

Posudek bakalářské práce

Matematicko-fyzikální fakulta Univerzity Karlovy

Autor práce	Viktória Brezinová	
Název práce	Automatic Extraction of the Main Characters from Books and Their Interactions	
Rok odevzdání	2022	
Studijní program	Informatika	
Studijní obor	Umělá inteligence	
Autor posudku	Mgr. Rudolf Rosa, Ph.D.	Oponent
Pracoviště	Ústav formální a aplikované lingvistiky	

K celé práci

lepší OK horší nevyhovuje

	lepší	OK	horší	nevyhovuje
Obtížnost zadání	X	X		
Splnění zadání		X		
Rozsah práce <i>... textová i implementační část, zohlednění náročnosti</i>		X		
<p>Zadání specifikuje několik netriviálních podúloh, které bylo potřeba v rámci práce vyřešit a vhodně vzájemně provázat. Zadání se však podařilo kompletně splnit – analyzovány jsou souvškyty postav i jejich interakce, řeší se sjednocení jmen postav i koreference. Celkově je rozsah práce z implementačního pohledu spíše podprůměrný, avšak z výzkumného pohledu naprůměrný.</p>				

Textová část práce

lepší OK horší nevyhovuje

	lepší	OK	horší	nevyhovuje
Formální úprava <i>... jazyková úroveň, typografická úroveň, citace</i>		X		
Struktura textu <i>... kontext, cíle, analýza, návrh, vyhodnocení, úroveň detailu</i>		X		
Analýza	X			
Vývojová dokumentace		X		
Uživatelská dokumentace		X		
<p>Text má dobrou úroveň, přes četné drobné chyby v angličtině je bez problémů srozumitelný a dostatečně podrobně popisuje celý postup.</p> <p>Velmi oceňuji kvalitní rozsáhlé rešerše, kdy pro každou podúlohu studentka nejprve prostudovala několik relevantních odborných prací a následně navrhla vlastní postup vycházející z těchto prací ale uzpůsobený potřebám řešené úlohy. Taktéž bych chtěl vyzdvihnout evaluaci, která je důsledná, detailní, pečlivě vyhodnocuje všechny části práce a zahrnuje například i chybovou analýzu. V těchto aspektech považuji tuto práci za nadprůměrně kvalitní.</p> <p>Zvolený postup je klasický, využívající automatické analýzy pomocí externích nástrojů, jádrem metody je pravidlové zpracování, podpořené několika ručně sestavenými slovníky. Pravidla jsou v textu detailně popsána, ale jde o text, který se nečte příliš dobře; vhodné by pro větší srozumitelnost bylo popis podpořit vizualizací těchto pravidel a/nebo konkrétními příklady. Zajímavým problémem se ukázala být otázka definice interakce postav, zvolenou definici považuji za přijatelnou a dobře zdůvodněnou.</p>				

Implementační část práce

lepší OK horší nevyhovuje

Kvalita návrhu ... architektura, struktury a algoritmy, použité technologie		X	X	
Kvalita zpracování ... jmenné konvence, formátování, komentáře, testování		X		
Stabilita implementace	X			

Program lze dle přiloženého návodu a s využitím instalačních skriptů snadno nainstalovat. Jeho použití je taktéž relativně snadné, rozhraní je velmi jednoduché a dostatečně popsáno v návodu, přiloženy jsou i ukázkové skripty, které program přímo spustí s výchozím/ukázkovým nastavením. Program funguje v souladu s popisem, nepozoroval jsem žádné problémy.

Zdrojové kódy mají asi 1400 LOC (88kB), kód je dostatečně vysvětlený jak pomocí komentářů v kódu, tak v textu práce, je přehledný a relativně srozumitelný.

Výhrady mám ke zvolené vizualizaci. Pro výskyty/souvýskyty je vizualizace dobrá: interpretace grafu je zřejmá, po kliknutí na bod v grafu se zobrazí dané místo v knize izolovaně a v rámci celého textu, a je zvýrazněno jméno postavy. Graf interakce postav je ale méně srozumitelný, interakce postav vlastně není přímo zobrazena, lze ji jen nepřímou vyznačit z rovnoběžných úseček či částí úseček odpovídajících jednotlivým postavám. Po kliknutí na bod v grafu se pak pouze zobrazí daný úsek textu, ale není zde nic vyznačeno – ani hranice dialogu, ani jednotlivé postavy. Věřím, že by vizualizaci zlepšilo zobrazení hranic dialogu, možná i nějaké barevné rozlišení; jakýmsi možným ideálem pak je například *XKCD Movie Narrative Charts* (<https://xkcd.com/657/>), kde použitá vizualizace není o mnoho komplexnější než vizualizace v bakalářské práci, avšak je výrazně přehlednější (vizualizace v BC práci je srovnatelná pouze v případě filmu *12 Angry Men*). Ohledně grafů přímo v textu práce pak nerozumím, proč ve Fig 4.6 vidíme interakci postav, které se ale v daném úseku podle Fig 4.5 vůbec nevyskytují, to mi připadá podivné.

Celkové hodnocení Výborně
Práci navrhuji na zvláštní ocenění Ne

Datum

Podpis