

Univerzita Karlova v Praze
Matematicko-fyzikální fakulta

ZÁZNAM O PRŮBĚHU OBHAJOBY
DISERTAČNÍ PRÁCE

Název práce: Linguistic issues in Machine Translation between Czech and Russian

Jazyk práce: angličtina

Jméno studentky: Natalia Klyueva

Studijní program: Informatika

Studijní obor: I-3 Matematická lingvistika

Školitel: doc. RNDr. Vladislav Kuboň, Ph.D. (ÚFAL MFF UK)

Oponenti: prof. PhDr. Jarmila Panevová, DrSc. (ÚFAL MFF UK)
doc. RNDr. Petr Strossa, CSc. (VŠE v Praze)

Členové komise: doc. Ing. Zdeněk Žabokrtský, Ph.D. (předseda, ÚFAL MFF UK)
prof. RNDr. Jan Hajič, Dr. (místopředseda, ÚFAL MFF UK)
prof. PhDr. Jarmila Panevová, DrSc. (ÚFAL MFF UK)
doc. RNDr. Markéta Lopatková, Ph.D. (ÚFAL MFF UK)
doc. RNDr. Karel Oliva, Dr. (ÚJČ AV ČR)
Mgr. Barbora Vidová Hladká, Ph.D. (ÚFAL MFF UK)
RNDr. Ondřej Bojar, Ph.D. (ÚFAL MFF UK)
RNDr. Martin Čmejrek, Ph.D. (IBM TJ. Watson Center)
Mgr. Jan Štěpánek, Ph.D. (Barclays Capital Services Limited)

Datum obhajoby: 22. září 2015

Průběh obhajoby: Obhajoba proběhla v úterý 22. 9. 2015 od 10:00 hodin v budově MFF UK v místnosti S1, Malostranské nám. 25, Praha 1.
Přítomní členové: viz prezenční listina
Hosté: viz prezenční listina

Předsedající, doc. Ing. Zdeněk Žabokrtský, Ph.D., přivítal přítomné, představil členy komise a konstatoval, že komise je usnášeníschopná. Proběhla krátká porada s členy komise, jejímž závěrem bylo, že obhajoba (kromě vlastní prezentace uchazečky) proběhne v češtině.

Předsedající stručně představil uchazečku a zkontroloval úplnost dokladů vyžadovaných k obhajobě. Konstatoval, že byly splněny všechny podmínky pro konání obhajoby a že nedošly žádné písemné připomínky. Přítomní dostali možnost nahlédnout do výtisků předkládané disertační práce a do seznamu publikovaných prací uchazečky.

Školitel, doc. RNDr. Vladislav Kuboň, Ph.D., vyjádřil své stanovisko k průběhu doktorského studia uchazečky a seznámil přítomné se svým hodnocením obhajované disertační práce.

Pokyny pro předsedy nebo místopředsedy komisi:

Práce v elektronické podobě musí být studentem vložena do SIS. Formulář vyplňte ve všech bodech v elektronické podobě. V bodě Členové komise se uvedou všichni členové komise a za jejich jména se uvede „(přítomen)“ nebo „(nepřítomen)“. Předseda nebo místopředseda komise je jejím členem. V bodě Průběh obhajoby by měly být uvedeny alespoň čtyři věty vystihující průběh obhajoby. Po vyplnění formuláře ho vytiskněte, dole formulář ještě vlastnoručně podepište a přiložte k zápisu o státní závěrečné zkoušce. Současně vložte formulář v elektronické podobě (bez vlastnoručního podpisu) do SIS.

Uchazečka dostala slovo a představila svou práci. Soustředila se na statistický a pravidlový přístup ke strojovému překladu (především mezi ruštinou a češtinou). Popsala použitá a vytvořená paralelní jazyková data a slovníky a podrobněji představila překladové systémy TectoMT a Moses. Do celkového srovnání později zahrнула rovněž systémy PC Translator a Google translate. Dále se uchazečka věnovala klasifikaci chyb ve strojovém překladu s přihlédnutím k různým rovinám jazykového popisu a představila úspěšnost jmenovaných systémů vzhledem k jednotlivým typům chyb. Na závěr se soustředila na využití povrchové valence v systému TectoMT a automatické extrakci valenčního slovníku pro ruštinu s využitím valenčního slovníku pro češtinu Vallex a morfologicky anotovaných paralelních dat.

Předsedající požádal přítomnou oponentku prof. Jarmilu Panevovou o přednesení posudku. Ta shrnula hlavní body svého posudku, kde zdůraznila přínosy obhajované práce, zmínila některé menší nedostatky a navrhla směry, kterými by se mohl zaměřit další výzkum. Uzavřela doporučením práce k obhajobě a následného udělení titulu Ph.D.

Předsedající dále shrnul posudek nepřítomného oponenta doc. Strossy, který považuje za jádro práce kapitoly 4 a 5, a citoval dotaz oponenta týkající se některých typů chyb pravidlových strojových systémů. Citoval rovněž závěr posudku, který potvrdil uchazeččinu schopnost vědecky pracovat a práci doporučil k obhajobě.

Uchazečka se detailně vyjádřila ke všem poznámkám a otázkám oponentů. Oponentka prof. Panevová vyjádřila svou spokojenost s odpověďmi uchazečky. Dr. Štěpánek a Dr. Bojar doplnili několik poznámek k využití žakovských korpusů, resp. ke klasifikaci chyb. Za nepřítomného oponenta doc. Strossu vyjádřil spokojenost s odpověďmi předsedající.

Předsedající následně otevřel veřejnou rozpravu. Konstatoval, že může probíhat v angličtině. Doc. Oliva měl otázku k transliteraci azbuky a poznámku k vidům sloves v jednom z uvedených příkladů. Dále upřesnil kvality systému Ruslan a zeptal se na možný dopad výsledků práce uchazečky na úspěšnost strojového překladu. Uchazečka připomněla tabulku, ve které ukázala dopad zapojení valenčního slovníku ze systému Ruslan v systému TectoMT (mírně negativní ve skóre Bleu, výrazně pozitivní v manuální anotaci). Doc. Oliva následně připomněl citát prof. Jelínka o zlepšení výsledků systému rozpoznávání řeči pokaždé, když propustil lingvistu. Doc. Lopatková upozornila na nespolehlivost skóre Bleu v dané úloze. Dr. Bojar se zeptal na využití valenčního slovníku vytvořeného uchazečkou – ta odpověděla, že ještě nebyl použit. Doc. Lopatková se zeptala na množství sloves v systému Ruslan, u kterých se lišila valence v češtině a ruštině. Doc. Žabokrtský upozornil, že TectoMT nelze považovat za čistě pravidlový překladový systém.

Následovala neveřejná diskuse členů komise a tajné hlasování.

Komise schválila udělení titulu Ph.D.

Počet publikací: 18

Výsledek hlasování:

Počet členů s právem hlasovacím: 9

Počet přítomných členů: 8

Odevzdáno hlasů kladných: 6

Odevzdáno hlasů neplatných: 1

Odevzdáno hlasů záporných: 1

Výsledek obhajoby: prospěl/a neprospěl/a

Předseda nebo místopředseda komise:

Pokyny pro předsedy nebo místopředsedy komisi:

Práce v elektronické podobě musí být studentem vložena do SIS. Formulář vyplňte ve všech bodech v elektronické podobě. V bodě Členové komise se uvedou všichni členové komise a za jejich jména se uvede „(přítomen)“ nebo „(nepřítomen)“. Předseda nebo místopředseda komise je jejím členem. V bodě Průběh obhajoby by měly být uvedeny alespoň čtyři věty vystihující průběh obhajoby. Po vyplnění formuláře ho vytiskněte, dole formulář ještě vlastnoručně podepište a přiložte k zápisu o státní závěrečné zkoušce. Současně vložte formulář v elektronické podobě (bez vlastnoručního podpisu) do SIS.