

I. IDENTIFIKAČNÍ ÚDAJE

Název práce:

Semantic Parsing of Questions

Jméno autora:

Martin Matulík

Typ práce:

Vedoucí:

Mgr. Petr Baudiš, katedra kybernetiky

Student měl za cíl prozkoumat oblast semantického parsování otázek formulovaných přirozenou řečí, reimplementovat nedávno publikovaný článek [XFZ14], který tuto úlohu řeší ve spojení s RDF znalostními bázemi novým způsobem, srovnat výsledky a vyvodit z nich závěry pro další výzkum.

Tento úkol se v podstatě podařilo splnit. Přitom je třeba zdůraznit, že zadání se ukázalo jako nadstandardně obtížné, protože ve vybraném článku student objevil řadu podstatných mezer, které vyžadovaly vlastní improvizaci a hledání alternativních cest k cíli. Podařilo se vytvořit open source systém, který provádí semantické parsování na základě strojového učení z anotovaného datasetu otázek (postupem inspirovaným [XFZ14]). Dalším přínosem systému je navíc rozšířená per-token anotace standardního akademického datasetu Free917.

Co se týče experimentů, výsledky měření sice nejsou příliš povzbudivé, student však provádí jejich základní analýzu a identifikuje hlavní a pochopitelné příčiny problému. Je trochu škoda, že se nepovedlo změřit přesnost systému na dalších datasetech, např. na QALD-4 (kde reportuje výsledky i [XFZ14]). Více mi je však líto nedotažených experimentů v rámci analýzy výsledků, které by umožňovaly důkladnější rozbor – např. student identifikuje problém se složitějšími otázkami na kombinace entit a navrhuje i možnou kategorizaci otázek (kterou aplikuje na trénovací množinu). Ale jak se liší výsledky podle daného rozdělení na testovací sadě? Kolik otázek je vlastně těch jednodušších? A proč konkr. přináší menší trénovací množina zlepšení?

Vyšší obtížnost zadání a pracnost doladování pipeline měla jistě svůj podíl na ne zcela optimálním rozvržení času v závěrečné fázi práce, proto je na ní bohužel patrné, že byla dokončována ve spěchu. Velmi stručný text není z mého hlediska na závadu v rešeršních a implementačních kapitolách, kde se čtenář nakonec dozví vše podstatné. Práci však chybí lepší zarámování problematiky v úvodu či perexech jednotlivých kapitol – když čtenář narazí na kapitolu o shift-reduce parsingu či strukturovaném perceptronu, měl by mít obecnou představu, co bude z hlediska práce výstupem algoritmu, místo aby se to dozvěděl až o pár vět až stránek později.

Nejvíce je však stručnost na škodu právě při analýze výsledků a v závěru, kde chybí odpověď na mnoho otázek, které si čtenář klade. Přitom třeba celkový přínos práce by šel snadno vyzdvihnout poukazem na výrazně rychlejší odpovídání na jednoduché otázky oproti jiným systémům a zdůrazněním, že na rozdíl od jiných prací zde pokládáme baseline pro systém, který pro mapování slov na relace nepoužívá žádné externí datové zdroje (jako např. další sady otázek či encyklopedické texty). Je nutné přičíst k dobru, že analýza identifikuje dva hlavní důvody nízké přesnosti (problém s mapováním relací a problém s kombinací více entit), ale chybí dodat i druhou stranu mince, tedy jaké cesty slibují překonání těchto problémů a jejich diskuze v kontextu původního článku [XFZ14].

Z formálního hlediska se práci rozhodně nedá mnoho vytknout. Angličtina není dokonalá, občas chybí např. předložka či čárka, text je však srozumitelný a odvahu studenta při volbě anglického jazyka je však přesto třeba vyzdvihnout. Student řádně cituje použité zdroje. Grafická úprava je v pořádku, obočí zdvihne jen občasné použití symbolu palců místo uvozovek a tag „ctuthesis“ v patičce každé stránky.

Student na práci během uplynulých semestrů pracoval pilně a vesměs pravidelně svůj postup konzultoval. Z technického hlediska práce přinesla užitečný výsledek. Nedostatkem je opravdu příliš stručný text - scházející hlubší analýza výsledků a lepší extrapolace pozorování do konkrétních závěrů a poučení. Proto navrhuji práci ohodnit jako **B - velmi dobře**.

Datum:

Podpis: