

GOVORNA POIZVEDOVANJA V ČEŠČINI, NEMŠČINI, SLOVAŠČINI IN SLOVENŠČINI

France Mihelič, Ivo Ipšič, Jerneja Gros, Karmen Pepelnjak, Simon Dobrišek, Nikola Pavešič

Laboratorij za umetno zaznavanje
Fakulteta za elektrotehniko
Tržaška 25, 1000 Ljubljana
Tel. + 386 61 1768 313, fax + 386 61 1264 631
e-mail: mihelicf@fe.uni-lj.si

POVZETEK

V članku predstavljamo mednarodni projekt SQEL (Spoken Queries in European Languages) in posebej opisujemo dosežke raziskovalne skupine na Fakulteti za elektrotehniko v okviru tega projekta. To so: posebne podatkovne zbirke slovenskega govora in besedil, udejanjenje razpoznavnika tekočega slovenskega govora, pomenska analiza govornih sporočil in sistem za samodejno tvorjenje slovenskega govora.

ABSTRACT

In the paper we present the SQEL (Spoken Queries in European Languages) project and outline the work done by the Speech Recognition Group at the Faculty of Electrical Engineering, especially on Slovenian speech corpora, realisation of a Slovenian continuous speech recogniser, semantic analysis of spoken messages and Slovenian text-to-speech synthesis.

1 UVOD

Govor je najbolj običajen in naraven način sporazumevanja med ljudmi. Z njim lahko človek izraža tudi najbolj zapletene misli in čustva. Že od začetka razvoja računalnikov v drugi polovici tega stoletja si raziskovalci prizadevajo, da bi se s strojem oziroma računalnikom sporazumevali kot z ljudmi – z govorom. Od prvih poskusov in prvih resnejših znanstvenih projektov s tega področja je preteklo že veliko časa. Preizkušenih je bilo mnogo zamisli in opravljenih veliko število sprva manj uspešnih poskusov, vendar so se raziskovalci v zadnjem desetletju s hitrim razvojem na področju zmogljivosti računalnikov in z novimi spoznanji s področij obdelave govornih signalov, razpoznavanja vzorcev, umetne inteligence in računalniškega jezikoslovja v marsičem že približali želenim ciljem. Na tržišču se že pojavljajo sistemi za narekovanje besedil za »velike« jezike, kot so angleščina, francoščina in nemščina, ki jih je mogoče uporabljati na osebnih računalnikih. Prav tako so na voljo sistemi za razpoznavanje ločeno izgovorjenih besed. Ti so od jezika neodvisni in uporabnik sam določi »zvoke« - besede, ki jih bo sistem razpoznaval. Sistemi, ki poleg razpoznavanja vključujejo tudi

razumevanje bolj zapletenih govornih sporočil, izraženih z enim ali več stavki, pa so še vedno predmet znanstvenega proučevanja. Trenutno so omejeni na eno samo, dovolj ozko, področje uporabe in jih zato ni preprosto prilagoditi drugemu namenu.

Uporabnost samodejnega razpoznavanja, razumevanja in tvorjenja govora je zelo raznovrstna, na primer:

- Prostorčno upravljanje s stroji in računalniki, brez uporabe tipkovnice, miške, ipd. Primere najdemo v industriji, prometu, vojski, pri delu invalidnih oseb.
- Uvedba samodejnih sistemov za govorno poizvedovanje prek telefona. Sem sodi poizvedovanje po prometnih, turističnih, vremenskih in drugih informacijah. Opise nekaterih takih sistemov najdemo v [1, 2, 3].
- Sama prevedba pisanih in drugih vidnih informacij v govor pa je pomembna pri branju besedil slepim in slabovidnim osebam, posredovanju podatkov v govorni obliki, ko smo prezaposleni s sprejemanjem drugih, predvsem vidnih podatkov (ob upravljanju vozila) ali ko vidni podatki niso dostopni (branje elektronske pošte in fax-a z uporabo telefona).

Uporaba omenjanih samodejnih postopkov je v marsičem odvisna od jezika, v katerem bo potekalo sporazumevanje. Tako je potrebno nekatere postopke posameznemu jeziku posebej prilagoditi, nekatere pa za izbrani jezik posebej zasnovati. Nobenega dvoma ni, da se bo uporaba jezikovnih tehnologij, med katere se ti postopki uvrščajo, s časom vse bolj uveljavila in bo marsikje poenostavila in pocenila dostop do informacij kakor tudi nekatere proizvodne industrijske postopke. Tako bomo tudi pri nas prej ali slej pristopili k uvajanju jezikovnih tehnologij. Če želimo, da se bomo s samodejnimi napravami lahko sporazumevali in jih uporabljali v slovenščini, je raziskovalno in razvojno delo na tem področju tudi pri nas nujno. Pričujoči prispevek govori o rezultatih dela na tem področju na mednarodnem projektu, pri katerem je sodelovala tudi raziskovalna skupina iz Slovenije.

2 PREDSTAVITEV PROJEKTA SQEL

Letos 24. aprila se je z znanstvenim delovnim srečanjem v okviru mednarodne konference Artificial Intelligence in Industry - AIII'98 [4] v Stari Lesni na Slovaškem končalo delo na mednarodnem triletнем projektu Spoken Queries in European Languages (SQEL – Govorjena poizvedovanja v evropskih jezikih) iz sklopa projektov Copernicus, ki jih denarno podpira Evropska skupnost. Na projektu so kot partnerji sodelovali raziskovalci iz Češke¹, Nemčije, Slovaške² in Slovenije. Slovenijo je na tem projektu predstavljala raziskovalna skupina šestih raziskovalcev pod vodstvom prof. Nikole Pavešiča iz Laboratorija za umetno zaznavanje na Fakulteti za elektrotehniko Univerze v Ljubljani.

Cilj projekta je bil razviti samodejni sistem za posredovanje informacij preko telefona, ki bo zmožen iz klicalečevega govora sprva ugotoviti, v katerem jeziku bo potekal pogovor in nato z njim voditi razgovor na vsebinsko omejenem področju poizvedovanj o letalskih oziroma železniških povezavah.

Tak sistem mora biti sposoben:

- ugotoviti, v katerem izmed štirih možnih jezikov (češki, nemški, slovaški, slovenski) je klicalec nagovoril sistem;
- razpoznati govor v omenjenih štirih jezikih;
- ugotoviti namen klicalečevega poizvedovanja;
- poiskati zelene informacije v podatkovni zbirki informacijskega sistema;
- postaviti morebitna vprašanja pri nepopolnih, slabo izraženih ali nepravilno razpoznanih zahtevah klicalca;
- tvoriti odgovore v obliki govornega sporočila v jeziku, ki ga govori klicalec.

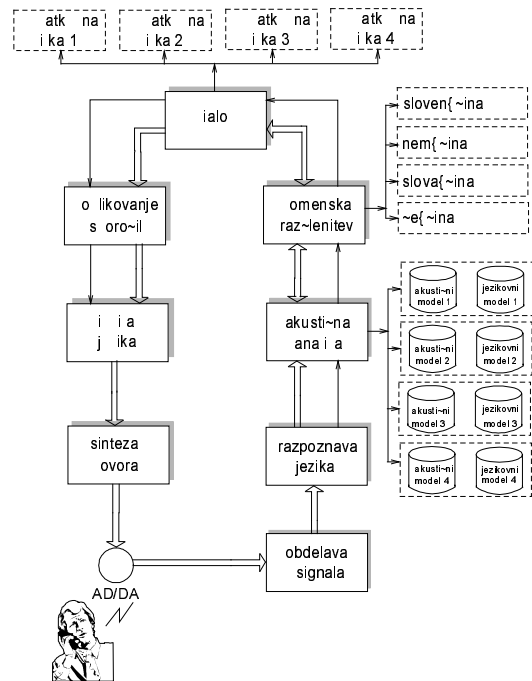
Projekt so vodili nemški sodelavci iz Univerze Friderika – Aleksandra Erlangen–Nürnberg iz Oddelka za razpoznavanje vzorcev (Lehrstuhl für Mustererkennung) pod vodstvom prof. Heinricha Niemann. Ta skupina je med sodelovanjem na predhodnih projektih že razvila delujoč sistem EVAR (Erkennen – Verstehen – Antworten – Rückfragen) za govorno telefonsko poizvedovanje o železniških povezavah v nemškem jeziku [1]. Na podlagi njihovih predhodnih izkušenj so si raziskovalne prizadevale zgraditi nemško – slovanski štirijezični sistem s prej naštetimi zmožnostmi. Raziskovalci iz slovanskih govornih področij so sisteme za svoje jezike zasnovali po obstoječem nemškem sistemu in razvili ustrezne manjkajoče komponente za svoj jezik [5, 6, 7, 8].

