

Modely sebeskládajících DNA nanostruktur

Vypracoval: Jakub Klemsa

Školitel: Ing. Štěpán Starosta, Ph.D.

Fakulta jaderná a fyzikálně inženýrská
Matematická informatika

19. června 2014

1. DNA vs. Chomského hierarchie
2. Modely založené na Wangovo dláždění
 - Wangovo dláždění
 - aTAM
 - Studované složitosti
 - Turingovská univerzalita aTAMu
 - Meze studovaných složitostí
 - Důsledky
3. Návrh řešení NP problémů
 - Přizpůsobení modelu NP
 - Problém k -kliky
 - Počítačová simulace
 - Další vyřešené problémy

Úvod

Výhody

- paralelizmus – ve zkumavce až 10^{18} „větších“ molekul
- energetická efektivita (Adleman [1])

Úvod

Výhody

- paralelizmus – ve zkumavce až 10^{18} „větších“ molekul
- energetická efektivita (Adleman [1])

Nevýhody

- pravděpodobnostní povaha
- chybovost
- na 10^{18} operací stačí hrubá síla (řádově dny na clusteru s tisíci jader)

Úvod

Výhody

- paralelizmus – ve zkumavce až 10^{18} „větších“ molekul
- energetická efektivita (Adleman [1])

Nevýhody

- pravděpodobnostní povaha
- chybovost
- na 10^{18} operací stačí hrubá síla (řádově dny na clusteru s tisíci jader)

Pole studia

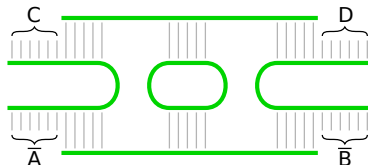
- kinetika reakcí
- **abstraktní modely pohledem matematické informatiky**

Souvislost s Chomského hierarchií (Winfree [5])

Lineární vlákna \leftrightarrow regulární jazyky

Větvená vlákna \leftrightarrow bezkontextové jazyky

Dvojkřížené molekuly \leftrightarrow Turingův stroj (TS)

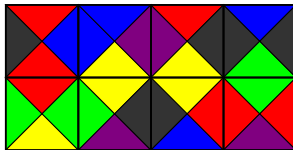


Obrázek: Dvojkřížená molekula, velmi stabilní (Seeman, Fu [4]).

Wangovo dláždění

Čtvercové dláždění roviny (její části), kde

- dlaždice mají na hranách barvu z konečné množiny barev (lepidel)
- jsou orientované (zakázáno rotovat nebo překlápět)
- sousedit smí pouze dlaždice se stejnou barvou na společné hraně



Obrázek: Wangovo dláždění.

aTAM

Rothemund, Winfree rozšířili definici

- každé lepidlo má přidružené přirozené číslo – síla lepidla
- existuje prázdné lepidlo se silou 0, které smí sousedit se všemi
- dláždění se utváří
 - z iniciální dlaždice
 - po jedné dlaždici
 - součet právě připojených lepidel musí být větší nebo roven zadané hodnotě (tzv. *teplota*, ozn. τ)

Studované složitosti I

Biostep complexity $Bs(n)$

- počet laboratorních procedur (popsané v Adleman [2], Winfree [5])
- jedna trvá až desítky minut
- za proveditelné budeme uvažovat pouze $Bs(n) \in O(1)$

Binding complexity $Bnd(n)$

- počet vazeb v koncovém dlážďení
- kvůli rostoucí psti chyby – proveditelné $Bnd(n)$ polynomiální

Studované složitosti II

Tile complexity $Ti(n)$

- počet různých dlaždic
- potřeba je syntetizovat – proveditelné $Ti(n)$ polynomiální

Glue complexity $Gl(n)$

- počet různých lepidel – sekvencí
- dlouhé se mohou vázat chybně – proveditelné $Gl(n)$ polynomiální

Studované složitosti II

Tile complexity $Ti(n)$

- počet různých dlaždic
- potřeba je syntetizovat – proveditelné $Ti(n)$ polynomiální

Glue complexity $Gl(n)$

- počet různých lepidel – sekvencí
- dlouhé se mohou vázat chybně – proveditelné $Gl(n)$ polynomiální

Turingovská univerzalita aTAMu

Tvrzení (Winfree [5])

aTAM je při teplotě $\tau = 2$ Turingovsky univerzální (TU).

- důkaz převodem na celulární automat
- nezjistím nic o spotřebě zdrojů

Vlastní důkaz

- přímočarý
- v práci str. 15-16

Meze studovaných složitostí

Lemma

Studované složitosti v tomto systému jsou omezené:

Biostep. $Bs(n) \in O(1)$.

Binding. $Bnd(n) \in O(s(n) \cdot t(n))$, kde $t(n)$ je čas a $s(n)$ prostor spotřebovaný simulovaným TS.

Tile. $Ti(n) \in O(n)$.

Glue. $Gl(n) \in O(n)$.

Proveditelnost BPP ve 2D při $\tau = 2$

BPP je třída jazyků rozhodnutelných pravděpodobnostním Turingovým strojem (PTS) v polynomiálním čase, považuje se za proveditelnou

Z předchozího lemmatu plyne:

Důsledek

BPP je proveditelná v modelu aTAM ve 2D při $\tau = 2$.

Poznámka

$P, ZPP, RP, \text{co-RP} \subseteq BPP$.

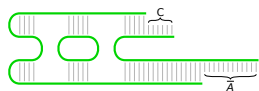
Přizpůsobení aTAMu řešení NP problémů

Odvozen z Winfreeho ukázky řešení problému Hamiltonovské cesty

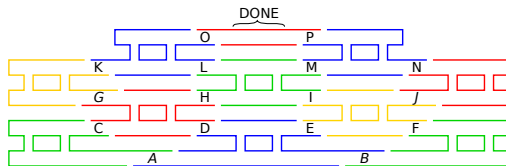
- $\tau = 2$
- 5 dalších typů dlaždic včetně daných sil lepidel
- pevně nastavená počáteční t_0

Poznámka

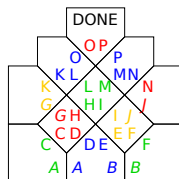
Tento model lze snadno simulovat klasickým aTAMem.



(a) Rohová molekula



(b) Schéma sebeskladu

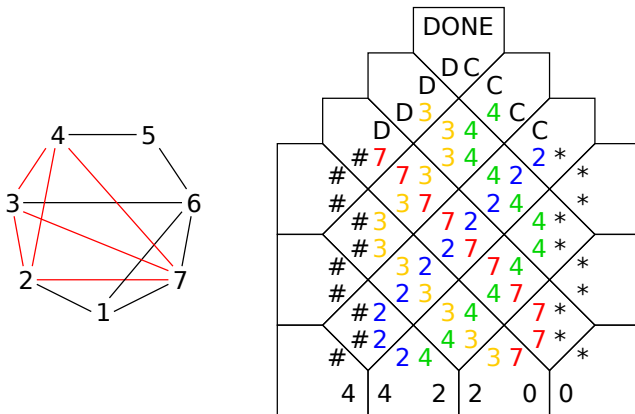


(c) Model

Obrázek: Evoluce modelu od molekul k dlaždicím. Zde lepidla A, B, G a J mají sílu 2, všechna ostatní mají sílu 1.

Problém k -kliky

NP-úplný problém, $Bnd \sim 5/4k^2$, $Ti \sim 2k^2e + 3kn$, $Gl \sim kn$.



Obrázek: Nalezení k -kliky. Řazení barev je dáno jejich vlnovou délkou.

Simulace v xgrow

xgrow je open-source simulátor jak kinetických tak abstraktních modelů.

- skriptem (na CD) k zadanému grafu generuji potřebné dlaždice
- se zapnutím kinetiky není jednoduché dosáhnout bezchybného dlážďení

Další vyřešené problémy

3-obarvení grafu

- NP-úplný problém

Grafový izomorfismus

- domnívá se, že není NP-úplný

Reference I



Leonard M Adleman.

Molecular computation of solutions to combinatorial problems.
Science - New York then Washington, pages 1021–1024, 1994.



Leonard M Adleman.

On constructing a molecular computer.
1995.



Matthew Cook, Yunhui Fu, and Robert Schweller.

Temperature 1 self-assembly: Deterministic assembly in 3d and probabilistic assembly in 2d.
In *Proceedings of the Twenty-Second Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 570–589. SIAM, 2011.



Tsu Ju Fu and Nadrian C Seeman.

Dna double-crossover molecules.
Biochemistry, 32(13):3211–3220, 1993.

Reference II



Erik Winfree.

Algorithmic self-assembly of DNA.

PhD thesis, California Institute of Technology, 1998.