

2013-06-23

Biologicky motivované výpočtové modely

Biologicky motivované výpočtové modely

Michal Kušil

mku@un.sk

24.6.2013

Vážená komisia, . . . , chcel by som vám prezentovať moje pokroky v dizertačnej práci.

2013-06-23

Biologicky motivované výpočtové modely

Outline

Prezentáciu začnem prehľadom existujúcich modelov, ktoré sú inšpirované biológiou. Potom budem hovoriť o P systémoch, pretože im som sa najviac venoval. Existuje množstvo variantov, o ktorých niečo poviem v ďalšej časti. Prezentáciu zavřím predstretím plánov na dizertačnú prácu.

2013-06-23

Biologicky motivované výpočtové modely

Prehľad problematiky

Prehľad modelov

Biologicky motivované výpočtové modely

Biologicky motivované výpočtové modely

Modely vznikajú v dvoch oblastiach

• modely biologických javov

• abstraktné informatické modely

Tieto modely majú dvojaké uplatnenie. Jednak v rámci biológie môžu slúžiť ako reálne modely správania sa živých systémov, na ktorých si možnosťami simulácie či verifikácie môžeme overovať správnosť nášho chápania ich biologickej činnosti, robiť virtuálne biologické experimenty. Na druhej strane môžu slúžiť ako nové inšpiratívne výpočtové modely otváraajúce rad teoretických informatických otázok (napr. výpočtová sila) alebo ako modely na popis aj iných ako biologických systémov.

2013-06-23

Biologicky motivované výpočtové modely

Prehľad problematiky

Prehľad modelov

Biologicky motivované výpočtové modely

Biologicky motivované výpočtové modely

• Neurónové siete (od 1943)

• Celulárne automaty (od 1942)

• Evolučné algoritmy (od 1960)

• L systémy (od 1962)

• P systémy (od 1968) [P]

Dlho skúmané modely ako neurónové siete, celulórne automaty, evolučné algoritmy, či L systémy, si už našli svoje uplatnenie v praxi, kým membránové systémy sú ešte len v začiatkoch svojho vývoja.

2013-06-23


Biologicky motivované výpočtové modely

Prehľad problematiky

P systémy

Membránová štruktúra

Membránová štruktúra



Membránové systémy sú inšpirované bunkami. Základom je preto membránová štruktúra, ktorá pozostáva z regiónov, ktoré sú oddelené membránami. Tvori to hierarchickú štruktúru, ktorá sa dá zobraziť ako strom.

2013-06-23

Biologicky motivované výpočtové modely

Prehľad problematiky

P systémy

Obsah membrány

Obsah membrány

• Multimnožina objektov

• $A \subseteq B$

• Pravidlo prepisovania

• $a \mid b \mid c \rightarrow a \mid bc \mid \epsilon$

• $a \mid b \mid c \rightarrow a \mid \epsilon \mid c$

V každej membráne je multimnožina objektov. Objekty predstavujú molekuly, alebo chemické zlúčeniny. Každá membrána má aj množinu prepisovacích pravidiel. Ľavá aj pravá strana pravidiel pozostáva z multimnožiny objektov, pričom ľavá strana nesmie byť prázdna. Posielanie objektov cez membránu sa uskutočňuje tak, že na pravej strane môžu mať objekty špecifikované, či ostanú v aktuálnom regióne, alebo sa pošlú cez membránu von, alebo dnu cez konkrétnu membránu. Pravidlo môže obsahovať špeciálny symbol delta. Po aplikácii takéhoto pravidla sa membrána rozpustí, jej pravidlá zaniknú, a objekty a prípadne membrány sa z vnútra vylejú von.

