

Oponentský posudok na dizertačnú prácu

RNDr. Petr Medek: **Computations of channels in protein molecules**

Spracoval Mgr. Alexej Kolcun, CSc.

Úvod

Predložená práca Dr. Medka pozostáva z ôsmich kapitol. Logicky možno prácu členiť na tri časti. Prvá obsahuje úvodné štyri kapitoly, kde autor definuje problém (kap. 2), stanovuje ciele (kap.1), popisuje formálny aparát ktorý je použitý (kap. 3) a dáva prehľad súčasných metód riešenia (kap.4). V druhej, nosnej časti práce, sú opísané autorove prístupy k riešeniu – statický prípad (kap.5) a dynamický prípad (kap.6). Tretia časť sumarizuje výsledky (kap.7) a predikuje ďalšie možné postupy (kap.8).

Také členenie považujem za správne, vcelku možno povedať že je obsahovo i objemovo vyvážené.

Dizertačná práca je napísaná v anglickom jazyku.

Dôležitosť práce pre daný obor

Hlavný prínos predloženej dizertačnej práce vidím v jej interdisciplinárnosti: sú tu vyvíjané metódy pre analýzu geometrie proteinových molekúl s cieľom identifikovať „tunely“ v takýchto molekulách. Úloha bola formulovaná v oblasti biochémie a vyriešenie tohto problému je dôležité pre pochopenie mechanizmov, ktoré dovoľujú ovplyvňovať vlastnosti proteinových molekúl. Z tohto pohľadu sa práca javí ako dôležitá. Skutočnú dôležitosť dosiahnutých výsledkov z hľadiska biochémie vzhľadom na svoje vzdelanie neviem kvalifikovane posúdiť. Očakával by som ale, že výsledky by mali byť publikované v spolupráci s odborníkmi v oblasti biochémie – toto som v priloženom zozname prací autora nenašiel.

Z pohľadu informatiky je úloha zaujímavá a netriviálna. Riešenia síce existujú, no ako autor demonštruje, jeho riešenie dáva lepšie výsledky. Ako je vidno zo zoznamu autorových prací, výsledky sú publikovateľné na medzinárodnom fóre. Očakával by som však aspoň jednu samostatnú publikáciu.

Riešenie problému a použité metódy

Pri hľadaní tunelu v bielkovinovej bunke autor vychádza zo zovšeobecnenia Voronoiovského diagramu (VD) – EVD (Euklidovský VD), hrany ktorého obsahujú váhy – vzdialenosti od najbližšej sféry. Tunel sa hľadá ako cesta v grafe s najmenšou sumárnou váhou.

Pre hľadanie cesty s minimálnou váhou je použitý štandardný prístup – Dijkstrov algoritmus. Autorov prínos riešenia spočíva v urýchlení procesu tvorby EVD, resp. generovaní váhovej funkcie, kde na základe dokázaného tvrdenia (Lemma 5.1) autor nemusí konštruovať celý EVD. Za prvý hlavný výsledok práce možno považovať metódu nájdenia tunelu v bielkovinovej molekule pre statický prípad molekuly (kap. 5). Pritom takto nájdené riešenie je lepšie, ako alternatívne prístupy (Tab. 7.1 – 7.9).

Autor rieši i dynamický prípad, tj. úlohu, kde pozície atómov v molekule sa postupne menia. Navrhnuté riešenie spočíva v nezávislej analýze statických prípadov a zhlukovaniu takto nájdených tunelov. Analyzované sú štyri metódy zhlukovania. Je analyzovaná i alternatívna metóda sledovania vývoja jednotlivých tunelov bez zhlukovania. Toto je možné považovať za druhý hlavný výsledok práce – v dostupnej literatúre sa porovnateľné riešenie nevyskytuje.

Výsledky sú implementované v SW Caver 2.0, ktorého spolutvorcom autor dizertácie je. Riešenie je konfrontované s inými prístupmi (kap. 7): Caver 1.0, (predošlá verzia projektu CAVER), MOLE, MolAxis.

Autor taktiež ukazuje, kde sú rezervy súčasného riešenia a navrhuje cesty ďalšieho vylepšenia použitej metódy pre dynamický prípad – využitie koherencie v postupnosti statických riešení.

Pripomienky a poznámky

Formálne pripomienky

1. Hlavnou formálnou pripomienkou je jazyková úroveň – myslím, že predložená práca by si zaslúžila precíznejšiu jazykovú korektúru.
2. Formálny popis algoritmov je nejednotný. Napr. algoritmy v kap. 3.1 sú vysvetlené súvislým textom, Dijkstrov algoritmus a algoritmy v kap. 3.4 sú vyjadrené výčtom jednotlivých krokov, algoritmy vlastného riešenia sú písané „pseudozdrojovým textom“. Práci by prospela jednotnosť vyjadrovania.
3. Predpokladám, že do Tab. 7.10 sa symboly „%“ dostali nedopatrením.

Faktické pripomienky

1. Práca je síce „informatická“, no vzhľadom na riešené problémy ju možno považovať za interdisciplinárnu. Preto si myslím, že by malo byť venované viac priestoru detailnejšiemu popisu geometrie: skutočne nie je možné bližšie charakterizovať vnútornú geometrickú stavbu skúmaných bielkovinových molekúl?
2. Podobne pri analýze dynamického prípadu by som očakával, že existujú akési – vopred známe obmedzenia vzájomného pohybu atómov v molekule.
Možno by toto pomohlo nájsť efektívnejšie riešenia pri hľadaní tunelov. Navyiac, v práci som nenašiel ani tak základnú informáciu, koľko atómov (rádovo) bielkovinová molekula obsahuje.
3. Lemma 5.1 je formulovaná „existenčným“ spôsobom. Nebolo by možné sformulovať silnejšie tvrdenie „konštruktívne“, t.j. nájsť, kde leží bod požadovaných vlastností? (Pritom v kap. 5.2 sú body požadovaných vlastností použité.)
4. Nie som si istý, či možno považovať variáciu veľkosti atómov v rozmedzí 1.2 – 1.85 Å za zanedbateľnú (je to vlastne pomer 2:3).
5. V kap. 5.2 (str. 54) je navrhnuté používať namiesto EDT len DT (viď predchádzajúca pripomienka), resp. RT. Je tak v konečnom dôsledku Lemma 5.1 využitá v autorovom riešení?
6. Bolo by asi vhodné detailnejšie analyzovať vzájomný vzťah DT, EDT a RT. Napr. na vybraných testovacích úlohách s vopred známymi presnými parametrami porovnať riešenia, ktoré dávajú príslušné konštrukcie.
7. Podobne pri porovnaní SW boli použité „reálne“ molekuly s neznámymi vlastnosťami a porovnávali sa riešenia, ktoré dané SW našli. Stálo by za to otestovať SW na modelových situáciách so známymi vlastnosťami (dĺžka a šírka tunelu) a zistiť naokoľko daný SW toto riešenie nájde.
8. Pre zhlukovanie tunelov pri analýze dynamického prípadu autor vychádzal z časopiseckých a konferenčných článkov [52, 21, 9, 42, 33]. Myslím, že by bolo rozumné vychádzať z vhodnej monografie.
9. Z obr. 7.2 vyplýva, že pre riešené úlohy je Dijkstrov algoritmus, aj keď má zložitosť $O(n^2)$ podstatne rýchlejší (zhruba 1:10), ako algoritmus pre DT, ktorý má zložitosť $O(n \log(n))$. Pre aké veľké n je tento pomer výsledných časov zachovaný?

Záverečné zhodnotenie

Predložená práca rieši aktuálny problém z oblasti 3D modelovania, ktorý nachádza praktické uplatnenie v biochémií. Podľa môjho názoru autor preukázal schopnosť tvorivého prístupu pri riešení problému. Práca spĺňa požiadavky štandardne kladené na dizertačné práce.

Prácu doporučujem k obhajobe a navrhujem hodnotenie B-C.

V Ostrave, 28.4.2010

Mgr. Alexej Kolcun, CSc.