

MASARYKOVA UNIVERZITA
FAKULTA INFORMATIKY



Názov diplomovej práce

DIPLOMOVÁ PRÁCA

Martin Demko

Brno, jar 2014

Prehlásenie

Prehlasujem, že táto diplomová práca je mojím pôvodným autorským dielom, ktoré som vypracoval samostatne. Všetky zdroje, pramene a literatúru, ktoré som pri vypracovaní používal alebo z nich čerpal, v práci riadne citujem s uvedením úplného odkazu na príslušný zdroj.

Vedúci práce: RNDr. David Šafránek, Ph.D.

Pod'akovanie

Zhrnutie

Klíčové slová

Obsah

1	Základné pojmy	3
1.1	<i>LTl - Lineárna Temporálna Logika</i>	3
1.2	<i>Büchiho automat</i>	4
1.3	<i>Model checking</i>	5
1.4	<i>Farebný model checking</i>	5
1.5	<i>Vstupný model</i>	5
1.6	<i>Michaelis-Mentenovej a Hillova kinetika</i>	8
1.7	<i>Abstrakcia</i>	8
2	Podobné nástroje	9
2.1	<i>BioDiVinE</i>	9
2.2	<i>PEPMC</i>	9
2.3	<i>RoverGene</i>	9
2.4	<i>ParaSim</i>	9
3	BioDiVinE 1.0	10
4	Nové vylepšenia	11
5	Implementácia	12
6	Použitie programu	13
7	Case study	14
8	Záver	15

Úvod

Kapitola 1

Základné pojmy

Než začneme zachádzať hlbšie do problematiky tejto práce a opisovať postupy a nástroje v nej použité, je potrebné vysvetliť na počiatku niekoľko pojmov. Tieto sa v práci mnohokrát opakujú a ich včasným uvedením predídeme nepochopiteľnosti textu.

1.1 LTL - Lineárna Temporálna Logika

Temporálna logika obecné je špeciálna vetva logiky zaoberajúca sa logickou štruktúrou výrokov v čase. Je to formalizmus vhodný pre overovanie vlastností formálnych dynamických systémov resp. matematických modelov.

Lineárna temporálna logika (ďalej len LTL) je najjednoduchšia verzia temporálnej logiky, ktorá neumožňuje vetvenie času ani kvantifikátory. Môžeme ju považovať tiež za konkrétny výpočtový kalkulus pracujúci s tzv. formulami, definovanými nasledujúcou syntaxou:

Atomické propozície (ďalej len AP)

$A > 0$
 $B \leq 5.834$
 $C \neq \text{"nie"}$
atd'...

Logické operátory

\neg, \vee	- základné logické operátory
$\wedge, \rightarrow, \leftrightarrow, \text{true}, \text{false}$	- odvodené logické operátory

Temporálne operátory

$X \phi$ - neXt, vyjadruje platnosť ϕ v ďalšom stave
 $G \phi$ - Global, vyjadruje trvalú platnosť ϕ

$\mathbf{F} \phi$ - **Future**, vyjadruje platnosť ϕ v niektorom z budúcich stavov
 $\psi \mathbf{U} \phi$ - **Until**, vyjadruje platnosť ψ až do kedy nezačne platiť ϕ
 $\psi \mathbf{R} \phi$ - **Release**, vyjadruje platnosť ϕ
 kde ϕ a ψ sú atomické propozície

Potom platí nasledujúce:

- Ak $p \in AP$, tak p je formula.
- Ak f a g sú formule, tak $\neg g$, $f \vee g$, $f \wedge g$, $\mathbf{X} g$, $\mathbf{F} g$, $\mathbf{G} g$, $f \mathbf{U} g$ a $g \mathbf{R} f$ sú formule.

Napriek tomu, že je LTL tou najprimitívnejšou temporálnou logikou, jej prevod do Büchiho automatu je v najhoršom prípade exponenciálne zložitý. Dôvod tohto prevodu bude vysvetlený v kapitole 1.3.

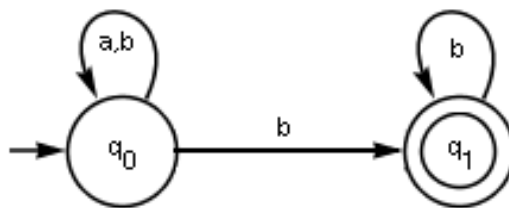
1.2 Büchiho automat

Automat obecné je matematický model stroja s konečným množstvom pamäte spracovávajúci vstup o neznámej veľkosti. Kvôli obmedzeniu pamäte ho nazývame konečným automatom. Vstup sa nazýva slovo a môže byť konečný aj nekonečný. Büchiho automat je potom najjednoduchším konečným automatom nad nekonečným slovom a preto patrí do skupiny ω -automatov.

Formálne je konečný automat \mathcal{A} päťica $(\Sigma, Q, \Delta, Q_0, F)$, pre ktorú platí:

- Σ je konečná abeceda
- Q je konečná množina stavov
- $\Delta \subseteq Q \times \Sigma \times Q$ je relácia nazývaná prechodová funkcia
- $Q_0 \subseteq Q$ je podmnožina množiny stavov, nazývaná iníciaľne stavy
- $F \subseteq Q$ je podmnožina množiny stavov, nazývaná akceptujúce stavy

Príklad jednoduchého automatu je daný na obrázku 1.1. Automat nad konečným slovom akceptuje toto slovo, ak po prejdení posledného znaku slova zodpovedajúceho prechodu medzi dvoma stavami, je tento posledný stav v množine F . Avšak automat nad nekonečným slovom nemôže nikdy prejsť cez posledný znak. Preto takýto automat akceptuje nekonečné slovo len v prípade, že počas prechádzania slova je aspoň jeden stav navštívený nekonečne často a zároveň tento stav patrí aj do množiny F .



Obr. 1.1: Príklad jednoduchého automatu

1.3 Model checking

1.4 Farebný model checking

1.5 Vstupný model

A. Multi-afinný ODE model

Vstupným modelom sa u nás myslí model biochemických reakcií, ktorý je v našom poňatí braný ako po častiach linearizovaný multi-afinný systém diferenciálnych rovníc. Ale začnime od počiatku a postupne sa dopracujeme k tomuto výsledku.

Na základe pravidiel o mass action kinetike [odkaz na wiki mass action] je možné modelovať ľubovoľnú biochemickú reakciu alebo dokonca sústavu takýchto reakcií pomocou sústavy nelineárnych obyčajných diferenciálnych rovníc (Ordinary Differential Equations alebo ODEs) [odkaz na wiki ODE].

Uvažujme multi-afinný systém vo forme $\dot{x} = f(x)$, kde $x = (x_1, \dots, x_n)$ je vektor premenných a $f = (f_1, \dots, f_n) : \mathbb{R}^n \rightarrow \mathbb{R}^n$ je vektor multi-afinných funkcií. Tieto funkcie sú vlastne polynómy, v ktorých je stupeň premenných x_1, \dots, x_n obmedzený na hodnotu 1. Každá premenná x_i , kde $i \in \{1, \dots, n\}$ predstavuje koncentráciu špecifickej chemickej látky a je interpretovaná ako $\mathbb{R}_+ = \{x \in \mathbb{R} \mid x \geq 0\}$. /*Mozno priklad*/

Z dôvodu, že premenné môžeme vyjadriť len ako nezáporné reálne čísla, je možné tiež spojiť stavový priestor nášho matematického systému obmedziť iba na prvý, resp. kladný kvadrant $\mathbb{R}_+^n = \{x \in \mathbb{R}^n \mid x \geq 0\}$.

Ak uvažujeme o premenných ako o nestabilných chemických látkach, ktoré sami od seba degradujú v čase, môžeme s kľudom obmedziť náš spojitý stavový priestor \mathcal{D} ešte viac. A síce na n -dimenzionálny obdĺžnik $\mathcal{D} = \prod_{i=1}^n [0, \max_i] \subset \mathbb{R}^n$, kde \max_i je horná hranica koncentrácie premennej x_i .

[odkaz na hibi2010 a hibi2009]

B. Po častiach linearizovaný multi-afinný ODE model

Multi-afinný systém diferenciálnych rovníc dokáže pokryť skoro celú mass action kinetiku s jedinou výnimkou. A tou sú homodiméry a reakcie s nimi spojené. Dôvodom je predchádzajúce obmedzenie multi-afinných funkcií f_1, \dots, f_n s ohľadom na stupeň premenných x_1, \dots, x_n .

Teoreticky sme schopný popísať akýkoľvek biochemický model pomocou pravidiel tejto kinetiky. V skutočnosti, ak sa pokúsime formulovať tieto pravidlá pre rozsiahly model, zistíme, že s narastajúcou veľkosťou rastie komplexita týchto pravidiel a navyše k úplnosti modelu je potrebné poznať veľké množstvo čo najpresnejšie vyčíslených parametrov. Práve tento druhý problém môže byť v niektorých prípadoch experimentálne neriešiteľný. Či už ide o veľké enzymatické komplexy, alebo (o látky s veľmi krátkou existenciou / o veľmi rýchlo degradujúce látky).