2013-06-23

Biologicky motivované výpočtové modely

Prehľad problematiky

P systémy

P systém

P systém

P systém definujeme ako:

$(P, \{V, \mu, w, \delta, \dots\}, M_0, R_1, R_2, \dots, R_n)$ kde

• V je abeceda objektov

• μ je membránová štruktúra

• M_0, R_1, \dots, R_n sú počiatočné multimnožiny v membránach

• δ je množina prepisovacích pravidiel v membránach

• δ je prázdna

$R_i \subseteq (V^+)^n \times (V^+)^n \times (V^+)^n \times (V^+)^n$

P systém sa definuje ako konštrukt pozostávajúci z abecedy objektov, membránovej štruktúry, počiatočnej multimnožiny objektov a prepisovacích pravidiel

2013-06-23

Biologicky motivované výpočtové modely

Prehľad problematiky

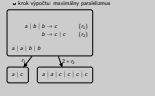
P systémy

Konfigurácia a krok výpočtu

Konfigurácia a krok výpočtu

• Konfigurácia = membránová štruktúra + aktuálny obsah

• krok výpočtu = aplikácia pravidla



Konfigurácia P systému v sebe zahŕňa membránovú štruktúru a obsah jednotlivých membrán. P systém má globálny časovač, v každom kroku, každá membrána aplikuje maximálnu multimnožinu pravidiel. V tomto príklade máme dve jednoduché pravidlá a multimnožinu aabb. Ak sa použije prvé pravidlo, žiadne ďalšie sa už v tomto kroku použiť nemôže, preto je to maximálna multimnožina. Ak sa použije druhé pravidlo, musí sa použiť dvakrát, aby to bola maximálna multimnožina. V tomto príklade sú teda dve maximálne multimnožiny pravidiel. O ďalšej konfigurácii sa rozhodne nedeterministicky.

2013-06-23

Biologicky motivované výpočtové modely

- Přehled problematik
 - P systémy
 - Jazyk
 - výpočetní aparát je matematickým objektem, není
 - jazyk objektivně prodlí čas výpočtu numericky
 - to není otázka o specifické numerice
 - gramotnost je abstraktní míra
 - přehledná zpráva: "PDE"

Ak je postupnosť krokov výpočtu konečná, výpočet je úspešný a výsledok výpočtu je multimnožina objektov, ktorá prešla cez vonkajšiu membránu. Podobne by sa dal definovať aj ako multimnožina objektov, ktorá ostane v špecifickej membráne. V akceptačnom móde sa do špecifickej membrány vloží multimnožina objektov, pričom ostatné membrány sú prázdne. Ak výpočet zastaví, pôvodná multimnožina patrí do jazyka, inak nepatrí.

Pre väčšinu známych modelov sú generatívny aj akceptačný mód rovnako silné, u P systémoch to nie je vždy tak, preto sa oplatí skúmať obidva módy.

Všetky možné úspešné výpočty definujú jazyk nad multimnožinami. Z hľadiska výpočtovej sily sa triedy týchto jazykov porovnávajú s Parikhovým zobrazením známych tried jazykov. P systémom zodpovedá trieda jazykov PsRE, čo je

2013-06-23

Biologicky motivované výpočtové modely

- └ Prehľad problematiky
 - └ Varianty
 - └ Varianty pravidiel
 - Inšpirácia (PIRE) [5]
 - Inšpirácia (PIRE) [5]
 - Inšpirácia
 - 2.2 Inšpirácia (PIRE) [5]
 - 2.1 Inšpirácia (PIRE) [5]
 - 2.1 Inšpirácia (PIRE) [5]
 - Inšpirácia (PIRE) [5]
 - Inšpirácia (PIRE) [5]

Pôvodná definícia definovala kontextové pravidlá. Ak by sme ľavú stranu obmedzili, že v nej môžu byť maximálne 2 objekty, sú to kooperatívne pravidlá. O nich Paun v pôvodnom článku ukázal, že sú tiež univerzálne.

Katalyzátory sa definujú ako podmnožina objektov. Keď sa katalyzátor nachádza na ľavej strane pravidla, musí sa aj na pravej strane. OKrem katalyzátorov sa na ľavej strane nachádza práve jeden objekt, ktorý nie je katalyzátor. Séria článkov skúmala, koľko minimálne katalyzátorov treba na univerzálnosť. Až v roku 2005 Freund ukázal, že stačia 2 a nastolil otvorený problém, ako je to s 1 katalyzátorom. Jediné, na čo sa prišlo, je, že keď sa použijú navše inhibitory, tak je to univerzálne.

Bezkontextové pravidlá majú na ľavej strane práve jeden objekt. Výpočtová sila zodpovedá bezkontextovým jazykom. A veľmi an inhibitory nepomôžu, bude to zodpovedať len špeciálnej triede L

2013-06-23

Biologicky motivované výpočtové modely

- Plány na dizertačnú prácu
 - Aktuálne riešené problémy
 - Aktuálne riešené problémy

Aktuálne riešené problémy

- maximálne parazitizácia (v akom stádiu) hostiteľa
 - akým sa dá rozšíriť reprodukčný model?
- v akom stádiu infekcie
 - zameriava sa parazitizácia?
 - zameriava sa parazitizácia na reprodukciu hostiteľa?
 - akékoľvek? - reprodukčné
 - nie reprodukčné (parazitizácia + eliminácia parazitických zosmarodkov, ...)
 - selektívne + vyčistenie infekcie formálnymi

Maximálny paralelizmus je veľmi silná featurea. Globálny časovač reakcií vo väčšine prípadov tvorí hranicu toho, čo je, a čo nie je univerálne. Ani v bunke sa nenachádza taký časovač, podľa ktorého by sa reakcie synchronizovali. Preto sa hľadajú spôsoby, ako túto featureu odľahčiť, prípadne, akými spôsobmi by sa dal rozšíriť sekvenčný móď, aby bol univerzálny.

Týmto problémom sa v poslednej dobe venujem aj ja.

Ukázalo sa, že existujú spôsoby, ako rozšíriť sekvenčný mód, aby sa stal univerzálnym. Napríklad pravidiel s prioritami, či s vytváraním nových membrán, pričom sa akákoľvek ohraničenie na počet vytvorených membrán nestačí, treba neobmedzené množstvo. Môj výsledok je, že použitie inhibítorov v sekvenčných P systémoch nám tiež zaručí univerzalitu. Článok bol zaslaný na konferenciu, obsahuje dôkaz pomocou simulácie maximálne paralelného P systému.

```

graph TD
    A[Biologicky motivované výpočtové modely] --> B[Plány na dizertačnú prácu]
    A --> C[Nové varianty]
    B --> D[Ďalšie plány]
    B --> E[Nové varianty]
    C --> F[Wright (1986) varianty]
    C --> G[Blaustein (1977) varianty]
    F --> H[variability]
    F --> I[heterozygosity]
    F --> J[heterozygosity]
  
```

Zaujímavú definíciu dobrého variantu uvádza Besozzi vo svojej PhD práci.

Mal by byť dostatočne realistický, vychádzať z reálnych pozorovaní biologických javov.

Mal by být dostatočně silný z výpočtového hlediska, ideálně univerzálný.

A mal by byť iredundantný, čiže definícia by mala byť dostatočne jednoduchá, nemala by obsahovať nič navyše.

2013-06-23

- Biologicky motivované výpočtové modely
 - └─ Přehled problematik
 - └─ Varianty
 - └─ Varianty objektov

Varianty objektov

- varianty objektov [5]
- varianty informálního objektu (v c. modelování)
- informální objekt (IO)
- implementace (IO)

Aby sme mohli definovať priamo jazyky nad stringami, Mate navrhlo variant P systems with worm objects. V membránach sú rovno multimnožiny stringov, čo je inšpirované reťazcami DNA. Takto dostal tiež univerzálny model.

