

Hodnocení vedoucího závěrečné práce

České vysoké učení technické v Praze

Fakulta informačních technologií

Student: Štěpán Sršeň
Vedoucí práce: Mgr. Jan Starý, Ph.D.
Název práce: Distribuovaná infrastruktura pro virtuální screening molekul
Obor: Informační technologie

Datum vytvoření: 30. 1. 2017

Hodnotící kritérium: 1. Náročnost a další komentář k zadání	Způsob hodnocení - následující škálou 1 až 5: 1=mimořádně náročné zadání, 2=náročnější zadání, 3=průměrně náročné zadání, 4=lehčí, ale ještě dostatečně náročné zadání, 5=nedostatečně náročné zadání
Popis kritéria: Podrobněji charakterizujte diplomovou (bakalářskou) práci a její případné návaznosti na předchozí nebo běžící projekty. Dále posuďte, čím je zadání této ZP náročné. (U obtížnější ZP lze dále tolerovat některé nedostatky, které by u ZP standardní obtížnosti tolerovány nebyly; a naopak u jednoduché ZP mohou být zjištěné nedostatky hodnoceny přísněji.) Komentář: Jedná se o kombinaci dvou oborů: kromě systémové práce s distribuovaným výpočetním clusterem (resp skriptování nad ním) je potřeba porozumět jistým partiím molekulární chemie, o kterou "vevnitř" jde.	
Hodnotící kritérium: 2. Splnění zadání	Způsob hodnocení - následující škálou 1 až 4: 1=zadání splněno, 2=zadání splněno s menšími výhradami, 3=zadání splněno s většími výhradami, 4=zadání nesplněno
Popis kritéria: Posuďte, zda předložená ZP splňuje zadání. V komentáři uveďte body zadání, které nebyly zcela splněny, případně rozšíření ZP oproti původnímu zadání. Nebylo-li zadání zcela splněno, pokuste se posoudit závažnost, dopady a případně i příčiny jednotlivých nedostatků. Komentář: Práce splňuje všechny body zadání, některé lépe, některé hůře.	
Hodnotící kritérium: 3. Rozsah písemné zprávy	Způsob hodnocení - následující škálou 1 až 4: 1=splňuje požadavky, 2=splňuje požadavky s menšími výhradami, 3=splňuje požadavky s většími výhradami, 4=nesplňuje požadavky
Popis kritéria: Zhodnoťte přiměřenost rozsahu předložené ZP vzhledem k obsahu, tj. zda všechny části ZP jsou informačně bohaté a ZP neobsahuje zbytečné části. Komentář: Rozsah práce je odpovídající. Kapitoly 2 a 3 se svou "chemickou" povahou vymykají zbytku textu, nicméně nutné minimum chemie a matematického aparátu je tu podáno zřejmě nejstručnějším možným způsobem.	
Hodnotící kritérium: 4. Věcná a logická úroveň práce	Způsob hodnocení - bodové hodnocení 0 až 100 bodů (známka A až F): 89 (B)
Popis kritéria: Posuďte, zda předložená ZP je po věcné stránce v pořádku, případně vyskytují-li se v práci věcné chyby nebo nepřesnosti. Zhodnoťte dále logickou strukturu ZP, návaznosti jednotlivých kapitol a pochopitelnost textu pro čtenáře.	

Komentář:

Práce je sepsána pečlivě. Po nutném molekulárním úvodu dává přehled nepoužívanějších formátů pro strukturu molekul a kvantových výpočetních programů. Pro vlastní použití pak volí SDF databázi Pubchem, konverzi do XYZ formátu pomocí OpenBabel, a chemický software ORCA. To vše jsou průmyslové standardy.

Formáty jsou ilustrovány na jednoduché molekule metanu.

V sekci 4.4 bych očekával více konkrétních softwarových balíčků a přehled práce s nimi.

Kapitola Návrh popisuje logické rozčlenění vytvořeného VMS (virtual molecule screener) na tři části: parser requestů a příprava vstupu, vlastní práce s výpočetním clusterem ORCA, a zpracování výsledků.

Důraz je kladen na fotochemické vlastnosti materiálů, ale zastřešující architektura je obecná a lze ji

bezbolestně rozšířit o jakýkoli další typ výpočtů, které ORCA podporuje.

Požadavky uživatele se formulují ve vlastním jednoduchém formálním jazyce, pro který je implementován parser (Například: "absorb 400 500 0 10" je příkaz pro nalezení molekul, které pohlcují nanejvýš 10% světla o vlnových délkách mezi 400 a 500 nm.)

Vlastní řídicí script je napsán v shellu, což je přirozená volba (spouštíme existující programy, připravujeme jim textový vstup, zpracováváme a archivujeme výstupy).

Shellem se rozumí standardní sh(1), výsledný kód je velmi přenositelný.

Výsledky ORCA výpočtů jsou pak zpracovány programem, který implementuje "chemické porozumění" těmto výsledkům (excitované stavy molekul, absorpce a emise světla). Výstup jednoho výpočtu může být zároveň vstupem pro další výpočet (podrobnější, náročnější, s přesnějšími ale dražšími metodami).

Zbytek kapitoly 5 popisuje podporované typy výpočtů a způsob jak je zadávat.

Kapitola 6 popisuje detaily implementace, použité vnitřní struktury a jejich zpracování.

Každý zdrojový soubor je okomentován, v textu jsou uvedeny stručné přehledy všech funkcí.

Některé samozřejmosti jako přidání atomu do seznamu by zřejmě bylo možné vypustit;

na druhou stranu takto je soupis kompletní.

Kapitola Testování popisuje měření výpočetní náročnosti v závislosti na různých parametrech

(distribuce na různý počet jader, velikost povolené paměti, použité ORCA metody, počet molekul, atd).

Ukazuje se (nepřekvapivě), že režie vytvořeného software je zanedbatelná vůči samotným chemickým výpočtům,

s jednou výjimkou: při velkých počtech molekul tato režie roste při použití přepínače -w (watch), který zapíná monitoring.

Naivní grep všech logů dobehých výpočtů může již při tisících molekul znamenat měřitelné zpomalení.

Tuto implementační chybu autor měřením odhalil, je v shellu opravitelná za cenu pečlivější práce

s množinou naforkovanými podvýpočtů. Bez monitoringu (-w) se tato chyba neprojevuje.

Podmínky testování jsou sepsány dosti podrobně, přesto však chybí některé detaily:

není-li počet molekul relevantní, měří se na "malém počtu" molekul (jakém počtu, které to jsou, je někde jejich seznam?).

Většina testů je na jednotkách tisíc molekul - v praxi to bude řádově více,

ale takové testy není možné v rozumném čase provést.

Hodnotící kritérium:

Způsob hodnocení - bodové hodnocení 0 až 100 bodů (známka A až F):

5. Formální úroveň práce

85 (B)

Popis kritéria:

Posuďte správnost používání formálních zápisů obsažených v práci. Posuďte typografickou a jazykovou stránku ZP, viz Směrnice děkana č. 14/2015, článek 3.

Komentář:

Práce je psána kultivovanou češtinou a dosti úspěšně se vyhýbá obvyklému ubíječím stylu technických textů.

Rovněž počet překlepů a gramatických nedostatků se zdá být nižší než je obvyklé.

Český i anglický abstrakt by se jistě vešly na jednu stránku, teď čtyři řádky přetékají; stejně tak Obsah;

seznamy implementovaných funkcí by zřejmě šly vysázet přitažlivěji; jiné typografické nedostatky jsem neodhalil.

