

PA054: Formální modely v systémové biologii

David Šafránek

2.5.2012

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



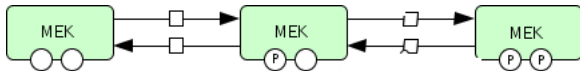
Obsah

Algebraický přístup k modelování biochemických reakcí

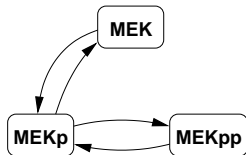
Princip modelování

Substráty – molekuly

- populace interagujících molekul
- každou molekulu (proteinu) lze popsat stavově
 - stav zachycuje konfiguraci vazebných míst
 - volná vazebná místa
 - vazebná místa obsazena jinou molekulou (dimerizující stavy)



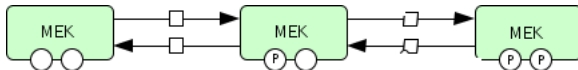
- formalizace prostřednictvím sekvenčního procesu:



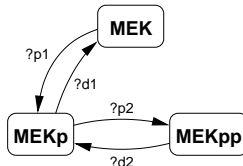
Princip modelování

Substráty – molekuly

- populace interagujících molekul
- každou molekulu (proteinu) lze popsat stavově
 - stav zachycuje konfiguraci vazebných míst
 - volná vazebná místa
 - vazebná místa obsazena jinou molekulou (dimerizující stavy)



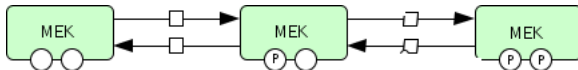
- formalizace prostřednictvím sekvenčního procesu:



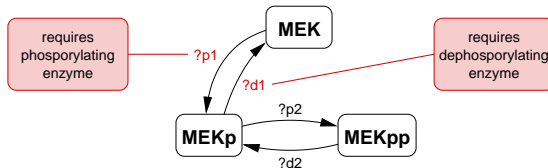
Princip modelování

Substráty – molekuly

- populace interagujících molekul
- každou molekulu (proteinu) lze popsat stavově
 - stav zachycuje konfiguraci vazebných míst
 - volná vazebná místa
 - vazebná místa obsazena jinou molekulou (dimerizující stavy)

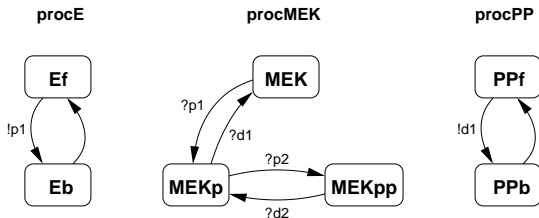


- formalizace prostřednictvím sekvenčního procesu:



Princip modelování

Interakce

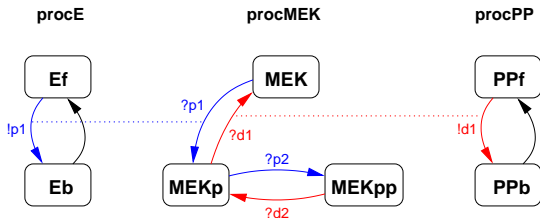


- paralelní chování procesů (molekul)
- binární synchronizace
- roztok lze modelovat jako populaci procesů:

$$procMEK | procMEK | procE | procE | procE | procPP | procPP$$

Princip modelování

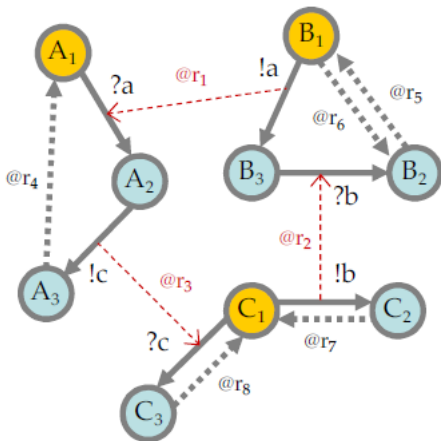
Interakce



- paralelní chování procesů (molekul)
- binární synchronizace
- roztok lze modelovat jako populaci procesů:

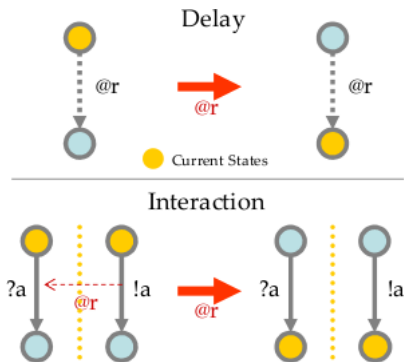
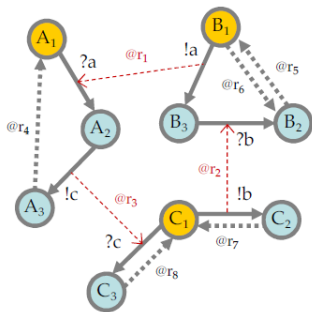
$$procMEK | procMEK | procE | procE | procE | procPP | procPP$$

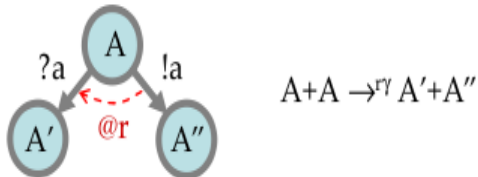
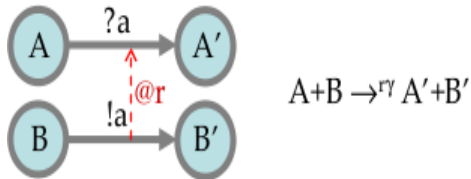
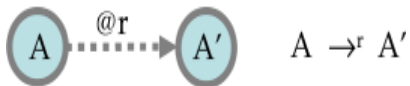
Molekuly jako komunikující automaty



individuální přechody a synchronizace lze modelovat stochasticky
 (provedení přechodu v čase $t \sim \text{Exp}(r_i)$)

Molekuly jako komunikující automaty



Elementární reakce jako komunikující automaty

Elementární kalkulus pro popis (bio)chemických systémů

Syntax

$$E ::= 0 \dot{:} X=M, E$$

Reagents

$$M ::= 0 \dot{:} \pi;P \oplus M$$

Molecule

$$P ::= 0 \dot{:} X | P$$

Solution

$$\pi ::= \tau_{(r)} \dot{:} ?n_{(r)} \dot{:} !n_{(r)}$$

Interaction prefix

$$\text{CGF} ::= E, P$$

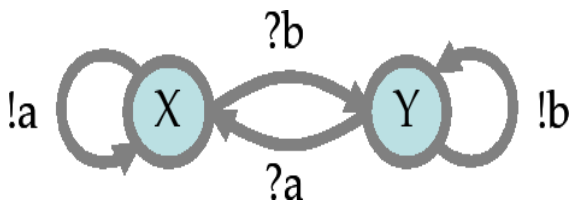
Chemical Ground Form

Příklad

$$X = !a_{(r)};X \oplus ?b_{(s)};Y$$

$$Y = !b_{(s)};Y \oplus ?a_{(r)};X$$

$$X \mid X \mid X \mid Y \mid Y$$



Elementární kalkulus pro popis (bio)chemických systémů

Sémantika

$$\begin{array}{l}
 \Rightarrow !x.P + M \xrightarrow{!x} P \\
 \Rightarrow ?x.P + M \xrightarrow{?x} P \\
 \Rightarrow \tau_r.P \xrightarrow{r} P \\
 P \xrightarrow{!x} P' \quad Q \xrightarrow{?x} Q' \Rightarrow P|Q \xrightarrow{x} P'|Q' \\
 M \xrightarrow{\alpha} P' \Rightarrow \pi.P + M \xrightarrow{\alpha} P' \\
 P \xrightarrow{\alpha} \Rightarrow P|Q \xrightarrow{\alpha} P'|Q \\
 X = P \quad P \xrightarrow{\alpha} P' \Rightarrow X \xrightarrow{\alpha} P'
 \end{array}$$

SOS pravidla definují fragment stochastického π -kalkulu

Rozšíření o předávání hodnot

Syntax

$E ::= 0 \dot{=} X(\mathbf{p})=M, E$	Reagents
$M ::= 0 \dot{=} \pi;P \oplus M$	Molecule
$P ::= 0 \dot{=} X(\mathbf{p}) \mid P$	Solution
$\pi ::= \tau_{(r)} \dot{=} ?n(\mathbf{p}) \dot{=} !n(\mathbf{p})$	Interaction prefix
$CPF ::= E, P$	Chemical Parametric Form

Rozšíření o předávání hodnot

Sémantika

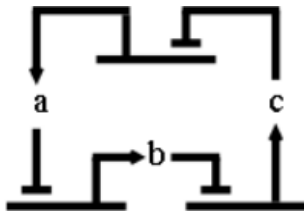
$$\begin{array}{l}
 P \xrightarrow{!x(n)} P' \quad Q \xrightarrow{?x(m)} Q' \\
 M \xrightarrow{\alpha} P' \\
 P \xrightarrow{\alpha} \\
 X(m) = P \quad P_{\{n/m\}} \xrightarrow{\alpha} P' \Rightarrow X(n) \xrightarrow{\alpha} P'
 \end{array}
 \Rightarrow
 \begin{array}{l}
 !x(n).P + M \xrightarrow{!x(n)} P \\
 ?x(m).P + M \xrightarrow{?x(m)} P_{\{n/m\}} \\
 \tau_r.P \xrightarrow{r} P \\
 P|Q \xrightarrow{x} P'|Q' \\
 \pi.P + M \xrightarrow{\alpha} P' \\
 P|Q \xrightarrow{\alpha} P'|Q'
 \end{array}$$

SOS pravidla definují fragment stochastického π -kalkulu

Příklad modelování genetické regulační sítě



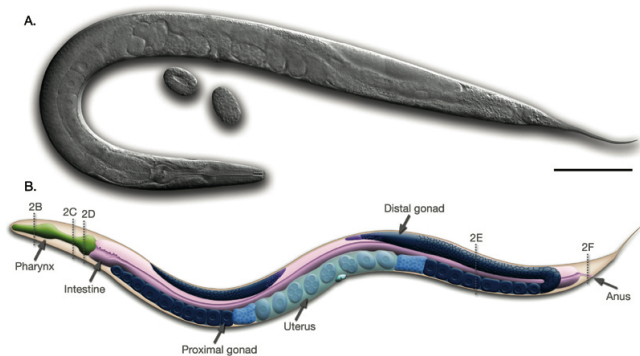
$$\begin{aligned}
 \text{Gene}(a, b) &= \tau_t.(\text{Gene}(a, b)|\text{Protein}(b)) + ?a.\text{Blocked}(a, b) \\
 \text{Blocked}(a, b) &= \tau_u.\text{Gene}(a, b) \\
 \text{Protein}(b) &= !b.\text{Protein}(b) + \tau_d.0
 \end{aligned}$$

Příklad modelování genetické regulační sítě

$Gene(c, a) | Gene(a, b) | Gene(b, c)$

High-level modelování

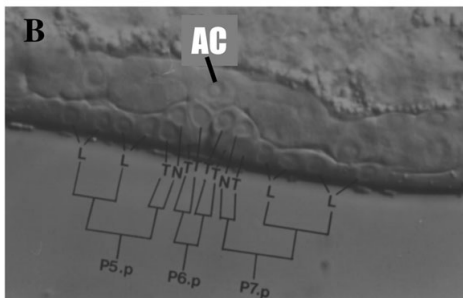
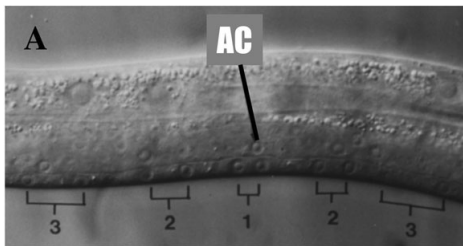
C. elegans



IntroFig1

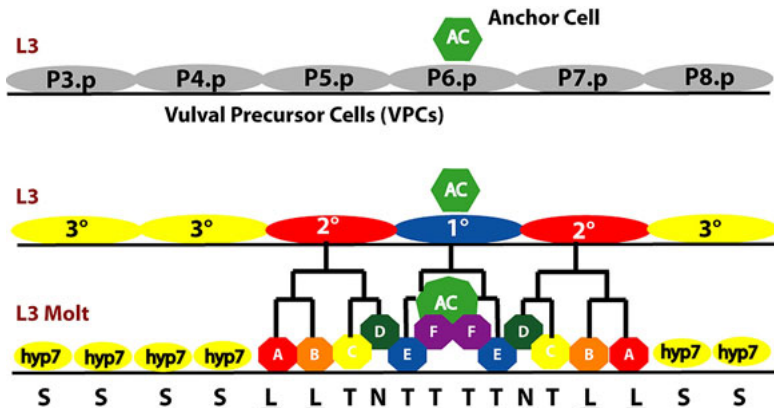
High-level modelování

C. elegans: vývoj vulvy



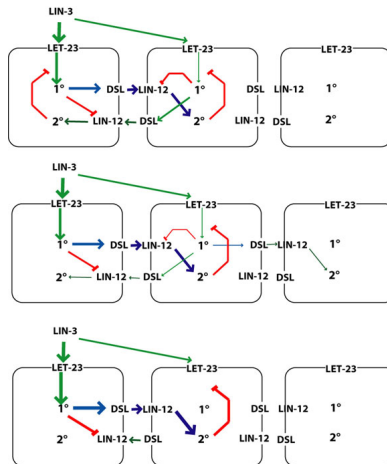
High-level modelování

C. elegans: vývoj vulvy



High-level modelování

C. elegans: možné varianty signálních drah

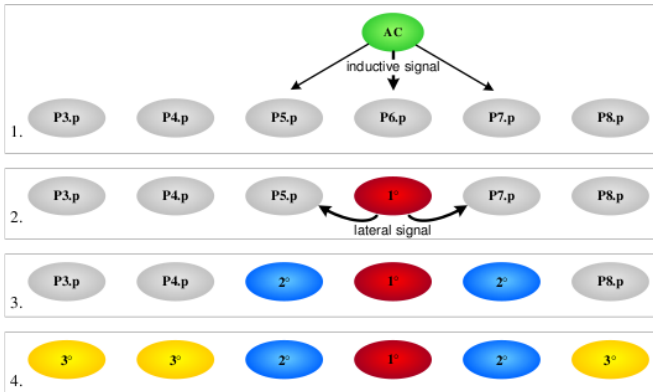


LIN3 ... signál od řídicí buňky (anchor cell, AC)

LIN12 ... laterální mezibuněčný signál

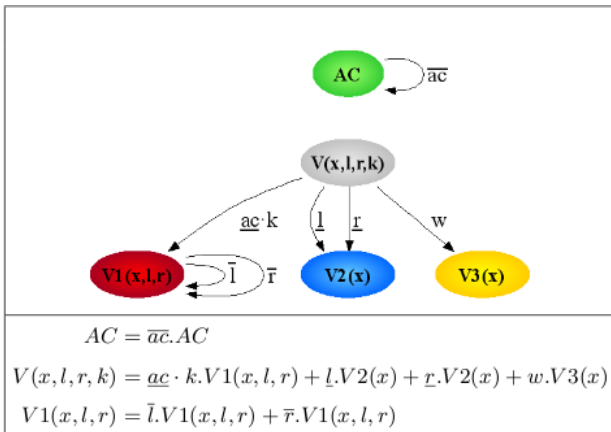
High-level modelování – interagující buňky

Vývoj vulvy *C. elegans*



High-level modelování – interagující buňky

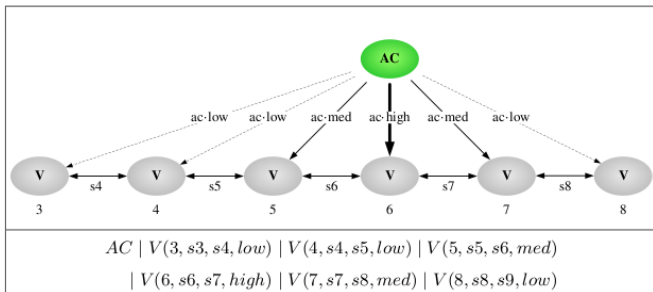
Vývoj vulvy *C. elegans*



x ... č. buňky, l, r ... kanály laterálního signálu,
 k ... vzdálenost od AC (frekvence interakce s AC)

High-level modelování – interagující buňky

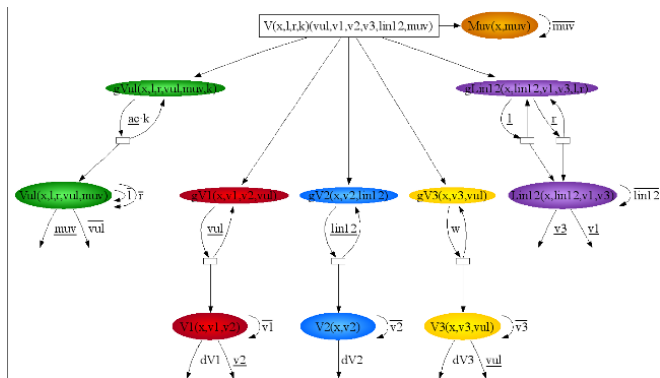
Vývoj vulvy *C. elegans*



Andrew Phillips, A Visual Process Calculus for Biology, in Symbolic Systems Biology: Theory and Methods, Jones and Bartlett Publishers, 2010.

High-level modelování – interagující buňky

Vývoj vulvy *C. elegans* – podrobný model



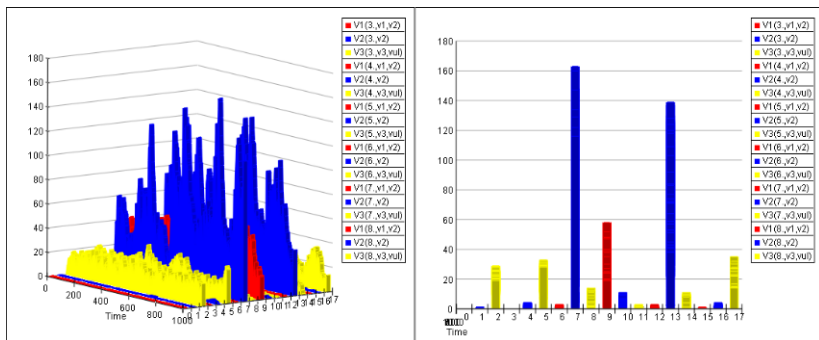
High-level modelování – interagující buňky