2013-06-23

- Biologicky motivované výpočtové modely
 - Prehľad problematiky
 - Varianty
 - Varianty kroku výpočtu
 - modelový paralelizmus (PARE)
 - schéma (slovo základov pomocou VHDL) [1]
 - schéma (slovo) = schéma [1] [2]
 - modelový paralelizmus (PARE) [2]

Podobne existuje množstvo variantov spôsobu aplikácie pravidiel v jednom kroku výpočtu.

Pri sekvenčnom móde sa v každom kroku nedeterministicky vyberie pravidlo, ktoré sa raz aplikuje. Ibarra ukázal, že sekvenčné P systémy sa dajú simulovať pomocou vector addition systems, čiže nie sú univerzálne.

V asynchrónnom móde sa v každom kroku vyberie ľubovoľná multimnožina pravidiel, ktorá sa aplikuje. Aspoň v jednej membráne táto multimnožina nesmie byť prázdna. Vo väčšine prípadov dáva asynchrónny mód rovnaké výsledky ako sekvenčný. V minimálnom paralelizme pre každú membránu platí, že ak sa môže použiť aspoň jedno pravidlo, tak sa aspoň jedno musí použiť. Je to kôby odvŕhnutie podmienky v maximálnom paralelizme, pričom sa zachováva univerzalita.

Dalšie plány

- Prírodné podmienky špecifikujúce ďalších účastníkov F systému z Mladého výpočtového modelu
 - rozloženie na úseky
 - územný / terénny
 - priestorový F systému
 - ...

Medzi ďalšie plány patrí preskúmanie, ako pomôže kombinovanie s ďalšími variantami z hľadiska výpočtovej sily.

Zaujímavé rozšírenia sú napríklad rozpadajúce sa objekty, kde má každý objekt určený čas rozpadu, a po danom počte krokov objekt zanikne.

Takisto inšpiratívne ja aj použitie energií. Každá reakcia spotrebuje nejakú energiu a po odpálení sa nejaká energia môže naspäť uvoľniť.

Pri symport / antiport pravidlách komunikácia medzi membránami je vždy párova, pričom vždy idú dva objekty buď súbežne, alebo oproti sebe.

Priestorové P systémy zavádzajú pre objekty pozíciu v priestore.

Variantov je veľké množstvo, no stále sa oplatí hľadať aj nové.

```

graph LR
    A[Biologicky motivované výpočtové modely] --> B[Plány na dizertačnú prácu]
    A --> C[Ďalšie plány]
    B --> D[Inšpirácie z výsledkov iných formalizmov]
    E[Inšpirácie z výsledkov iných formalizmov] --> C
  
```

Biologicky motivované výpočtové modely

- Plány na dizertačnú prácu
 - Inšpirácie z výsledkov iných formalizmov
- Ďalšie plány
 - Inšpirácie z výsledkov iných formalizmov

Inšpirácie z výsledkov iných formalizmov

- Protein atlas
 - via to interactions
 - 3 biological data
 - via software Protein atlas
- CIS (Complex of Learning Interactions)
 - statistical model, via analysis of systems [7]
 - reaction system [Brennig, 2007]

Ďalšie plány zahŕňajú aj porovnanie s výsledkami iných formalizmov napríklad Petriho sietí, reaction systems, či CLS. Petriho siete nie sú univerzálne, no niektoré rozšírenia, napríklad s inhibítormi sú. Chcel by som preskúmať aj iné rozšírenia Petriho sietí, ktoré ešte nikto nevyskúšal aplikovať v P systémoch. Petriho siete sú oveľa viac preskúmané ako P systémy, takže sa oplatí nimi inšpirovať.

Zaujímavý formalizmus je CLS, ktorý je sekvenčný, ale napriek tomu vie simulovať max. paralelné P systémy. Výhodou je, že pracuje so stringami. Článok písali Barbuti a Milazzo, ktorý je mojím oponentom.

Okrem Petriho sieťi som dostal odporúčanie pozrieť sa aj na formalizmus Reaction Systems. Zaoberá sa interakciami medzi reakciami, ktoré sa môžu navzájom ovplyvňovať - inhibovať alebo posilňovať.