Za češčino, slovaščino in slovenščino je bilo potrebno udejaniti razpoznavalnike govora, ki bodo sposobni dovolj zanesljivo razpoznavati spontani govor s področja letalskih oz. železniških povezav prek telefona. Za te tri jezike je bilo treba na novo razviti jezikovne razčlenjevalnike, ki bodo lahko samodejno razbirali pomen uporabnikovih poizvedovanj. Prav tako je bilo potrebno razvijati postopke

za samodejno tvorjenje govora za podajanje informacij. Zasnovati smo morali osnovni podsistem za razpoznavanje jezika, ki lahko dovolj hitro – že na podlagi nekaj izgovorjenih besed – ugotovi, za kateri jezik gre [9]. Podsistem za dialog, ki je od jezika neodvisen, smo vsi prevzeli od nemških sodelavcev s tem, da so češki sodelavci tudi ta modul v svojem podsistemu nekoliko spremenili [6]. To je tisti del sistema, ki na podlagi pomena razpoznanega govora omogoča iskanje zelenih informacij in ugotavlja morebitne manjkajoče ali protislovne podatke, ki jih je potrebno v pogovoru z uporabnikom razjasniti. Groba struktura sistema je razvidna iz slike 1.

3 SKUPNI DOSEŽKI

Na zadnjem skupnem srečanju smo ugotovili, da smo zastavljene cilje projekta dosegli in marsikje tudi presegle ter skupaj pridobili veliko novega znanja in izkušenj. Večjezični eksperimentalni sistem, ki deluje in je trenutno vzpostavljen le v Nemčiji – po telefonu (št. 0049 9131 16287), bo po nekaterih prilagoditvah dostopen zopet v začetku jeseni – še preboleva nekatere otroške bolezni. Tako zaradi svoje velike kompleksnosti (na enem računalniku – delovni postaji HP 9000 735/125 – v nekaterih delih procesa sodelujejo hkrati štirje razpoznavalniki govora) sistem deluje občutno počasneje od primerljivih enojezičnih sistemov.



Slika 1. Struktura večjezičnega sistema za govorna poizvedovanja

Prav tako se je med preizkušanjem na Slovaškem pokazalo, da razpoznavalnik jezika ne deluje dobro, ko klicalec uporablja mobilni telefon, za kar je krivo specifično razmerje signala in šuma in oblika motenj, ki pri taki

¹ Tehnična univerza v Pilznu, prof. Vaclav Matoušek.

² Tehnična univerza v Košicah, prof. Dušan Krokavec.

povezavi nastanejo in na katere sistem ni bil naučen. Vendar so te težave bolj tehnične kot vsebinske narave in jih bo moč z izboljšavo postopkov in prilagoditvami na različne vrste telefonskih signalov odpraviti. Pričakujemo, da bomo večjezični sistem v kratkem vzpostavili tudi v Ljubljani, kjer bo na voljo za preizkušanje tudi klicalcem iz Slovenije.

4 PRISPEVEK SLOVENSКИH SODELAVCEV

Povejmo še, kako smo se dela na tem projektu lotili v Sloveniji in kateri dosežki in spoznanja so pomembni predvsem za naše okolje. Prej omenjena skupina raziskovalcev v Laboratoriju za umetno zaznavanje se že dalj časa intenzivno ukvarja s problematiko samodejne obdelave, razpoznavanja in tvorjenja slovenskega govora. Prav tako že več let formalno in neformalno sodelujemo z Oddelkom za razpoznavanje vzorcev na Univerzi Erlangen – Nürnberg preko skupnih projektov, izmenjav raziskovalcev, znanstvenih srečanj in skupnih objav. Tako je bila vloga za skupni projekt pri Evropski skupnosti in delo na tem projektu nekako logično nadaljevanje skupnih prizadevanj. V našem Laboratoriju se posvečamo prav od govorca neodvisnemu razpoznavanju spontanega govora. Tovrstno razpoznavanje potrebujemo tudi pri sistemu za govorno poizvedovanje prek telefona, kjer od uporabnika ne moremo zahtevati, da govori na vnaprej predpisan način - na primer tako, da bo med besedami delal daljše premore, še manj pa, da bomo sistem vnaprej prilagojen klicalcem glasu.

Za udejanjenje takega sistema potrebujemo obsežne zbirke govornih podatkov čim več različnih govorcev, na podlagi katerih samodejni sistem učimo in tudi preizkušamo. Za tematsko omejeno področje dialoga pa potrebujemo tudi posebej izbran govor, ki se nanaša na izbrano področje uporabe. Nadalje lahko določimo potrebno število besed, ki jih bomo morali razpoznavati, oblike vprašanj in odgovorov in možne poteke pogovorov. Pomenski obseg pogovorov, ki jih bo sistem obvladoval, natančno definiramo. V ta namen smo v sodelovanju s slovenskim letalskim prevoznikom Adrio Airways spremljali več ur pogovorov v njihovem centru za rezervacije letalskih poletov, zbrali značilne pogovore in se tematsko omejili le na poizvedovanja o urnikih poletov z Brniškega letališča. Gre za vprašanja o obstoju določene letalske povezave, letalskem prevozniku, času prihoda in odhoda ter trajanju poleta. Poizvedovanja smo nato smiselno uredili, sestavili seznam stavkov pogovora ter z uporabo generativne slovnice³ množico stavkov razširili.

V laboratorijskem okolju smo s petdeset govorcev prek telefona in neposredno v računalnik v digitalni obliki posneli govorno zbirko 8850 stavkov, kar predstavlja približno 28 ur govora. Zbirko smo dopolnili s seznamom izgovorjenih stavkov, leksikonom uporabljenih besed, ki je vseboval fonetični prepis besed in njihove skladišne in pomenske

opise, ter z nekaterimi podatki o govornih in njihovem načinu govora [11]. Ta podatkovna zbirka predstavlja eno prvih obsežnejših dokumentiranih govornih zbirk za slovenski govor in ni primerna zgolj pri postavljanju in preizkušnji sistemov za razpoznavanje tekočega slovenskega govora, marveč tudi za druge študije s področja slovenskega jezika. Ker gre tu za fonetično označeno zbirko podatkov, je na njej možno proučevati tudi fonetične lastnosti slovenskega govornega jezika in dialektov govorcev. Prav to zbirko podatkov smo s pridom uporabili za določanje nekaterih prozodičnih karakteristik slovenskega govora, kar smo potrebovali pri izgradnji podsistema za samodejno tvorjenje govora [12, 13]. Posebno pozornost smo posvetili izbiri množice osnovnih govornih enot za razpoznavanje slovenščine ter izbiri postopkov za določanje značilnosti slovenskega govora, na katerih temelji kasnejše razpoznavanje. Pri tem smo se lahko naslonili na spoznanja predhodnih in tekočih raziskav [14, 15, 16, 17].

Poseben in od jezika močno odvisen del v okviru celotnega sistema je predstavljal pomenski razčlenjevalnik razpoznanega govora. Njegova naloga je, da razpoznan niz besed pomensko opredeli in ga zapiše v tako imenovanem pomensko predstavitevem jeziku SIL (Semantic Interpretation Language). Ker je razpoznan niz besed zaradi napak pri razpoznavanju in ohlapnega izražanja uporabnikov pogosto skladišniško nepravilen, mora biti razčlenjevalnik na take napake čim bolj neobčutljiv, iz razpoznanega niza pa mora izločiti le za sistem pomembne pomenske podatke. Sistema na primer prav nič ne zanima, ali želi uporabnik v kraju pristanka letala obiskati svojo staro mamo, in ali je uporabnik že polnoleten. Ker je bil problem razumevanja govora tematsko močno omejen, smo se te naloge lotili z ugotavljanjem manjšega števila pomenskih kategorij, kot so kraj prihoda, kraj odhoda, letalski prevoznik, časovni izrazi, potrditve in uvodne ter zaključne izjave. Vse te kategorije smo z uporabo leksikona možnih besed in ustreznimi programskimi postopki skušali odkriti med razpoznanimi besedami in jih ovrednotiti [18].

Pomemben izvorni prispevek naše skupine na tem projektu predstavlja tudi podsistem za samodejno tvorjenje slovenskega govora, ki smo ga zasnovali tako na široko, da predstavlja zaključeno celoto [19]. Sistem je sposoben tvoriti slovenski govor na podlagi poljubnega slovenskega besedila in tako ni uporaben le v sistemih za posebej določena govorna poizvedovanja. Tak sistem je na primer zelo uporaben za branje vidno prizadetim osebam in sploh vedno, ko uporabnik ne more prejemati informacij v pisni obliki. Sistem smo v skladu z mednarodnimi standardi s tega področja ovrednotili in predstavili domači in tuji strokovni javnosti. Bralci z dostopom do interneta pa si lahko podatke o sistemu in nekaj vzorcev samodejno tvorjenega govora ogledajo in poslušajo na naslovu <http://luz.fe.uni-lj.si/english/SQEL/synthesis-eng.html>. Prav tako si nekatere podatke v zvezi s projektom SQEL ogledate na naslovu <http://luz.fe.uni-lj.si/english/SQEL/home-eng.html>.

³ V stavkih smo spreminjali imena letalskih družb, krajev in časov prihoda in odhoda in jih širili oziroma krčili z dodajanjem ali izpuščanjem nekaterih besed [10].

O obsegu opravljenega dela naše skupine priča na to vezana bibliografija 39 del. Na tematiki projekta pa so v Sloveniji v triletnem obdobju sodelovali mladi raziskovalci z magisteriji in doktorati ter mnogi študenti dodiplomskega študija.

5 ZAKLJUČEK

Triletno delo na skupnem projektu je zaključeno, seveda pa se delo na tem področju pri nas in na tujem nadaljuje. Preizkusiti in udejaniti nameravamo še mnogo novih idej in izboljšav s tega področja in vsekakor vsaj neformalno nadaljevati s skupnimi raziskavami in sodelovanjem z drugimi raziskovalnimi skupinami, še zlasti s slovenskimi slavisti in jezikoslovci. Ovire pri našem nadaljnjem raziskovalnem in razvojnem delu pa gotovo predstavljajo restriktivne razmere na raziskovalnem področju v naši državi. Čeprav je jasno, da so raziskave s področja jezikovnih tehnologij za Slovence s stališča naše nacionalne identitete še kako pomembne in da jih namesto nas ne bo opravil nihče drug, se pri podpori delu na tem področju pri nas izkazuje drugače. Bo zopet prišlo do razmer, da bomo kupovali ustrežno, večkrat tudi manj kvalitetno, znanje in izdelke v tujini, ker smo po nemarnem zavrgli tisto, česar smo se naučili in kar smo naredili doma?