Vývoj vulvy *C. elegans* – podrobný model

$$\begin{aligned}
 V(x, l, r, k) &= (vul, v1, v3, v3, lin12, muv) \\
 &\quad (gVul(x, l, r, vul, muv, k) \mid gLin12(x, lin12, v1, v3, l, r) \\
 &\quad \mid gV1(x, v1, v2, vul) \mid gV2(x, v2, lin12) \mid gV3(x, v3, vul) \mid Muv(x, muv)) \\
 gVul(x, l, r, vul, muv, k) &= \underline{ac} \cdot k.(gVul(x, l, r, vul, muv, k) \mid Vul(x, l, r, vul, muv)) \\
 gLin12(x, lin12, v1, v3, l, r) &= \underline{l}.(gLin12(x, lin12, v1, v3, l, r) \mid Lin12(x, lin12, v1, v3) \\
 &\quad + \underline{r}.(gLin12(x, lin12, v1, v3, l, r) \mid Lin12(x, lin12, v1, v3)) \\
 gV1(x, v1, v2, vul) &= \underline{vul}.(gV1(x, v1, v2, vul) \mid V1(x, v1, v2)) \\
 gV2(x, v2, lin12) &= \underline{lin12}.(gV2(x, v2, lin12) \mid V2(x, v2)) \\
 gV3(x, v3, vul) &= \underline{w}.(gV3(x, v3, vul) \mid V3(x, v3, vul)) \\
 Vul(x, l, r, vul, muv) &= \bar{l}.Vul(x, l, r, vul, muv) + \bar{r}.Vul(x, l, r, vul, muv) + \bar{vul} + \underline{muv} \\
 Lin12(x, lin12, v1, v3) &= \bar{lin12}.Lin12(x, lin12, v1, v3) + \underline{v3} + \underline{v1} \\
 Muv(x, muv) &= \bar{muv}.Muv(x, muv) \\
 V1(x, v1, v2) &= \bar{v1}.V1(x, v1, v2) + \underline{v2} + dV1 \\
 V2(x, v2) &= \bar{v2}.V2(x, v2) + dV2 \\
 V3(x, v3, vul) &= \bar{v3}.V3(x, v3, vul) + \underline{vul} + dV3
 \end{aligned}$$

High-level modelování – simulace

Vývoj vulvy *C. elegans*

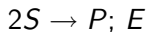


Kalkuly pro biologické systémy

- Stochastic π -calculus (SPiM, BioSPI)
- BioPEPA
- κ -calculus
- Brane-Calculus, BetaBinders, BlenX
- ...

BioPEPA

- na rozdíl od π -based kalkulů je BioPEPA orientovaný blíže SBML
- rozlišení není až na úroveň molekul, ale na úroveň substrátů a reakcí
⇒ proces není molekula, ale substrát
- výhodou je kompatibilita sémantik s SBML (srovnej s Petriho sítěmi)
- možnost plného modelování netriviální kinetiky (např. regulace, enzymová kinetika)

BioPEPA – příklad

$$S = (\alpha, 2) \downarrow S$$

$$E = (\alpha, 1) \oplus E$$

$$P = (\alpha, 1) \uparrow P$$

$$(S(I_{S_0}) \bowtie_{\alpha} (E(I_{E_0}))) \bowtie_{\alpha} P(I_{P_0}))$$

Sémantika dynamiky je definována pravidlem $f_{\alpha} = \frac{v \cdot E \cdot S}{K + S^2}$.

BioPEPA – syntax

Algebra modelových komponent

$$S ::= (\alpha, k) \text{ op } S \mid S + S \mid C$$

$$\text{op} = \downarrow \mid \uparrow \mid \oplus \mid \ominus \mid \odot$$

$$P ::= P \bowtie_{\mathcal{L}} P \mid S(I)$$

- \mathcal{L} je množina reakcí
- I je iniciální podmínka

BioPEPA – syntax

Definice elementárních komponent – substráty

Každá komponenta C je charakterizována 5-ticí $\langle H, N, M_0, M, V \rangle$, kde:

- $H \in \mathbb{R}^+$ je rozlišení kvantity (velikost jedné “úrovně” koncentrace),
- $N \in \mathbb{N}$ je maximální hodnota kvantity,
- $M_0 \in \mathbb{R}^+ \cup \{-\}$ je iniciální koncentrace,
- $M \in \mathbb{R}^+ \cup \{-\}$ je maximální koncentrace,
- V označuje kompartment do něhož je komponenta přiřazena.

Pozn. (1): $-$ značí “nedefinovanou hodnotu”

Pozn. (2): M, M_0 jsou uvedeny pro kalibraci se spojitým modelem.