Práve preto sa ponúkajú možnosti aproximácie, ktoré nie len zmenšujú systém a tým aj dimenzionalitu matematického modelu, ale zjednodušujú aj výpočtovú zložitosť. Takouto možnosťou je aj aproximácia kvázistacionárneho stavu (Quasi-steady state approximation) [odkaz na Quasi-steady-state.pdf a LectureNotes.pdf]. Napríklad Michaelis-Mentenovej kinetika (viď. 1.6), či obecnější Hillova kinetika (viď. 1.6) a tiež sigmoidálne prepínače publikované na konferencii CAV (Grosu et al. 2011) [], ktoré náš nástroj ponúka. Takto redukované diferenciálne rovnice majú formu racionálnych polynómov, získaných ako lineárna kombinácia týchto regulačných funkcií, medzi ktoré patria aj Heavisideove alebo schodové funkcie []. V skutočnosti výsledný matematický model nie je multi-afinný, ale na druhú stranu je ho možné aproximovať v zmysle po častiach multi-afinného systému. A to tak, že nahradíme všetky regulačné funkcie sústavou vhodných po častiach lineárnych rampových funkcií. Tieto sú definované nasledovne:

$$r^+_{\text{coor}}(x_i, \theta_i, \theta'_i, x_j, x'_j) = \begin{cases} x_j, & \text{pre } x_i < \theta_i, \\ x_j + (x'_j - x_j) \frac{x_i - \theta_i}{\theta'_i - \theta_i}, & \text{pre } \theta_i < x_i < \theta'_i, \\ x'_j, & \text{pre } x_i > \theta'_i. \end{cases} ;$$

$$r^+(x_i, \theta_i, \theta'_i, a, b) = r^+_{\text{coor}}(x_i, \theta_i, \theta'_i, a\theta_i + b, a\theta'_i + b);$$

$$\begin{aligned} \text{kde } i, j &\in \{1, \dots, n\} \text{ a } i \neq j, \\ \theta_i, \theta'_i &\in \mathbb{R}^+, \theta_i < \theta'_i \leq \max_i, \\ a, b &\in \mathbb{R}. \end{aligned}$$

Potom klesajúce rampové funkcie sú definované ako kvantitatívny doplnok rastúcich:

$$\begin{aligned} r^- \text{coor}(x_i, \theta_i, \theta'_i, x_j, x'_j) &= 1 - r^+ \text{coor}(x_i, \theta_i, \theta'_i, x_j, x'_j) \\ r^-(x_i, \theta_i, \theta'_i, a, b) &= 1 - r^+(x_i, \theta_i, \theta'_i, a, b) \end{aligned}$$

Už zmienené regulačné funkcie majú nasledujúce formy:

$$\begin{aligned} \text{hill}^+(x_i, d, \theta_i, a, b) &= a + (b - a) \frac{[x_i]^d}{[\theta_i]^d + [x_i]^d}; \\ \text{hill}^+(x_i, d, \theta_i, a, b) &= 1 - \text{hill}^-(x_i, d, \theta_i, a, b); \end{aligned}$$

$$\begin{aligned} s^+(x_i, e, \theta_i, a, b) &= a + (b - a) \frac{1 + \tanh(e(x_i - \theta_i))}{2}; \\ s^-(x_i, e, \theta_i, a, b) &= 1 - s^+(x_i, e, \theta_i, a, b); \end{aligned}$$

$$\begin{aligned} h^+(x_i, \theta_i, a, b) &= a, \text{ ak } x_i < \theta_i; \text{ b inak;} \\ h^-(x_i, \theta_i, a, b) &= 1 - h^+(x_i, \theta_i, a, b); \end{aligned}$$

kde $\text{hill}^+, \text{hill}^-$ sú funkcie Hillovej kinetiky,

s^+, s^- sú sigmoidálne prepínače,

h^+, h^- sú Heavisideove (schodové) funkcie,

$$\theta_i \in \mathbb{R}^+, \theta_i \leq \max_i,$$

$$i \in \{1, \dots, n\},$$

$$a, b \in \mathbb{R}_0^+,$$

$$e, d \in \mathbb{R}^+.$$

Špeciálnym prípadom je recipročná hodnota sigmoidálnej funkcie:

