

Milan Stamenković
 II. EVROPSKA KONFERENCIJA O GOVORNOJ KOMUNIKACIJI I
 TEHNOLOGIJI

Druga evropska konferencija o govornoj tehnologiji održana je u Parizu od 26. do 28. rujna 1989. u prostorijama Grada znanosti i industrije (Cité des sciences et de l'industrie). Glavna tema konferencije bila je govorna tehnologija koja obuhvaća sisteme za automatsku (računarsku) analizu, sintezu i prepoznavanje govora, razumijevanje i strojno prevodenje prirodnog jezika te različite sintaksičke i semantičke modele. Bilo je prijavljeno oko 600 referata iz 32 zemlje, koji su prikazani u nekoliko simultanih sekcija. Istodobno se održavala izložba opreme i gotovih aplikacija (komercijalni sistemi za sintezu i prepoznavanje govora), a organiziran je posjet Francuskom centru za istraživanje tehnologije govora LIMSI-CNRS.

Konferencija je imala dva osnovna cilja:

- pregled dostignuća govorne tehnologije u svijetu, strukture istraživačkih timova i opreme

- udruživanje u zajednička istraživanja u sklopu projekata EUREKA (SPRIT) radi podjele područja istraživanja, uvođenje stand-

ardizacije opreme i govornog (digitaliziranog) materijala i pripreme za multilingvalni aspekt govornih komunikacija za Evropu bez granica.

Predstavlja su i rezultati pojedinih projekata ESPRIT (u sklopu projekta EUREKA definiranih na prethodnoj konferenciji održanoj u Edinburgu prije dvije godine). Konferenciji je dan velik publicitet, a za ograničen broj sudionika organizirano je i primanje kod gradonačelnika Pariza.

Radovi na konferenciji bili su podijeljeni u desetak globalnih kategorija: analiza i sinteza govora; obrada signala; obrada jezika; percepcija govora; prepoznavanje govora; fonetika; konektivni modeli; pomoć hendikepiranima itd. Za analizu i sintezu govora gotovo se uvijek upotrebljavaju različiti modaliteti LPC kodiranja govora, a bio je predstavljen i sistem za prijenos govora brzinom od samo 100 bita/sekundu. Na području obrade signala najviše radova odnosilo se na kompresiono kodiranje i ugradnju mehanizama za filtraciju radi eliminiranja šumova.

Posebnu pozornost izazvali su

radovi za analizu-prepoznavanje govora (nekoliko tisuća riječi) temeljni na konektivnim modelima. Općenito, konektivni je model skup elementarnih jedinica- procesora (ćelija), koji su međusobno povezani neposredno ili u slojevima. Za svaku ćeliju definira se lokalna energija (aktivacija, razina pobudenosti) koja ovisi o težini veza i stanja okolnih ćelija i interno stanje ćelije. Ovisno o algoritmu učenja veze se ažuriraju da bi se postigao globalni minimum mreže (funkcija cilja, energije itd), što se manifestira kao "učenje". Medutim, inherentna struktura neuronskih mreža odgovara učenju samo na percepcijskoj razini, odnosno na razini signala, što znači da konektivni modeli uglavnom zahtijevaju naknadnu simboličnu obradu.

Prateća izložba obilovala je komercijalnim proizvodima govorne tehnologije (sistemi za prepoznavanje i sintezu govora na osnovi teksta). Najveća gužva bila je oko štanda DRAGONDICTATE sistema za automatsko diktiranje, koji može, prema izjavama udruženja APRICOTE, prepoznavati do 30.000 riječi, tj. može neposredno služiti kao prvi komercijalni uređaj za automatsko diktiranje. DRAGONDICTATE je posebno prilagođen za englesko govorno područje (i to američka varijanta) a cijena mu je oko 20.000 USA \$. Sistem je prilagođen za unaprijed pripremljenog govornika, a točnost prepoznavanja može biti i do 95%.

Najviši domet u automatskoj sintezi govora postignut je za francuski jezik u CNET-u. Opisana su i načela automatskog informacijskog sistema MAIREVOX, koji funkcionira preko telefonske mreže u Lannionu. Naime, sistem za prepoznavanje govora neovisno o govorniku (cca 100 riječi)

te razrađena sinteza na osnovi teksta omogućili su da je u Lannionu nedavno stavljen u pogon kompjutorizirani sistem za davanje informacija o važnijim događajima u gradu (dežurne bolnice, apoteke, kazališne predstave itd.) kojeg opslužuje samo stroj - opremljen automatskom telefonskom centralom i sistemom za sintezu govora. naravno, način vođenja "razgovora" je formaliziran tkao d stroj uvijek ponudi nekoliko mogućnosti za informiranje, a time se smanjuju pogreške prepoznavanja. Takav način komunikacije znatno usporava mnoštvo potpitanja prije nego što se dobije željena informacija, ali radi bez prisustva čovjeka!

Ima li se na umu, da su od 600 članaka pripremljenih za konferenciju iz Jugoslavije bila samo 2, tada je izložbeni štand "Rudi Čajavca" iz Banjaluke pravo iznenađenje. U nedavno osnovanom laboratoriju za govornu tehnologiju LOGOS, počela je proizvodnja sklopova za kompresiono kodiranje govora. nakon digitalizacije poruka (u trajanju od 3 minute) sklop restaurira govorni signal ovisno o kombinaciji na adresnoj sabirnici. Zgodnu dosjetku predstavljao je i sat koji izgovara vrijeme nakon pritiska na određenu tipku. Iako, s teorijske strane, ti uređaji upotrebljavaju samo rudimentarnu sintezu govora (digitalno kodiranje i restauriranje signala) na toj je konferenciji učinjen značajan napredak za jugoslavensku govornu tehnologiju.

Ovdje treba istaknuti da je govorna tehnologija jedna od rijetkih tehnologija koja je čvrsto vezana za kulturu (jezik, govor) svake nacije pojedinačno. To znači d se ona ne može uvesti preko transfera tehnologije, jer npr. sistem za prpoznavanje govora za

engleski neće dobro raditi za hrvatski ili srpski i obratno. Naravno, postoje zajednička načela, ali određene specifičnosti znatno su prisutne naročito kada je riječ o sintezi govora na osnovi teksta. S obzirom na sve rečeno, jasno je da govornu tehnologiju za jugoslavenske jezike tek treba razvijati kao zadatak koji u bliskoj budućnosti trebaju rješavati

isključivo jugoslavenski stručnjaci.

Literatura:

1. "European Conference on Speech Communication and Technology" izdavači: J.P.Tubach, J. Mariani, CEP, Edinburgh 1989, vol. I/II.