BioPEPA – syntax

Definice elementárních komponent – substráty

- diskretizace koncentrace pomocí úrovní stejné délky (v počtu molekul)
- diskrétní sémantika tedy umožňuje kalibraci se spojitou (aproximace)
- počet úrovní pro komponentu C_i je $N_i + 1$, kde N_i je max. úroveň
- platí: $N_i = \lceil \frac{M_i}{h} \rceil$
- pro aktuální úroveň $0 \leq l_i \leq N_i$ je koncentrace definována vztahem $x_i = l_i \cdot h$

BioPEPA – definice biologického procesu

BioPEPA systém je šestice $\langle \mathcal{V}, \mathcal{N}, \mathcal{K}, \mathcal{F}_R, Comp, P \rangle$ kde:

- \mathcal{V} ... množina kompartmentů,
- \mathcal{N} ... množina kvantit popisujících substráty,
- \mathcal{K} ... množina definic parametrů,
- \mathcal{F}_R ... množina definic dynamiky,
- $Comp$... definice komponent,
- P ... komponenta popisující systém.

BioPEPA – sémantika

- každá akce α má přiřazenu kinetickou funkci f_α nad kvantitami substrátů
- kvantitativní substrátů jsou určeny z kontextu akce α
- kinetické funkce mohou obsahovat kinetické parametry, které musí být asociovány v \mathcal{K}
- příklady typů kinetických funkcí:
 - $fMA(k) = k \cdot \prod_{i=1}^{n_j} (C_i)^{\kappa_i}$ kde n_j je počet reaktantů reakce α_j a κ_i je příslušný stechiometrický koeficient reaktantu C_i
 - $fH(v, K, n) = v \cdot C^n / (K + C^n)$ je Hillova kinetika

BioPEPA – operační sémantika

Předpokládáme \mathcal{C} spočetná množinu zahrnující všechny možné modelové komponenty.

BioPEPA systém má asociovanou stochastickou sémantiku, která je definována prostřednictvím *odvozující relace* $\rightarrow_c \subseteq \mathcal{C} \times \Omega \times \mathcal{C}$, kde $\theta \in \Omega$ vyjadřuje kvantitativní informaci potřebnou k vyhodnocení kinetické funkce,

$$\theta := (\alpha, w)$$

kde $w := [S : op(l, \kappa)] \mid w@w$, $S \in \mathcal{C}$, l hodnota kvantity komponenty S , κ stechiometrický koeficient komponenty S v α .

Relace \rightarrow_c je definována jako minimální relace splňující pravidla na násl. slidu.

BioPEPA – odvozující relace

$$\text{prefixReac} \quad ((\alpha, \kappa)\downarrow S)(l) \xrightarrow{(\alpha, [S:\downarrow(l, \kappa)])}_c S(l - \kappa) \quad \kappa \leq l \leq N$$

$$\text{prefixProd} \quad ((\alpha, \kappa)\uparrow S)(l) \xrightarrow{(\alpha, [S:\uparrow(l, \kappa)])}_c S(l + \kappa) \quad 0 \leq l \leq (N - \kappa)$$

$$\text{prefixMod} \quad ((\alpha, \kappa)op S)(l) \xrightarrow{(\alpha, [S:op(l, \kappa)])}_c S(l) \quad \text{with } op = \ominus, \oplus, \Theta \text{ and } 0 \leq l \leq N$$

$$\text{choice1} \quad \frac{S_1(l) \xrightarrow{(\alpha, w)}_c S'_1(l')}{(S_1 + S_2)(l) \xrightarrow{(\alpha, w)}_c S'_1(l')} \quad \text{choice2} \quad \frac{S_2(l) \xrightarrow{(\alpha, w)}_c S'_2(l')}{(S_1 + S_2)(l) \xrightarrow{(\alpha, w)}_c S'_2(l')}$$

$$\text{constant} \quad \frac{S(l) \xrightarrow{(\alpha, S':[op(l, \kappa)])}_c S'(l')}{C(l) \xrightarrow{(\alpha, C:[op(l, \kappa)])}_c S'(l')} \quad \text{with } C \stackrel{def}{=} S$$

$$\text{coop1} \quad \frac{P_1 \xrightarrow{(\alpha, w)}_c P'_1}{P_1 \boxtimes_{\mathcal{L}} P_2 \xrightarrow{(\alpha, w)}_c P'_1 \boxtimes_{\mathcal{L}} P_2} \quad \text{with } \alpha \notin \mathcal{L}$$

$$\text{coop2} \quad \frac{P_2 \xrightarrow{(\alpha, w)}_c P'_2}{P_1 \boxtimes_{\mathcal{L}} P_2 \xrightarrow{(\alpha, w)}_c P_1 \boxtimes_{\mathcal{L}} P'_2} \quad \text{with } \alpha \notin \mathcal{L}$$

$$\text{coop3} \quad \frac{P_1 \xrightarrow{(\alpha, w_1)}_c P'_1 \quad P_2 \xrightarrow{(\alpha, w_2)}_c P'_2}{P_1 \boxtimes_{\mathcal{L}} P_2 \xrightarrow{(\alpha, w_1 @ w_2)}_c P'_1 \boxtimes_{\mathcal{L}} P'_2} \quad \text{with } \alpha \in \mathcal{L}$$

BioPEPA – stochastická sémantika

Stochastická sémantika BioPEPA systému je určena relací:

$$\rightarrow_s \subseteq \mathcal{P} \times \Gamma \times \mathcal{P}$$

kde $\gamma \in \Gamma$, $\gamma := (\alpha, r_\alpha)$, $r_\alpha \in \mathbb{R}^+$ určuje parametr exponenciální distribuce události v čase (rate).

Relace \rightarrow_s je definována pomocí pravidla:

$$\text{Final} \frac{P \xrightarrow{(\alpha_j, w)}_c P'}{\langle \mathcal{V}, \mathcal{N}, \mathcal{K}, \mathcal{F}, \text{Comp}, P \rangle \xrightarrow{(\alpha_j, r_\alpha[w, \mathcal{N}, \mathcal{K}])} \rightarrow_s \langle \mathcal{V}, \mathcal{N}, \mathcal{K}, \mathcal{F}, \text{Comp}, P' \rangle}$$

kde

$$r_\alpha[w, \mathcal{N}, \mathcal{K}] = \frac{f_\alpha[w, \mathcal{N}, \mathcal{K}]}{h}$$

- h je velikost kroku
- $f_\alpha[w, \mathcal{N}, \mathcal{K}]$ značí kin. funkci f_α vypočítanou v přísl. kontextu

BioPEPA – stochastická sémantika

Evaluace $f[w, \mathcal{N}, \mathcal{K}]$:

- pro každou komponentu C_i je odvozena koncentrace $l_i \cdot h$
- volný výskyt C_i je nahrazen $(l_i \times h)^{\kappa_{ij}}$, kde κ_{ij} je stechiometrický koeficient substrátu C_i v reakci R_j
- pozn.: pro $h = 1$ a vhodné omezení kinetických funkcí dostáváme přímou stochastickou sémantiku (relace \rightarrow_s definuje CTMC, který bychom získali interpretací stochastické Petriho sítě odpovídající BioPEPA systému)

BioPEPA – příklad

Uvažujme reakci $\alpha : 2X + Y \xrightarrow{k} 3Z$. Kinetika je definována funkcí $v = k \cdot X^2 \cdot Y$. Definujeme komponenty BioPEPA:

$$X \stackrel{\text{def}}{=} (\alpha, 2)\downarrow X \quad Y \stackrel{\text{def}}{=} (\alpha, 1)\downarrow Y \quad Z \stackrel{\text{def}}{=} (\alpha, 3)\uparrow Z$$

Reakční systém je definován:

$$X(l_{X_0}) \bowtie_{\{\alpha\}} Y(l_{Y_0}) \bowtie_{\{\alpha\}} Z(l_{Z_0})$$

kde $l_{X_0}, l_{Y_0}, l_{Z_0}$ značí iniciální úrovně koncentrace, kin. funkce je definována $f_\alpha = fMA(k)$.

Pro reakci bude odvozením stochastické sémantiky vypočítán rate

$$r_\alpha = \frac{k \cdot (l_X \cdot h)^2 (l_Y \cdot h)}{h}$$

Reakce bude provedena pouze v situaci: N_X je nejméně 3, N_Z je nejméně 4.

κ -calculus