

U N I V E R Z A V L J U B L J A N I

Fakulteta za elektrotehniko

Melita Hajdinjak

**VODENJE DIALOGA
MED ČLOVEKOM IN RAČUNALNIKOM
V NARAVNEM JEZIKU**

MAGISTRSKO DELO

Mentor: prof. dr. France Mihelič

Ljubljana, 2004

Ključne besede: sistemi za dialog, sistemi za podajanje informacij, podatkovna zbirka, pomenska analiza, semantični okvir, eksperiment Čarovnik iz Oza, modul za tvorjenje naravnega jezika, enosmerna cevovodna struktura, Griceove maksime, Shneidermanova zlata pravila oblikovanja sistemov, modul za vodenje dialoga, tehnike modeliranja dialoga, tehnika končnih avtomatov, tehnika polnjenja predalčkov, tehnika množice domen, tehnika usmerjanja k cilju, tehnika razumskih agentov, konverzacijska teorija iger, konverzacijske igre, konverzacijske poteze, rekurzivne mreže prehodov, predstavitev znanja, modalna logika, intuicionistična logika, modalni operatorji, Kripkejevi modeli, evalvacija modula za vodenje dialoga, ogrodje PARADISE, multipla linearna regresija

Zahvala

Prva zahvala gre najboljšemu mentorju sploh, prof. dr. Francetu Miheliču, ki me je toplo sprejel v raziskovalno skupino Laboratorija za umetno zaznavanje, sisteme in kibernetiko, bdel nad mojim delom, se čudil neizgovorljivim izrazom in me kljub vsemu podpiral pri vseh mojih odločitvah.

Hvala tudi vodji laboratorija, prof. dr. Nikoli Pavešiću, za številne prijazne in spodbudne besede, nenazadnje pa tudi za njegovo neustrašljivost in razigranost pri športnem udejstvovanju laboratorijske ekipe.

Zahvalila bi se še doc. dr. Andreju Bauerju, katerega dragoceni nasveti in opazke so bili pri pisanju tega dela nepogrešljivi.

Prav posebna hvala 'sosedu' Janezu, ki je 'oživljal' mene in moj računalnik, ko naju je (enega zaradi izčrpanosti, drugega zaradi moje neumnosti) napadel virus. V času podiplomskega študija in raziskovalnega dela mi je poleg Janeza bil v izjemno podporo tudi Boštjan, ki je znatno prispeval k prijetnemu delovnemu vzdušju. Hvala Boštjan! Hvala tudi Simonu, nepogrešljivemu in edinemu pravemu 'električistu', ki mi je kljub obtoževanju, da živim samo od zraka in ljubezni (ali pa ravno zaradi tega), vselej rad priskočil na pomoč. Tisočkrat hvala tudi Ankici, katere prihoda sem se vselej razveselila, pa ne (samo) zaradi daril in mehiških kosil, temveč (tudi) zaradi njene dobrosrčnosti in pozitivne energije.

(P)osebna zahvala pa gre mojemu Borisu ter Urški, Nevi, Simoni in bratu Dušanu, ki so (tako ali drugače) izkazali izjemen posluh, razumevanje in potrpežljivost, ne le za vse težave, s katerimi sem se srečevala, ampak tudi za moje prve (pol)maratonske podvige.

Nenazadnje sem zahvalo dolžna tudi svojim staršem, ki so mi ves čas stali ob strani in me vzpodbujali pri delu.

Povzetek

Vsebina magistrskega dela temelji na razvoju modula za vodenje dialoga in na predstavitvi znanja v dvojezičnem sistemu za podajanje vremenskih informacij, ki ga razvijamo raziskovalci Laboratorija za umetno zaznavanje, sisteme in kibernetiko Fakultete za Elektrotehniko v Ljubljani ter raziskovalci Oddelka za računalništvo in informatiko Filozofske fakultete na Reki.

Podali bomo pregled zgodovine nastanka in razvoja te najbolj razširjene oblike govornih vmesnikov, namreč sistemov za podajanje informacij, pri čemer uporabo govora upravičimo z njegovo naravnostjo, učinkovitostjo in fleksibilnostjo.

Vir znanja nastajajočega sistema za podajanje informacij o vremenu in vremenskih napovedih predstavljajo spletne strani *Hidrometeorološkega zavoda Agencije Republike Slovenije za okolje*. Del podatkov, ki je na teh spletnih straneh podan v tekstovni obliki, s pomensko analizo pretvorimo v bolj pregledno tabelarno obliko. Zaradi časovno odvisne strukture teh vremenskih informacij, tabelarno obliko kasneje pretvorimo v množico svetov in v njih veljavnih izjav. Takšna predstavitev znanja, ki jo poimenujemo *dialog mIPC Kripkejev model*, omogoča sistemu, da uporabnika usmerja k izbiri relevantnih, dosegljivih delčkov informacij.

Preden se lotimo gradnje modula za vodenje dialoga, ki je osrednji del vsakega sistema za dialog, s pomočjo *eksperimenta Čarovnik iz Oza* pridobimo zbirko podatkov, ki odraža jezik dialogov človek – računalnik. Po pregledu uveljavljenih tehnik modeliranja dialoga se odločimo za tehniko usmerjanja k cilju, ki jo delno formaliziramo s pomočjo *konverzacijske teorije iger*, podatke pa pridobivamo s polnjenjem predalčkov. V tem modelu je dialog sestavljen iz *konverzacijskih iger* in *konverzacijskih potez*, ki predstavljajo različne strategije in so udejanjene kot *rekurzivne mreže prehodov*. Da bi dosegli čim večjo kooperativnost in čim večjo neodvisnost od domene, množico konverzacijskih iger in potez, ki se pojavljajo v literaturi, ustrezno razširimo.

Po dograditvi modula za vodenje dialoga, v namen evalvacije, ponovno izvedemo eksperiment Čarovnik iz Oza. Na tako pridobljenih podatkih uporabimo potencialno splošno metodologijo evalvacije sistemov za dialog, namreč ogrodje PARADISE, ki omogoča izpeljavo ocene učinkovitosti sistema kot uteženo linearno kombinacijo od domene odvisnih *parametrov uspešnosti naloge* in *cen dialoga*. Pri evalvaciji podatkov iz obeh eksperimentov Čarovnik iz Oza pridemo do zaključka, da lahko učinkovitost modula za vodenje dialoga opazno izboljšamo, če zagotovimo čim večjo fleksibilnost sistema pri iskanju dosegljivih, relevantnih podatkov. To fleksibilnost sicer v precejšnji meri že zagotavlja dialog mIPC Kripkejev model znanja, zaradi velikega vpliva na zadovoljstvo uporabnikov pa na koncu predlagamo še nekatere izboljšave.

Abstract

The content of this thesis is based on the construction of the dialogue manager and the development of the knowledge representation in a bilingual spoken dialogue system for weather-information retrieval, which represents a joint project of the Laboratory of Artificial Perception, Systems, and Cybernetics of the Faculty of Electrical Engineering in Ljubljana and the Department of Information Science of the Faculty of Philosophy in Reka.

We give a brief overview of the evolution of the most widely used type of speech interfaces, i.e. information-providing dialogue systems. The use of speech is justified by its naturalness, effectiveness, and flexibility.

The knowledge source of the developing dialogue system for weather-information retrieval are the web pages of the *Environmental Agency of the Republic of Slovenia*. The textual part of this data is transformed into table data structure to ensure greater readability. Due to time-dependency of the data structure this data is further transformed into several sets of valid statements, each corresponding to one of the possible worlds. This knowledge representation, i.e. *dialogue mIPC Kripke model*, enables the system to direct the user to select relevant, available data.

Before constructing the dialogue manager, which is the central module of any dialogue system, we conduct the *Wizard-of-Oz experiment* to collect dialogues that reflect the language of human-computer communication. Considering the most accepted dialogue-modeling techniques we decide to implement the goal-directed technique, which is partly formalized by using *conversational game theory*. The data, on the other hand, is acquired using the slot-filling approach. In this model, the dialogues consists of exchanges, called *conversational games*, which are analyzed as *conversational moves* and implemented as *recursive transition networks*. To achieve better cooperativity and a greater portability we expand the set of conversational games and conversational moves that have been used before.

After completing the construction of the dialogue manager we once more conduct the Wizard-of-Oz experiment. A potential general methodology for evaluating spoken dialogue systems, i.e. the PARADISE framework, is used to derive a combined performance metric as a weighted linear combination of a task-based success measure and dialogue costs. The evaluation of data collected during the Wizard-of-Oz experiments leads to the conclusion that user satisfaction can be remarkably improved by assuring the system to be as flexible as possible in directing the user to select relevant, available data. Although, this flexibility is in no small measure supported by the non-traditional dialogue mIPC Kripke model, some additional improvements of this model are proposed.

Kazalo

Zahvala	iii
Povzetek	v
Abstract	vii
1 Uvod	1
1.1 Motivacija	2
1.2 Sistemi za dialog	3
1.3 Zgodovina sistemov za podajanje informacij	4
1.4 Razvoj dvojezičnega sistema za podajanje informacij o vremenu	5
2 Zbirka vremenskih podatkov	7
2.1 Vir podatkov	8
2.2 Zajemanje spletnih podatkov	8
2.3 Pomenska analiza	12
2.4 Zbirka podatkovnih tabel	15
3 Eksperiment Čarovnik iz Oza	18
3.1 Zakaj eksperiment Čarovnik iz Oza	19
3.2 Sistem Čarovnik iz Oza	20
3.3 Prvi eksperiment Čarovnik iz Oza	24
3.4 Drugi eksperiment Čarovnik iz Oza	28
4 Oblikovanje in razvoj sistemov za dialog	31

4.1	Kdaj je sistem za dialog uporaben	32
4.2	Vodila pri oblikovanju sistemov za dialog	33
4.3	Tehnike modeliranja dialoga	35
5	Modeliranje dialoga z uporabo konverzacijske teorije iger	40
5.1	Konverzacijska teorija iger	41
5.2	Modeliranje dialoga v sistemu za podajanje informacij o vremenu . . .	44
6	Predstavitev znanja v sistemu za dialog	52
6.1	Modalna logika	53
6.2	Intuicionistična modalna logika	58
6.3	Predstavitev vremenskih podatkov	62
7	Evalvacija modula za vodenje dialoga	76
7.1	Ogrodje PARADISE	77
7.2	Rezultati evalvacije dialogov iz prvega eksperimenta Čarovnik iz Oza .	80
7.3	Rezultati evalvacije dialogov iz drugega eksperimenta Čarovnik iz Oza .	84
8	Zaključek	87
8.1	Razširitev predstavitve znanja	88
8.2	Nadaljnje delo	90
	Viri in literatura	92
	Terminološki slovar	101

Slike

1.1	Zgradba sistema za dialog	4
1.2	Dvojezični sistem za podajanje informacij o vremenu	6
2.1	Naraščanje števila različnih besed v zbirki <i>Vremenska napoved in obeti za Slovenijo</i>	9
2.2	Naraščanje števila različnih besed v zbirki <i>Vremenska napoved za gorski svet</i>	9
2.3	Naraščanje števila različnih besed v zbirki <i>Napoved za letalstvo</i>	10
2.4	Naraščanje števila različnih besed v zbirki <i>Napoved vodostajev</i>	10
2.5	Naraščanje števila različnih besed v zbirki <i>Snežne razmere</i>	10
2.6	Model pomenskih analizatorjev vremenskih podatkov	13
2.7	Primer podatkovne tabele, zgrajene s pomočjo analizatorja <i>Vremenska napoved in obeti za Slovenijo</i>	16
2.8	Grafični vmesnik do zbirke podatkovnih tabel	16
3.1	Grafični vmesnik sistema Čarovnik iz Oza – primer delovnega okna 1	22
3.2	Grafični vmesnik sistema Čarovnik iz Oza – primer delovnega okna 2	22
4.1	Preprost primer dialoga, modeliranega s končnim avtomatom	36
4.2	BDI model	38
5.1	Primer zemljevida, uporabljenega v spontanah, govorjenih dialogih človek – človek	42
5.2	Konverzacijske igre, udejanjene v modulu za vodenje dialoga v sistemu za podajanje informacij o vremenu	47
6.1	Hassejev diagram relacije <i>biti manj specifičen</i> med krajevnimi delčki informacij, ki se navezujejo na Slovenijo	65

6.2	Hassejev diagram relacije <i>biti manj specifičen</i> med krajevnimi delčki informacij, ki se navezujejo na ostalo Evropo	65
6.3	Hassejev diagram relacije <i>biti manj specifičen</i> med časovnimi delčki informacij. Zaradi majhnega števila možnih dosegljivih časovnih trenutkov, so vzpostavljene vse povezave, zaradi česar relacija <i>biti manj specifičen</i> v tem primeru izgubi pomen, ki izvira iz imena.	66
6.4	Hassejev diagram relacije <i>biti manj specifičen</i> med podatkovnimi delčki informacij	66
7.1	Model učinkovitosti sistema za dialog, ki ga zajema ogrodje PARADISE	78
7.2	Primer kontingenčne tabele dialoga	81
8.1	Hassejev diagram relacije <i>biti manj specifičen</i> (povezave črne barve) in relacije sosednosti (povezave sive barve) med krajevnimi delčki informacij, ki se navezujejo na Slovenijo	88
8.2	Hassejev diagram relacije <i>biti manj specifičen</i> (povezave črne barve) in relacije sosednosti (povezave sive barve) med krajevnimi delčki informacij, ki se navezujejo na ostalo Evropo	89
8.3	Hassejev diagram relacije <i>biti manj specifičen</i> (povezave črne barve) in relacije sosednosti (povezave sive barve) med podatkovnimi delčki informacij	89

Tabele

2.1	Primeri vnosov v slovar <i>Vremenska napoved in obeti za Slovenijo</i>	11
2.2	Rezultati testiranja pomenskega analizatorja <i>Vremenska napoved in obeti za Slovenijo</i>	14
3.1	Primer tvorjenja povedi z modulom za tvorjenje naravnega jezika	23
3.2	Del vprašalnika, s katerim so udeleženci eksperimenta ocenjevali sistem Čarovnik iz Oza	25
3.3	Povprečne ocene sistema Čarovnik iz Oza	25
3.4	Primer dialoga med uporabnikom in sistemom Čarovnik iz Oza	26
3.5	<i>Zgradba SSSQ zbirke spontanega govora</i>	27
3.6	<i>Rezultati analize dialogov, pridobljenih v prvem eksperimentu Čarovnik iz Oza</i>	28
3.7	Povprečne ocene v drugem eksperimentu Čarovnik iz Oza	29
3.8	<i>Zgradba SSSQ2 zbirke spontanega govora</i>	29
3.9	<i>Rezultati analize dialogov, pridobljenih v drugem eksperimentu Čarovnik iz Oza</i>	30
5.1	<i>Primer dialoga med uporabnikom (U) in sistemom za dialog (S)</i>	48
6.1	<i>Interpretacija modalnega operatorja \diamond pri dani interpretaciji \square</i>	56
6.2	<i>Pomen relacije R pri različnih interpretacijah izjave $\square A$</i>	57
7.1	<i>Srednje vrednosti parametrov v prvem (WOZ1) in v drugem (WOZ2) eksperimentu Čarovnik iz Oza</i>	82
7.2	<i>Korelacijski koeficienti parametrov v prvem eksperimentu</i>	83
7.3	<i>Korelacijski koeficienti parametrov v drugem eksperimentu</i>	84

1 Uvod

1.1 Motivacija

1.2 Sistemi za dialog

1.3 Zgodovina sistemov za podajanje informacij

1.4 Razvoj dvojezičnega sistema za podajanje informacij o vremenu

V uvodnem poglavju bomo podali glavne motive razvoja govornih vmesnikov in opisali njihovo zgradbo.

Nadaljevanje poglavja, kakor tudi celotnega magistrskega dela, bo namenjeno predvsem sistemom za podajanje informacij. Podali bomo pregled zgodovine nastanka in razvoja takšnih sistemov ter opisali dvojezični sistem za podajanje vremenskih informacij, ki ga razvijamo raziskovalci Laboratorija za umetno zaznavanje, sisteme in kibernetiko Fakultete za Elektrotehniko v Ljubljani ter raziskovalci Oddelka za računalništvo in informatiko Filozofske fakultete na Reki.

Vsebina magistrskega dela temelji na razvoju modula za vodenje dialoga in na predstavitvi znanja v tem dvojezičnem sistemu za podajanje informacij.

1.1 Motivacija

Že pred nekaj desetletji se je med ljudmi pojavila želja, da bi s stroji komunicirali z govorom. Znanstvena fantastika daje številčen vir takšnih govornih vmesnikov, ki sega od izjemno inteligentnega in zelo komunikativnega računalnika HAL v Stanley Kubrickovem filmu *2001: A Space Odyssey* (1968) do govorečega avta K.I.T.T., ki je v nanizanki *Knight Rider* (1982-1986) Davidu Hasselhoffu pomagal pri boju proti zločinu.

Po drugi strani pa komunikacija s stroji že dolgo predstavlja izziv tudi raziskovalcem na področju umetne inteligence. To zanimanje izvira iz znamenitega članka o testu računalniške inteligence [Turing-1950], ki ga je leta 1950 objavil Alan Turing. Ta test, ki je trenutno znan pod imenom *Turingov test*, je v bistvu igra, v kateri se sodnik in dva agenta pogovarjata o poljubno izbrani temi, na podlagi katere mora sodnik presoditi, kateri izmed agentov je človek in kateri je računalnik. Osnovna ideja Turingovega testa je, da je vsak računalniški program, ki se ga po njegovem obnašanju med dialogom da zamenjati za človeka, inteligenten. Zanimiva je Turingova napoved, da bo večina računalnikov leta 2000 sposobna ta test uspešno prestati.

Eden izmed prvih računalniških programov, ki je uspešno prestal Turingov test, je računalniški terapevt ELIZA [Weizenbaum-1966], ki ga je razvil Joseph Weizenbaum v sredini šestdesetih let. Primer dialoga med pacientko in ELIZO je podan spodaj:

Pacientka: Moški so vsi enaki.
 ELIZA: V kakšnem smislu?
 Pacientka: Vedno nas prosijo za kaj.
 ELIZA: Ali se spomnite kakšnega primera?
 Pacientka: Moj fant me je pripravil do tega, da sem prišla sem.
 ELIZA: Nadaljujte.
 Pacientka: Mogoče bi se lahko naučila shajati z mamo.
 ELIZA: Povejte mi kaj več o svoji družini.

Ta dialog na prvi pogled zgleda precej impresiven, čeprav je program presenetljivo enostaven. Vse ELIZINE konverzacijske sposobnosti so shranjene v urejeni zbirki ključnih besed, katere vhodi so na primer:

beseda	stopnja	vzorec	izhod
mama	10	?X	"Povejte mi kaj več o svoji družini."
so	3	?X so ?Y	"Kaj pa, če ne bi bili ?Y."

Torej, vsaka beseda je povezana s stopnjo (maksimalna vrednost je 10), z vhodnim vzorcem in primernim izhodom. Celoten algoritem ELIZE lahko opišemo takole:

- V vhodni povedi poišči ključne besede, katerih vzorci v zbirki ključnih besed se ujema z vzorci na vhodu.
- Če je takšnih besed več, izberi besedo z največjo stopnjo.
- S pomočjo definicije izhoda tvori odziv.

- Če v zbirki ni ustreznih ključnih besed, povej "povejte mi več", "hm, nadaljujte" ali podobno.

Zaradi preprostosti tega algoritma ELIZI ne uspe preslepiti dovolj skeptičnih ljudi. Nekateri odzivi namreč niso smiselni, čez nekaj časa pa postane tudi jasno, da sistem ne sledi zgodovini in razvoju dialoga, saj nima prav nobenega znanja niti o poteku dialoga niti o jeziku samem.

Uporaba govornih vmesnikov v znanstveni fantastiki pa ne izvira le iz želje, da bi se stroji obnašali kot ljudje, da bi bili inteligentni, ampak je globlje zasidrana. Govor je naraven – govoriti se naučimo še preden znamo brati in pisati. Govor je tudi učinkovit – večina ljudi je sposobnih govoriti petkrat hitreje kot tipkati in verjetno celo desetkrat hitreje kot pisati. Še več, govor je fleksibilen – med komunikacijo se nam ni treba ničesar dotikati in ne opazovati. To so le nekateri razlogi, zakaj uporaba govornih vmesnikov prispeva k izboljšanju kvalitete življenja.

1.2 Sistemi za dialog

Sistem za dialog ali govorni vmesnik imenujemo računalniški sistem, ki uporabniku omogoča, da z govorom dostopa do določenih aplikacij. Poznamo več vrst sistemov za dialog, najbolj razširjeni pa so *sistemi za podajanje informacij*, *sistemi za govorno upravljanje naprav*, *sistemi za interaktivne govorne odzive* in *sistemi za reševanje problemov*. Slikovit pregled najpogostejših sistemov za dialog je podan v [Krahmer-2001]. Različne vrste sistemov za dialog so povezane z različnimi tehničnimi ovirami (npr. velikost slovarja in perpleksnost učnega besedila) in kar je bolj pomembno, zahtevajo uporabo različnih strategij vodenja dialoga.