$$s^+(x_i, e, \theta_i, a, b)^{-1} = s^-(x_i, e, \theta_i + \frac{\ln(\frac{a}{b})}{2e}, b^{-1}, a^{-1}),$$

ktorú označujeme ako $s^+ \text{inv}(x_i, e, \theta_i, a, b)$. Potom klesajúcu recipročnú funkciu označíme obdobne ako doplnok rastúcej:

$$s^- \text{inv}(x_i, e, \theta_i, a, b) = 1 - s^+(x_i, e, \theta_i, a, b).$$

Dôkaz možno nájsť v článku *From cardiac cells to genetic regulatory network* na strane 6 [1].

Teraz už môžeme zdefinovať úplný formát nášho po častiach linearizovaného multi-afinného ODE modelu (ďalej len PMA model z ang. piece-wise multi-affine ODE model). PMA model \mathcal{M} je daný ako $\dot{x} = f(x)$, kde x je stále vektor premenných (x_1, \dots, x_n) , ale $f = (f_1, \dots, f_n) : \mathbb{R}^n \rightarrow \mathbb{R}^n$ je tentokrát vektor po častiach multi-afinných funkcií. Nevyhnutnou súčasťou modelu \mathcal{M} je množina prahových hodnôt (z ang. threshold) $\theta_m^i \in \mathbb{R}^+$ spĺňajúca $\min_i = \theta_1^i < \theta_2^i < \dots < \theta_{\eta_i}^i = \max_i$, kde $i \in \{1, \dots, n\}$, $m \in \{1, \dots, \eta_i\}$ a platí, že $\eta_i \geq 2$.

Uvažujme Ω ako časť modelu \mathcal{M} tak, že $\Omega = \prod_{i=1}^n \{1, \dots, \eta_i - 1\}$. Funkcia $g : \mathbb{R}^n \rightarrow \mathbb{R}^+$ je vtedy po častiach multi-afinná, ak je multi-afinná na každom n -dimenzionálnom intervale $(\theta_{j_1}^1, \theta_{j_1+1}^1) \times \dots \times (\theta_{j_n}^n, \theta_{j_n+1}^n)$, kde $(j_1, \dots, j_n) \in \Omega$ a zároveň $\forall i, 1 \leq i \leq n, j_i < \max_i$. Potom dostávame n -dimenzionálny PMA model pozostávajúci z funkcií f v nasledujúcom tvare:

$$f_i(x) = \sum_{j \in I^+} \rho_i^j(x) - \sum_{j \in I^-} \rho_i^j(x),$$

kde $i \in \{1, \dots, n\}$, I^+ a I^- sú konečné množiny indexov, pre ktoré platí $I^+ \cap I^- \neq \emptyset$ a ρ_i^j je ľubovoľná hodnota z

$$\rho(x_l) = \begin{cases} c, & c \in \mathbb{R} \\ p, & p \in \langle g, h \rangle; g, h \in \mathbb{R} \\ x_k, & k \in \{1, \dots, n\} \\ r^*(x_k, \theta_m^k, \theta_{m+1}^k, a', b'), & m \in \{1, \dots, \eta_k - 1\}; a', b' \in \mathbb{R} \\ r^*_{coor}(x_k, \theta_m^k, \theta_{m+1}^k, x_l, x'_l), & l \in \{1, \dots, n\} \\ s^*(x_k, e, \theta_m^k, a, b), & e \in \mathbb{R}^+; a, b \in \mathbb{R}_0^+ \\ s^*_{inv}(x_k, e, \theta_m^k, a, b), & \\ hill^*(x_k, d, \theta_m^k, a, b), & d \in \mathbb{R}^+ \\ h^*(x_k, \theta_m^k, a, b), & \end{cases},$$

kde $*$ $\in \{+, -\}$. Príklad PMA modelu možno nájsť na obrázku ...

1.6 Michaelis-Mentenovej a Hillova kinetika

[odkaz na wiki a LectureNotes.pdf]

1.7 Abstrakcia

Kapitola 2

Podobné nástroje

2.1 BioDiVinE

2.2 PEPMC

2.3 RoverGene

2.4 ParaSim

Kapitola 3

BioDiVinE 1.0

Kapitola 4

Nové vylepšenia

Kapitola 5

Implementácia

Kapitola 6

Použitie programu

Kapitola 7

Case study

Kapitola 8

Záver

Príloha