6 VIRI

1. W. Eckert, T. Kuhn, H. Niemann, S. Rieck, A. Scheuer, E.G. Schukat-Talamazzini. A spoken Dialogue System for German Intercity Train Timetable Inquiries. *Eurospeech '93*, Berlin 1993, str. 1871-1874.
2. V. Zue, S. Seneff, J. Glass, L. Hetherington, E. Hurley, H. Meng, C. Pao, J. Polifroni, R. Schloming, P. Schmid. From Interface to Content: Translingual Access and Delivery of on-line information. *Proc. Eurospeech 1997*, Rodos 1997, str. 2227-2230.
3. L. Haaren, M. Blasband, M. Gerritsen, M. Schijdel. Evaluating Quality of Spoken Dialogue Systems: Comparing a Technology-focused and User-focused Approach, *First International Conference on Language Resources and Evaluation*. izd. A. Rubio, N. Gallardo, R. C. y A. Tehada. Granada 1998, str. 655-660.
4. *Proceedings of the Scientific Conference Artificial Intelligence in Industry*, izdali Jan Sarnovsky, Peter Sinčak, Marian Mach, Marek Hatala, Visole Tatré, Slovaška, april 1998, 479 str.
5. V. Matoušek, P. Mautner. Spontaneous Speech Recognition, *Proceedings of the 2nd SQEL Workshop on Multi-Ligual Information Retrieval Dialogs*, Plzen 1997, str. 84 – 89.
6. Jana Krutišova, Vaclav Matuošek, Jana Ocelikova. Development of a Dialogue Strategy, *Proceedings of the Scientific Conference Artificial Intelligence in Industry*, izdali Jan Sarnovsky, Peter Sinčak, Marian Mach, Marek Hatala, Visole Tatré, Slovaška, april 1998, str. 457 – 468.
7. D. Krokavec, J. Ivanecky. Slovak Spoken Dialog System. *Proceedings of the Scientific Conference Artificial Intelligence in Industry*, izdali Jan Sarnovsky, Peter Sinčak, Marian Mach, Marek Hatala, Visole Tatré, Slovaška, april 1998, str. 447 – 456.
8. I. Ipšič, F. Mihelič, S. Dobrišek, J. Gros, N. Pavešič. An Overview of the Spoken Queries in European Languages Project: The Slovenian Spoken Dialog System, *Proceedings of the Scientific Conference Artificial Intelligence in Industry*, izdali Jan Sarnovsky, Peter Sinčak, Marian Mach, Marek Hatala, Visole Tatré, Slovaška, april 1998, str. 431–438.
9. S. Harbeck, E. Noeth, H. Niemann. Multilingual Speech Recognition. *Proceedings of the 2nd SQEL Workshop on Multi-Ligual Information Retrieval Dialogs*, Plzen 1997, str. 9 – 15.
10. J. Gros, F. Mihelič, N. Pavešič. Sentence Hypothesis using Ng-grams, *Proceedings of the Eurospeech95*, Madrid, september 1995, str. 1759 – 1762.
11. S. Dobrišek, J. Gros, F. Mihelič, N. Pavešič. Recording and Labelling of the GOPOLIS Slovenian Speech Database. *First International Conference on Language Resources and Evaluation*. izd. A. Rubio, N. Gallardo, R. C. y A. Tehada. Granada 1998, str. 1089-1096.
12. J. Gros, N. Pavešič, F. Mihelič. Speech Timing in Slovenian TTS, *EUROSPEECH'97, Proceedings of the 5th European Conference On Speech Communication and Technology*, Rodos, Grčija, september 1997, Vol. 1, str. 323-326.
13. J. Gros, N. Pavešič, F. Mihelič. Syllable and Segment Duration at Different Speaking Rates in the Slovenian Language, *EUROSPEECH'97, Proceedings of the 5th European Conference On Speech Communication and Technology*, Rodos, Grčija, september 1997, Vol. 2, str. 951–954.
14. F. Mihelič, I. Ipšič, S. Dobrišek, N. Pavešič. Feature Representations and Classification Procedures for Slovene Phoneme Recognition, *Pattern Recognition Letters 13*, North-Holland, Nizozemska, december 1992, str. 879–891.
15. S. Dobrišek, F. Mihelič, N. Pavešič. Merging of Time Delayed Feature Vectors into Extended Vector in Order to Improve Phoneme Recognition, *Adaptive Methods and Emergent Techniques for Signal Processing and Communications, Proceedings of the 4th COST 229 Workshop*, izdal Jurij F. Tasič, Ljubljana, Slovenija, april 1994, str. 145 – 150.
16. I. Ipšič, F. Mihelič, E.G. Schukat-Talamazzini, N. Pavešič. Generating word hypotheses in the Slovene continuous speech recognition. *Visual Modules*, izdala F. Solina in W.G. Kropatsch, Maribor, maj 1995, str. 77-85.
17. S. Dobrišek, F. Mihelič, N. Pavešič. A Multiresolutionally Oriented Approach for Determination of Cepstral Features in Speech Recognition, *EUROSPEECH'97, Proceedings of the*

5th European Conference On Speech Communication and Technology, Rodos, Grčija, september 1997, Vol. 3, str. 1367—1370.

18. K. Pepelnjak, F. Mihelič, N. Pavešić. Semantic Decomposition of Sentences in the System Supporting Flight Services, *CIT - Journal of Computing and Information Technology*, Vol. 4, No. 1, Zagreb 1996, str. 17-24.
19. J. Gros, N. Pavešić, F. Mihelič. Text-to-Speech Synthesis: A complete system for the Slovenian Language, *CIT - Journal of Computing and Information Technology*, Vol. 5, No. 1, Zagreb 1997, str. 11-19.