Sistem za dialog je ponavadi modularno zgrajen, pri čemer je izhod iz prejšnjega modula vhod v naslednji modul. Osnovni moduli takega sistema (slika 1.1) so:

Modul za razpoznavanje govora

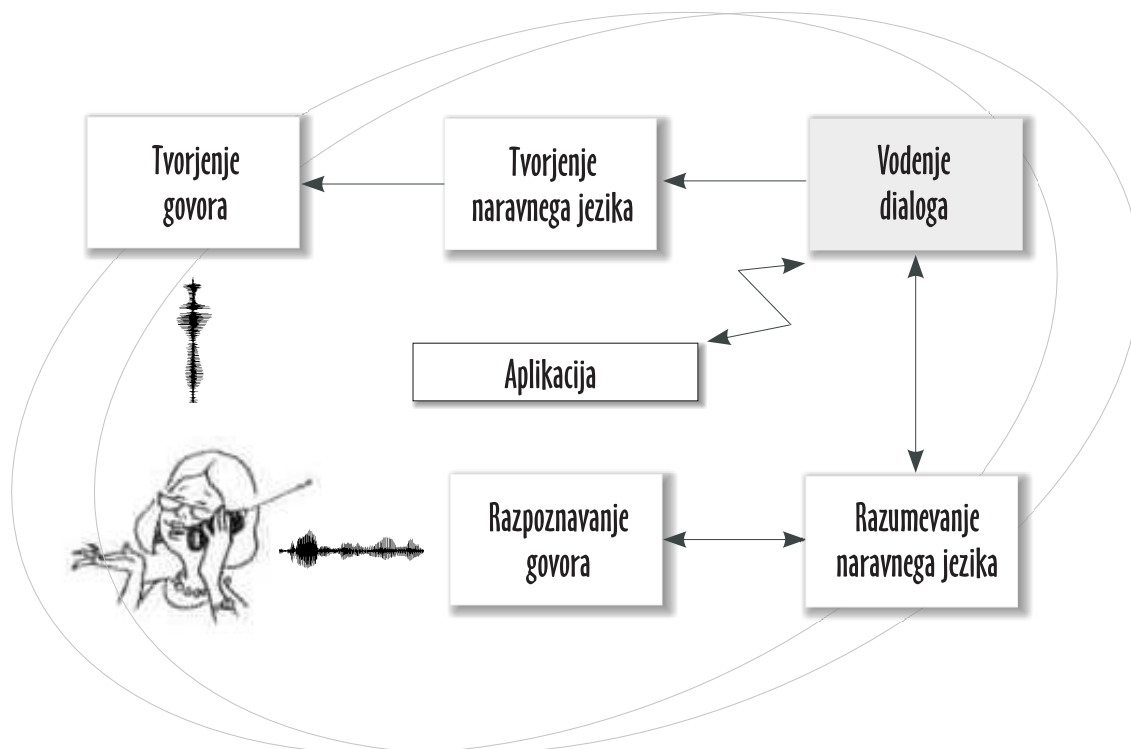
Vhod v modul za razpoznavanje govora je govorni signal uporabnika, njegova naloga pa je določiti zaporedje besed, ki jih je uporabnik izgovoril. V prvi fazi lahko to pomeni določiti mrežo hipotez o besedah, iz katere modul za razpoznavanje govora šele s pomočjo pomenske analize posameznih zaporedij, ki jo opravlja modul za razumevanje naravnega jezika, izloči zaporedje besed, ki se zdi najverjetnejše.

Modul za razumevanje naravnega jezika

Vhod v modul za razumevanje naravnega jezika je zaporedje besed, izhod pa njegova pomenska predstavitev. Pri razumevanju je lahko v veliko pomoč tudi informacija o poteku dialoga, ki jo dobimo iz modula za vodenje dialoga.

Modul za vodenje dialoga

Modul za vodenje dialoga je zagotovo osrednji modul vsakega sistema za dialog, ki vodi in uravnava interakcijo med uporabnikom in aplikacijo. Vhod v modul je pomenska predstavitev, ki strukturirano podaja vsebino uporabnikove izjave. Ker ima neposreden dostop do aplikacije, lahko modul za vodenje dialoga na podlagi vhodne pomenske



Slika 1.1: Zgradba sistema za dialog

predstavitve na primer spreminja stanje aplikacije (v primeru govornega upravljanja neke naprave) ali pa iz podatkovne zbirke vrača določene informacije (v primeru podajanja informacij).

Modul za tvorjenje naravnega jezika

Vhod v modul za tvorjenje naravnega jezika je pomenska predstavitev, ki jo tvori modul za vodenje dialoga, njegova naloga pa je, da jo pretvori v naravni jezik.

Modul za tvorjenje govora

Naloga modula za tvorjenje govora je, da odziv sistema, podan v naravnem jeziku, pretvori v govor. Ker je uporabnik z govornim izhodom v neposrednem stiku, ta močno vpliva na uporabnikovo oceno učinkovitosti sistema.

V nadaljevanju se bomo osredotočili na sisteme za podajanje informacij, t.j. govorne vmesnike za dostop do podatkovnih zbirk.

1.3 Zgodovina sistemov za podajanje informacij

Prototipi sistemov za podajanje informacij so se pojavili že v poznih šestdesetih in zgodnjih sedemdesetih letih. Najbolj znan sistem iz tega obdobja je LUNAR [Woods-1972], ki podaja informacije o kemijskih analizah kamnov z Lune. Ker so bili LUNAR in drugi zgodnji sistemi za podajanje informacij zgrajeni za specifično podatkovno zbirko, jih ni bilo mogoče na preprost način prirediti za druge podatkovne zbirke. Še več sistemov

za podajanje informacij se je pojavilo v poznih sedemdesetih letih. Omenimo samo sistem RENDEZVOUS [Codd-1974], ki je uporabnika vodil pri zastavljanju vprašanj, in sistem LADDER [Hendrix-1978], namenjen obsežnim podatkovnim zbirkam.

Eden izmed najbolj znanih sistemov iz zgodnjih osemdesetih let je v prologu udejanjen sistem CHAT-80 [Warren-1982]. V sredini osemdesetih let pa so postali sistemi za podajanje informacij zelo priljubljeno raziskovalno področje. Posledica tega je bilo veliko število udejanjenih sistemov, pri čemer je bil večji del raziskav namenjen prenosljivosti. Tako se je na primer sistem ASK [Thompson-1983], ki je dostopal do več zunanjih podatkovnih zbirk, od uporabnika učil nove besede in nove koncepte na različnih točkah interakcije, sistem JANUS [Resnik-1989] pa je dostopal do različnih aplikacij (podatkovne zbirke, ekspertni sistemi, grafične naprave, ipd.).

Kljub nekaterim izjemnim lastnostim številnih sistemov za podajanje informacij, ki so bili razviti v sredini osemdesetih let, si ti niso uspeli pridobiti pričakovanega hitrega in širokega javnega priznanja. Čeprav je za nekatere sisteme veljalo, da so bili komercialno uspešni, so sistemi za podajanje informacij še dandanes prej eksotična kot pa standardna izbira dostopa do podatkovne zbirke. To je delno posledica tako razvoja učinkovitih alternativ h govornim vmesnikom, kot je grafični vmesnik, kot preslabe zanesljivosti razumevanja naravnega govora.

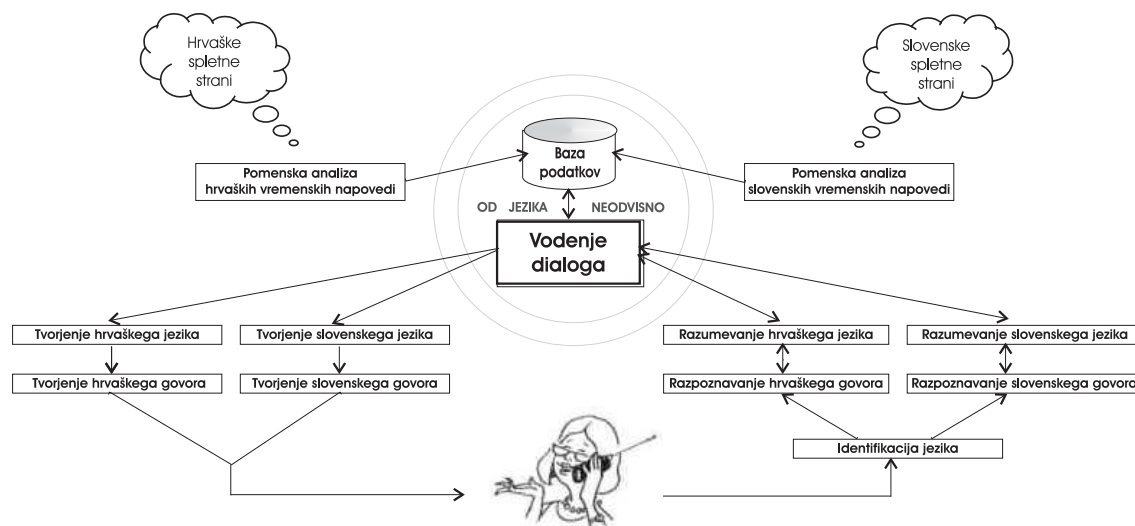
Že v poznih osemdesetih, predvsem pa v devetdesetih letih, je število objavljenih člankov o sistemih za dialog drastično upadlo. Kljub temu so raziskovalci sisteme za dialog še naprej izboljševali, vključevali so napredke in nova spoznanja na področju obdelave naravnega govora ter raziskovali prednosti vključitve dveh modalnosti, t.j. govora in grafike. Postal je jasno, da prihodnost govornih vmesnikov temelji na potrebi po dostopu do informacij preko telefona, na potrebi po uporabi govora, ko so roke ali oči kako drugače zaposlene, ter kot pomoč invalidnim osebam. Sodobnejši sistemi za podajanje informacij preko telefona so tako bili zgrajeni za različne domene, npr. za podajanje informacij o restavracijah [Jurafsky-1994], o gledališčih [Hoeven-1995], o potovanjih z železnico [Allen-1995, Sturm-1999], o letalskih prevozih [Ipšić-1999, Stallard-2000], o vremenskih napovedih [Zue-2000] in prispeli elektronski pošti [Walker-2000].

Cilj sodobnejših sistemov za dialog je strategije interakcije čimbolj približati naravnim strategijam, t.j. strategijam komunikacije človek – človek. To pomeni, da se je po eni strani potrebno omejiti le na najpogostejše zaznane naravne strategije, po drugi strani pa je potrebno razvijati in uporabljati najnovejše oz. najučinkovitejše tehnologije na področju obdelave govornega signala in razumevanja naravnega jezika.

1.4 Razvoj dvojezičnega sistema za podajanje informacij o vremenu

Raziskovalci Laboratorija za umetno zaznavanje, sisteme in kibernetiko Fakultete za Elektrotehniko v Ljubljani ter raziskovalci Oddelka za računalništvo in informatiko Filozofske fakultete na Reki smo si zastavili cilj razviti sistem [Žibert-2003], ki bo preko telefona v slovenskem in v hrvaškem jeziku podajal informacije o vremenu in

vremenskih napovedih. Eden izmed razlogov za to odločitev je veliko število turističnih izmenjav med obema državama, drugi pa čedalje manjši delež prebivalcev obeh držav, ki govorijo in razumejo oba jezika. Načrtovana zgradba sistema je prikazana na sliki 1.2. Sistem naj bi bil sposoben odgovarjati na vprašanja o vremenu in vremenski napovedi (splošna vremenska napoved, biovremenska napoved, temperatura, smer in hitrost vetra, zračni tlak, vidljivost, višina in vrsta snega, vzhod in zahod sonca, ipd.) za različne kraje in pokrajine v Sloveniji in na Hrvaškem ter za večja mesta ostalih evropskih držav.



Slika 1.2: Dvojezični sistem za podajanje informacij o vremenu

Ker naj bi sistem omogočal hkratno komunikacijo v obeh jezikih, t.j. v slovenskem in v hrvaškem jeziku, bo uporabnikova izjava najprej potovala v modul za identifikacijo jezika, ki bo ugotovil, v katerem jeziku uporabnik govori. V skladu s to ugotovitvijo bo nato izjava posredovana v ustrezno zaporedje modulov za razpoznavanje govora [Martinčič-2003] in razumevanje naravnega jezika, pri čemer bosta oba modula za razumevanje naravnega jezika pri pretvorbi izjave v njeno pomensko predstavitev upoštevala tudi trenutno stanje dialoga. Modul za vodenje dialoga bo (če bo potrebno), neodvisno od jezika, v zbirki vremenskih podatkov poiskal ustrezne podatke in jih, primerno strukturirane, poslal ustreznemu modulu za tvorjenje naravnega jezika. Ta bo to pomensko predstavitev odziva pretvoril v ustrezen (slovenski ali hrvaški) naravni jezik, ki bo nato še sintetiziran [Vesnicher-2003], t.j. umetno pretvorjen v govor, in posredovan uporabniku preko telefonske linije.

V magistrskem delu se bomo poleg modula za vodenje dialoga osredotočili tudi na zbirko vremenskih podatkov, ki predstavlja takorekoč celotno znanje našega sistema. Vir slovenskega dela podatkovne zbirke so spletne strani *Hidrometeorološkega zavoda Agencije Republike Slovenije za okolje*¹, vir hrvaškega dela pa spletne strani *Državnega hidrometeorološkega zavoda Republike Hrvaške*².

¹Hidrometeorološki zavod Agencije Republike Slovenije za okolje, http://www.arso.gov.si/podrocja/vreme_in_podnebje/

²Državni hidrometeorološki zavod Republike Hrvaške, <http://www.tel.hr/dhmz/>

2 Zbirka vremenskih podatkov

- 2.1 Vir podatkov
 - 2.2 Zajemanje spletnih podatkov
 - 2.3 Pomenska analiza
 - 2.4 Zbirka podatkovnih tabel
-

Opisali bomo nastanek zbirke podatkovnih tabel, ki predstavlja znanje nastajajočega sistema za podajanje informacij o vremenu in vremenskih napovedih. Pri tem se bomo omejili na slovenski del podatkov.

Podatke o vremenu in vremenski napovedi zajemamo s spletnih strani *Hidrometeorološkega zavoda Agencije Republike Slovenije za okolje*. Del podatkov je na teh spletnih straneh podan v tabelarni, del pa v tekstovni obliki. Ker želimo, da bo sistem sposoben jedrnato odgovarjati na uporabnikova vprašanja, je potrebno tudi tekstovno podan vir informacij predstaviti čimbolj strukturirano in pregledno. V ta namen smo zgradili pomenske analizatorje besedil, ki tekstovno podane informacije pretvorijo v tabelarno strukturo.

Pri izdelavi pomenskih analizatorjev smo si pomagali z zbirko spletnih podatkov, pridobljeno s skoraj enoletnim zajemanjem vremenskih podatkov s spletnih strani *Hidrometeorološkega zavoda Agencije Republike Slovenije za okolje*.

2.1 Vir podatkov

Po zgledu podobnih sistemov kot je JUPITER [Zue-2000], bodo vir podatkov našega dvojezičnega sistema za podajanje vremenskih informacij spletne strani *Hidrometeorološkega zavoda Agencije Republike Slovenije za okolje* ter spletne strani *Državnega hidrometeorološkega zavoda Republike Hrvaške*. V magistrskem delu se bomo omejili le na slovenski vir podatkov.

Del vremenskih podatkov je na svetovnem spletu podan v tabelarni, del pa v tekstovni obliki. Tipičen primer tekstovno podane vremenske napovedi za Slovenijo izgleda takole:

Napoved za Slovenijo

Danes bo sončno, burja na Primorskem bo ponehala. Najvišje dnevne temperature bodo od 20 do 26, na Primorskem do 29 stopinj C. Tudi jutri bo sončno, popoldne bo predvsem na zahodu spremenljive oblačnosti več. Proti večeru ter ponoči lahko nastanejo posamezne nevihte. Najnižje jutranje temperature bodo od 8 do 15, najvišje dnevne od 22 do 29 stopinj C.

Obeti

V nedeljo bo sprva še delno do spremenljivo oblačno, občasno so še možne manjše krajevne padavine, popoldne pa se bo postopno zjasnilo. V ponedeljek bo spet sončno in še topleje.

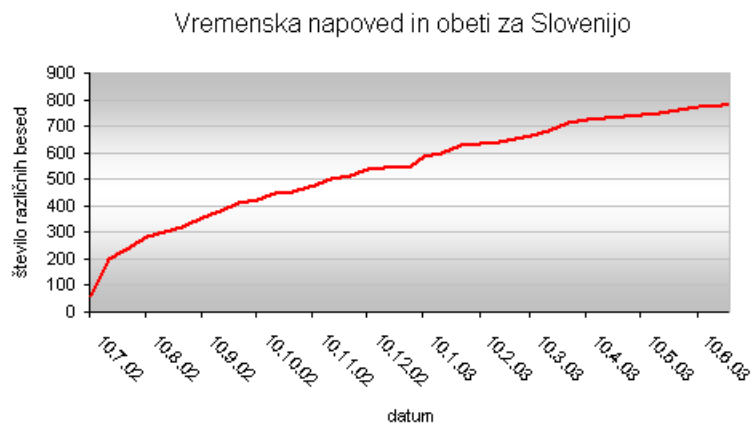
Takšni tekstovno podani podatki pogosto vsebujejo veliko podrobnejše informacije, kot pa uporabnika zanimajo. Če želimo, da bo sistem za podajanje informacij sposoben ponuditi le zahtevano informacijo in uporabnika ne bo dolgočasil z dolgimi monologi, mora takšno besedilo *razumeti*. S pojmom *razumeti* ne mislimo, da bo računalnik razmišljal tako, kot razmišlja človek. Mislimo le, da bo informacije pravilno uporabljal in se pravilno odzival. V ta namen smo po zgledu [Pepelnjak-1996a, Pepelnjak-1996b] zgradili pomenske analizatorje, ki vremenske napovedi analizirajo in jih zapišejo v za računalnik bolj pregledni in berljivi tabelarni strukturi [Hajdinjak-2002a].

Pri izdelavi pomenskih analizatorjev za tematsko omejena področja *Vremenska napoved in obeti za Slovenijo*, *Vremenska napoved za gorski svet*, *Napoved za letalstvo*, *Napoved vodostajev* in *Snežne razmere* smo si pomagali z zbirko spletnih podatkov, ki smo jo pridobili s skoraj enoletnim zajemanjem vremenskih podatkov s spletnih strani *Hidrometeorološkega zavoda Agencije Republike Slovenije za okolje*.

2.2 Zajemanje spletnih podatkov

Zajemanje vremenskih podatkov s spletnih strani *Hidrometeorološkega zavoda Agencije Republike Slovenije za okolje* je potekalo od sredine julija leta 2002 do sredine junija leta 2003.

Pet HTML datotek, za katere smo se odločili izdelati pomenske analizatorje, smo zajemali večkrat na dan (glede na pogostost osveževanja podatkov na spletnih straneh). Te datoteke smo obdelovali s programom, zapisanim v skriptnem jeziku PERL, ki je iz njih odstranil vse posebnosti HTML formata. Tako smo dobili tekstovne datoteke, iz katerih smo izluščili vse besede, jih prešteli in uredili po abecednem vrstnem redu. To nam je omogočilo spremljati naraščanje števila vseh različnih besed na tematsko omejenih področjih *Vremenska napoved in obeti za Slovenijo* (slika 2.1), *Vremenska napoved za gorski svet* (slika 2.2), *Napoved za letalstvo* (slika 2.3), *Napoved vodostajev* (slika 2.4) in *Snežne razmere* (slika 2.5).

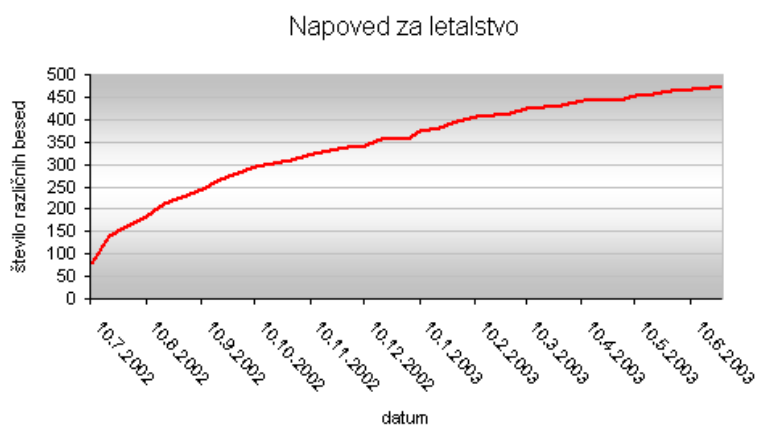
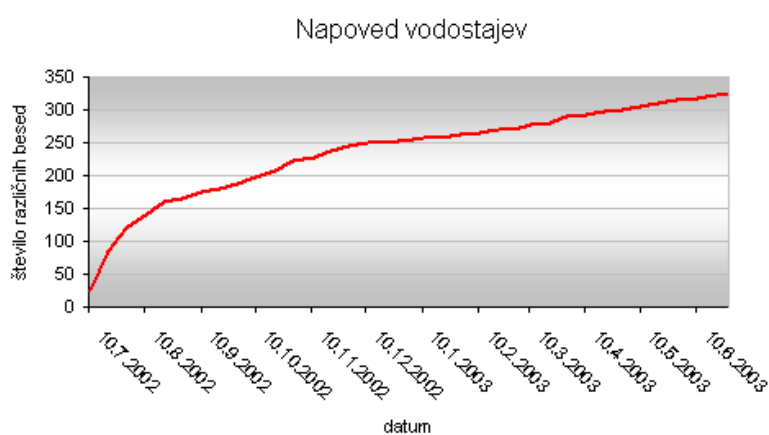
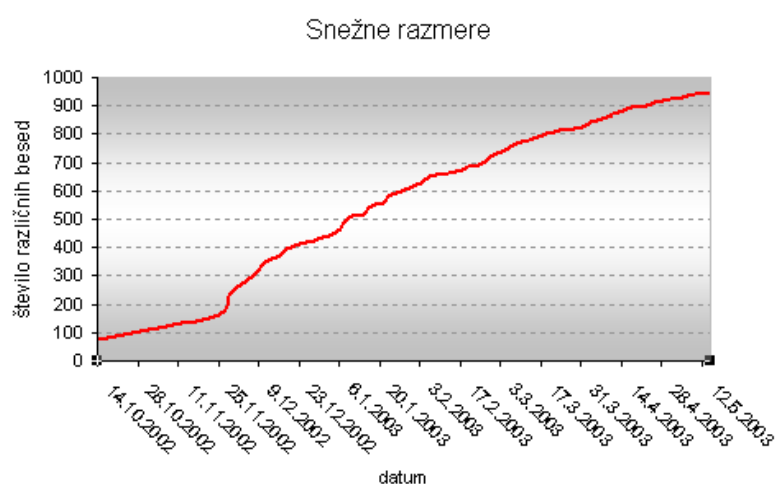


Slika 2.1: Naraščanje števila različnih besed v zbirki *Vremenska napoved in obeti za Slovenijo*



Slika 2.2: Naraščanje števila različnih besed v zbirki *Vremenska napoved za gorski svet*

Število vseh besed v posameznih podatkovnih zbirkah je z razliko od števila različnih besed po pričakovanju naraščalo približno linearno. Krivulje, ki prikazujejo naraščanje števila različnih besed v posameznih podatkovnih zbirkah, so si, z izjemo krivulje za *Snežne razmere*, sicer vse zelo podobne, vendar daleč od linearosti. Po nekaj mesecih zajemanja podatkov so se namreč vse začele opazno umirjati, kar je še najbolj izrazito

Slika 2.3: Naraščanje števila različnih besed v zbirki *Napoved za letalstvo*Slika 2.4: Naraščanje števila različnih besed v zbirki *Napoved vodostajev*Slika 2.5: Naraščanje števila različnih besed v zbirki *Snežne razmere*

pri krivulji za *Napoved vodostajev*. Če bi z zajemanjem vremenskih podatkov še nadaljevali, bi umirjanje naraščanja teh krivulj zelo verjetno postalo še bolj izrazito, saj bi podatkovne zbirke vsebovale dovolj vremenskih podatkov iz vseh letnih časov.

Beseda	Kategorija	Tip
jutri	jutri	rel_dan
sredo	sreda	dan_teden
jutro	zjutraj	del_dneva
Prekmurju	prekmurje	pokrajina
države	slovenija	pokrajina
hribih	hribi	lege
severovzhodu	severovzhod	stran_neba
dežuje	dež	vreme
burjo	burja	veter
toplejše	topleje	temperatura
ponehale	bodo ponehale	stanje
občutno	precej	lastnost
-5	-5	stevilo
do	do	rel_izraz
zapihala	pihati	keyword
C	celzija	keyword
zaradi	x	ignore
kakšna	x	ignore

Tabela 2.1: Primeri vnosov v slovar *Vremenska napoved in obeti za Slovenijo*

Za vsako izmed petih podatkovnih zbirk smo izdelali *slovar besed*, v katerem smo vsaki besedi iz zbirke (ne glede na spol, sklon in število) priredili njeno *semantično kategorijo* in *semantični tip*. Besedam z istim pomenom smo priredili isto semantično kategorijo, tej pa smo priredili še enega izmed naslednjih semantičnih tipov:

- ↪ *časovni* semantični tipi (rel_dan, dan_teden ali del_dneva),
- ↪ *krajevni* semantični tipi (pokrajina, lege, stran_neba, voda ali predel),
- ↪ semantični tipi s *ključno informacijo* (vreme, veter, temperatura, pretok, pojavi, turbulenca ali oblaki),
- ↪ *dopolnilni* semantični tipi (stanje, lastnost ali stevilo),
- ↪ *relativni izraz* (rel_izraz),
- ↪ *ključna beseda* (keyword) ali
- ↪ *nepomembna informacija* (ignore).

V posameznih slovarjih je med 300 in 1000 vnosov. Primeri vnosov v slovar *Vremenska napoved in obeti za Slovenijo* so podani v tabeli 2.1.

Semantični tip *keyword* poleg semantičnih tipov s *ključno informacijo* najbolj zanesljivo določa pomen stavka, medtem ko semantični tip *ignore* združuje tiste kategorije, ki za samo razumevanje besedila niso ključnega pomena in jih lahko že na nivoju slovarja besed izločimo.

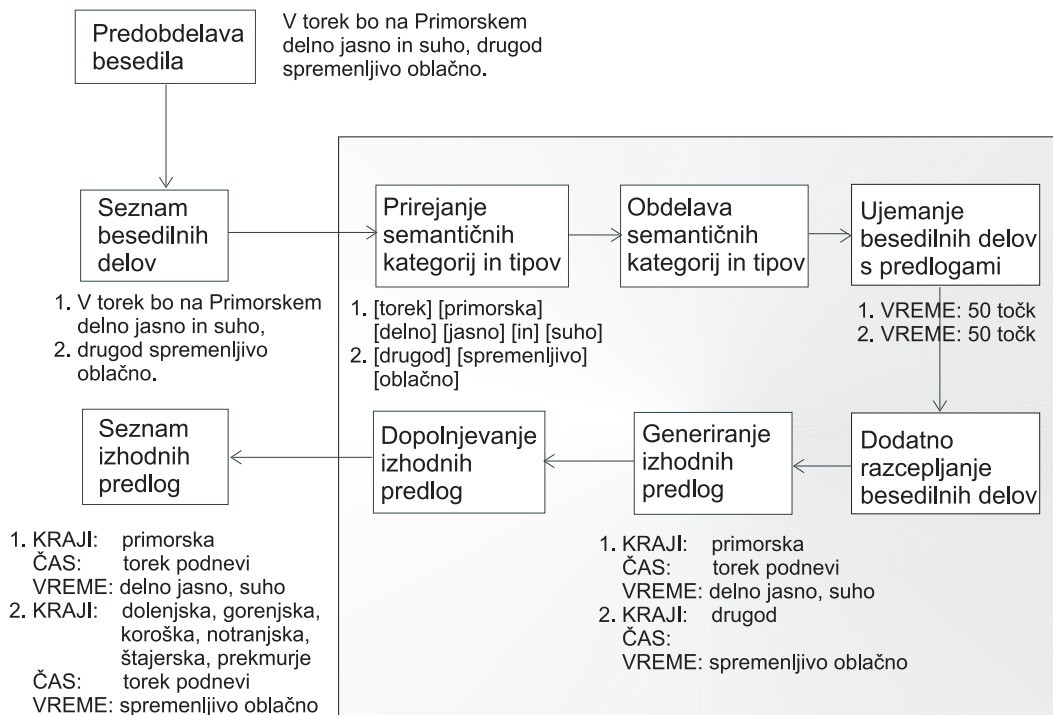
2.3 Pomenska analiza

Sistem za dialog JUPITER [Zue-2000], po katerem se zgledujemo, uporablja za razumevanje naravnega govora in besedila sistem TINA [Seneff-1992], ki povedi preoblikuje v jezikovno transparentne *semantične okvirje*. JUPITER uporablja sistem TINA v dva različna namena – za pomensko analizo uporabnikovih zahtev ter za pomensko analizo vremenskih napovedi. TINA pomensko analizira povedi, sestavne dele vremenske napovedi, tako da posamezni povedi priredi zaporedni indeks in ustrezne semantične kategorije oz. semantične okvirje. Podatkovno zbirko sistema JUPITER tako sestavljajo semantičnim kategorijam prirejeni sezname indeksov povedi. Na nivoju gradnje podatkovne zbirke se s tem pomenska analiza zaključí. Šele ko uporabnik sistemu JUPITER zastavi vprašanje, se znova vključi TINA, ki s pomočjo ustreznih povedi in njim prirejenih semantičnih kategorij oblikuje ustrezen odziv.

Mi smo se odločili za precej drugačen pristop. Že na nivoju gradnje podatkovne zbirke vremenske podatke pomensko analiziramo [Hajdinjak-2002b] in jih zapišemo v tabelarni obliki, s čimer sistemu omogočimo dostop do podatkov brez interakcije s kakšnim podsistemom. Pri tem si pomagamo s pomenskimi analizatorji, katerih model je prikazan na sliki 2.6. Trenutno smo udejanili štiri pomenske analizatorje, in sicer analizatorje za tematsko omejena področja *Vremenska napoved in obeti za Slovenijo*, *Vremenska napoved za gorski svet*, *Napoved za letalstvo* ter *Napoved vodostajev*. Nedokončan je ostal še pomenski analizator za *Snežne razmere*.

Vhod v analizatorje vremenskih podatkov je seznam *besedilnih delov*, t.j. delov povedi, ki so ločeni s pikami ali z vejicami. Del vhodnega seznama je tako lahko po eni strani samo stavčni člen, po drugi strani pa tudi priredno zložena poved.

Naš namen je bil posameznemu besedilnemu delu določiti pomen, t.j. vrsto informacije, na katero se ta nanaša. Pri analizi podatkov *Vremenska napoved in obeti za Slovenijo* se na primer izkaže, da lahko besedilne dele pomensko razdelimo na *vreme*, *veter* in *temperaturo*. Pomen posameznih besedilnih delov pomenski analizator določi s pomočjo vnaprej definiranih predlog (vsakemu pomenu je prirejena ena predloga), t.j. točkovanih seznamov semantičnih kategorij in tipov. Točke kategorij in tipov, ki predloge sestavljajo, smo določili eksperimentalno. Podoben pristop je bil uporabljen pri pomenski analizi stavkov v sistemu za podajanje informacij o letalskih poletih [Pepelnjak-1996a, Pepelnjak-1996b]. Vrednost posamezne predloge pri analizi vhodnega besedilnega dela predstavlja seštevek točk, prirejenih semantičnim kategorijam in tipom posameznih besed danega besedilnega dela. V skladu z možnim



Slika 2.6: Model pomenskih analizatorjev vremenskih podatkov

pomenom besedilnih delov, ki sestavljajo besedila iz zbirke *Vremenska napoved in obeti za Slovenijo*, smo definirali tri osnovne predloge, namreč *VREME*, *VETER* in *TEMPERATURA*. Če največje število točk, ki ga besedilni del dobi glede na vnaprej definirane predloge, ne doseže določenega praga, ga analizator pripiše dodatni predlogi *NEPOMEMBNA INFORMACIJA*, ki je skupna vsem pomenskim analizatorjem.

Ker se v pridobljenih zbirkah vremenskih podatkov nekateri besedilni deli nanašajo na dva različna pomena, prvotno pripisovanje pomenov ni dovolj dobro. S pikami in vejicami ločeni deli besedila namreč ne zagotavljajo pomenskega ločevanja priredno zloženih povedi. Edini priredni veznik, ki se je pojavil v naših zbirkah vremenskih podatkov, in pred katerim v povedi ni vejice, je veznik *in*. Priredno zložene povedi, v katerih veznik *in* nastopa, zato na mestu tega veznika še enkrat razcepimo in preverimo, ali sta pomena obeh tako nastalih delov res različna. Če ugotovimo, da sta pomena različna, in se nobeden izmed njiju ne nanaša na pomen *NEPOMEMBNA INFORMACIJA*, ju ločimo.

Semantični analizator vremenskih napovedi izvede analizo v šestih korakih:

1. korak : PRIREJANJE SEMANTIČNIH KATEGORIJ IN TIPOV besedilnim delom iz vhodnega seznama. To prirejanje poteka na osnovi *slovarjev besed*, v katerih smo vsaki besedi priredili semantično kategorijo in semantični tip.

2. korak : OBDELAVA SEMANTIČNIH KATEGORIJ IN TIPOV. Analizator nekatera zaporedja semantičnih kategorij in tipov združi, nekatera pa preoblikuje. Obdelava vključuje predvsem semantični tip *rel_izraz*, ki vpliva na pomen konteksta.

3. korak : UJEMANJE BESEDILNIH DELOV S PREDLOGAMI oz. določanje pomena posameznih besedilnih delov poteka na osnovi vnaprej definiranih predlog, ki vsebujejo točkovane sezname semantičnih kategorij in tipov.

4. korak : DODATNO RAZCEPLJANJE BESEDILNIH DELOV. Na mestu veznika *in* analizator dodatno razcepi tiste besedilne dele, za katere ugotovi, da vsebujejo dva različna pomena.

5. korak : GENERIRANJE IZHODNIH PREDLOG, t.j. seznamov časovnih in krajevnih ter za pomen stavka ključnih informacij. V primeru *Vremenska napoved in obeti za Slovenijo* smo v skladu z možnimi pomeni besedilnih delov definirali tri različne izhodne predloge, za vsakega od možnih osnovnih pomenov besedilnega dela (*VREME*, *VETER* in *TEMPERATURA*) po eno. Velikokrat pa vhodni besedilni del ne nosi vseh potrebnih informacij – največkrat manjkata oz. nista popolni časovna in/ali krajevna informacija. Posledica tega so nepopolne izhodne predloge.

6. korak : DOPOLNJEVANJE IZHODNIH PREDLOG. V zadnjem koraku analizator nepopolne izhodne predloge dopolni. Manjkajoče informacije poišče v izhodnih predlogah prejšnjih besedilnih delov.

Rezultat pomenskega analizatorja je seznam izhodnih predlog. Vsaka izmed izhodnih predlog vsebuje časovno informacijo, ki je natanko določena z dnem in delom dneva, in krajevno informacijo, ki je natanko določena s pokrajino, lego ali stranjo neba. Izhodne predloge besedilnih delov s pomenom *VREME* vsebujejo še opis vremena, izhodne predloge besedilnih delov s pomenom *VETER* še opis vetra, izhodne predloge besedilnih delov s pomenom *TEMPERATURA* še opis temperature, itd. Besedilnim delom, katerih pomen je *NEPOMEMBNA INFORMACIJA*, analizator izhodne predloge ne priredi. Primer analize je prikazan na sliki 2.6.

Analizator tematsko omejenega področja *Vremenska napoved in obeti za Slovenijo* smo testirali na 32 vremenskih napovedih, ki niso bile del podatkovne zbirke, na osnovi katere smo analizator gradili. Po predobdelavi testnih napovedi smo dobili 340 besedilnih delov, katerih seznam je bil vhod v analizator. Analizator je popolnoma pravilno pomensko razčlenil 332 besedilnih delov, kar znaša 97,65%. V 4. koraku je dodatno razcepil štiri besedilne dele, katerih nadaljno razčlenjevanje je bilo prav tako popolnoma uspešno. Rezultati so prikazani v tabeli 2.2.

	Število besedilnih delov	Odstotki
uspešno	332	97,65 %
delno uspešno	1	0,29 %
neuspešno	7	2,06 %
skupaj	340	

Tabela 2.2: Rezultati testiranja pomenskega analizatorja *Vremenska napoved in obeti za Slovenijo*

Pomanjkljivost analizatorja, ki smo jo pri testiranju odkrili, se odraža v neuspešno razčlenjenem stavku *V torek in sredo se bo nadaljevalo nestanovitno vreme s popoldanskimi plohami in nevihtami*. Analizator je ta besedilni del analiziral kot:

KRAJI: -
ČAS: torek podnevi, sredo podnevi, danes popoldne
VREME: nestanovitno, plohe, nevihte

Popolnoma pravilno pa bi bilo določiti dve izhodni predlogi, namreč:

KRAJI: -
ČAS: torek podnevi, sredo podnevi
VREME: nestanovitno

KRAJI: -
ČAS: torek popoldne, sredo popoldne
VREME: plohe, nevihte

Analizator torej ni popolnoma pravilno izluščil niti časovne niti ključne informacije, ki se v danem stavku oz. besedilnem delu prepletata.

Vse zgrajene analizatorje nameravamo še testirati in odpravljati ugotovljene pomanjkljivosti.

2.4 Zbirka podatkovnih tabel

Prvi del zbirke *MS Access*-ovih podatkovnih tabel je sestavljen iz vremenskih podatkov, ki so na spletnih straneh že podani v tabelarni obliki, zato jih od tam le prepisemo v tabele z enako podatkovno strukturo. Takih tabel je osem:

Vreme po Sloveniji,

Podatki z avtomatskih meteoroloških postaj,

Vreme po Evropi,

Višina snežne odeje,

Podatki z lavinskih postaj,

Poročilo o stopnji požarne ogroženosti naravnega okolja,

Stanje voda in

Napoved za slovensko Primorje.

Drugi del zbirke *MS Access*-ovih podatkovnih tabel je sestavljen iz štirih tabel, dobljenih s pomensko analizo naslednjih besedil:

Vremenska napoved in obeti za Slovenijo,

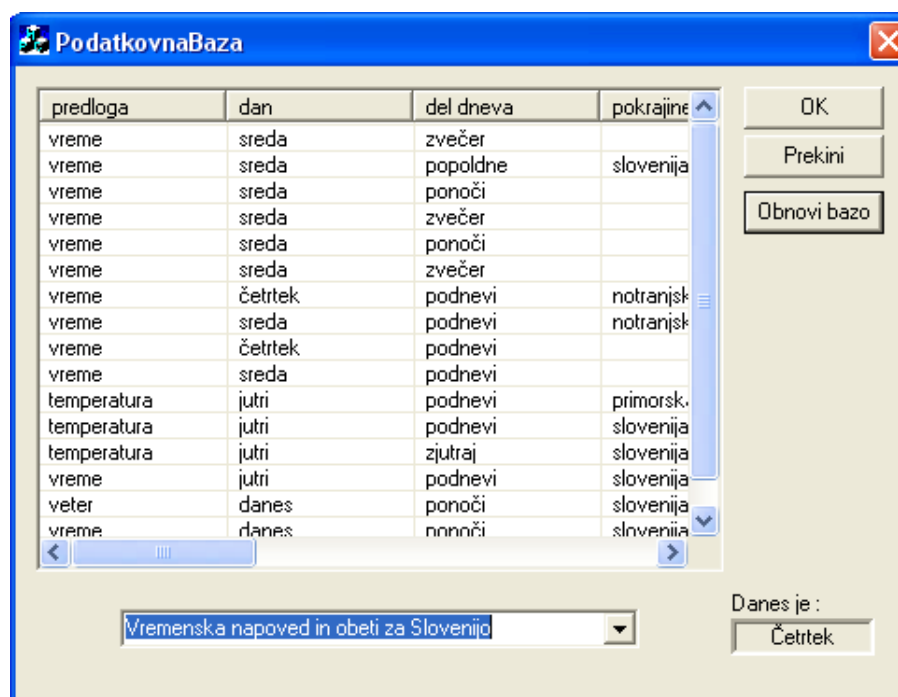
Vremenska napoved za gorski svet,

Napoved za letalstvo in

Napoved vodostajev.

predloga	dan	del dneva	pokrajine	lege	strani_neba	vreme	veter	temperatura
vreme	danes	podnevi	slovenija			oblačno		
vreme	danes	podnevi			zahod,notranjost	rahle padavine		
vreme	danes	podnevi			notranjost	sneg		
vreme	danes	podnevi	primorska			dež		
veter	danes	podnevi	primorska				zmerna do močna burja	
veter	danes	podnevi			notranjost		severovzhodni veter	
temperatura	danes	podnevi	slovenija					od -2 do 4
temperatura	danes	podnevi	primorska					okoli 6
vreme	jutri	podnevi	slovenija			padavine bodo ponehale		
veter	jutri	podnevi	primorska				zmerna burja	
temperatura	jutri	zjutraj	slovenija					najnižje od -4 do 0
temperatura	jutri	zjutraj		gore,hribi				do -7
temperatura	jutri	zjutraj	primorska					okoli 2
temperatura	jutri	podnevi	slovenija					najvišje malo nad 0
temperatura	jutri	podnevi	primorska					do 8
vreme	sobota	podnevi	slovenija			pooblačitve		
vreme	sobota	podnevi			notranjost,zahod	padavine		
vreme	sobota	podnevi		nižine		rahel dež		
vreme	sobota	zvečer	slovenija			padavine		
vreme	nedelja	podnevi	slovenija			oblačno,občasno rahel dež		

Slika 2.7: Primer podatkovne tabele, zgrajene s pomočjo analizatorja *Vremenska napoved in obeti za Slovenijo*



Slika 2.8: Grafični vmesnik do zbirke podatkovnih tabel

Struktura izhodnih predlog omogoča, da so podatki v teh tabelah časovno, krajevno in pomensko ločeni. Primer take tabele je na sliki 2.7, na sliki 2.8 pa je grafični vmesnik za dostop do zbirke vseh podatkovnih tabel, preko katerega lahko zbirko osvežimo.

Ker je večina podatkovnih tabel na spletnih straneh spremenljive velikosti (števili krajev in podanih informacij se iz dneva v dan spreminjata), napovedi pa so razmeroma jedrnate in skope, za našo zbirko podatkovnih tabel ob danem trenutku velja naslednje:

- ↪ vsebuje podatke o majhni podmnožici vseh pokrajin, mest in krajev v Sloveniji in Evropi,
- ↪ ne vsebuje enakih vrst informacij za vse pokrajine, mesta in kraje ter
- ↪ struktura informacij je časovno odvisna.

Taka struktura znanja sistem zelo omejuje pri želji, da bi uporabnika med dialogom vodil in mu pomagal ter porodi potrebo po ponujanju relevantnih, dosegljivih podatkov.

3 Eksperiment Čarovnik iz Oza

3.1 Zakaj eksperiment Čarovnik iz Oza

3.2 Sistem Čarovnik iz Oza

3.3 Prvi eksperiment Čarovnik iz Oza

3.4 Drugi eksperiment Čarovnik iz Oza

Opisali bomo eksperiment *Čarovnik iz Oza*, s katerim lahko s pomočjo človeka (čarovnika) simuliramo delovanje še nedokončanega sistema, in podali razloge za uporabo eksperimenta pri razvoju sistemov za dialog.

Pri razvoju sistema za podajanje informacij o vremenu smo ta eksperiment uporabili dvakrat. Pri izvajanju prvega eksperimenta smo uporabili sistem *Čarovnik iz Oza*, katerega sestavni deli so bili: grafični vmesnik, zbirka tabel vremenskih podatkov, modul za tvorjenje naravnega jezika in modul za tvorjenje slovenskega govora. V drugem eksperimentu smo v sistem vključili še modul za vodenje dialoga, tako da je naloga čarovnika ostala le nadomestiti delovanje modula za razpoznavanje govora in modula za razumevanje naravnega jezika.

Rezultat obeh eksperimentov sta zbirki spontanega slovenskega govora SSSQ in SSSQ2, sestavljeni iz dialogov, ki odražajo jezik komunikacije človek – računalnik.

3.1 Zakaj eksperiment Čarovnik iz Oza

Številni raziskovalci menijo, da naj bi dialogi človek – računalnik v čim večji meri sledili dialogom človek – človek. Toda dialogi človek – računalnik in dialogi človek – človek se razlikujejo v tolikšni meri, da podatki iz slednjih hitro postanejo neustrezen in nezanesljiv vir informacij za nekatere pomembne vidike oblikovanja sistemov za dialog kot sta na primer način in kompleksnost interakcije. Fraser in Gilbert [Fraser-1991] ugotavljata, da čeprav se je bolje zanašati na analizo dialogov človek – človek kot pa le na intuicijo, dejstvo, da se dialogi človek – človek in dialogi človek – računalnik razlikujejo, ostane in presenetljivo bi bilo, če bi oboji sledili enakim pravilom. Smith in Gordon [Smith-1997] zato trdita, da analiza dialogov človek – človek sicer predstavlja dobro osnovo za določitev nalog sistema in slovarja besed, ki naj jih sistem razume, ni pa preprosto ugotoviti, kateri vidiki te analize bodo ustrezali dialogom človek – računalnik. Zanimiva je njuna ugotovitev, da so raziskovalci ujeti v začaran krog – za konstrukcijo sistema za dialog po eni strani potrebujejo značilnosti dialogov človek – računalnik, po drugi strani pa je, dokler sistem ne obstaja, nemogoče vedeti, kako bodo dialogi potekali.

Trenutno najboljša alternativa za zbiranje podatkov, ki odraža jezik dialogov človek – računalnik, je tako imenovan *eksperiment Čarovnik iz Oza*. Ime tehnike se nanaša na moža za zaveso v knjigi in filmu *Čarovnik iz Oza*. V teh eksperimentih so uporabniki prepričani, da se pogovarjajo z računalnikom, kar pa ni res. V resnici za računalnikom sedi človek (čarovnik), ki simulira delovanje sistema za dialog. V nekaterih primerih [Whittaker-1989, Eskenazi-1999] čarovnik simulira celoten sistem, v drugih [Dahlbäck-1993, Kim-1997] pa le del sistema. Ugotovljeno je bilo [Whittaker-1989, Fraser-1991, Dahlbäck-1993], da podatki, pridobljeni z eksperimentom Čarovnik iz Oza, bolj natančno odražajo jezik komunikacije človek – računalnik kot dialogi človek – človek. Glavni razlog za to je prilagajanje udeležencev dialoga jezikovnim sposobnostim sogovornika.

Tehnika eksperiment Čarovnik iz Oza pa ima tudi nekatere slabosti:

1. Konstrukcija kvalitetnega sistema za simulacijo in učenje čarovnikov je zahtevna naloga. Dahlbäck, Jönsson in Ahrenberg [Dahlbäck-1993] to utemeljujejo z dejstvom, da smo ljudje fleksibilni, računalniki pa konsistentni; da ljudje počasi tipkamo, računalniki pa se odzivajo hitro; da računalniki nikoli ne delajo majhnih napak (tipkarske napake), ljudje pa jih delamo pogosto. Torej če želimo, da bodo uporabniki ves čas dialoga ostali v prepričanju, da se pogovarjajo z računalnikom, čarovnik ne sme le sedeti za računalnikom in tipkati svoje odgovore, ki so nato umetno pretvorjeni v govor. Vmesnik, pred katerim sedi, mora čarovniku pomagati pri njegovem delu, tako da čimbolj zmanjša potrebo po tipkanju in čimbolj poveča konsistentnost njegovih odzivov.
2. Udeleženci eksperimenta igrajo le vloge resničnih uporabnikov, zato med dialogom le stežka pokažejo svoje potrebe, motivacije, izhode in ovire. To težavo lahko do neke mere odpravimo s pomočjo uvedbe vnaprej določenih scenarijev ali situacij. Ker jih ponavadi zastavijo raziskovalci sami, to ni najbolj idealen način

ugotavljanja nalog, ki naj jih bo sistem sposoben opraviti.

3. Pojavljajo se tudi etična vprašanja. To težavo lahko delno zmanjšamo, če udeležencem eksperimenta na koncu povemo, kaj smo počeli, zakaj smo tako ravnali in nenazadnje, da bomo podatke nemudoma uničili, če sami tako želijo.

Eksperimentalni del razvoja sistema za dialog običajno poteka v treh korakih. V prvem koraku opravimo analizo dejanskih dialogov človek – človek. Na osnovi te analize konstruiramo enega ali celo več sistemov Čarovnik iz Oza, ki služijo kot ogrodje za izvedbo istoimenskega eksperimenta. Zadnji korak je izboljševanje in dopolnjevanje sistema za dialog s pomočjo podatkov dejanskih uporabnikov.

Pri razvoju sistema za podajanje informacij o vremenu, ki je predmet tega dela, smo prva dva koraka že naredili. Najprej smo vzpostavili osebno komunikacijo s *Hidrometeorološkim zavodom Agencije Republike Slovenije za okolje*, ki nudi telefonske pogovore z dežurnim prognostikom, na podlagi katerih so nam znali povedati, kaj dejanske uporabnike sploh zanima in kako po teh podatkih poizvedujejo. Te informacije so nam bile vodilo pri oblikovanju prvega sistema Čarovnik iz Oza [Hajdinjak-2003a, Hajdinjak-2003b]. Kasneje, ko smo modul za vodenje dialoga že razvili, smo, z namenom vrednotenja delovanja tega modula, eksperiment še enkrat ponovili.

3.2 Sistem Čarovnik iz Oza

Sistem Čarovnik iz Oza, ki preko telefona podaja informacije o vremenu in vremenski napovedi [Hajdinjak-2003a], s katerim smo izvajali eksperimente Čarovnik iz Oza, je bil modularno zasnovan, vsi moduli pa so bili povezani z grafičnim vmesnikom, za katerim je sedel človek (čarovnik) in simuliral del sistema za dialog. Sistem smo udeležili na osebem računalniku z vgrajeno ISDN DIVA Server BRI-2M PCI kartico, s katero smo komunicirali preko posebnega računalniškega programa, ki je omogočal vzpostavljanje in prevzemanje telefonskih pogovorov, poslušanje in snemanje pogovorov, kodiranje in dekodiranje ISDN formata zvočnih datotek, pošiljanje wave datotek po telefonski liniji, ipd.

V zgodnjih fazah razvoja sistema za dialog je primerneje, če avtomatičnega razumevanja govora ne vključimo v sistem, saj v tem primeru ne bomo omejeni z naglasi, spolom in starostjo uporabnikov. V literaturi najdemo dva načina simuliranja razumevanja govora – napake razumevanja govora lahko sistematsko vključimo v besedilo, na katerega se čarovnik odziva [Pirker-1999], ali pa simuliramo takorekoč popolno razumevanje govora [Dahlbäck-1993, Kim-1997, Eskenazi-1999]. Za slednji način smo se odločili tudi mi, čeprav se zavedamo, da so napake razumevanja govora izjemno pomemben vidik uporabnikove interakcije s sistemom. Za ta pristop smo se odločili, ker je napake razumevanja govora, še posebej pa napake razpoznavanja govora, zelo težko ustrezno simulirati. Naš čarovnik je zato poslušal uporabnika, brez da bi ga pri tem kakorkoli omejevali.

Sistem Čarovnik iz Oza sestavljajo:

- ↪ platforma za ISDN telefonijo,
- ↪ zbirka podatkovnih tabel (poglavje 2), do katerih dostopa čarovnik,
- ↪ grafični vmesnik, ki čarovniku omogoča vodenje dialoga,
- ↪ modul za tvorjenje naravnega jezika in
- ↪ modul za tvorjenje slovenskega govora.

Eksperiment je potekal tako, da je čarovnik sedel za grafičnim vmesnikom, poslušal izjave uporabnika, v skladu s katerimi je po lastni presoji izbiral odzive na grafičnem vmesniku in s tem po potrebi dostopal do vremenske zbirke podatkovnih tabel. Odzivi so potovali v modul za tvorjenje naravnega jezika, katerega izhod (naravno besedilo) je potoval v modul za tvorjenje slovenskega govora, tega pa je sistem po telefonski liniji poslal do uporabnika.

Grafični vmesnik, ki čarovniku omogoča vodenje dialoga, smo želeli oblikovati tako, da bi bili odzivi čarovnika po hitrosti in konsistentnosti čimbolj podobni odzivom računalnika. Zato smo ta grafični vmesnik zasnovali kot uporabniku (čarovniku) prijazno spletno aplikacijo, ki vsebuje

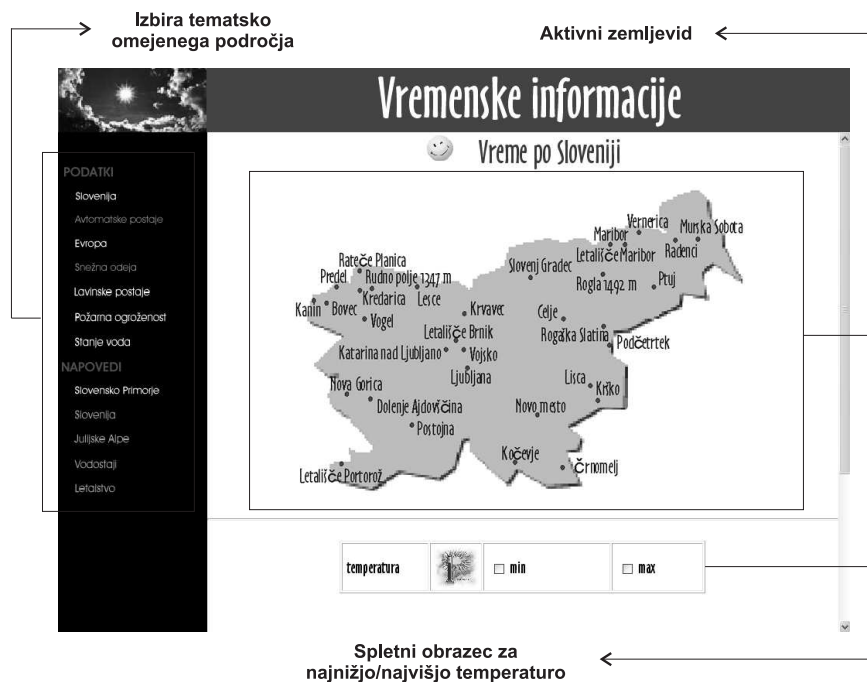
- ↪ spletne obrazce,
- ↪ aktivne slike,
- ↪ slikovna polja,
- ↪ spustne sezname,
- ↪ ...

in omogoča

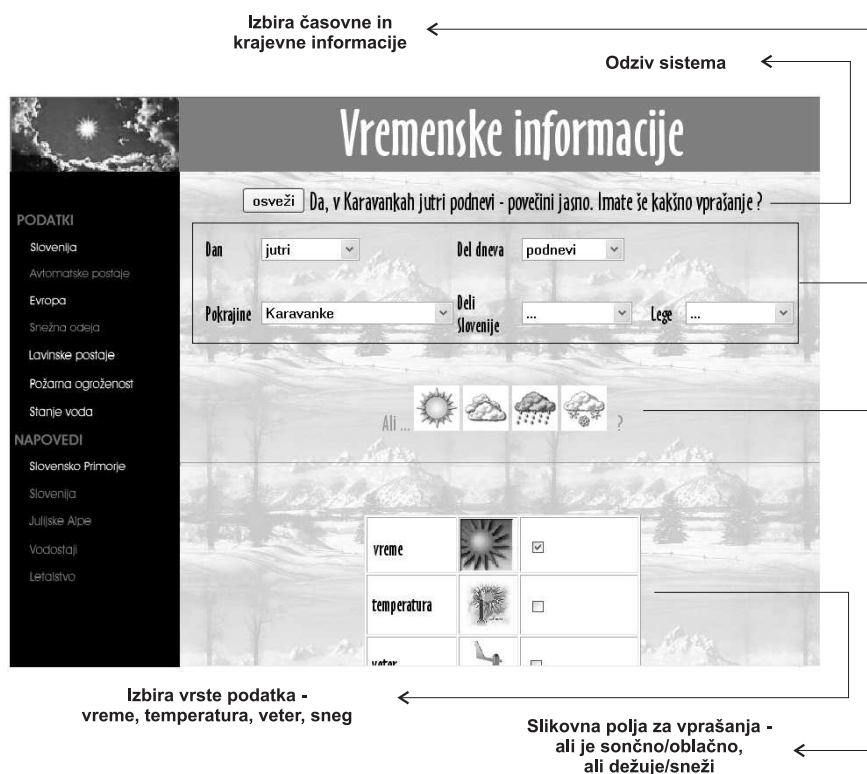
- ↪ izbiro krajev s pomočjo aktivnih zemljevidov,
- ↪ predvajanje vnaprej pripravljenih zvočnih datotek,
- ↪ uporabo bližnjic,
- ↪ ...

Za obdelavo spletnih obrazcev uporabljamo programe, zapisane v skriptnem jeziku CGI, ki jih čarovnik izvaja s pritiski tipk na miški oz. tipkovnici. Primer delovnih oken je prikazan na slikah 3.1 in 3.2.

Cilj modula za tvorjenje naravnega jezika je računalniško predstavitev podatkov pretvoriti v naravno besedilo. Dandanes je večina modulov za tvorjenje naravnega jezika zelo podobno zgrajenih – proces tvorjenja naravnega jezika je pri večini razdeljen na



Slika 3.1: Grafični vmesnik sistema Čarovnik iz Oza – primer delovnega okna 1



Slika 3.2: Grafični vmesnik sistema Čarovnik iz Oza – primer delovnega okna 2

več zaporednih podmodulov, kjer se podatki pretakajo le v eni smeri, t.j. iz predhodnega podmodula v naslednji podmodul. Ta način povezave podmodulov je Reiter (1994) poimenoval *enosmerna cevovodna struktura*. Številni raziskovalci so nasprotovali deljenju procesa tvorjenja naravnega jezika na več podmodulov [Appelt-1985], drugi [Danlos-1984] so sprejeli podmodule, a so trdili, da bi struktura povezovanja podmodulov morala omogočati tudi vzvratni pretok podatkov. Čeprav so dvomi in utemeljitve teh raziskovalcev razumljivi, je argument, s katerim je Marr [Marr-1976] utemeljil deljenje procesa tvorjenja naravnega jezika na podmodule, gledano z inženirskega stališča, veliko močnejši. Argumentiral je, da podmoduli omogočajo učinkovitejšo in bolj strukturirano zgradbo ter so, zaradi podobnosti zgradbi človeških možganov, smiselni. Tudi psiho-lingvistične raziskave [Levelt-1989] jasno kažejo, da so visoko modularizirani človeški možgani verjetno nekoliko kompleksnejši od enosmerne cevovodne strukture, vendar ta razlika najverjetneje ni tako zelo velika.

Ti argumenti so nas pripeljali do zaključka, da smo proces tvorjenja naravnega jezika v sistemu Čarovnik iz Oza razdelili na naslednje podmodule:

1. *Podmodul za določanje vsebine* spremeni čarovnikove zahteve v strukturirano predstavitev atomarnih komunikacijskih ciljev, ki jih lahko dosežemo s tvorjenjem ene same povedi.
2. *Podmodul za oblikovanje povedi* spremeni strukturirano predstavitev atomarnih komunikacijskih ciljev, ki je izhod iz podmodula za določanje vsebine, v abstraktno slovnično predstavitev povedi, sestavljeno iz besed in nekaterih slovničnih povezav.
3. *Podmodul za slovnične popravke* dokonča proces tvorjenja naravnega jezika, s tem da abstraktno slovnično predstavitev povedi, ki je izhod iz podmodula za oblikovanje povedi, s pomočjo znanja o slovenskem jeziku preoblikuje v slovnično pravilne povedi.

Primer tvorjenja povedi *V Ljubljani ta trenutek – jasno, temperatura zraka je -9 stopinj Celzija* je prikazan v tabeli 3.1.

<i>Podmodul za določanje vsebine</i>	(kraj:Ljubljana, čas:trenutno, vreme:jasno) (temperatura:-9)
<i>Podmodul za oblikovanje povedi</i>	(:Ljubljana: :trenutno: - jasno) (temperatura zraka :biti: -9 stopinj celzija)
<i>Podmodul za slovnične popravke</i>	V Ljubljani ta trenutek – jasno, temperatura zraka je -9 stopinj celzija.

Tabela 3.1: Primer tvorjenja povedi z modulom za tvorjenje naravnega jezika

Modul za tvorjenje slovenskega govora [Gros-1997a, Gros-1997b], ki je bil vključen v sistem Čarovnik iz Oza, je bil zgrajen v *Laboratoriju za umetno zaznavanje, sisteme in kibernetiko*.

3.3 Prvi eksperiment Čarovnik iz Oza

Naloga čarovnika v prvem eksperimentu Čarovnik iz Oza je bila igrati vlogo modulov za razpoznavanje govora in razumevanje naravnega jezika, s pomočjo grafičnega vmesnika pa tudi vlogo modula za vodenje dialoga. Za ta eksperiment smo se odločili v prvi fazi razvoja sistema za podajanje informacij o vremenu, da bi pridobili podatke, na osnovi katerih bi potem gradili omenjene module, ki jih je čarovnik v eksperimentu nadomeščal.

V eksperimentu je sodelovalo 76 uporabnikov, in sicer 38 žensk in 38 moških. Pri izbiri smo pazili, da bi uporabniki predstavljali čimbolj reprezentativen vzorec (spol, starost, izobrazba, narečje, telefonska linija, okolje telefonskega pogovora). Povprečna starost uporabnikov je bila 34 let, povprečna izobrazba pa srednja šola. Zastopanih je bilo vseh osem osnovnih slovenskih narečij, pazili pa smo tudi na vrste telefonskih linij (analogna, ISDN, GSM) in na okolja (tiho okolje, šolski hodnik, ulica, prostor z več ljudmi, menza, vklopljen radio/TV-sprejemnik, ipd.), v katerih so se uporabniki v času telefonskega pogovora s sistemom Čarovnik iz Oza nahajali.

Pred začetkom eksperimenta smo udeležencem povedali, da se bodo pogovarjali z računalnikom, t.j. s sistemom za podajanje informacij o vremenu, in jim dali ustna navodila o splošni funkcionalnosti sistema. Ker smo želeli po eni strani pridobiti čimveč posnetkov, po drugi strani pa, kljub igranju vloge dejanskih uporabnikov, zagotoviti čimbolj realne in raznolike dialoge, smo vsakemu izmed udeležencev dodelili dve nalogi. Prva naloga je bila pridobiti določeno informacijo, druga pa je zajemala določen scenarij oz. situacijo, ki naj bi si jo udeleženec poskušal zamisliti. Primera takšnih nalog sta:

1. Poskušajte ugotoviti, ali v Ljubljani sije sonce.
2. Ob koncu tedna bi radi šli v hribe. Kaj vas zanima?

in

1. Poskušajte ugotoviti, kakšne temperature lahko pričakujemo jutri.
2. Načrtujete izlet s kolesom. Kaj vas zanima?

Udeležencem smo omogočili tudi lastno izbiro vprašanj – povedali smo jim, da lahko po opravljenih nalogah dialog s sistemom nadaljujejo.

Po pogovoru s sistemom Čarovnik iz Oza smo uporabnike prosili, da izpolnijo vprašalnik, katerega prvi del je vseboval vprašanja o spolu, starosti, izobrazbi, zaposlitvi, narečju, vrsti telefonske linije in okolju, v katerem so se v času dialoga s sistemom nahajali. Drugi del vprašalnika se je nanašal na dialog s sistemom in je, poleg vprašanja, ali so dobili odgovor na prvo nalogo, zajemal različne vidike njihove interakcije s sistemom (tabela 3.2). Vprašanja, s pomočjo katerih so udeleženci ocenjevali sistem Čarovnik iz Oza, smo povzeli po [Walker-1997, Kamm-1999, Litman-2002], sprašujejo pa po učinku modula za tvorjenje govora, učinku modula za razpoznavanje govora, težavnosti pridobivanja informacij, hitrosti interakcije, izkušenosti uporabnikov, usreznosti odzivov sistema, pričakovanem obnašanju sistema in načrtovani rabi sistema v prihodnosti. Odgovori so bili podani z lestvico od 1 (nikakor se ne strinjam) do 5 (popolnoma se strinjam).

-
-
1. Ali ste sistem brez težav razumeli?
 2. Ali vas je sistem razumel?
 3. Ali ste brez težav prišli do odgovora na vaša vprašanja?
 4. Ali je bila hitrost interakcije s sistemom primerna?
 5. Ali ste na vsakem koraku dialoga vedeli, kaj morate povedati?
 6. Ali se je sistem na vaše izjave odzival hitro (brez pojasnilnih vprašanj)?
 7. Ali se je sistem obnašal tako, kot ste med dialogom od njega pričakovali?
 8. Glede na vašo izkušnjo s sistemom za podajanje informacij o vremenu, ali boste (ko bo to mogoče) poklicali sistem, če vas bo zanimalo vreme?
-
-

Tabela 3.2: Del vprašalnika, s katerim so udeleženci eksperimenta ocenjevali sistem Čarovnik iz Oza

Povprečne ocene uporabnikov so podane v tabeli 3.3. Najslabše je bila ocenjena hitrost interakcije s sistemom, in sicer 3.94, kar pa niti ni tako slabo, če vemo, da je bil povprečen čas čakanja na odziv sistema 5.57 sekund. Večji del tega časa je porabil čarovnik, da je posredoval izbran odziv. Vse ostale ocene, od katerih se večina vsaj delno nanaša na učinek čarovnika, ležijo med 3.99 in 4.51.

