

## POSUDEK NA DIPLOMOVOU PRÁCI R. VITOVJÁKA 'DYNAMICKÉ HAŠOVACÍ TABULKY'.

Cílem předložené práce bylo zhodnotit metody používající dynamické hašovací postupy. Splnění tohoto cíle vyžaduje napsat programy, které realizují tyto dynamické metody (konkrétně lineární hašování a spirální hašování), vytvořit pro ně testovací prostředí a provést a vyhodnotit množství experimentů s nimi. Navíc je potřeba porovnat efektivitu těchto metod se standardním postupem, tj. hašováním se separovanými řetězci. Tento velmi ambiciózní program asi nelze beze zbytku splnit a jeho rozsah výrazně překračuje množství práce a výsledků očekávaných od standardní diplomové práce. Tato práce, i když plně nesplnila svůj cíl, tak obsahuje kvantum výsledků, které odpovídají dobré diplomové práci a jsou prezentovány na odpovídající úrovni. Uvedu podrobnější rozbor.

Předložená diplomová práce má neformálně řečeno dvě odlišné části. V první z nich je dán popis studovaných metod a hrubý teoretický rozbor jejich efektivity. Tato část práce je přehledového charakteru a prezentuje výsledky získané z literatury. Zde bych si přál podrobnější teoretický rozbor kombinovaný s porovnáním s jinými metodami. Jeho cílem by bylo nalezení teoretických podmínek, za jakých je výhodné použít studované metody. Toto mi chybí a vnučuje se mi tak otázka po smyslu těchto metod. Na druhé straně mám pocit, že splnění tohoto požadavku je na okraji našich možností (a mimo autorův zájem). Předložené výsledky ukazují, že tyto metody mají své opodstatnění, jen když faktor naplnění je větší než 1 (autor v experimentech pracuje s faktorem naplnění kolem 5). Vnučuje se mi tak otázka, zda efektivita těchto metod není přímo úměrná velikosti faktoru naplnění. Naproti tomu stojí otázka po praktickém smyslu použití standardních metod s velkým faktorem naplnění. Protože ale tato část je přehledová, směřují tyto otázky spíše k autorům studovaných metod než k diplomantovi. Jinak tato část práce je dobře a čitelně napsaná a obsahuje jen málo nedostatků. Zmíním se o několika důležitějších. Výraz pro  $U(\alpha, x)$  na str. 14 není podle mého názoru v pořádku (myslím, že tam přičtení 1 nemá být). Autor při popisu spirálního hašování používá funkce  $f$  a  $g$ , které nezavedl. Je pravda, že z kontextu lze jejich význam pochopit, ale může to vést u čtenáře k nepochopení. Autor by měl věnovat větší pozornost zavedení aproximačních metod pro spirální hašování, měl by alespoň trochu zdůvodnit použitý tvar aproximace.

Druhá část textu popisuje testovací prostředí a provedené experimenty a diskutuje výsledky a závěry těchto experimentů. Oceňuji, že některé závěry autor použil při návrhu programů realizujících vyšetřované metody. Vlastní výsledky experimentů jsou prezentovány na příloženém CD. Tato část je vlastním obsahem autorovy práce. Získané výsledky přesvědčivě ukazují, že za vyšetřované situace je výhodné používat snadno spočítatelnou hašovací funkci (největší nevýhoda spirálního hašování). Prezentované výsledky také ukazují, že za vyšetřované situace není jasné, proč používat dynamické metody hašování (standardní metoda vypadá, že je efektivnější). Avšak na úplné zhodnocení těchto metod je potřeba daleko větší množství experimentů prováděných za různých parametrů specifikujících vyšetřovanou situaci. Předložená práce tak motivuje několik otázek, které je třeba zodpovědět před konečným zhodnocením vyšetřovaných metod. Jako příklad uvedu některé z nich: Jak závisí efektivita vyšetřovaných metod na velikosti faktoru naplnění? Mohou být vyšetřované metody efektivnější než klasické při dostatečně velkém faktoru naplnění? Jak závisí efektivita vyšetřovaných metod na rozdělení vstupních dat? Nezmění se hodnocení, když použijeme jiný druh paměti než je interní paměť? Tyto otázky nemůže autor zodpovědět bez další práce, která svou velikostí možná i hodně překračuje množství

práce vložené do tohoto textu. Některé tyto otázky už položil sám v závěrečné části práce.

Zhodnocení: Předložená práce se zabývá zajímavou tematikou. Presentované výsledky motivují řadu dalších otázek a úroveň jejich prezentace je dobrá. Plné zhodnocení vyšetřovaných metod vyžaduje zodpovězení řady otázek, z nichž některé inicializovala tato práce (tedy cíl práce nebyl a ani nemohl být beze zbytku splněn). Pro tato fakta **doporučuji uznat tuto práci jako diplomovou.**



doc. RNDr. Václav Koubek, DrSc.