

Posudek bakalářské práce

Matematicko-fyzikální fakulta Univerzity Karlovy v Praze

Autor práce Marek Polák
Název práce Využití celulárních automatů pro kompresi dat
Rok odevzdání 2016
Studijní program Informatika **Studijní obor** Programování a softwarové systémy

Autor posudku Otakar Trunda **Role** Vedoucí
Pracoviště KTIML

Prosím vyplňte hodnocení křížkem u každého kritéria. Hodnocení *OK* označuje práci, která kritérium vhodným způsobem splňuje. Hodnocení *lepší* a *horší* označují splnění nad a pod rámec obvyklý pro bakalářskou práci, hodnocení *nevyhovuje* označuje práci, která by neměla být obhájena. Hodnocení v případě potřeby doplňte komentářem. Komentář prosím doplňte všude, kde je hodnocení jiné než *OK*.

K celé práci

	lepší	OK	horší	nevyhovuje
Obtížnost zadání	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Splnění zadání	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Rozsah práce ... <i>textová i implementační část, zohlednění náročnosti</i>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<p>Komentář Práce se zabývá netriviální problematikou celulárních automatů a řešitel musel prokázat také dobré znalosti v oblasti komprese dat, a to jak teoretické, tak praktické. Práce si klade poměrně ambiciózní cíl: vylepšit některý ze současně používaných kompresních algoritmů. Z těchto důvodů lze obtížnost zadání považovat za nadstandardní.</p> <p>Zadání bylo splněné v dostatečné míře. Řešitel prokázal schopnost samostatně zpracovat složitější problém, formálně definovat potřebné pojmy a správně s nimi pracovat, orientovat se v odborné literatuře a srozumitelně prezentovat svoje závěry.</p> <p>Rozsahem textu a zdrojových kódů práce mírně převyšuje standard obvyklý pro bakalářskou práci. Kladně lze hodnotit také rozsáhlé a výpočetně náročné experimenty.</p>				

Textová část práce	lepší	OK	horší	nevyhovuje
Formální úprava ... <i>jazyková úroveň, typografická úroveň, citace</i>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Struktura textu ... <i>kontext, cíle, analýza, návrh, vyhodnocení, úroveň detailu</i>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Analýza	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Vývojová dokumentace	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Uživatelská dokumentace	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<p>Komentář Práce je psaná česky a má dobrou jazykovou úroveň. Text je bez viditelných pravopisných či stylistických chyb a citované práce jsou dobře zvolené a relevantní k danému tématu.</p> <p>Text je členěný do pěti kapitol, které na sebe logicky navazují. Většina pasáží je dobře srozumitelná, občas se však vyskytují složité slovní popisy některých pojmů, například u definice „zpětného koeficientu“, popisu omega kódu, i jinde. Slovní popis je sice správný a přesný, ale k lepšímu pochopení jeho významu by pomohlo přidat formální definici pomocí matematických vzorců. Obecně lze říct, že autor v celé práci upřednostňuje spíše slovní popisy před matematickými formulami.</p> <p>Na některých místech by pomohlo také větší množství grafických příkladů či ukázek. Například u pojmů „Trace and backtrack“ a „BiEntropy“.</p> <p>Kapitola „Analýza“ podává pouze úvod do problematiky a vymezení některých pojmů. Skutečná analýza problému je provedená až v kapitole „Návrh“, kde se například srovnává použití dopředného a zpětného běhu automatu a jejich výhody a nevýhody vzhledem ke kompresi dat.</p> <p>Autor zvolil konkrétní způsob použití CA pro kompresi a tento způsob dobře prozkoumal. Do kategorie „analýza“ lze řadit i první sadu experimentů, která zkoumá vlastnosti různých elementárních CA, které jsou relevantní vzhledem ke kompresi a pozitivně lze hodnotit také snahu autora o interpretaci a hlubší pochopení získaných výsledků. V tomto ohledu je analýza provedená rozsáhle a pečlivě a svou metodikou se blíží skutečné vědecké práci.</p> <p>Slabší stránkou je naopak nedostatečné zhodnocení některých alternativ. Například: autor bez přesvědčivého zdůvodnění zvolil pouze elementární CA a o jiných typech CA se zmiňuje jen okrajově. Dále: popsání problémů s dopředným a zpětným chodem automatu by bylo možné řešit použitím reversibilních automatů, které v práci nejsou vůbec zmiňované.</p> <p>Další nedostatek spatřuji v nepřilíš důsledné teoretické analýze problému, zejména co se týče vlivu hodnot některých parametrů na celkovou efektivitu. Autor sice zmiňuje, že například „s rostoucí velikostí automatu roste jak paměťová, tak především časová složitost algoritmu“, nebo „Zdá se, že kompresní poměr se zlepšuje s rostoucí velikostí bloku“ a podobně, ale není poskytnutá žádná matematická formule vyjadřující tyto závislosti. V ideálním případě by měla být jasně formulovaná složitost algoritmu v závislosti na velikosti bloku a počtu generací, mělo by být popsáno, kolik bitů je zapotřebí pro zápis čísla generace a jak přesně závisí kompresní poměr na počtu generací a velikosti bloku. Na základě těchto vztahů by pak mělo být možné lépe odhadnout vhodné kombinace parametrů algoritmu, například: pokud máme k dispozici omezený čas, je lepší použít malý počet generací a větší velikost bloku, nebo větší počet generací a malou velikost bloku? Jaká je nejefektivnější kombinace? A podobně. Kompletní řešení tohoto problému jde sice daleko nad rámec bakalářské práce, bylo by však vhodné některé z těchto postupů v práci aspoň naznačit nebo se o ně pokusit.</p> <p>Závěrečné experimenty na reálných datech jsou prováděné pouze na dvou souborech. Pro získání relevantních výsledků je toto příliš malý vzorek.</p> <p>Uživatelská dokumentace je poskytnutá v podobě přiloženého souboru. Dokumentace je</p>				

smysluplná a úplná a poskytuje všechny informace potřebné k plnohodnotnému používání obou přiložených aplikací.
 Vývojová dokumentace však poskytnutá není, a to ani formou komentářů ve zdrojových kódech.

Implementační část práce

	lepší	OK	horší	nevyhovuje
Kvalita návrhu ... <i>architektura, struktury a algoritmy, použité technologie</i>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Kvalita zpracování ... <i>jmenné konvence, formátování, komentáře, testování</i>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Stabilita implementace	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Komentář Navržený software má rozumnou architekturu a implementace ctí všechna základní paradigmatu objektově orientovaného programování. Z hlediska softwarového inženýrství lze doporučit ještě přidání unit testů a používání standardizovaných návrhových vzorů. Hlavním problémem zpracování je však absence jakýchkoli komentářů a to i u hlaviček veřejných metod.

Implementace je stabilní, pro lepší praktickou použitelnost doporučuji ještě přidat ukazatel průběhu (některé činnosti programu trvají dlouho a uživatel neví, jestli bude čekat 10 sekund nebo 10 minut) a přepracovat způsob zadávání počátečního stavu automatu – v současné podobě je potřeba ručně zadat přesný počet symbolů, což je pro 64 nebo 128 znaků nepraktické.

Celkové hodnocení Výborně

Práci navrhuji na zvláštní ocenění Ne

Datum 26. srpna 2016

Podpis