	Povprečne ocene
Tvorjenje govora	4.42
Razpoznavanje govora	4.51
Pridobivanje informacij	4.27
Hitrost interakcije	3.94
Izkušnost uporabnikov	4.40
Ustreznost odzivov	4.23
Pričakovano obnašanje	4.31
Raba v prihodnosti	3.99

Tabela 3.3: Povprečne ocene sistema Čarovnik iz Oza

V eksperimentu smo opazili nekaj zanimivosti. Prva je ta, da je večina uporabnikov svoja vprašanja oblikovala zelo podobno, ponavadi *Zanima me ...*, *Rad/Rada bi vedel/vedela ...*, *Mi lahko (prosim) poveste ...*, *Ali mi lahko poveste ...*, ipd. Druga zanimivost je prilagajanje uporabnikovega vedenja na pričakovane jezikovne sposobnosti

sistema. V našem eksperimentu smo ugotovili, da je bilo prvo vprašanje, ki so ga uporabniki zastavili, ponavadi veliko daljše in veliko manj jedrnato kot vprašanja, ki so sledila. Primeri dolgih in nejedrnatih začetnih vprašanj so na primer:

- Jaz moram danes na Primorsko, pa me je malo strah burje. Zdaj me pa zanima, kje začenja pihati, a na vetrišču ali šele s Svete gore navzdol proti Gorici.
- Glejte. Trenutno se peljem z Jesenic proti Ljubljani. Razmere na cesti so obupne. Vseskozi piha močan veter s smeri vzhoda. Zdaj me pa zanima, kakšen veter piha v Ljubljani. Se pravi, moč pa smer, od kod piha.
- Danes bi radi prišli v Ljubljano na izlet. Malo bi se sprehajali, pa na ljubljanski grad bi šli radi. Pa nas zanima, če je v Ljubljani kaj sonca, če je lep razgled z gradu.

V primerih, ko je sistem uporabnike prosil, naj vprašanje ponovijo, pa je njihov govor postal bolj artikuliran, počasnejši in/ali glasnejši. Že [Zoltan-1991] je prišel do podobnih ugotovitev, namreč da jezikovne sposobnosti sistema za dialog opazno vplivajo na vedenje uporabnikov. Cohen in Oviatt (1995) pa trdita, da bodo uporabniki svoje besedišče zelo verjetno prilagodili sistemovemu, če bodo v njem zaznali omejitve.

SISTEM:	<i>Dober dan. Pogovarjate se s sistemom za podajanje informacij o vremenu. Kakšno je vaše vprašanje?</i>
UPORABNIK:	Mi lahko prosim poveste, kakšna bo danes vidljivost na Štajerskem?
SISTEM:	<i>Počakajte trenutek, prosim ... [čarovnik izbira svoj odziv]</i> <i>Na Štajerskem bo danes podnevi vidljivost nad 10 km. Imate še kakšno vprašanje?</i>
UPORABNIK:	A bo kaj snežilo?
SISTEM:	<i>Kateri kraj vas zanima?</i>
UPORABNIK:	Štajerska.
SISTEM:	<i>Počakajte trenutek, prosim ... [čarovnik izbira svoj odziv]</i> <i>Ne, na Štajerskem danes podnevi – razjasnitve. Imate še kakšno vprašanje?</i>
UPORABNIK:	Kakšno pa bo vreme v prihodnjih dneh na Poljskem?
SISTEM:	<i>Imam podatke za danes.</i>
UPORABNIK:	Torej, kako bo danes na Poljskem?
SISTEM:	<i>Počakajte trenutek, prosim ... [čarovnik izbira svoj odziv]</i> <i>V Varšavi na Poljskem je oblačno, temperatura zraka je -6 stopinj celzija. Imate še kakšno vprašanje?</i>
UPORABNIK:	A je kaj snega na Poljskem?
SISTEM:	<i>Te informacije ne ponujam. Imate morda kakšno drugo vprašanje?</i>
UPORABNIK:	Ne, najlepša hvala. Nasvidenje.
SISTEM:	<i>Hvala lepa za sodelovanje. Nasvidenje.</i>

Tabela 3.4: Primer dialoga med uporabnikom in sistemom Čarovnik iz Oza

Tretja zanimivost, ki smo jo opazili v prvem eksperimentu Čarovnik iz Oza, so smeh, šale in duhovite opazke, ki so jih nekateri uporabniki delali na račun sistema, medtem ko je čarovnik posredoval svoj odziv. Primeri takih opazk so:

↪ Kakšen glas ... Grozno ...

↪ Išče po računalniku ...

↪ Razmišlja ...

Takšne reakcije in stranske opazke so najbrž bolj verjetne v testiranjih sistemov za dialog kot pa v komercialnih sistemih, saj v takih testiranjih uporabniki ponavadi igrajo le vloge dejanskih uporabnikov, niso pod časovnimi omejitvami in ne odgovarjajo za stroške telefonskega pogovora.

Primer dialoga med uporabnikom in sistemom Čarovnik iz Oza je podan v tabeli 3.4.

Zbirka spontanega govora, ki smo jo pridobili v prvem eksperimentu Čarovnik iz Oza, smo poimenovali SSSQ [Hajdinjak-2003b]. Kratica SSSQ se nanaša na angleško besedno zvezo *Slovenian Spontaneous Speech Queries*. Zbirko smo označili z orodjem TRANSCRIBER [Barras-2001], pri čemer smo poleg oznak besed uporabili še posebne oznake za narečne besede, premore in negovorne zvoke (vdih, izdih, vzdihljaj, tlesk, smeh, pih, kašelj in različne zvoke, ki nastajajo pri obotavljanju oz. razmišljanju).

Zgradba te zbirke spontanega govora je prikazana v tabeli 3.5. Sestavljena je iz 76 dialogov, ki vsebujejo 662 izjav uporabnikov. Skupna dolžina dialogov znaša 103.02 minute, dolžina človeškega govora pa 41.18 minut.

število dialogov	76
število izjav uporabnikov	662
število različnih besed	920
število vseh besed	5442
dolžina vseh dialogov	103.02 min
dolžina človeškega govora	41.18 min

Tabela 3.5: Zgradba SSSQ zbirke spontanega govora

Dialoge, pridobljene z eksperimentom Čarovnik iz Oza, ki odražajo strukturo dialogov človek – računalnik, smo podrobneje analizirali in ugotovili, da je povprečna dolžina pogovora bila 81.33 sekund, povprečni čas trajanja uporabnikovega govora pa 32.51 sekund. Ta razlika je precej velika zaradi nekoliko daljšega (5.57 sekund v povprečju) odzivnega časa sistema, t.j. časa, ki ga je čarovnik porabil, da je sistemu sporočil, kako naj se na uporabnikovo izjavo odzove. Povprečno število izjav uporabnikov v dialogih je bilo 8.71, povprečno število besed v izjavah uporabnikov pa 7.96. V tabeli 3.6 so poleg povprečnih vrednosti podane tudi mediane posameznih parametrov.

V poglavju 7 so podani rezultati izračuna dodatnih parametrov, ki jih uporabljamo za evalvacijo sistema Čarovnik iz Oza in modula za vodenje dialoga.

	povprečje	mediana
število izjav na dialog	8.71	8.00
število besed na izjavo	7.96	7.00
dolžina dialoga	81.33 s	68.22 s
dolžina človeškega govora	32.51 s	28.80 s

Tabela 3.6: Rezultati analize dialogov, pridobljenih v prvem eksperimentu Čarovnik iz Oza

3.4 Drugi eksperiment Čarovnik iz Oza

Ko smo razvili modul za vodenje dialoga (poglavje 5) in ga vgradili v sistem, smo eksperiment Čarovnik iz Oza ponovili. Cilj in namen drugega eksperimenta je bil pridobiti podatke, ki bi poleg podatkov iz prvega eksperimenta tvorili osnovo za evalvacijo modula za vodenje dialoga.

Naloga čarovnika v drugem eksperimentu Čarovnik iz Oza je, v primerjavi s prvim eksperimentom, bila igrati vlogo modulov za razpoznavanje govora in razumevanje naravnega jezika, ne pa tudi vloge modula za vodenje dialoga. V tem eksperimentu je čarovnik sedel pred vmesnikom modula za vodenje dialoga in preko tipkovnice vnašal pomensko predstavitev uporabnikove izjave, ki bo v končnem sistemu izhod iz modula za razumevanje naravnega jezika. Vse nadaljnje delo (vodenje dialoga, iskanje podatkov, tvorjenje naravnega jezika, tvorjenje govora) je opravljal sistem.

V eksperimentu je sodelovalo 68 uporabnikov, 29 žensk in 39 moških, katerih povprečna starost je bila 32 let. Čeprav je 17 izmed njih že sodelovalo v prvem eksperimentu Čarovnik iz Oza in so zato bili malo bolj izkušeni od ostalih, med vedenjem obeh skupin nismo opazili nobene razlike.

Uporabnike (tudi tiste, ki so že sodelovali v prvem eksperimentu) smo, tako kot v prvem eksperimentu, prepričali, da se bodo pogovarjali z računalnikom, t.j. s sistemom za podajanje informacij o vremenu, in jim dali ustna navodila o splošni funkcionalnosti sistema. Vsakemu izmed udeležencev smo dodelili dve nalogi, podobno kot v prvem eksperimentu, od katerih je bila prva pridobiti določeno informacijo, druga pa scenarij oz. situacija, ki naj bi si jo udeleženec poskušal zamisliti. Po končanem pogovoru s sistemom smo uporabnike prosili, da izpolnijo enak vprašalnik kot v prvem eksperimentu. Del vprašalnika, ki zajema vprašanja, nanašajoča se na ocenjevanje sistema, je podan v tabeli 3.2.

Povprečne ocene udeležencev drugega eksperimenta Čarovnik iz Oza so podane v tabeli 3.7. Te ocene so nekoliko slabše od prvih, najslabše pa so bili ocenjeni: preprostost pridobivanja informacij, hitrost interakcije sistema, ustreznost odzivov sistema in raba v prihodnosti.

Zbirko spontanega govora, ki smo jo pridobili v drugem eksperimentu Čarovnik iz Oza, smo iz podobnega razloga kot v prvem eksperimentu poimenovali SSSQ2 in jo označili v skladu z oznakami zbirke SSSQ. Zgradba SSSQ2 zbirke spontanega govora je prikazana

	Povprečne ocene
Tvorjenje govora	4.29
Razpoznavanje govora	4.29
Pridobivanje informacij	3.74
Hitrost interakcije	3.76
Izkušnost uporabnikov	4.28
Ustreznost odzivov	3.76
Pričakovano obnašanje	4.04
Raba v prihodnosti	3.78

Tabela 3.7: Povprečne ocene v drugem eksperimentu Čarovnik iz Oza

v tabeli 3.8. Sestavljena je iz 68 dialogov, ki vsebujejo 885 izjav uporabnikov. Skupna dolžina dialogov znaša 148.92 minut, dolžina človeškega govora pa 50.51 minut. Skupno število različnih besed (zajete so tudi govorne različice) v obeh zbirkah spontanega govora (SSSQ, SSSQ2) je 1229.

število dialogov	68
število izjav uporabnikov	885
število novih besed	309
število različnih besed	716
število vseh besed	5403
dolžina vseh dialogov	148.92 min
dolžina človeškega govora	50.51 min

Tabela 3.8: Zgradba SSSQ2 zbirke spontanega govora

Dialoge, pridobljene z drugim eksperimentom Čarovnik iz Oza, smo podrobneje analizirali in ugotovili, da je bila povprečna dolžina pogovora 117.57 sekund, povprečni čas trajanja uporabnikovega govora pa 39.87 sekund. Ta razlika je spet precej velika, in sicer zaradi nekoliko daljšega (6.61 sekund v povprečju) odzivnega časa sistema, t.j. čas, ki ga je čarovnik porabil, da je modulu za vodenje dialoga posredoval pomensko predstavitev uporabnikove izjave. Povprečno število izjav uporabnikov v dialogih je bilo 11.64, povprečno število besed v izjavah uporabnikov pa 6.33. V tabeli 3.9 so poleg povprečnih vrednosti podane tudi mediane posameznih parametrov.

Vidimo (tabeli 3.6 in 3.9), da so drugi eksperiment zaznamovali predvsem daljši dialogi, in sicer tako po številu izjav kot tudi po časovnem trajanju. To je v glavnem posledica v drugem eksperimentu izvajane strategije vodenja dialoga (poglavje 5), ki se od strategije čarovnika v prvem eksperimentu razlikuje predvsem po prošnjah sistema za potrditev posredovanih informacij. Te potrditve, zavrnitve oz. popravki so zaradi svoje jedrnatosti tudi glavni razlog v povprečju krajših izjav, saj so velikokrat sestavljeni le iz ene ali dveh besed.

Zanimivo je, da so udeleženci drugega eksperimenta slabše ocenili prav vse vidike svoje

	povprečje	mediana
število izjav na dialog	11.64	11.00
število besed na izjavo	6.33	6.00
dolžina dialoga	117.57 s	108.42 s
dolžina človeškega govora	39.87 s	37.78 s

Tabela 3.9: Rezultati analize dialogov, pridobljenih v drugem eksperimentu Čarovnik iz Oza

interakcije s sistemom, tudi učinek modulov za tvorjenje in razpoznavanje govora, ki sta v obeh eksperimentih ostala enaka. Glede na to, da so tako storili tudi tisti, ki so sodelovali v obeh eksperimentih Čarovnik iz Oza, razloga ne moremo pripisati večji kritičnosti udeležencev. Zelo verjetno pa so slabše ocene posledica manjšega zadovoljstva z nekaterimi posameznimi vidiki učinkovitosti sistema. Najbolj opazna razlika je v ocenah težavnosti pridobivanja informacij (4.27 oz. 3.74) in ustreznosti odzivov sistema (4.23 oz. 3.76), pri čemer je treba poudariti, da so bile slabše ocene nekaterih vidikov učinkovitosti sistema vsekakor pričakovane, saj bi v nasprotnem primeru modul za vodenje dialoga v drugem eksperimentu bolje opravljal svojo nalogo kot človeški operater v prvem eksperimentu.

Da bi vsaj delno odpravili etično vprašljivost eksperimentov Čarovnik iz Oza, smo po obeh eksperimentih udeležencem povedali, kaj smo počeli, zakaj smo tako ravnali, in jih vprašali za dovoljenje uporabe pridobljenih podatkov v raziskovalne namene. Prav vsi so pokazali razumevanje in odobravanje teh eksperimentov, dovolili pa so nam tudi uporabo posnetkov in drugih podatkov, ki smo jih v eksperimentih pridobili.

4 Oblikovanje in razvoj sistemov za dialog

4.1 Kdaj je sistem za dialog uporaben

4.2 Vodila pri oblikovanju sistemov za dialog

4.3 Tehnike modeliranja dialoga

Opredelili bomo pojem uporabnosti govornih vmesnikov, ki zaobsega *učljivost*, *učinkovitost*, *zapomnljivost*, *napake* in *zadovoljstvo* uporabnikov.

Opisali bomo Griceove maksime naravne komunikacije in našeli Shneidermanovih 8 zlatih pravil oblikovanja sistemov, ki pogosto služijo kot vodilo pri oblikovanju sistemov za dialog.

Pregledali bomo tudi trenutno uveljavljene tehnike modeliranja dialoga ter podali njihove dobre in slabe lastnosti.

4.1 Kdaj je sistem za dialog uporaben

Ena izmed najpomembnejših lastnosti sistemov za dialog je *uporabnost*. To je širok pojem, ki ima na področju sistemov za dialog več različnih definicij. Nielsen (1993) je na primer zapisal, da je uporabnost sistema večdimenzionalen pojem, ki zaobsega *učljivost*, *učinkovitost*, *zapomnljivost*, *napake* in *zadovoljstvo* ter opisuje načine, kako jih merimo. V nadaljevanju bomo opisali, kakšen je pomen teh pojmov v primeru govornih vmesnikov.

Govorni vmesnik naj bo *učljiv*, kar pomeni, da bi se ga naj dalo brez težav naučiti uporabljati. To je pri govornih vmesnikih vsekakor lažje doseči kot pri ostalih sistemih, saj temelji na naravnem načinu komunikacije – na govoru. Zaradi omejitev pri razpoznavanju govora pa je število različnih vhodov v tak sistem omejeno, posledica česar je, da se morajo uporabniki naučiti, kaj lahko v določenem kontekstu povedo in kaj ne. Najenostavnejši način določanja stopnje učljivosti je merjenje povprečnega časa, ki ga novi uporabniki porabijo, dokler ne dosežejo določene stopnje izkušenosti.

Naslednji pojem je *učinkovitost* govornega vmesnika, ki je definiran kot stopnja storilnosti sistema pri interakciji z izkušenimi uporabniki, kot imenujemo uporabnike, ki so si nabrali s sistemom določeno količino izkušenj. Eden izmed načinov, kako meriti učinkovitost govornega vmesnika, je določiti značilne naloge in meriti čas, ki ga izkušeni uporabniki porabijo, da te naloge izpolnijo.

Pojem *zapomnljivosti* je pomemben predvsem za uporabnike, ki govorni vmesnik uporabljajo le občasno. Stopnjo zapomnljivosti lahko merimo s spominskim testom, kjer uporabnike po končani interakciji s sistemom prosimo, naj razložijo, kako bi se sistem odzival na določene govorne ukaze, in/ali povedo, kateri govorni ukazi so potrebni za izvedbo določene akcije. Drug način merjenja stopnje zapomnljivosti je merjenje časa interakcije s sistemom, ki ga uporabniki potrebujejo za izvedbo določene naloge.

Četrta dimenzija pojma uporabnosti je povezana z *napakami*. V splošnem bi naj uporabniki sistema delali čim manj napak, čeprav je v primeru govornih vmesnikov veliko bolj verjetno, da bo večino napak naredil sistem. Število in pogostost napak, ki jih delajo uporabniki in sistem, ugotovimo tako, da jih preštejemo.

Zadnja dimenzija je subjektivni pojem *zadovoljstva*, ki pove, ali uporabniki dani govorni vmesnik radi uporabljajo. Obstaja več načinov merjenja zadovoljstva, najpogostejši pa je način z vprašalniki, ki jih uporabniki po interakciji s sistemom izpolnijo. Alternativa so merjenja psiho-fizičnih lastnosti uporabnikov, kot sta srčni utrip in krvni tlak. Ta merjenja so sicer bolj objektivna in zanesljiva, vendar so nepriljubljena, ker ponavadi vodijo do nesproščenosti uporabnikov. Swerts in Krahmer (2000) pa sta predlagala merjenje zadovoljstva s pomočjo govornega signala uporabnikov; če interakcija ni tekoča, uporabniki namreč težijo k spremembam govora: k boljši artikulaciji, glasnejšemu in/ali počasnejšemu govoru, ipd.

4.2 Vodila pri oblikovanju sistemov za dialog

Če želimo razvijati sisteme za podajanje informacij, ki bodo ustrezali zahtevam uporabnikov, moramo najprej ugotoviti, kako naj se ti obnašajo. V ta namen si oglejmo osnovni princip in štiri maksime naravne komunikacije, ki jih je Grice (1975) oblikoval kot vodilo sodelujoče komunikacije med ljudmi. Osnovni princip, ki ga naj ne bi kršil nobeden udeleženec dialoga, je **princip sodelovanja**:

Prispevek k dialogu naj ustreza pričakovanjem v trenutnem stanju dialoga.

Njegove maksime pa karakterizirajo *sodelujoč odgovor*, t.j. odgovor, ki je pravilen, uporaben in ni dvoumen.

1. *Maksima kvalitete* zahteva, da naj bodo vsi prispevki k dialogu veljavni. Izjave in podatki udeležencev dialoga morajo biti resnični in ne pretirani. Kaplan (1982) je maksimo kvalitete priredil sistemom za podajanje informacij. Odzivi, ki jih daje takšen sistem, naj uporabnika ne bi zapeljali, in če lahko jedrnat odgovor uporabnika pripelje do napačnih zaključkov, bi naj sistem z dodatnimi informacijami to preprečil.
2. *Maksima kvantitete* pravi, da naj bodo vsi prispevki k dialogu ravno prav informativni, da ne smejo biti niti preskopi niti preveč nabiti s podrobnostmi.
3. *Maksima odnosa* pravi, da naj se udeleženec dialoga odziva v skladu s cilji in namerami sogovornika.
4. *Maksima načina* pravi, da naj se udeleženec dialoga izogiba dvoumnostim, da naj bo jedrnat. Za sistem za podajanje informacij to pomeni, da naj bodo njegove izjave zasnovane v obliki, ki jo bo uporabnik brez težav razumel. Sistem naj daje čimbolj jedrnate izjave, ne da bi kršil katero izmed ostalih maksim.

Ker je naš namen razviti sistem za dialog, ki bo sodeloval z uporabnikom, nam bodo te maksime služile kot vodilo pri oblikovanju sistema.

Številni raziskovalci, ki se ukvarjajo z uporabnostjo sistemov, so predstavili delno prekrivajoča se vodila za oblikovanje vseh vrst sistemov (npr. Normanovi 4 principi dobre konstrukcije [Norman-1988], Nielsenovih 11 hevristik uporabnosti [Nielsen-1993] in Shneidermanovih 8 zlatih pravil [Shneiderman-1986]), nekateri pa so ta vodila priredili govornim vmesnikom ali zasnovali nova vodila [Hapeshi-1993, Leiser-1993, Lea-1994, Cosky-1995, Gardner-1999, Karat-1999]. Mnoge hevristike in vodila lahko razumemo kot pomoč uporabniku pri snovanju *pojmovnega modela sistema*. Uporabnikov pojmovni model, ki ga ta ponavadi gradi s pomočjo izkušenj, zajema mišljenje oz. predpostavke uporabnika o delovanju sistema in omogoča miselno simulacijo delovanja sistema. Dober pojmovni model bo uporabniku omogočal napoved posledic svojih dejanj. Tudi razvijalci imajo ponavadi oblikovan pojmovni model sistema, pri čemer se v idealnem primeru modela uporabnika in razvijalcev ujemata. Ker pa uporabnik in razvijalci ponavadi ne sodelujejo neposredno (vmesni člen je sistem), ta ideal ni nikoli dosežen.