Hodnotící kritérium:

Způsob hodnocení - bodové hodnocení 0 až 100 bodů (známka A až F):

6. Práce se zdroji

70 (C)

Popis kritéria:

Vyjádřete se k aktivitě studenta při získávání a využívání studijních materiálů k řešení ZP. Charakterizujte výběr studijních pramenů. Posuďte, zda student využil všechny relevantní zdroje nebo zda se pokoušel řešit již vyřešené problémy. Ověřte, zda jsou všechny převzaté prvky řádně odlišeny od vlastních výsledků a úvah, zda nedošlo k porušení citační etiky a zda jsou bibliografické citace úplné a v souladu s citačními zvyklostmi a normami.

Komentář:

Po formální stránce jsou citace v pořádku. Práce nicméně obsahuje 54 referencí - mnohé z těch,

kteřé popisují detaily chemických výpočtů, by se myslím daly bez újmy vypustit.

Hodnotící kritérium:

Způsob hodnocení - bodové hodnocení 0 až 100 bodů (známka A až F):

7. Hodnocení výsledků, publikační výstupy a ocenění

90 (A)

Popis kritéria:

Vyjádřete se k úrovni dosažených hlavních výsledků ZP, např. k úrovni teoretických výsledků, nebo k úrovni a funkčnosti technického nebo programového vytvořeného řešení, apod. Případně také zhodnoťte, zda software nebo zdrojové texty, které nevytvořil sám student, byly v ZP použity v souladu s licenčními podmínkami a autorským právem. Popište případnou publikační činnost a získaná ocenění související s řešením této ZP.

Komentář:

Výsledkem je funkční software použitelný v chemické praxi, ostatně to bylo původní motivací. Návrh je dostatečně obecný, aby umožnil rozšíření o další typy výpočtů bez zásahu do celkové architektury. Kód je přenositelný, relativně čistý, a až na výjimky dobře členěný. Všechny vytvořené programy mají svou manuálovou stránku - v některých detailech je tato dokumentace pozadu za funkcionalitou.

Hodnotící kritérium:

Způsob hodnocení - nehodnotí se

8. Komentář o využitelnosti výsledků

Popis kritéria:

Uvedte, zda hlavní výsledky ZP rozšiřují již publikované známé výsledky a/nebo přinášející zcela nové poznatky. Uvedte možnosti využití výsledků ZP v praxi.

Komentář:

Software bude rutinně používán na VŠCHT, tak byl od začátku zamýšlen. Komerčním produktům na poli chemických výpočtů (farmacie atd) nemůže konkurovat, ale pro svou zvolenou oblast (fotochemie materiálů) je adekvátní.

Hodnotící kritérium:

Způsob hodnocení - následující škálou 1 až 5:

9. Aktivita a samostatnost studenta v průběhu řešení

9a:

1=výborná aktivita,
2=velmi dobrá aktivita,
3=průměrná aktivita,
4=slabší, ale ještě dostatečná aktivita,
5=nedostatečná aktivita

9b:

1=výborná samostatnost,
2=velmi dobrá samostatnost,
3=průměrná samostatnost,
4=slabší, ale ještě dostatečná samostatnost,
5=nedostatečná samostatnost

Popis kritéria:

Posuďte, zda byl student během řešení aktivní, zda dodržoval dohodnuté termíny, jestli své řešení průběžně konzultoval a zda byl na konzultace dostatečně připraven (9a). Posuďte schopnost studenta samostatné tvůrčí práce (9b).

Komentář:

Po úvodním rozběhu pracoval řešitel zcela samostatně.

Hodnotící kritérium:

Způsob hodnocení - bodové hodnocení 0 až 100 bodů (známka A až F):

10. Celkové hodnocení

89 (B)

Popis kritéria:

Shrňte stránky ZP studenta, které nejvíce ovlivnily Vaše celkové hodnocení. Celkové hodnocení **nemusí** být aritmetickým průměrem či jinou hodnotou vypočtenou z hodnocení v předchozích jednotlivých kritériích 1 až 9.

Text hodnocení:

Text i vytvořený software je nad průměrem bakalářských prací, bodové hodnocení volím záměrně na hraně.

Podpis vedoucího práce: