

Sadržaj

1. Uvod	3	2. Teoretske osnove.....	4
.....	4	2.1. Govor	4
.....	4	2.2. Sinteza govora.....	5
.....	5	2.2.1. Načini sinteze govora.....	6
.....	6	3. Pregled postojećih alata za sintezu govora.....	7
.....	7	3.1. MBrola.....	7
.....	7	3.2. Festival.....	7
.....	7	3.3. WinSpeech.....	8
.....	8	3.4. BaBel Technologies	8
.....	8	3.4.1. Infovox Desktop	9
.....	9	3.5. GnuSpeech.....	9
.....	9	3.6. VoiceText	10
.....	10	3.7. Loquendo Text-to-Speech (TTS)	11
.....	11	3.8. TextAloud.....	12
.....	12	3.9. TextToSpeech Kit	13
.....	13	3.10. AlfaNum	13
.....	13	3.11. FIS	14
.....	14	4. Naš doprinos	16
.....	16	4.1. Program SintezaGovora.....	17
.....	17	5. Literatura.....	22

2

1. Uvod

3

2. Teoretske osnove

2.1. Govor

Govor predstavlja osnovni način komunikacije me u ljudima. Govorom čovjek prenosi svoje misli i osjećanja drugim ljudima. Čovjekov vokalni aparat je složen i obuhvata mnoge organe: pluća, dušnik, grkljan, glasne žice, usnu šupljinu sa jezikom, zubima i usnama, i nosnu šupljinu. Govor, posmatrano kao zvučni signal, nosi u sebi mnoštvo informacija. Osim onoga što je rečeno, u njemu su sadržane informacije o govorniku, njegovom polu, starosti, porijeklu, a nerijetko i zdravstvenom i emocionalnom stanju. Pri razumjevanju onog što je rečeno, čovjek koristi poznavanje jezika i razumjevanje konteksta. Percepcija govora nije čisto auditorni proces, odnosno, leksička segmentacija nije uro ena, već naučena sposobnost. Registrovanje dobijenih zvukova prvi je korak u percepciji govora. Mozak mora da odredi da li se zaista radi o zvukovima koji pripadaju govoru, jer se govor u mozgu obra uje na drugačiji način, nego što je to slučaj sa muzikom ili ambijentalnim šumom. Mozak tako e mora da odredi da li je u pitanju jezik koji slušalac razumije. Potom se vrši fonetska analiza sadržaja, i to u realnom vremenu, odnosno, ne čeka se da govornik završi da bi se dobila informacija o tome koji su glasovi izgovoreni. Pri ovome treba odbaciti negovorne elemente, kao što su: uzdasi, nakašljanje i sl. Rekonstrukcija sadržaja govorne cjeline vrši se na osnovu niza dobijenih glasova, ali vodeći računa i o semantičkom kontekstu, tako da će govorna cjelina najvjerojatnije biti ispravno rekonstruisana po značenju, čak i ako odre eni glasovi u njoj nedostaju ili su manje ili više oštećeni prilikom artikulacije, što je u spontanom govoru čest slučaj.

4

2.2. Sinteza govora

Sinteza govora predstavlja operaciju pretvaranja pisanog ulaza u govorni izlaz. Ulaz može biti u obliku grafemske, ortografske ili fonemske skripte, zavisno od izvora. Jednostavnije rečeno, sinteza govora je umjetno generisanje ljudskog govora. Sistemi koji se za to koriste nazivaju se sintetizatori govora, a mogu biti implementirani kao softver ili hardver. Sinteza govora često se kraće naziva Textto-Speech (TTS), s

obzirom na to da upravo i pretvaraju tekst u govor. Zadatak sinteze govora je da na osnovu ulazne informacije u tekstualnom obliku generiše govorni signal razumljiv čovjeku. To podrazumjeva i da bi sintetizovan govor morao da zvuči prirodno, odnosno, da bi trebalo da posjeduje intonaciju karakterističnu za prirodan ljudski govor. Primjene sintetizatora govora su mnogobrojne. Ovakvi sistemi su u aplikacijama računarske telefonije nezamjenljivo sredstvo za pružanje informacija koje se često mijenjaju i nepraktično je angažovati govornika koji bi trebalo da ih pročita (npr. e-mail). Sintetizatori govora su od izuzetne pomoći i osobama sa invaliditetom – pogotovo onima sa oštećenim govorom, koje zahvaljujući ovakvim sistemima mogu gotovo normalno da komuniciraju – čak i telefonom, kao i slijepim osobama, kojima je omogućeno da se samostalno služe računarom i na taj način se lakše uključe u svakodnevni život. Postoji nekoliko algoritama za sintezu govora. Izbor algoritma ovisi o operaciji koju želimo izvršiti. Najjednostavniji način je jednostavno snimiti glas osobe koja govori željene izraze, ali to predstavlja samo ograničen izvor fraza i rečenica. Kvaliteta zavisi o načinu snimanja. Sofisticiraniji, ali lošije kvalitete su algoritmi koji dijele govor u manje jedinice. Najčešće korištena jedinica je fonem, najmanja lingvistička jedinica. Zavisno od jeziku, postoji oko 35-50 fonema u zapadnoevropskim jezicima. Problem je u kombinovanju fonema, jer tečan govor zahtjeva tečan prelaz izme u elemenata (fonemskih jedinica). Razumljivost je stoga manja, no mala je i zahtjevnost memorije. Rješenje ovog problema je korištenje difona. Umjesto dijeljenja u prelazima, stanka se radi u sredini fonema, što ostavlja prelaze netaknute. To daje oko 400 elemenata i kvaliteta raste. Što su duže te jedinice, postoji više elemenata, ali uz potrebnu memoriju raste i kvaliteta. Ostale jedinice koje su u širokoj primjeni koriste su: poluslogovi, slogovi, riječi ili njihova kombinacija.

**----- OSTATAK TEKSTA NIJE PRIKAZAN. CEO RAD MOŽETE
PREUZETI NA SAJTU. -----**

www.maturskiradovi.net

MOŽETE NAS KONTAKTIRATI NA E-MAIL: maturskiradovi.net@gmail.com