

## ***ESSLLI 2006 – 18. evropska poletna šola logike, jezika in informacij, Malaga, 31. 7.–11. 8. 2006***

Spoznanja sodobne znanosti vse bolj upoštevajo načelo, da se da kompleksne probleme najbolje obvladati z interdisciplinarnim preučevanjem. *Evropska poletna šola logike, jezika in informacij (ESSLLI)* predstavlja rezultate interdisciplinarnega sodelovanja med jezikoslovjem, logiko in računalništvom.

*ESSLLI* ima svoje korenine v Groningenu, kjer je leta 1989 potekala prva tovrstna poletna šola. Leta 1991 je pokroviteljstvo prevzela *Zveza logike, jezika in informacij (FoLLI)*. Vsako leto se v 14 dneh zvrsti približno 50 različnih tečajev, ki jih obišče približno 500 udeležencev s celega sveta. Združevanje, predstavljanje in komentiranje dosežkov in spoznanj jezikoslovja, logike in računalništva se načeloma vsako leto dogaja v drugem evropskem mestu. Do sedaj so to priložnost dobili Leuven (1990), Saarbrücken (1991), Essex (1992), Lizbona (1993), København (1994), Barcelona (1995), Praga (1996), Aix en Provence (1997), Saarbrücken (1998), Utrecht (1999), Birmingham (2000), Helsinki (2001), Trento (2002), Dunaj (2003), Nancy (2004) in Edinburg (2005).

Letos so se študentje, raziskovalci in strokovnjaki s področij jezikoslovja, logike in jezikovnih tehnologij zbrali na Šoli za računalniške vede v Malagi. Organizacijo je prevzel *GISUM*, skupina za izdelavo programske opreme Univerze v Malagi. Izbirati je bilo mogoče med 41 predavanji, ki so bila namenjena tako tistim, ki se šele začinjajo ukvarjati s tovrstnim raziskovanjem, kot tudi že uveljavljenim raziskovalcem in strokovnjakom, šestimi delavnicami, študentsko sekcijo in štirimi večernimi predavanji.

Predavanja in delavnice so bili razdeljeni v tri tematske sklope, in sicer *Jezik in računalništvo*, *Jezik in logika* ter *Logika in računalništvo*. V sklopu *Jezik in logika* so bili predstavljeni različni logični pristopi in operacije kot pomoč pri skladenjsko-

semantičnih dinamičnih jezikovnih opisih. Tovrstni opisi so nadgradnja statičnih jezikovnih opisov, ki temeljijo na formalnih slovnica. Problemi so bili obravnavani na diskurzivni ravni, predvsem z upoštevanjem pragmatike in logike védenja na ravni informacij, v celoti dostopnih iz besedila. Sklop *Jezik in računalništvo* je bil posvečen različnim načinom reševanja logičnih problemov kot pomoč pri oblikovanju računalniških programov. Predavanja po sklopih so potekala sočasno, zato bo v nadaljevanju nekoliko natančneje predstavljena vsebina tematskega sklopa *Jezik in računalništvo*.

Tematski sklop *Jezik in računalništvo* je bil večinoma v znamenju načinov in metod kompleksnejšega označevanja korpusov, ki je v slovenskem prostoru vsaj zaenkrat še v zametkih. Predavatelji iz Nemčije, Italije, Francije, Japonske, Amerike, Velike Britanije, Avstralije, Nizozemske in Brazilije so se večinoma sklicevali na t. i. globinske slovnice, na osnovi katerih naj bi se razložile povezave med informacijami v besedilu in med posameznimi deli besedila na osnovi uporabe posameznikovega védenja oz. verjetnostnih modelov. Pri tem so navajali predvsem spoznanja logike in kognitivnega jezikoslovja. Največja pozornost je bila posvečena metodologiji in problemom semantičnega in diskurzivnega označevanja korpusov.

Prvi teden sta **Sabine Schulte im Walde** (Saarlandska univerza) in **Heike Zinsmeister** (Univerza v Tübingenu) na primerih posameznih korpusov predstavili različne načine njihovega označevanja. **Sebastian Pado** in **Katrin Erk** (Saarlandska univerza) sta napravila zgodovinski pregled skladenjsko-semantičnega označevanja korpusov. Največ pozornosti sta posvetila problemom korpusnega označevanja semantičnih vlog s pomočjo okvirjev in shem. Pri tem sta si pomagala s *FrameNetom*. Z leksikalnosemantičnimi problemi pri označevanju korpusa sta se

ukvarjala tudi **Carlo Strapparava** in **Alfio Gliozzo** (ITC-irst, Trento). Predstavila sta koncept semantičnih področij (*semantic domains*), s pomočjo katerega naj bi bilo dvoumne pomene in variante pomenov računalniško mogoče razrešiti na osnovi tematske podobnosti. Pri iskanju tematske podobnosti besed sta si pomagala z *WordNetom*. **Yusuke Miyao** (Univerza v Tokiu), **Josef van Genabith** (Univerza v Dublinu) in **Julia Hockenmaier** (IRCS, Univerza v Pennsylvaniji/ICCS, Univerza v Edinburgu) so predstavili načine in prednosti diskurzivno (tj. na osnovi védenja oz. verjetnostnih modelov) označenega korpusa *Penn* v primerjavi z nediskurzivnimi metodami označevanja. **Johan Bos** (Univerza v Rimu »La Sapienza«) in **Patrick Blackburn** (INRIA Lorraine) pa sta podrobneje predstavila en segment diskurzivno označenega korpusa *Penn*, in sicer povezovanje besedila na osnovi besedilnih referenc z uporabo »teorije prikazovanja diskurza« (DRT), ki temelji predvsem na razreševanju anafor in besedilnih inferencah. **Anna Korhonen** (Univerza v Cambridgu), **Timothy Baldwin** (Univerza v Melbournu) in **Aline Villavicencio** (Zvezna univerza v Riu Grande du Sol/Univerza v Bathu) so prikazali različne načine pridobivanja podatkov iz korpusov, posebej pa so se posvetili korpusnemu pridobivanju skladiškovno-semantičnih podatkov na osnovi uporabe jezikovnega védenja. Osredotočili so se zlasti na besede, ki jih ni v slovarju, in na nenormirane skladiškovne konstrukcije. **Damir Čavar** (Univerza v Indiani) pa je poleg osnov programiranja predstavil tudi različne načine pridobivanja podatkov iz korpusa s pomočjo programov *MzScheme* in *DrScheme*.

Drugi teden sta **Dan Flickinger** (Univerza v Oslu/CSLI Stanford) in **Ann Copestake** (Univerza v Cambridgeu) predstavila različne načine globinskega in površinskega semantičnega označevanja korpusa, na osnovi katerih bi bilo mogoče s pomočjo računalniških algoritmov pridobivati semantične informacije iz struktur. **Rada Mihalcea** (Univerza v Severnem Texasu) je govorila o različnih načinih oz. metodah

računalniškega razdvoumljanja besed v korpusu, ki so velikokrat odvisni od ravnih označenosti korpusa, ter o problemih pri razdvoumljanju besednega pomena. **Nicholas Asher** (Univerza Texas – Austin) je v okviru diskurzivne semantike predstavil nekaj problematičnih področij pri označevanju korpusa, kot so anafore, modalni glagoli in množinski samostalniki. **Annie Zaenen** (PARC) se je po kratkem zgodovinskem in teoretičnem uvodu posvetila praktičnim primerom korpusnega razreševanja anafor. **Roger Levy** (Univerza v Edinburgu/Univerza v Kaliforniji, San Diego) je teoretično in praktično predstavil psiholingvistično-statistične metode za razreševanje dvoumnih anafor v stavku, pri čemer se je opiral na verjetnostno teorijo in eksperimentalne podatke. **Tracy Holloway King** (PARC) in **Miriam Butt** (Univerza v Konstanzu) sta v okviru leksikalne in logične semantike govorili o kompleksnih predikatnih konstrukcijah v posameznih jezikih ter o operacijah za oblikovno spreminjanje argumentov in njihovem vključevanju v računalniške sisteme. **Markus Egg** (Rijksuniversiteit Groningen) in **Valia Kordoni** (Saarlandska univerza) sta predstavila problematična področja pri računalniškem povezovanju morfosintakse in leksikalne semantike. **Kemal Oflazer** (Univerza Sabanci) je podrobneje predstavil proces morfološkega označevanja korpusa ter načine pridobivanja jezikovnih informacij iz besednih morfemov v različnih jezikih. **Marco Baroni** (Univerza v Bolonji, SITLEC) in **Stefan Evert** (Univerza v Osnabrücku, Inštitut za kognitivne vede) sta se posvetila leksikalni statistiki oz. statističnemu pridobivanju podatkov o frekvenci besed in morfemov s pomočjo programa *ZipfR*. **Oliver Lemon** in **James Henderson** (Univerza v Edinburgu, HCRC) pa sta teoretično in praktično predstavila različne bolj ali manj fleksibilne dialoške sisteme, njihovo zgradbo, lastnosti in možnosti uporabe. Posebno pozornost sta namenila modelom dialoške komunikacije in njihovim omejitvam ter metodam za zmanjševanje napak pri prepoznavanju govora v kontekstu.

Delavnice so bile namenjene predstavitvam dela študentov in raziskovalcev v povezavi z določeno temo, izpostavitvi glavnih problemov ter medsebojni izmenjavi mnenj v obliki diskusije. V tematskem sklopu *Jezik in računalništvo* je bila delavnica posvečena različnim koherentnim modelom. V tematskem sklopu *Jezik in logika* so bile delavnice posvečene pojmu skladnosti na meji med skladnjo in semantiko, formalnim ontologijam v komunikaciji ter dvoumnim anaforam. V tematskem sklopu *Logika in računalništvo* pa sta bili delavnici posvečeni logiki védenja ter razumnosti in védenju.

V študentski sekciji so svoje raziskovalno delo profesionalnemu občinstvu predstavili dodiplomski in podiplomski študentje, z namemom dobiti povratno informacijo o svojem delu. Vsak dan sta v eni uri svoje delo, ki po tematiki ni bilo povezano, predstavila dva študenta.

Na štirih večernih predavanjih so nastopili štirje priznani strokovnjaki na svojem področju. **Yde Venema** (Univerza v Amsterdamu) je predaval o koalgabri, **José Meseguer** (Univerza v Illinoisu) o logičnih in metalogičnih okvirjih, **Nicholas Asher** (Univerza Texas – Austin) o dinamični semantiki in pragmatiki, **Annie Zaenen** (PARC) pa o dostopnosti podatkov kot zunajskladenjskem psiholingvističnem dejavniku za motivacijo skladenjske izbire.

*ESSLLI* je poletna šola, ki prinaša veliko zanimivih informacij o metodologiji interdisciplinarnega sodelovanja med jezikoslovjem, logiko in računalništvom. Večina metodoloških pristopov je zaenkrat osredotočenih na angleški jezik. Sodelovanje med jezikom in računalništvom je usmerjeno predvsem v različne računalniške programe in teorije, s pomočjo katerih bi bilo mogoče najboljše računalniško označiti besedilo kot celoto in iz besedila dobiti čim več informacij. Konkretno jezikovne analize na osnovi predstavljenih metod pa so velikokrat narejene z očmi logika ali računalniškega strokovnjaka, zaradi česar v primerjavi z diskurzivnimi pristopi pri oblikovanju računalniških algoritmov velikokrat ostanejo na formalni ravni enega stavka ali dveh. Zaradi kompleksnosti uporabe takšnih programov in operacij je to po eni strani razumljivo, vendar bi se mogoče vseeno bilo dobro vsaj na ravni interpretacije podatkov v večji meri povezati z jezikoslovci, ki se ukvarjajo z analizami diskurza kot celote.

Naslednje evropsko mesto, ki bo prizorišče *ESSLLI* od 6. do 17. avgusta 2007, je Dublin. Vodstvo organizacijskega odbora je prevzel **Carl Vogel**, vodstvo programskega odbora pa **Tomaž Erjavec**.

Urška Jarnovič  
Univerza v Ljubljani, Filozofska fakulteta  
urska.jarnovic@ff.uni-lj.si

## Slovenski slavistični kongres *Preseganje meje*, Zagreb, 5.–7. oktober 2006

Prva postaja »slovenističnega misijona«, kot je »serijo slovenskih slavističnih kongresov zunaj domačih meja: v Zagrebu, Trstu in Celovcu« v uvodu v 17. zbornik SDS poimenoval njegov urednik in dosedanji predsednik SDS, Miran Hladnik (na občnem zboru v Zagrebu je predsedovanje prevzel Miran Košuta), je za nami. Ni naključje, da je odločitev zanjo med drugim ilustriral tudi s Prešernovimi besedami iz *Krsta pri*

*Savici*: »?/M/ed svoje rojake / Slovence gre, in **dálej čez njih mejo**, / do smrti **tam** preganja zmot oblake' (poudarki M. H.)«, predvsem pa je izbiro Zagreba utemeljil s promoviranjem slovenistike zunaj slovenskih meja in podpiranjem njenega razvoja na sosednji univerzi, z uzaveščanjem potrebe po prilagoditvi izobraževalnega sistema govorcem slovenščine kot drugega/tujega jezika, z izražanjem podpore in pomoči zagrebški slovenski skupnosti (med 1500 in