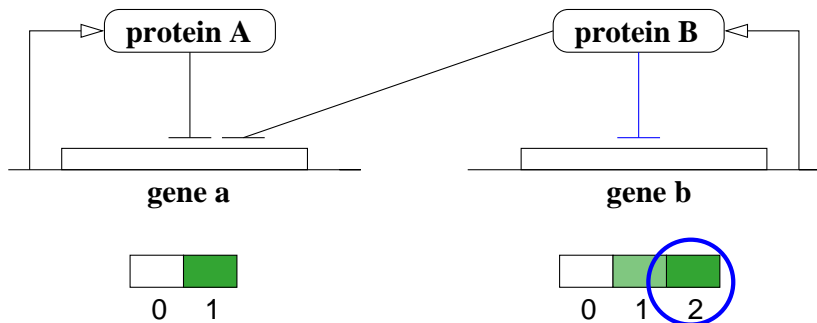
- identifikace diskretních úrovní exprese

Diskrétní charakteristika dynamiky



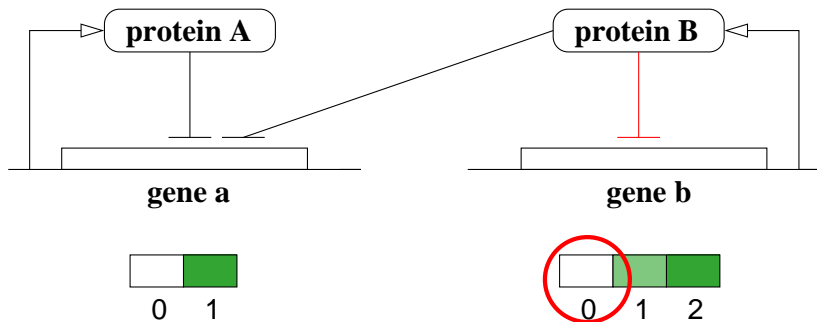
- spontánní (tzv. bázová) transkripce: $A \rightarrow 1$, $B \rightarrow 2$

Charakteristika regulace – autoregulace



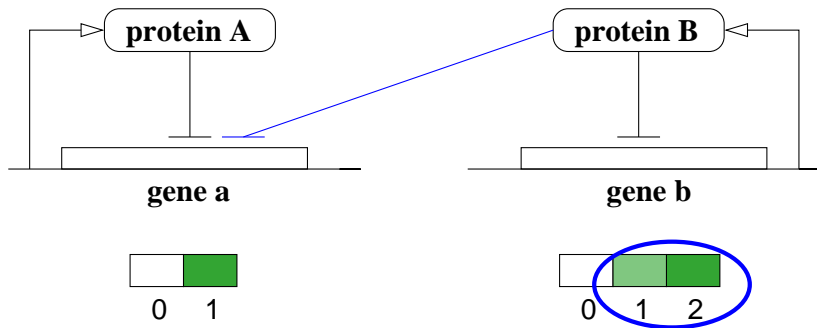
- místo projevu regulace $B \rightarrow^- B$ ($B = 2 \Rightarrow$ regulace aktivní)

Charakteristika regulace – autoregulace



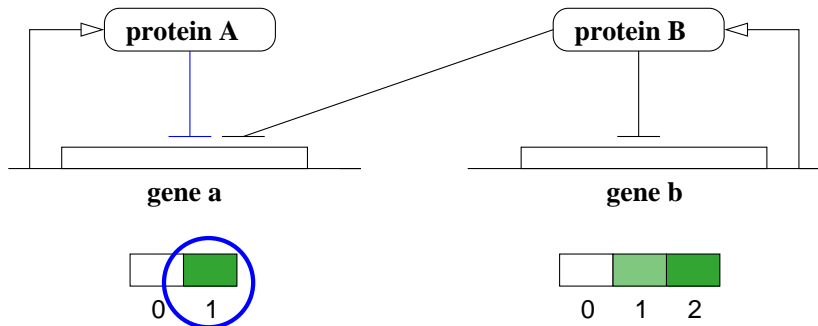
- cílový bod regulace $B \rightarrow^- B$ ($B = 2 \Rightarrow B \rightarrow 0$)