Krahmer (2001) zato trdi, da mora pri gradnji dobrega pojmovnega modela sistem uporabniku pomagati. To je še posebej težko pri govornih vmesnikih, kjer uporabniki sistema ne vidijo. Gradnjo dobrega pojmovnega modela v govornih vmesnikih bi zagotovo olajšali, če bi vzporedno z govorno komunikacijo vzpostavili tudi grafično komunikacijo, kjer bi sistem na zaslonu, pred katerim bi sedel uporabnik, prikazoval trenutno dogajanje. Okoliščine pa so tiste, ki take multimodalne komunikacije pogosto ne dovoljujejo.

V nadaljevanju bomo opisali verjetno najširše sprejeta vodila pri oblikovanju sistemov, namreč *Shneidermanovih 8 zlatih pravil oblikovanja sistemov* [Shneiderman-1986], in jih, kjer bo to potrebno, s pomočjo literature dopolnili do vodil za govorne vmesnike.

1. *Težiti h konsistentnosti*: Konsistentnost je temelj uporabnosti sistema. Uporabnik si lažje zgradi dober pojmovni model sistema in njegova dejanja lažje napove, če so odzivi sistema konsistentni. Kljub temu pa je včasih, ko želimo uporabnika na kaj opozoriti ali pritegniti njegovo pozornost, bolje uporabiti nekonsistentne in nestandardne odzive.
2. *Rednim uporabnikom omogočiti bližnjice*: Med izkušeno uporabniki sistema so ponavadi zelo velike razlike. Novi, neizkušeni uporabniki ponavadi potrebujejo pomoč, izkušeni uporabniki pa bi svoj cilj radi dosegli čim enostavneje in čim hitreje, zato je treba takšne razlike omogočiti. V govornih vmesnikih to dosežemo z dialogom mešanih iniciativ.
3. *Ponuditi informativno podporo*: Uporabniki naj vedo, katero dejanje so sprožili in kakšen rezultat so dosegli. To je še posebej pomembno pri govornih vmesnikih, ker uporabnik sistema ne vidi in ker zaradi omejitev razpoznavanja govora obstaja možnost, da sistem uporabnika ni pravilno razumel. Odzivi sistema naj bodo zato razumljivo jedrnat, enostaven in konsistentni, vsebujejo pa naj tudi nadaljevalno vprašanje ali navodilo kako nadaljevati [Lea-1994].
4. *Dialog naj bo zaprt*: Zaporedje dejanj naj bo urejeno po skupinah z začetkom, sredino in koncem, s čimer interakcijo strukturiramo in tako omogočimo, da se uporabnik hitreje uči in si zgradi boljši pojmovni model sistema. V primeru govornih vmesnikov je uporabniku treba ponuditi pomoč, da se v dialogu znajde [Lea-1994]. Omogočiti je treba prekinitvev in osvežitvev predhodnega stanja dialoga [Leiser-1993] ter preskoke [Lea-1994].
5. *Preprečevati napake in omogočiti preprosto ravnanje z njimi*: Število napak v veliki meri določa uporabnost sistema, kar je še posebej pomembno pri govornih vmesnikih. Ko se napaka pojavi, naj krivdo prevzame sistem, ki naj uporabniku omogoči, da napako popravi, brez da bi moral vse še enkrat ponoviti [Lea-1994].
6. *Omogočiti enostaven preobrat dejanj*: Uporabnik naj se zaveda, da bo morebitne napake brez težav popravil, kar bo zmanjšalo njegovo zaskrbljenost med interakcijo s sistemom.
7. *Podpreti notranji nadzor*: Uporabniki ponavadi ne marajo kontrole sistema, zato naj bodo pobudniki dejanj. Izjema so situacije, ko pride do napak, saj je v tem primeru ugodneje, če pobudo, ki vodi do rešitve, prevzame sistem.

8. *Zmanjšati kratkočasovna spominska bremena*: To pravilo temelji na omejitvah kratkočasovnega spomina pri človeku, pravi pa, da je zaporedja dejanj in ključnih besed, ki si jih mora uporabniki zapomniti, potrebno minimizirati. To vodi do zanimivega paradoksa pri naravnih govornih vmesnikih. Po eni strani si uporabniku takega vmesnika ni treba zapomniti načina interakcije, po drugi strani pa ga omejitve razpoznavanja govora prisilijo, da uporablja omejen slovar besed, ki jih sistem pozna.

Množice pravil oz. vodil kot so Shneidermanova zlata pravila, so uporabne iz dveh razlogov. Uporabljamo jih lahko že pri razvoju sistema ali pa šele pri hevristični evalvaciji sistema [Nielsen-1993], ki jo ponavadi izvaja manjša skupina strokovnjakov, ki na osnovi danih pravil ovrednoti njegovo uporabnost. Take množice vodil pa imajo ponavadi slabosti. Osnovna težava je, da je večino vodil lažje oblikovati kot izvesti. V primeru Shneidermanovih zlatih pravil so nekatera pravila že na prvi pogled osnovna, druga (npr. 6 in 7) pa se zdijo manj osnovna. Kraemer (2001) si celo zastavlja vprašanje, zakaj je pravilo 8 sploh potrebno. Zakaj se to pravilo ne glasi *zagotoviti dober pojmovni model* [Norman-1988] ali *zagotoviti pomoč in navodila* [Nielsen-1993]. Tudi to sta pomembni pravili, specifični za govorne vmesnike, ki pa ne spadata med osem zlatih pravil.

4.3 Tehnike modeliranja dialoga

Potek dialoga med uporabnikom in sistemom uravnava modul za vodenje dialoga. Ideal mnogih raziskovalcev je strategija vodenja dialoga, ki bi omogočala uporabniku, da se s sistemom pogovarja tako kot s človekom, in znotraj katere bi tako uporabnik kot sistem imela možnost prevzeti iniciativo, ko bi to želela. Tako strategijo pa je težko doseči iz najmanj dveh razlogov; ker predvideva, da bi naj uporabnik imel možnost povedati takorekoč karkoli kadarkoli, kar predstavlja veliko težavo za razumevanje govora, in ker bi moral sistem ves čas slediti iniciativi in se glede na njo fleksibilno odzivati.

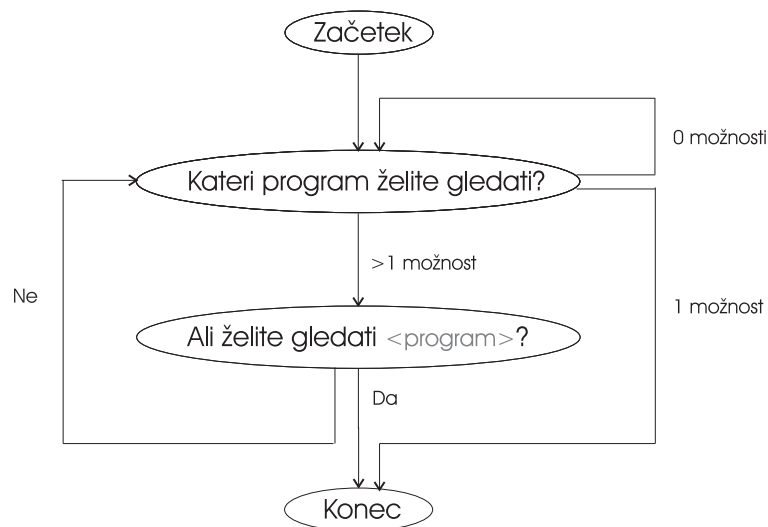
Vsak modul za vodenje dialoga skupaj s strategijo in tehniko modeliranja dialoga predstavlja nekakšen model dialoga. Tehnika, s katero je strategija vodenja dialoga udejanjena, pa mora zadoščati vsaj eni zahtevi, namreč računski obvladljivosti. Opisali bomo naslednje temeljne tehnike modeliranja dialoga:

- ↪ *tehnika končnih avtomatov*,
- ↪ *tehnika polnjenja predalčkov*,
- ↪ *tehnika množice domen*,
- ↪ *tehnika usmerjanja k cilju* in
- ↪ *tehnika razumskih agentov*.

Najpreprostejša tehnika modeliranja dialoga temelji na *končnih avtomatih*, dialog pa je definiran kot pot med začetnim in končnim stanjem avtomata. Ker je število stanj

in število prehodov med stanji v takšnih avtomatih končno in vnaprej določeno, lahko z njimi modeliramo samo preproste dialoge, ki so povsem odvisni od domene, t.j. od komunikacijskega področja. Vzrok je v tem, da ta tehnika ne omogoča sistematičnega ločevanja med nalogo in strategijo vodenja dialoga. Prav zaradi tega je, če želimo model dialoga prenesti na drugo domeno, končni avtomat potrebno ponovno zgraditi. Nadaljnja slabost je hitra eksplozija stanj in prehodov med njimi. Prednosti tehnike končnih avtomatov pa so zagotovo: preprostost, razumljivost in časovna učinkovitost računalniškega algoritma [Larson-1992].

Na sliki 4.1 je prikazan primer modela dialoga za ugotavljanje želenega televizijskega programa, ki temelji na tehniki končnih avtomatov. Dialog se začne v začetnem stanju *Začetek*, od koder se takoj premakne v naslednje stanje (takšen prehod v končnem avtomatu ponavadi imenujemo ϵ -prehod). V tem stanju sistem uporabnika vpraša, kateri televizijski program bi rad gledal. Nadaljevanje dialoga je odvisno od števila hipotez izgovorjenega televizijskega programa, ki jih vrne modul za razpoznavanje govora. Če modul za razpoznavanje ne vrne nobene hipoteze (npr. v primeru, ko uporabnik ni odgovoril na vprašanje), sistem vprašanje ponovi. Če vrne eno samo hipotezo (npr. SLO1), je cilj dialoga dosežen, zato se premakne v končne stanje avtomata *Konec*. Če pa modul za razpoznavanje govora vrne več hipotez izgovorjenega televizijskega programa (npr. SLO1, SLO2), sistem uporabnika vpraša, ali želi gledati enega izmed teh programov (npr. *Ali želite gledati SLO1?*). V primeru, ko razpoznavalnik govora vrne *Da*, je cilj dialoga dosežen, v nasprotnem primeru pa sistem vprašanje ponovi.



Slika 4.1: Preprost primer dialoga, modeliranega s končnim avtomatom

Zaradi slabosti in omejitev dialogov, modeliranih s končnimi avtomati, so se pojavile nove tehnike modeliranja dialoga. *Tehnika polnjenja predalčkov* in *tehnika množice domen* še vedno temeljita na končnih avtomatih, a dialog modelirata na višjem, splošnejšem nivoju. Nekatere tehnike poskušajo končne avtomate obogatiti, npr. z verjetnostnim sklepanjem [Krahmer-2001], ostale tehnike, kot sta *tehnika usmerjanja k cilju* in *tehnika razumskih agentov*, pa gledajo na dialog z drugačne perspektive in se osredotočajo na cilje in namere uporabnika.

Tehnika polnjenja predalčkov [Goddeau-1996] rešuje številne težave dialogov, modeliranih s tehniko končnih avtomatov, in predstavlja ta čas najobičajnejši način modeliranja dialogov v govornih vmesnikih. Vhod v modul za dialog je v tem primeru pomenška predstavitev uporabnikove izjave, ki je podana s predalčki in njihovimi elementi. Napolniti predalčke pomeni pridobiti ustrezne tipe podatkov, ki bodo omogočili dostop do podatkovne zbirke ali izvršitev kake druge akcije. Prednost te tehnike v primerjavi s tehniko končnih avtomatov je boljša prenosljivost na druge domene, kar je omogočeno z ločitvijo naloge in strategije vodenja dialoga; naloga je napolniti predalčke, strategija pa predstavlja način, kako to naredimo. Slabost tehnike polnjenja predalčkov, ki se je sicer izkazala za zelo uporabno v sistemih za podajanje informacij, pa je, da sprememba strategije vodenja dialoga zahteva oblikovanje nove množice pravil.

Tipični primeri sistemov, kjer je tehnika polnjenja predalčkov zelo ustrezna, so sistemi za podajanje informacij o železniških povezavah [Sturm-1999] in sistemi za podajanje informacij o letalskih prevozih [Ipšič-1999, Stallard-2000]. Ti dostopajo do voznega reda, ko od uporabnika pridobijo za to potrebne podatke (vstopna in izstopna postaja, čas in datum potovanja). Primer pomenske predstavitve uporabnikove izjave *Ali leti v torek, 15. junija, kakšno letalo v London?* je:

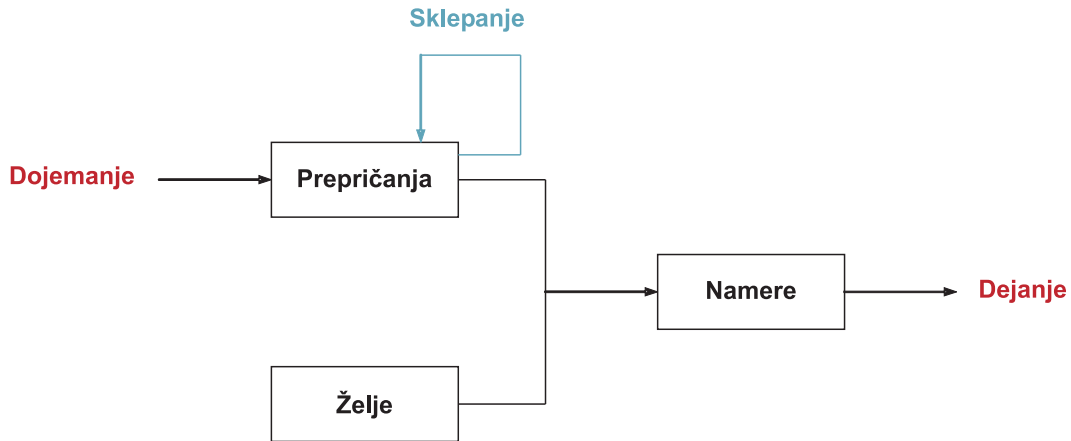
vstopna postaja	–
izstopna postaja	London
datum potovanja	15/6/2004
čas potovanja	00:00»24:00

Naslednja tehnika modeliranja dialoga je *tehnika množice domen* [Xu-2000], ki omogoča preklapljanje med različnimi domenami oz. komunikacijskimi področji (npr. rezervacije hotelov in najemi avtomobilov), modeliranimi s tehniko polnjenja predalčkov. Prednosti in slabosti so podobne kot pri tehniki polnjenja predalčkov, s pomembno razliko, da tehnika množice domen že zajema več različnih domen.

Prva od tehnik nestrukturiranega modeliranja dialoga je *tehnika usmerjanja k cilju* [Allen-1995], ki je z razliko od prej omenjenih tehnik primerna predvsem za sisteme za reševanje problemov, v katerih uporabnik in sistem sodelujeta pri reševanju problema oz. doseganju določenega cilja. Tu se uporablja *teorija govornih dejanj* [Searle-1969], ki trdi, da je izjave uporabnika potrebno obravnavati kot dejanja (npr. vprašanje, zahteva, obvestilo), ki jih je mogoče opredeliti s pomočjo potrebnih in zadostnih pogojev za izvedbo. Te pogoje in njihove posledice lahko definiramo s pomočjo prepričanj, želja in namer poslušalcev in govorcev.

Na sliki 4.2 je prikazan BDI model [Bratman-1988], ki podaja zvezo med izjavami na eni ter prepričanji, željami in namerami uporabnika na drugi strani. Model predpostavlja, da sistem na osnovi predpostavk o trenutnih prepričanjih in željah uporabnika določi komunikacijske namere, ki predstavljajo temelj konstrukcije načrta govornih dejanj, za katerega verjame, da bo pripeljal do cilja. Sistem začne načrt izvajati tako, da sproži začetno dejanje. Izvajanje govornih dejanj, še bolj pa reakcije uporabnika, ponavadi vodijo do novega stanja sveta, t.j. do novih predpostavk o prepričanjih in željah uporabnika. V skladu z njimi, če je potrebno, sistem načrt govornih dejanj spremeni. Prednost te tehnike je nizka odvisnost od domene in dialog, ki je podoben človekovemu,

slabost pa je poleg v splošnem časovno zahtevnega računalniškega algoritma tudi dejstvo, da pojma prepričanje in želja nista natančno formalizirana. To pomeni, da ne obstajajo neodvisni razumski principi, ki bi pojasnjevali obnašanje sistema.



Slika 4.2: BDI model

Obnašanje človeka med dialogom vsekakor najbolje modelira *tehnika razumskih agentov* [Smith-1998], ki s pomočjo logične formalizacije prepričanj, želja in namer tehniko usmerjanja k cilju izboljša. Sadek (1994), ki je eden izmed glavnih zagovornikov uporabe te tehnike modeliranja dialoga v govornih vmesnikih, je uvedel vrsto sprememb in razširitev osnovne tehnike razumskih agentov, predstavljene v [Cohen-1990]. Formalizacijo prepričanj, želja in namer, ki zaobsega tudi množico aksiomov, imenujemo *razumska enota*. Razumska enota temelji na dveh razumskih principih:

- ↪ agent izbira dejanja v skladu s svojim namerami in
- ↪ agentova namera po nekem dejanju vsebuje namero uresničitve predpogojev tega dejanja.

V namene formalizacije z $B(i, p)$ označimo prepričanje i -tega agenta v resničnost izjave p , $G(i, p)$ pa naj pomeni, da je p želja i -tega agenta. Prepričanja lahko formaliziramo z logiko NKD45 [Montague-1960], ki zaobsega naslednje aksiome sklepanja:

- (N) če p vedno velja, potem velja tudi $B(i, p)$
- (K) $B(i, p) \wedge B(i, p \Rightarrow q) \Rightarrow B(i, q)$
- (D) $B(i, p) \Rightarrow \neg B(i, \neg p)$
- (4) $B(i, p) \Rightarrow B(i, B(i, p))$
- (5) $\neg B(i, p) \Rightarrow B(i, \neg B(i, p))$

Aksiom (N) trdi, da agent verjame vse, kar vedno velja, aksiom (K) pravi, da so prepričanja zaprta za implikacijo, aksiom (D) pa, da so prepričanja konsistentna (če je agent i prepričan v resničnost izjave p , potem ne more verjeti tudi $\neg p$). Aksioma (4) in (5) opisujeta pozitivno oz. negativno introspekcijo (če je agent i prepričan v resničnost izjave p , potem verjame v svoje prepričanje o resničnosti izjave p ; če agent i ni prepričan v resničnost izjave p , potem verjame, da sam ni prepričan v resničnost p).

Podobno lahko formaliziramo tudi želje:

$$(K) \quad G(i, p) \wedge G(i, p \Rightarrow q) \Rightarrow G(i, q)$$

$$(D) \quad G(i, p) \Rightarrow \neg G(i, \neg p)$$

Povezava med prepričanji in željami je v [Cohen-1990] podana z dvema omejitvama; z *razumno omejitvijo*

$$\neg B(i, p) \Rightarrow \neg G(i, p),$$

ki trdi, da si agent ne more želeli nekaj, za kar verjame, da ni res, in s *pričakovano omejitvijo implikacije*

$$G(i, p) \wedge B(i, p \Rightarrow q) \Rightarrow G(i, q),$$

ki trdi, da so tudi pričakovane posledice ciljev cilji agenta. Čeprav so vsa podana pravila sklepanja preprosta in intuitivna, pa se ta, če želimo dovoliti, da agent obstoječo željo tudi opusti, hitro zakomplicirajo.

Modul za vodenje dialoga kot razumski agent uravnava dialog tako, da na osnovi pomenske predstavitve uporabnikove izjave in pravil sklepanja prihaja do predpostavk o prepričanjih in željah uporabnika. Prednost te tehnike, poleg boljšega modeliranja človekovega obnašanja med dialogom, je, da je teorija od domene in jezika takorekoč neodvisna. Ko želimo govorni vmesnik prirediti za novo domeno, je potrebno le aksiomatizirati znanje te domene. Slabost pa je v kompleksnosti in moči uporabljene logike, ki logiko 1. reda kombinira z epistemično logiko (za sklepanja o prepričanjih), s časovno logiko (za sklepanja glede časa) in z dinamično logiko (za sklepanja o dejanjih). Nadaljna slabost je časovno zahteven računalniški postopek osveževanja množice prepričanj in želja ter dejstvo, da je formalizacija relevantnega znanja domene v splošnem težka naloga.

Za primer vzemimo izjavo uporabnika, ki jo sistem razume kot vprašanje *Ali velja p* (npr. *Ali je na sporedu večernega televizijskega programa SLO2 detektivka?*). Sistem izpelje sklep, da je želja uporabnika izvedeti, ali velja p , in da uporabnik trenutno ni prepričan v resničnost nobene izmed obeh možnosti (p ali $\neg p$). Po principu sodelovanja sistem to namero prevzame in na osnovi prvega razumskega principa se odloči uporabnika obvestiti o tem, katera izmed možnosti p in $\neg p$ velja. Na osnovi drugega razumskega principa sistem ugotovi, za katero od obeh možnosti so predpogoji izpolnjeni. Če sistem (ne pa tudi uporabnik) verjame, da velja p , to pomeni, da so izpolnjeni predpogoji prvega dejanja. Sistem to dejanje izbere in o tem obvesti modul za tvorjenje naravnega jezika (*Da, na SLO2 bo danes zvečer detektivka.*) ter osveži svoje znanje. Ena izmed posledic tega dejanja je prepričanje uporabnika, da velja p .

Tehnike modeliranja dialoga, ki smo jih navedli, kljub svojim pomanjkljivostim pokrivajo večino možnih aplikacij dialoga med človekom in računalnikom (*uporabni dialogi*), ne pa tudi celotnega obsega naravne človeške komunikacije.

5 Modeliranje dialoga z uporabo konverzacijske teorije iger

5.1 Konverzacijska teorija iger

5.2 Modeliranje dialoga v sistemu za podajanje informacij o vremenu

Za modeliranja dialoga v sistemu za podajanje informacij o vremenu smo se odločili uporabiti tehniko usmerjanja k cilju, delno formalizirano s pomočjo *konverzacijske teorije iger*, pri čemer podatke pridobivamo s polnjenjem predalčkov. V takem modelu je dialog sestavljen iz *konverzacijskih iger* in *konverzacijskih potez*, ki predstavljajo različne strategije in so udejanjene kot *rekurzivne mreže prehodov*.

Da bi dosegli čim večjo kooperativnost in čim večjo neodvisnost od domene, smo množico konverzacijskih iger in potez, ki se pojavljajo v literaturi, ustrezno razširili.

5.1 Konverzacijska teorija iger

Pri modeliranju dialoga v sistemu za podajanje informacij o vremenu smo uporabili *konverzacijsko teorijo iger*, ki izhaja iz množice raziskav [Power-1979, Houghton-1987, Kowtko-1992, Allen-1997, Poesio-1997, Stolcke-2000], kjer so dialogi sestavljeni iz iger, ki jih načrtujejo razumski agenti. Namen [Power-1979] in [Houghton-1987] je bil izdelati teorijo, ki bo opisovala, kako cilji sogovornika vplivajo na potek dialoga. Na osnovi teh ugotovitev sta Kowtko in Isard (1992) dialoge strukturirala na dva funkcijska nivoja:

↪ *konverzacijske igre* in

↪ *konverzacijske poteze*.

Konverzacijske igre povezujemo z željami oz. konverzacijskimi cilji, kot je na primer cilj pridobiti določeno informacijo, in so sestavljene iz zaporedja izjav, ki se začno z iniciativo in končajo, ko je cilj igre dosežen ali igra prekinjena. Sestavne dele konverzacijskih iger imenujemo konverzacijske poteze. To so izjave, deli izjav ali množice izjav, ki izražajo isto namero kot je na primer potrditev ali preverjanje. Konverzacijske igre lahko poljubno gnezdimo eno v drugi, če le cilj notranje konverzacijske igre služi izpolnitvi cilja zunanje igre.

Najbolj znana uporaba konverzacijske teorije iger je kodiranje obsežne (15 ur) zbirke spontanih, k cilju usmerjenih, govornih dialogov človek – človek [Carletta-1996a], kjer je dodan še en funkcijski nivo, namreč nivo *dogovorov*, t.j. delov dialoga, ki predstavljajo večje korake pri doseganju končnega cilja. Naloga enega izmed udeležencev teh dialogov je bila podvojiti zemljevid, ki ga je dobil drugi udeleženeec. Primer takega zemljevida je na sliki 5.1. Dolžina in oblika dogovorov v teh dialogih je bila v veliki meri odvisna od zemljevida, ki so ga udeleženci dobili in na podlagi katerega so si nalogo razdelili na obvladljive kose. Tipičen primer dogovora je del dialoga, v katerem udeleženeec, ki zemljevid podvaja, uspe narisati nek njegov odsek. Študija zbirke teh dialogov je pokazala, da v konverzacijskih igrah vprašanjem v veliki večini sledijo odgovori, izjavam pa odobritve ali zavrnitve.

Na osnovi te študije avtorji razlikujejo tri osnovne tipe konverzacijskih potez:

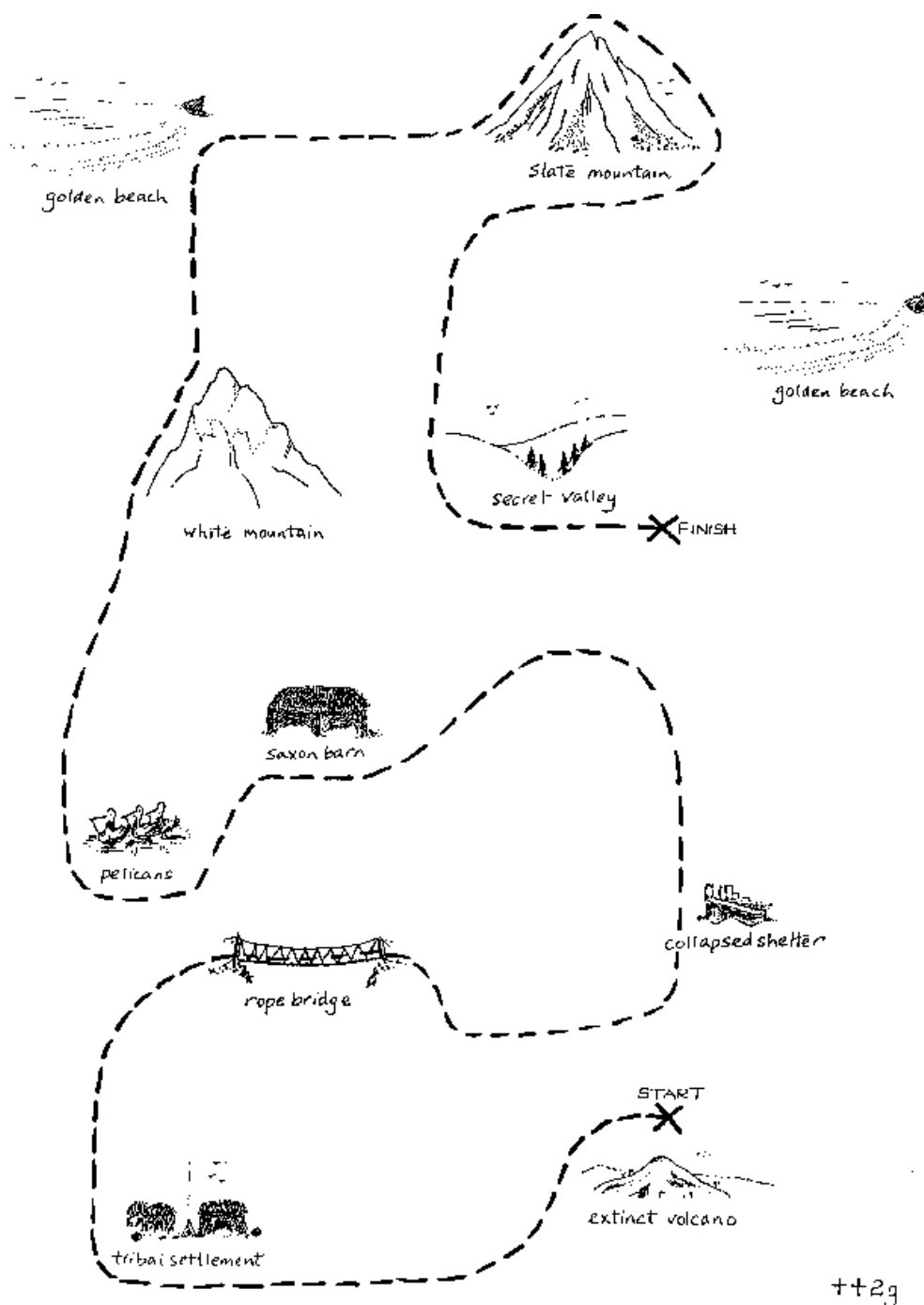
↪ *začetne poteze*,

↪ *odzivne poteze* in

↪ *poteze pripravljenosti*.

Začetne poteze se pojavljajo na začetku igre, so iniciatorji. Odzivne poteze se pojavljajo znotraj iger in so namenjene izpolnitvi pričakovanj oz. cilja, poteze pripravljenosti pa se pojavljajo potem, ko se igra zaključi, in pripravljajo dialog na začetek nove igre. Zbirka govornih dialogov [Carletta-1996a] je zakodirana z množico 12 kategorij konverzacijskih potez; z začetnimi potezami

- INSTRUCT (navodilo),



Slika 5.1: Primer zemljevida, uporabljenega v spontanah, govornih dialogih človek – človek

- EXPLAIN (razlaga),
- CHECK (preverjanje posredovane informacije),
- ALIGN (preverjanje pozornosti ali strinjanja sogovornika),
- QUERY-YN (vprašanje, na katero je mogoče odgovoriti z da ali ne) in
- QUERY-W (vprašanje, ki ne spada v nobeno drugo kategorijo),

z odzivnimi potezami

- ACKNOWLEDGE (razumevanje in sprejemanje prejšnje poteze),
- REPLY-Y (odziv z da na QUERY-YN potezo),
- REPLY-N (odziv z ne na QUERY-YN potezo),
- REPLY-W (odziv na katerokoli potezo, ki ne pomeni le da ali ne) in
- CLARIFY (odziv na katerokoli vprašanje, ki vsebuje več informacij kot vprašanje zahteva),

ter s potezo pripravljenosti

- READY (zaključek trenutne in priprava na začetek nove igre).

Ideja pojmovanja konverzacijskih iger in potez se je pri modeliranju dialoga izkazala za zelo uporabno, in sicer tako z opisnega [Power-1979, Houghton-1987, Kowtko-1992, Carletta-1996a] kakor tudi z algoritemskega [Lewin-2000] vidika. Konverzacijska teorija iger se tako predvsem ukvarja z modeliranjem dialogov človek – človek in dialogov človek – računalnik. Uporaba te teorije za modeliranje dialoga v sistemih, kjer konverzacijske poteze osvežujejo, dopolnjujejo in urejajo množico prepričanj in želja, je bila predlagana v projektu TRINDI (Task Oriented Instructional Dialogue) [Larsson-2000, Lewin-2000]. Za modeliranje dialoga v sistemu AUTOROUTE DEMONSTRATOR, ki je bil razvit v tem projektu, je bilo udejanjenih pet različnih konverzacijskih iger:

- *query game*, ki jo iniciira ena od potez povpraševanja, *qw* ali *qw-r*, in ji sledi odziv *rw*, ki sproži potezo sprejemanja *ack* takoj ali pa šele po zaporedju prošnje za potrditev *cnf* in odziva na njo, *ryes*, *rno* ali *rmod*,
- *information-giving game*, ki je sestavljena le iz poteze podajanja informacije *inf* in poteze sprejemanja tega odgovora *ack*,
- *pardon game*, ki je sestavljena iz sogovorniku nerazumljive izjave *unrec* in prošnje za ponovitev *pdm*,
- *interruption game*, ki je sestavljena iz poteze podajanja nepomembne informacije *unimp* in igre *information-giving game*, ter
- *hello game*, ki je sestavljena iz dveh pozdravnih potez *hello*.

Naštete konverzacijske igre so formalizirane kot *rekurzivne mreže prehodov*, t.j. kot diagrami oz. načrti, sestavljeni iz prehodov, ki predstavljajo konverzacijske poteze, te pa ustrezajo dejanjem. Rekurzivne mreže prehodov so bolj kompleksne od končnih avtomatov, saj lahko eno mrežo poljubno gnezdimo v drugi, če cilj konverzacijske igre, ki je predstavljena z notranjo mrežo, služi izpolnitvi cilja igre, ki je predstavljena z zunanjo mrežo.

5.2 Modeliranje dialoga v sistemu za podajanje informacij o vremenu

Eden izmed ciljev pri razvoju sistema za podajanje informacij o vremenu je bil razviti modul za vodenje dialoga, ki bi bil od jezika neodvisen in čim enostavneje prenosljiv na druge domene podajanja informacij. Pretehtali smo učinkovitost uveljavljenih tehnik modeliranja dialoga in po temeljitem razmisleku prišli do zaključka – dialog modelirati s tehniko usmerjanja k cilju, pri čemer smo v namen strukturiranja uporabili konverzacijsko teorijo iger, t.j. konverzacijske igre in poteze. Podatke, potrebne za dostop do podatkovne zbirke, pridobivamo s polnjenjem predalčkov. Kot temelj definicije ustreznih konverzacijskih iger in potez so nam služili podatki iz prvega eksperimenta Čarovnik iz Oza (poglavje 3) ter sistem kodiranja govorjenih dialogov človek – človek [Carletta-1996a], ki smo ga opisali v prejšnjem podpoglavju.

Tudi mi, podobno kot [Carletta-1996a], ločimo tri osnovne tipe konverzacijskih potez:

- ↪ *začetne poteze* se ponavadi pojavljajo na začetku konverzacijskih iger in v dialog uvajajo nove namere;
- ↪ *odzivne poteze* se pojavljajo znotraj konverzacijskih iger in služijo izpolnitvi pričakovanj in želja, ki se v dialogu pojavijo;
- ↪ *poteze pripravljenosti* se pojavijo potem, ko je bila neka konverzacijska igra zaključena, in dialog pripravijo na začetek nove igre.

Množici konverzacijskih iger in konverzacijskih potez, ki so bile udejanjene v projektu TRINDI [Larsson-2000], smo precej razširili, da bi zagotovili večjo prenosljivost sistema na druge domene. Definirali smo 14 kategorij začetnih in 10 kategorij odzivnih potez ter poteze pripravljenosti READY, ki naznanjajo konec prejšnje igre in dialog pripravljajo na začetek nove konverzacijske igre.

Kategorije začetnih konverzacijskih potez, ki smo jih udejanjili v modulu za vodenje dialoga v sistemu za podajanje informacij o vremenu, so:

- GREET
Pozdravi (npr. *Dober dan.* ali *Nasvidenje.*).
- INDECIPHERABLE
Izjave uporabnika, ki jih sistem ne razpozna ali ne razume.
- PARDON
Izjave, s katerimi uporabnik ali sistem prosita sogovornika, naj ponovi prejšnjo izjavo (npr. *Prosim?* ali *Ali lahko ponovite vprašanje, prosim?*).
- HELP
Izjave, s katerimi uporabnik prosi za pomoč, ko se v dialogu ne znajde oz. ne ve, kako nadaljevati (npr. *Kaj pa znaš?*).

- **TIMEOUT**
Zaobsega tiste odzive uporabnika, ko sistem v vnaprej določenem času ne zazna govora.
- **INTERRUPT**
Izjave, s katerimi uporabnik prekine tekočo konverzacijsko igro (npr. *Ne, to me pa ne zanima.*).
- **ALIGN**
Izjave, s katerimi uporabnik preveri, ali se njegova prepričanja ujemajo s prepričanji sistema.
- **CHECK**
Izjave, s katerimi sistem prosi za potrditev informacij, za katere sicer že verjame, da držijo, vendar v to ni povsem prepričan. Te poteze pokrivajo sklepanje na podlagi zgodovine dialoga (npr. *Ali sprašujete za ta trenutek?* ali *Ali sprašujete za ponedeljek o vremenu?*).
- **END**
Izjave, s katerimi uporabnik zaključi dialog (npr. *Ne, nimam več vprašanj.* ali *Hvala, to bi bilo vse.*).
- **QUERY-YN**
Vprašanja, na katera je mogoče odgovoriti z da ali z ne in ne sodijo v kategorijo CHECK ali ALIGN potez (npr. *Ali bo deževalo?* ali *Potem me pa zanima, če je burja v Novi Gorici.*).
- **QUERY-WR**
Vprašanja po določenih informacijah ali pojasnilih. V to kategorijo sodijo tudi vsa vprašanja, ki ne spadajo v nobeno od ostalih kategorij (npr. *Mi lahko poveste, kakšno bo v naslednjih dneh v hribih vreme?* ali *Kateri kraj oziroma katera pokrajina vas zanima?*).
- **QUERY-WI**
Izjave, s katerimi uporabnik prosi sistem, da mu našteje vse vrste podatkov za določen kraj in določen časovni trenutek, do katerih trenutno dostopa (npr. *Zanima me, kakšne informacije imate o razmerah na smučišču Vogel.*).
- **QUERY-WL**
Izjave, s katerimi uporabnik prosi sistem, da mu našteje vse pokrajine in kraje v nekem območju, za katere lahko ponudi določeno vrsto podatka ob določenem časovnem trenutku (npr. *Zanima me, za katere kraje v Sloveniji imate podatke o višini snega.*).
- **QUERY-WT**
Izjave, s katerimi uporabnik prosi sistem, da mu našteje vse časovne trenutke, za katere lahko ponudi določeno vrsto podatka za določeno pokrajino ali kraj (npr. *Zanima me, za koliko naprej lahko napoveste vreme.*).

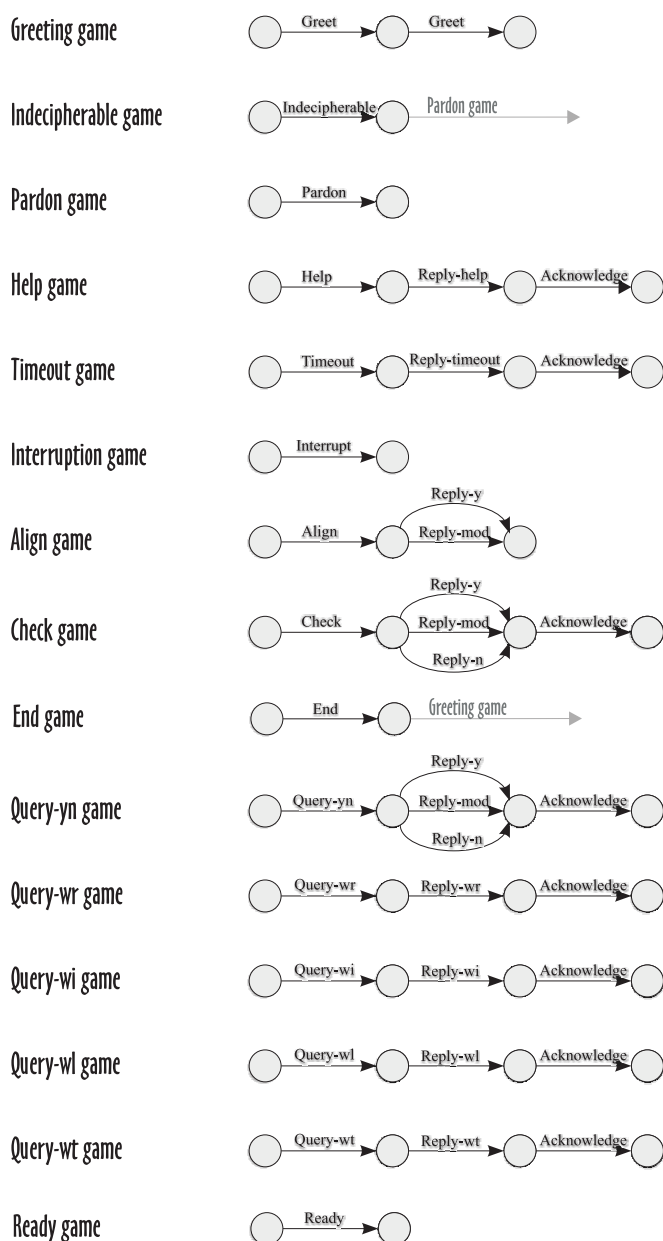
V skladu s kategorijami začetnih konverzacijskih potez smo udejanjili naslednje kategorije odzivnih potez:

- ACKNOWLEDGE
Odzivi, s katerimi uporabnik ali sistem minimalno pokažeta, da sta potezo, na katero se odzivata, slišala in razumela (npr. *A ne?* ali *Hvala.*).
- REPLY-HELP
Odzivi sistema na HELP potezo (npr. *Sedaj me lahko vprašate še kaj v zvezi z mojim odgovorom ali pa mi zastavite novo vprašanje, če ga imate. ali Odgovarjati znam na vprašanja o vremenu, vetru, temperaturi in zračnem tlaku za različne kraje v Sloveniji in Evropi, o času vzhoda in zahoda sonca, o snežnih razmerah, o splošni vremenski napovedi, o biovremenski napovedi, o napovedi za letalstvo ter hidrološki napovedi za Slovenijo, o napovedi za slovensko primorje ... in mogoče še o čem.*).
- REPLY-TIMEOUT
Odzivi sistema na TIMEOUT potezo (npr. *Ali imate še kakšno vprašanje? ali Niste mi povedali vseh potrebnih podatkov ali pa jih nisem razumel. Če vas zanima kaj drugega, vprašajte, sicer pa prosim odgovorite na naslednje vprašanje: Ali sprašujete za jutri čez dan?*).
- REPLY-Y
Odzivi z da na vprašanja, na katera je mogoče odgovoriti z da ali z ne (QUERY-YN, CHECK, ALIGN), ki pomenijo potrditev (npr. *Da.* ali *Ja, za ta trenutek ... če piha ali ne piha.*).
- REPLY-N
Odzivi z ne na vprašanja, na katera je mogoče odgovoriti z da ali z ne (QUERY-YN, CHECK), ki pomenijo zanikanje (npr. *Ne.*).
- REPLY-MOD
Odzivi s popravkom na vprašanja, na katera je mogoče odgovoriti z da ali z ne (QUERY-YN, CHECK, ALIGN) (npr. *Ne, sprašujem za Kope, Pohorje. ali Ne, za prejšnji teden.*).
- REPLY-WR
Odzivi na QUERY-WR poteze (npr. *V Sloveniji jutri čez dan – najvišje dnevne temperature bodo od 2 do 5 stopinj Celzija. Imate še kakšno vprašanje? ali Kranjska Gora.*).
- REPLY-WI
Odzivi sistema na QUERY-WI poteze (npr. *Za ta trenutek na Voglu vam lahko ponudim podatke o vremenu, temperaturi zraka, višini novozapadlega snega, skupni višini snega in hitrosti vetra. Ali vas kakšna od teh informacij zanima?*).
- REPLY-WL
Odzivi sistema na QUERY-WL poteze (npr. *Podatke o skupni višini snega za Slovenijo za ta trenutek imam za Vojsko, za Rateče, za Vernerico, za Krovavec, za*

Kanin, za Kredarico, za Predel in za Vogel. Torej, ali vas sedaj zanima kakšen od teh krajev?).

- REPLY-WT

Odzivi sistema na QUERY-WT poteze (npr. *Podatke o temperaturi zraka za Slovenijo imam za danes čez dan, za jutri zjutraj in za jutri čez dan. Želite, da vam podam informacijo za kakšnega od naštetih časovnih trenutkov?*).



