

# Oponentský posudek diplomové práce

## Autor a název předložené práce

Adéla Limburská: *Semantic information from FrameNet and the possibility of its transfer to Czech data*

---

## Téma práce

Předložená práce se zabývá automatickým zpracováním sémantické informace obsažené v databázi FrameNet. Explicitně zadaným úkolem studentky bylo navrhnout, implementovat a podrobně evaluovat automatický systém, který využije databázi FrameNet a extrahovanou informaci o anglických sémantických rámcích aplikuje pro automatické rozpoznání sémantické struktury v českých větách, přičemž mají být automaticky rozpoznány jak sémantické rámce, tak jejich elementy. Předpokládalo se při tom využití existujících nástrojů pro automatickou analýzu jak anglických, tak českých jazykových dat.

Na základě vlastních zkušeností s podobnými projekty považuji toto zadání za dosti obtížné. Kritickým předpokladem pro úspěšnost metod strojového učení, které zde připadají do úvahy, bývá dostatek kvalitních trénovacích dat, což se myslím ukázalo jako hlavní problém, se kterým se studentka musela utkat.

## Rešerše – data – experimenty – evaluace

Krátká úvodní kapitolka po velmi stručném úvodu do problematiky de facto představuje stručné shrnutí celé práce, což se mi nejeví jako příliš vhodné. Čtenář totiž k porozumění uvedeným výsledkům potřebuje znalosti detailů, které jsou v práci uvedeny až později. V úvodu mi naopak chybí detailnější a jasné vysvětlení zadání.

Struktura dalších kapitol předložené práce je vcelku standardní. V kapitolách 1 a 2 se vysvětluje teorie sémantických rámců včetně praktické implementace v dostupné databázi FrameNet a popisuje se anglicko-český paralelní korpus PCEDT, který autorka využívá jako primární a nejspolehlivější zdroj trénovacích dat pro rozpoznávání rámců v češtině. V kapitole 3 je popsán postup, jak bylo získáno větší množství trénovacích dat pomocí automatického strojového překladu. Bohužel ani jedna z uvedených metod neposkytuje příliš spolehlivé trénovací/testovací příklady.

V kapitole 4 se popisují experimenty se dvěma metodami pro automatické rozpoznání rámců: pomocí rozhodovacích stromů a pomocí SVM. Žádný pokus však nevedl k překročení baseline nad úroveň statistické chyby (ve skutečnosti není příliš jasné, o jaký “plus-mínus údaj”

se vlastně jedná, viz připomínka níže). Přestože úspěšnost obou použitých metod strojového učení je značně závislá na nastavení parametrů trénování, v práci není žádné ladění parametrů zmíněno. Není zde ani žádná analýza chyb.

V následující kapitole 5 autorka činí logický krok, který však měla učinit mnohem dříve, rozhodně již v rámci přípravy trénovacích dat. Teprve nyní začíná zkoumat trénovací data z hlediska množství příkladů dostupných pro jeden pár sémantický rámec/lemma. A při tom (nikoliv nečekaně) zjišťuje, že její data obsahují nezanedbatelně velké množství v tomto smyslu unikátních příkladů. Rozhodne se proto vývojová data mírně redukovat, aby se tak vyhnula nutnosti predikovat příklady, pro které je to v rámci strojového učení a na základě daných dat velmi obtížné nebo nemožné. Tím se celková naměřená úspěšnost klasifikace posouvá o několik procentních bodů nahoru, což však není žádné překvapení, to nastává z nutnosti.

Závěrečná kapitolka 6 poměrně podrobně shrnuje celou práci a dosažené výsledky srovnává s podobnými projekty, přičemž ale toto srovnání není nijak relevantní pro posouzení kvality práce samotné autorky. Srovnávané projekty totiž pracují se zcela odlišným přístupem k rozlišování sémantiky lexikálních jednotek a tudíž také se zcela odlišnými sadami příkladů, jejich numerické výsledky jsou proto víceméně nesrovnatelné. Přesto se domnívám, že přehled podobných projektů a jejich srovnání je užitečné a inspirující.

## Připomínky a otázky

- Dokážete na základě zpracovaných a analyzovaných dat kvantifikovat, do jaké míry je smysluplné aplikovat anglické sémantické rámce na češtinu? Muselo být přece od počátku jasné, že rámce nebudou zcela “pasovat”. Co jste se v tomto ohledu při analýze dat dověděla?
- Jaký smysl má vyhodnocování experimentů pracujících také s daty “none annotation”, viz např. str. 30?
- U výsledků přesnosti predikce se uvádí jakási “maximal deviation” – z poznámky pod čarou na str. 30 mi není jasné, co to přesně je a jaký to má význam. Pokud se jedná o nějaký způsob odhadu statistické chyby, jaký vliv má pak fakt, že testovací data jsou poměrně nespolehlivá? To se v práci vůbec nediskutuje!
- Proč se výběr příznaků omezil pouze na vlastní lemma a jeho automaticky rozpoznané syntaktické rodiče a následníky? Testovala jste úspěšnost automatické extrakce těchto příznaků?
- Pokoušela jste se ladit parametry trénování rozhodovacích stromů a SVM?
- Nebylo by možné k testování úspěšnosti klasifikace použít ručně vyhodnocený vzorek vět popsáný v kap. 2.4 očištěný od chyb zavlečených při automatickém transferu do češtiny?

## **Celkové hodnocení**

Předložená práce zřejmě splňuje zadání pouze částečně. Autorka ukazuje několik pokusů o automatickou detekci sémantických rámců, ale vůbec nepřistoupila k rozpoznávání jejich elementů, což však bylo výslovným požadavkem zadání. Popsané experimentování s metodami strojového učení navíc skončilo na poměrně triviální úrovni. Provedené experimenty jsou sice popsány dosti dobře a srozumitelně, celkově ale práce působí poněkud nedokončeně. To se projevuje mj. i rozsahem textu. Pokud odhlédneme od částí čistě rešeršních (ani ty nejdou příliš do hloubky), rozsah textu, který lze brát jako vlastní příspěvek autorky předložené práce, je relativně malý.

Formální požadavky kladené na diplomovou práci jsou však splněny. Autorka dostatečně cituje použitou literaturu, jejíž seznam na konci práce obsahuje cca 35 zpravidla cizojazyčných položek. Práce je napsána slušnou angličtinou prakticky bez chyb a dobře se čte. Hlavní přínos práce lze dle mého názoru spatřovat v tom, že předložené dílo představuje první systematický pokus v dané oblasti provedený pro češtinu a ukazuje obtížnost problému v rámci pokusu o jeho přímočaré řešení standardními či spíše základními prostředky.

Předloženou práci doporučuji k obhajobě.

V Praze, 1. září 2016

RNDr. Martin Holub, Ph.D.