Charakteristika regulace – vstupní funkce



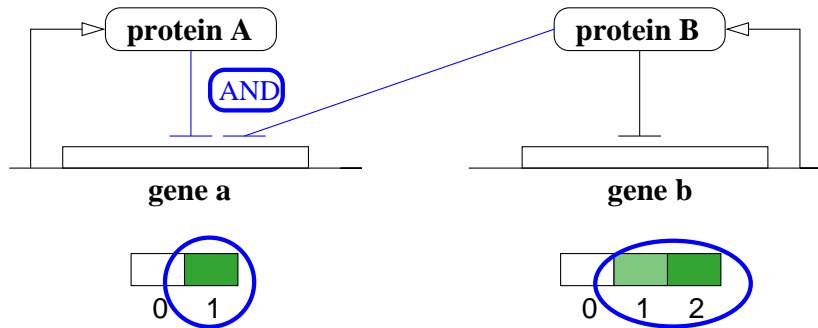
- místo projevu regulace $B \rightarrow^- A$ ($B \in \{1, 2\} \Rightarrow$ reg. aktivní)

Charakteristika regulace – vstupní funkce



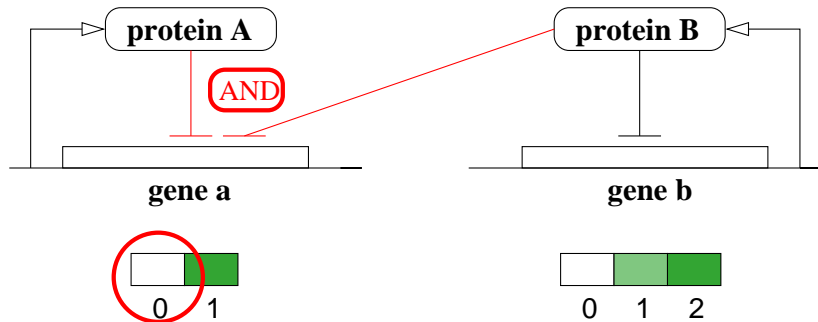
- místo projevu regulace $A \rightarrow^- A$ ($A = 1 \Rightarrow$ reg. aktivní)

Charakteristika regulace – vstupní funkce



- AND-kompozice regulací $A \rightarrow^- A \wedge B \rightarrow^- A$:
 $A = 1 \wedge B \in \{1, 2\} \Rightarrow$ regulace aktivní

Charakteristika regulace – vstupní funkce



- cílový bod složené regulace $A \rightarrow^- A \wedge B \rightarrow^- A$:
 $A = 1 \wedge B \in \{1, 2\} \Rightarrow A \rightarrow 0$

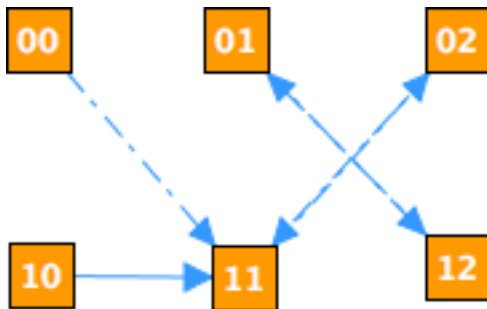
Stavový prostor – synchronní sémantika

- přechodový systém $\langle S, T, S_0 \rangle$
 - $S \equiv \{0, 1\} \times \{0, 1, 2\}$
 - $S_0 \subseteq S$, uvažujeme $S_0 = S$
 - $T \subseteq S \times S$ přechodová relace (zobrazení):

zdrojový stav	aktivní regulace	cílový stav
[0, 0]	$\emptyset; [A \rightarrow 1, B \rightarrow 2]$	[1, 1]
[0, 1]	$B \rightarrow^- A; [A \rightarrow 1, B \rightarrow 2]$	[1, 2]
[0, 2]	$B \rightarrow^- B \wedge B \rightarrow^- A; [A \rightarrow 1, B \rightarrow 0]$	[1, 1]
[1, 0]	$A \rightarrow^- A; [A \rightarrow 1, B \rightarrow 2]$	[1, 1]
[1, 1]	$A \rightarrow^- A \wedge B \rightarrow^- A; [A \rightarrow 0, B \rightarrow 2]$	[0, 2]
[1, 2]	$A \rightarrow^- A \wedge B \rightarrow^- A \wedge B \rightarrow^- B; [A \rightarrow 0, B \rightarrow 0]$	[0, 1]

Stavový prostor – synchronní sémantika

přechodový systém $\langle S, T, S_0 = S \rangle$:



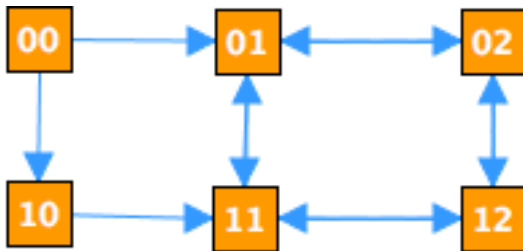
Stavový prostor – asynchronní sémantika

- přechodový systém $\langle S, T, S_0 \rangle$
 - $S \equiv \{0, 1\} \times \{0, 1, 2\}$
 - $S_0 \subseteq S$, uvažujeme $S_0 = S$
 - $T \subseteq S \times S$ přechodová relace:

zdroj. stav	aktivní regulace	cílové stavy
[0, 0]	$\emptyset; [A \rightarrow 1, B \rightarrow 2]$	[1, 0], [0, 1]
[0, 1]	$B \rightarrow^- A; [A \rightarrow 1, B \rightarrow 2]$	[1, 1], [0, 2]
[0, 2]	$B \rightarrow^- B \wedge B \rightarrow^- A; [A \rightarrow 1, B \rightarrow 0]$	[0, 1], [1, 2]
[1, 0]	$A \rightarrow^- A; [A \rightarrow 1, B \rightarrow 2]$	[1, 1]
[1, 1]	$A \rightarrow^- A \wedge B \rightarrow^- A; [A \rightarrow 0, B \rightarrow 2]$	[0, 1], [1, 2]
[1, 2]	$A \rightarrow^- A \wedge B \rightarrow^- A \wedge B \rightarrow^- B; [A \rightarrow 0, B \rightarrow 0]$	[0, 2], [1, 1]

Stavový prostor – asynchronní sémantika

přechodový systém $\langle S, T, S_0 = S \rangle$:



Vlastnosti diskrétních sémantik

- synchronní sémantika
 - efekt aktivních regulací uplatněn pro všechny proteiny ve stejný okamžik
 - nerealistická aproximace, dává však deterministický přechodový systém
- asynchronní sémantika
 - efekt aktivních regulací uplatněn pro každý protein individuálně (interleaving)
 - nutno uvažovat všechny možné souběhy
 - věrnější aproximace, dává však nedeterministický přechodový systém
 - možnost definovat priority

Nástroj GINsim

- nástroj Gene Interaction Network simulation (GINsim)
<http://gin.univ-mrs.fr/GINsim/accueil.html>
- umožňuje asynchronní i synchronní simulaci transkripční regulace
- inherentně diskrétní model (vícehodnotová logika)
 - místo přesné hodnoty koncentrace rozlišujeme několik diskrétních úrovní
 - s každou regulací spjat aktivační interval diskrétních úrovní specifikující kdy je regulující protein aktivní
 - u každého proteinu je specifikován individuální/kompozitní projev vstupních regulací
 - možnost neregulované (bázové) transkripce
- grafové algoritmy pro transkripční síť i přechodový systém

Obsah

Boolovské síťe

Rekonstrukce boolovských sítí

Rekonstrukce boolovských sítí z microarray dat

- diskretizace microarray dat (časová řada, množina řad)
- clustering genů do skupin s podobným profilem
- inference logiky v dynamice klastrů

Rekonstrukce boolovských sítí z microarray dat

Diskretizace

- metody řízené absolutními hodnotami
- metody citlivé na variace mezi jednotlivými časovými body
- matice expresních dat $A_{ij} \in \mathbb{R}$
- cílem je získat matici diskretizovanou $A_{ij} \in \{0, 1\}$

Rekonstrukce boolovských sítí z microarray dat

Diskretizace

metody řízené absolutními hodnotami:

- dle průměrné hodnoty exprese

$$A_{ij} = \begin{cases} 1, & A'_{ij} \geq A'_{iJ} \\ 0, & \text{jinak} \end{cases}$$

A_{iJ} zn. průměrnou hodnotu exprese (lze nahradit A_{iJ} , A_{iJ})

Rekonstrukce boolovských sítí z microarray dat

Diskretizace

metody řízené absolutními hodnotami:

- dle průměrné hodnoty exprese

$$A_{ij} = \begin{cases} 1, & A'_{ij} \geq A'_{iJ} \\ 0, & \text{jinak} \end{cases}$$

A_{iJ} zn. průměrnou hodnotu exprese (lze nahradit A_{iJ} , A_{iJ})

Rekonstrukce boolovských sítí z microarray dat

Diskretizace

metody řízené absolutními hodnotami:

- dle mediánu exprese (mezi nejnižší a nejvyšší hodnotou)

$$A_{ij} = \begin{cases} 1, & A'_{ij} \geq M_{IJ} \\ 0, & \text{jinak} \end{cases}$$

M_{IJ} zn. medián exprese (lze nahradit M_{Ij} , M_{iJ})

Rekonstrukce boolovských sítí z microarray dat

Diskretizace

metody řízené absolutními hodnotami:

- dle maxima exprese

$$A_{ij} = \begin{cases} 1, & A'_{ij} \geq H_{IJ}(1 - X) \\ 0, & \text{jinak} \end{cases}$$

H_{IJ} zn. maximum exprese (lze nahradit H_{Ij} , H_{jJ})

$X \in [0, 1]$ je procento maxima uvažované jako prahová hodnota

Rekonstrukce boolovských sítí z microarray dat

Diskretizace

metody řízené absolutními hodnotami:

- TOP $X\%$

$$A_{ij} = \begin{cases} 1, & \text{pokud } \text{pozice}(A'_{ij}, \text{sort}(A')) \geq \frac{X \times n \times m}{100} \\ 0, & \text{jinak} \end{cases}$$

$\text{sort}(A)$ zn. seznam prvků matice A setříděných vzestupně

$\text{pozice}(x, L)$ zn. pozici prvku x v seznamu L

Rekonstrukce boolovských sítí z microarray dat

Diskretizace

metody citlivé na variaci časové řady:

- diskriminační metoda (Transitional State Discrimination)
- využívá normalizace hodnot prostřednictvím z-skóre:

$$z = \frac{x - \mu}{\sigma}$$

kde σ je standardní odchylka, μ je průměrná hodnota exprese

$$A_{ij} = \begin{cases} 1, & |A'_{ij} - A'_{i(j-1)}| \geq t \\ 0, & \text{jinak} \end{cases}$$

t je diskriminační počet std. odchylek

Rekonstrukce boolovských sítí z microarray dat





- vstup: množina uzlů (odp. klastrů) $V = \{v_1, v_2, \dots, v_n\}$
- pro každý uzel v_i urči, jakým způsobem podmnožina uzlů $U = \{u_1, \dots, u_q\} \subseteq V$ vysvětluje expresní profil uzlu v_i (detekce způsobu regulace)
- k tomu je nutno projít všechny časové body všech časových řad
- výsledkem předch. kroku je aktivační-inhibiční funkce tvaru $v_i(t+1) = (u_1(t) \vee u_2(t) \vee \dots) \wedge \neg(u_j(t) \vee u_{j+1}(t) \vee \dots)$ kde $u_1, u_2, \dots \subseteq U$ jsou aktivátory, $u_j, u_{j+1}, \dots \subseteq U$ jsou inhibitory
- procedura je spuštěna pro každé možné nastavení q , tedy $O(n^q)$
- typicky stačí $q \leq 5$
- výstup: množina potenciálních modelů (sítí)

Rekonstrukce boolovských sítí z microarray dat

- $V = \{v_1, v_2, v_3\}$
- výstupem pro $q = 3$ a uzel v_3 je funkce
 $v_3(t + 1) = v_1(t) \wedge \neg v_2(t)$

		Time					
		1	2	3	4	5	6
Gene	v_1	0	0	1	1	0	0
	v_2	1	0	0	1	1	0
	v_3	1	0	0	1	0	0

Literatura

-  Bower, J.M. & Bolouri, H. *Computational Modeling of Genetic and Biochemical Networks*. Bradford Book, 2001.
-  A.G. Gonzalez, A. Naldi, L. Sánchez, D.Thieffry, C. Chaouiya. *GINsim: a software suite for the qualitative modelling, simulation and analysis of regulatory networks*. Biosystems (2006), 84(2):91-100
-  Kauffman, S. A. (1969). Metabolic stability and epigenesis in randomly constructed genetic nets. *Journal of Theoretical Biology*, 22:437-467
-  Martin, S. et al. (2007) *Boolean dynamics of genetic regulatory networks inferred from microarray time series data*. *Bioinformatics*, 23(7):866-874