Slika 5.2: Konverzacijske igre, udejanjene v modulu za vodenje dialoga v sistemu za podajanje informacij o vremenu

Na osnovi teh konverzacijskih potez smo definirali 15 konverzacijskih iger, za vsako od 14 začetnih potez in potezo pripravljenosti po eno, namreč: GREETING GAME, INDECIPHERABLE GAME, PARDON GAME, HELP GAME, TIMEOUT GAME, INTERRUPTION

GAME, ALIGN GAME, CHECK GAME, END GAME, QUERY-YN GAME, QUERY-WR GAME, QUERY-WI GAME, QUERY-WL GAME, QUERY-WT GAME in READY GAME. Te igre so sestavljene iz konverzacijskih potez, udejanjene pa so kot rekurzivne mreže prehodov (slika 5.2), ki dovoljujejo poljubna gnezdenja, če le cilj notranje igre služi izpolnitvi cilja zunanje igre.

<i>S1:</i>	<i>Pozdravljeni. Pogovarjate se s sistemom za podajanje informacij o vremenu. Kakšno je vaše vprašanje?</i>
<i>U1:</i>	<i>Zanima me, kakšno vreme bo jutri v Sloveniji.</i>
<i>S2:</i>	<i>V Sloveniji jutri čez dan – padavine bodo ponehale, delno razjasnitve. Želite vedeti še kaj?</i>
<i>U2:</i>	<i>Ja. Zanima me še vreme v gorah.</i>
<i>S3:</i>	<i>Ali sprašujete za jutri čez dan?</i>
<i>U3:</i>	<i>Kako?</i>
<i>S4:</i>	<i>Niste mi povedali vseh potrebnih podatkov ali pa jih nisem razumel. Če vas zanima kaj drugega, vprašajte, sicer pa prosim odgovorite na naslednje vprašanje. Ali sprašujete za jutri čez dan?</i>
<i>U4:</i>	<i>Ne, zanima me vreme za ta trenutek.</i>
<i>S5:</i>	<i>O vremenu v gorah ta trenutek žal nimam podatkov. Imam pa podatke o vremenu ta trenutek na Kredarici, na Voglu, na Kaninu, na Kravcu in v Predelu. Potem lahko izbirate med podatki o vremenu v gorah za danes čez dan in jutri čez dan. Izbirate lahko tudi med podatki o vremenski napovedi v gorah ta trenutek.</i>
<i>U5:</i>	<i>Povejte mi prosim za Vogel.</i>
<i>S6:</i>	<i>Ali sprašujete za ta trenutek o vremenu?</i>
<i>U6:</i>	<i>Ja.</i>
<i>S7:</i>	<i>Na Voglu ta trenutek – megla. Vam lahko še kako pomagam?</i>
<i>U7:</i>	<i>Ja. Povejte mi, kakšna je temperatura zraka in kakšna je višina snega na Voglu.</i>
<i>S8:</i>	<i>Na Voglu ta trenutek – skupna višina snega je 180 centimetrov, temperatura zraka je -7 stopinj Celzija. Imate še kakšno vprašanje?</i>
<i>U8:</i>	<i>Ja. Kaj vse mi lahko poveste o trenutnem vremenskem stanju v Ljubljani?</i>
<i>S9:</i>	<i>Za ta trenutek v Ljubljani vam lahko ponudim 9 vrst informacij kot na primer podatek o vremenu, vlažnosti zraka, temperaturi zraka, hitrosti vetra, smeri vetra, vrednosti zračnega tlaka, indeksu požarne ogroženosti, vrsti požarne ogroženosti in tako naprej. Ali vas kakšna od teh informacij zanima?</i>
<i>U9:</i>	<i>Ne. To bi bilo vse.</i>
<i>S10:</i>	<i>Hvala lepa za sodelovanje. Nasvidenje.</i>

Tabela 5.1: Primer dialoga med uporabnikom (U) in sistemom za dialog (S)

Lastnost, ki izrazito vpliva na uporabnost in posledično tudi na uspešnost sistemov za dialog, je sposobnost obvladovanja raznolikih konverzacijskih strategij. Ta vsebujejo osnovna obnašanja kot so preverjanja in potrditve, prevzemanje iniciative, nuditi ustrezno pomoč uporabniku, ponovitev zadnje izjave, prekinitev dialoga in ponovna vzpostavitev konteksta. Predstavljen modul za vodenje dialoga, ki je del sistema za

podajanje informacij o vremenu, podpira vse omenjene konverzacijske strategije, te pa so udejanjene kot konverzacijske igre.

Zue in sodelavci [Zue-2000] trdijo, da je eden od odločilnih vidikov uporabnosti sistemov za dialog sposobnost uporabnika informirati o obsegu znanja, ki ga sistem ima. Uporabnik preprosto mora imeti možnost vprašati in celo biti usmerjen k izbiri relevantnih, dosegljivih podatkov, ko sistem ni sposoben odgovoriti na njegovo vprašanje. Le tako bo namreč sposoben zgraditi dober pojmovni model sistema. V našem modulu za vodenje dialoga to potrebo delno zadovoljimo z udejanjenjem konverzacijskih iger QUERY-WI GAME, QUERY-WL GAME in QUERY-WT GAME, ki poleg tega dodatno strukturirajo dialog. Po drugi strani pa usmerjanje uporabnika k izbiri relevantnih, dosegljivih podatkov v veliki meri podpira tudi predstavitev časovno odvisne strukture znanja sistema, predstavljene v poglavju 6.

Nepogrešljive pri vzpostavljanju konteksta so konverzacijske igre CHECK GAME, ki se nanašajo na zgodovino dialoga. V teh igrah sistem uporabnika prosi, naj potrdi predpostavke, do katerih je prišel na podlagi prejšnjih konverzacijskih dejanj.

Vhod v modul za vodenje dialoga je pomenska predstavitev uporabnikove izjave, ki je sestavljena iz imena konverzacijske poteze in elementov ustreznih predalčkov. Tipi predalčkov so zelo odvisni od domene oz. sestave podatkovne zbirke, do katerih dostopamo. V primeru domene vremenskih podatkov smo definirali tri tipe predalčkov, namreč *kraj*, *čas* in *podatek*, pri čemer je množica predalčkov, ki jih sistem mora zapolniti, preden se lahko odzove, odvisna od pomena uporabnikove izjave oz. od tipa konverzacijske poteze, ki jo uporabnik naredi. Pri izjavah, ki ustrezajo potezam QUERY-YN in QUERY-WR, je potrebno zapolniti vse tri predalčke, pri izjavah, ki ustrezajo potezam QUERY-WI, le predalčka *kraj* in *čas*, pri QUERY-WL le predalčka *čas* in *podatek*, če se želimo omejiti na določeno območje, pa tudi *kraj*, pri QUERY-WT pa le *kraj* in *podatek*. Pri potezi ALIGN, razen zapolnitve vsaj enega predalčka, ni omejitev, saj se uporabnik sam odloča, kaj želi preveriti. Ko pa se uporabnik s potezo REPLY-WR odziva na vprašanje QUERY-WR, ki mu ga je zastavil sistem, so število in tipi potrebnih predalčkov odvisni od vprašanja. Vse ostale možne poteze uporabnika ne zahtevajo zapolnitve nobenega predalčka.

Primer dialoga med uporabnikom in sistemom je v tabeli 5.1 podan v naravnem jeziku, spodaj pa je ta dialog preveden v pomenske predstavitve izjav, ki so vhod v modul za vodenje dialoga, s katerim uporabnik komunicira preko grafičnega vmesnika.

GREETING GAME

S1: (GREET) *Dober dan. Pogovarjate se s sistemom za podajanje informacij o vremenu. Kakšno je vaše vprašanje?*

QUERY-WR GAME

U1: QUERY-WR

kraj: Slovenija

čas: jutri podnevi

podatek: vreme

S2: (REPLY-WR) *V Sloveniji jutri čez dan – padavine bodo ponehale, delno razjasnitve. Želite vedeti še kaj?*

READY GAME

U2: READY

QUERY-WR GAME

U2: QUERY-WR

kraj: gore

čas: –

podatek: vreme

CHECK GAME

S3: (CHECK) *Ali sprašujete za jutri čez dan?*

HELP GAME

U3: HELP

S4: (REPLY-HELP) *Niste mi povedali vseh potrebnih podatkov ali pa jih nisem razumel. Če vas zanima kaj drugega, vprašajte, sicer pa prosim odgovorite na naslednje vprašanje. Ali sprašujete za jutri čez dan?*

U4: REPLY-MOD

čas: trenutno

S5: (REPLY-WR) *O vremenu v gorah ta trenutek žal nimam podatkov. Imam pa podatke o vremenu ta trenutek na Kredarici, na Voglu, na Kaninu, na Krvavcu in v Predelu. Potem lahko izbirate med podatki o vremenu v gorah za danes čez dan in jutri čez dan. Izbirate lahko tudi med podatki o vremenski napovedi v gorah ta trenutek.*

QUERY-WR GAME

U5: QUERY-WR

kraj: Vogel

čas: –

podatek: –

CHECK GAME

S6: (CHECK) *Ali sprašujete za ta trenutek o vremenu?*

U6: REPLY-Y

S7: (REPLY-WR) *Na Voglu ta trenutek – megla. Vam lahko še kako pomagam?*

READY GAME

U7: READY

QUERY-WR GAME

U7: QUERY-WR

kraj: Vogel

čas: trenutno

podatek: temperatura, skupna višina snega

S8: (REPLY-WR) *Na Voglu ta trenutek – skupna višina snega je 180 centimetrov, temperatura zraka je -7 stopinj Celzija. Imate še kakšno vprašanje?*

READY GAME

U8: READY

QUERY-WI GAME

U8: QUERY-WI

kraj: Ljubljana

čas: trenutno

S9: (REPLY-WI) *Za ta trenutek v Ljubljani vam lahko ponudim 9 vrst informacij kot na primer podatek o vremenu, vlažnosti zraka, temperaturi zraka, hitrosti vetra, smeri vetra, vrednosti zračnega tlaka, indeksu požarne ogroženosti, vrsti požarne ogroženosti in tako naprej. Ali vas kakšna od teh informacij zanima?*

END GAME

U9: END

Greeting game

S10: (GREET) *Hvala lepa za sodelovanje. Nasvidenje.*

Ta primer dialoga ponazarja udejanjenje številnih konverzacijskih strategij: dve igri GREETING GAME, ki se začneta na pobudo izjav S1 in S10; štiri igre QUERY-WR GAME, ki se začno na pobudo izjav U1, U2, U5 in U7; dve igri CHECK GAME, ki se začneta na pobudo izjav S3 in S6; HELP GAME se začne na pobudo izjave U3; QUERY-WI GAME se začne na pobudo izjave U8; END GAME se začne na pobudo izjave U9; tri igre READY GAME pa uvajajo začetki izjav U2, U7 in U8 ter dialog pripravljajo na začetek nove igre. Odzivna konverzacijska poteza REPLY-WR, pripisana izjavi S5, kaže, na kakšen način sistem uporabnika usmerja k izbiri relevantnih, dosegljivih podatkov, ko ne najde točnega odgovora na uporabnikovo vprašanje.

V naslednjem poglavju 6 bomo opisali predstavitev znanja, ki takšno usmerjanje k izbiri relevantnih, dosegljivih podatkov omogoča.

6 Predstavitev znanja v sistemu za dialog

6.1 Modalna logika

6.2 Intuicionistična modalna logika

6.3 Predstavitev vremenskih podatkov

Podali bomo motivacijo za uporabo modalne logike ter definirali *Kripkejeve modele* osnovne modalne logike, ki poleg običajnih izjavnih veznikov vsebujejo še modalna operatorja *škatla* in *diamant*. Iz klasičnih modalnih logik bomo prešli na intuicionistične logike oz. na intuicionistične modalne logike, ki ne vsebujejo Aristotelovega zakona o izključeni tretji možnosti ($A \vee \neg A$).

Intuicionistično modalno logiko bomo uporabili za predstavitev znanja v sistemu za podajanje informacij o vremenu, za kar smo se odločili predvsem zaradi časovno odvisne strukture vira vremenskih podatkov. Uvedli bomo dva nova modalna operatorja, imenovana *spodnji* in *zgornji diamant*, ter razred *dialog mIPC Kripkejevih modelov*.

Takšna predstavitev znanja sistemu omogoča, da se na uporabnikova vprašanja odziva z usmerjanjem k izbiri relevantnih, dosegljivih podatkov, pri čemer relevantnost temelji na relaciji *biti manj specifičen*, definirani med posameznimi delčki informacij (kraji, časovni trenutki in vrste podatkov).

6.1 Modalna logika

Ljudje v vsakdanjem življenju razlikujemo več oblik resnic, kot na primer *biti nujno res*, *vedeti, da je res*, *verjeti, da je res* in *biti res v prihodnosti*. Izjava

Janez Drnovšek je predsednik Republike Slovenije.

je trenutno sicer resnična, vendar se bo to v prihodnosti spremenilo. Še bolj nazoren primer je izjava

Sončni sistem sestavlja devet planetov.

Čeprav smo še pred kratkim v resničnost te izjave bili precej prepričani, se je to občutno spremenilo, ko so astronomi odkrili telo, imenovano Sedna, ki bi lahko bil deseti planet našega osončja. Tudi če se to ne bi zgodilo, ta izjava, zaradi omejitev sodobne znanosti, ne bi bila nujno resnična. Po drugi strani pa izjava

Kvadratni koren iz 25 je 5.

ni samo resnična, ampak tudi nujno resnična in resnična v prihodnosti. Vseeno pa resničnost te izjave nekaterim ljudem (npr. majhnim otrokom) ni znana, drugi (npr. tisti, ki imajo zmotno znanje) pa v njeno resničnost ne verjamejo.

Ker se v umetni inteligenci pogosto pojavljajo sistemi več agentov, od katerih ima lahko vsak agent drugačno znanje o svetu (okolju) in o znanju drugih agentov, brez *modalne logike* [Huth-2000] o znanju agentov takorekoč ne bi mogli sklepati. Modalna logika poleg klasičnih izjavnih veznikov uporablja *modalne operatorje* (enomestne izjavne veznike), s pomočjo katerih je sposobna izraziti eno ali več oblik resnice. Cilj je zato pri dani obliki resnice razviti logiko, ki bo ta koncept sposobna izražati in formalizirati. To storimo tako, da določimo lastnosti, ki naj jih ta logika ima, in primere sklepanja, ki naj jih bo sposobna izražati.

Modalna logika je torej razširitev klasične logike z novimi, modalnimi operatorji. Ključna oseba pri snovanju modalne logike je bil C.I. Lewis, idejo uporabljati modalno logiko za sklepanje o znanju pa je uvedel Jaakko Hintikka [Hintikka-1975].

Najbolj znana modalna operatorja sta \Box in \Diamond , ki ju v *osnovni modalni logiki* beremo *škafra* in *diamant*. Ko modalno logiko uporabimo za izražanje različnih oblik resnic, omenjena modalna operatorja imenujemo dani resnici primerno, npr. *nujno* in *mogoče*, ali v primeru izražanja znanja agenta, *agent ve* in *agent ne izključuje veljavnosti*.

Na eni strani imamo dano logiko, na drugi pa primere oz. *modele* te logike. V modelih veljajo vsa pravila sklepanja, ki definirajo logiko, kar imenujemo *pravilnost*. Če pa je vsaka enačba, ki ji zadoščajo vsi modeli dane logike, v logiki dokazljiva oz. sledi kot logična posledica, pravimo, da je logika *polna*. Včasih je dokazljiva že vsaka enačba, ki ji zadošča neka podmnožica modelov. V tem primeru pravimo, da je logika *polna* za

dano podmnožico modelov. Poseben primer polnosti pa predstavlja (če seveda obstaja) tako imenovani *univerzalni model* \mathcal{U} , v katerem je dana enačba veljavna natanko tedaj, ko je v logiki dokazljiva.

Definicija 1 Modeli \mathcal{M} osnovne modalne logike so določeni z:

1. množico W , katere elemente imenujemo svetovi,
2. relacijo R na množici W , ki jo imenujemo relacija dosegljivosti, wRv pa beremo kot svet v je dosegljiv iz sveta w ,
3. funkcijo $V : W \rightarrow \mathcal{P}(\text{izjavne spremenljivke})$, ki jo imenujemo funkcija pripisovanja veljavnosti, pri čemer $p \in V(w)$ pomeni, da ima izjavna spremenljivka p v svetu w resnično vrednost.

V čast Saula Kripkeja, ki je 50-ih in 60-ih letih prejšnjega stoletja močno zaznamoval modalno logiko, takšne modele ponavadi imenujemo Kripkejevi modeli ali Kripkejeve semantike možnih svetov [Perry-www].

Dana izjava je lahko v nekem svetu resnična, v drugem pa neresnična. Resničnost izjave, ki vsebuje samo klasične logične veznike ($\wedge, \vee, \neg, \Rightarrow$ in \Leftrightarrow), je v klasični modalni logiki določena lokalno in neodvisno od ostalih svetov. Ko pa izjava vsebuje katerega izmed modalnih operatorjev (npr. \Box ali \Diamond), je njena resničnost v posameznem svetu odločilno odvisna od ostalih možnih svetov.

Veljavnost izjav v svetovih Kripkejevih modelov $\mathcal{M} = (W, R, V)$ opišimo z relacijo \Vdash_V .

Definicija 2 Naj bo $\mathcal{M} = (W, R, V)$ model osnovne modalne logike, $w \in W$, \top in \perp običajni konstanti resnica in laž, p izjavna spremenljivka, A in B pa izjavi. Relacija \Vdash_V ($w \Vdash_V A$ beremo kot v svetu $w \in W$ velja A) je najmanjša relacija med svetovi in izjavami, ki zadošča naslednjim pogojem:

1. $w \Vdash_V \top$
2. $w \not\Vdash_V \perp$
3. $w \Vdash_V p$ čee $p \in V(w)$
4. $w \Vdash_V \neg A$ čee $w \not\Vdash_V A$
5. $w \Vdash_V A \wedge B$ čee $w \Vdash_V A$ in $w \Vdash_V B$
6. $w \Vdash_V A \vee B$ čee $w \Vdash_V A$ ali $w \Vdash_V B$
7. $w \Vdash_V A \Rightarrow B$ čee iz $w \Vdash_V A$ sledi $w \Vdash_V B$
8. $w \Vdash_V A \Leftrightarrow B$ čee $w \Vdash_V A$ natanko tedaj, ko $w \Vdash_V B$
9. $w \Vdash_V \Box A$ čee $\forall v \in W : (iz\ wRv\ sledi\ v \Vdash_V A)$
10. $w \Vdash_V \Diamond A$ čee $\exists v \in W : (wRv\ in\ v \Vdash_V A)$

Prva dva pogoja pravita, da \top vedno velja, \perp pa nikoli. Iz naslednjega pogoja sledi, da je $V(w)$ množica vseh izjavnih spremenljivk, ki so v svetu w veljavne. Pogoji, ki se nanašajo na nemodalne izjavne veznike ($\neg, \wedge, \vee, \Rightarrow$ in \Leftrightarrow), zrcalijo klasično

razumevanje teh veznikov v danem svetu. Neobičajna sta le pogoja za \Box in \Diamond . Prvi od njiju pravi, da je izjava $\Box A$ v svetu w resnična, če je izjava A resnična v vsakem svetu, ki je z relacijo R dosegljiv iz sveta w . Drugi pa pravi, da je izjava $\Diamond A$ v svetu w resnična, če obstaja kakšen svet, z relacijo R dosegljiv iz sveta w , kjer je izjava A resnična.

Zanimivo je, da kljub veljavnosti \top v vsakem svetu, $\Diamond \top$ ni nujno res. Velja namreč

$$w \Vdash_V \Diamond \top \quad \text{če} \quad w \text{ je v relaciji } R \text{ z vsaj enim svetom.}$$

Tudi $\Box \perp$ ni nujno neveljavna, čeprav \perp ni veljavna v nobenem svetu. Velja namreč

$$w \Vdash_V \Box \perp \quad \text{če} \quad w \text{ ni v relaciji } R \text{ z nobenim svetom.}$$

Naj ekvivalenca $A \equiv B$ velja natanko tedaj, ko v poljubnem svetu poljubnega modela veljata ali obe ali pa nobena izmed izjav A in B . Primeri ekvivalence sta de Morganovi pravili za \Box in \Diamond ,

$$\neg \Box A \equiv \Diamond \neg A,$$

$$\neg \Diamond A \equiv \Box \neg A,$$

ter pravili za konjunkcijo in disjunkcijo teh modalnih operatorjev,

$$\Box(A \wedge B) \equiv \Box A \wedge \Box B,$$

$$\Diamond(A \vee B) \equiv \Diamond A \vee \Diamond B.$$

Omenimo še, da je $\Box \top$ ekvivalentno \top ,

$$\Box \top \equiv \top,$$

ne pa tudi $\Diamond \top$, in da je $\Diamond \perp$ ekvivalentno \perp ,

$$\Diamond \perp \equiv \perp,$$

ne pa tudi $\Box \perp$.

Definicija 3 *Pravimo, da je izjava A osnovne modalne logike veljavna, če je veljavna v vsakem svetu vsakega modela, in izjava A je veljavna v modelu $\mathcal{M} = (W, R, V)$, če je veljavna v vsakem svetu $w \in W$.*

Očitno veljavne so na primer izjave:

$$\neg \Box A \Leftrightarrow \Diamond \neg A$$

$$\neg \Diamond A \Leftrightarrow \Box \neg A$$

$$\Box(A \wedge B) \Leftrightarrow \Box A \wedge \Box B$$

$$\Diamond(A \vee B) \Leftrightarrow \Diamond A \vee \Diamond B$$

$$\Box \top \Leftrightarrow \top$$

$$\diamond \perp \Leftrightarrow \perp$$

V modalni logiki zelo pomembna pa je veljavna izjava

$$\mathbf{K} : \quad \Box(A \Rightarrow B) \wedge \Box A \Rightarrow \Box B,$$

ki je dobila ime po Saulu Kripkeju. Za dokaz veljavnosti izjave \mathbf{K} je treba pokazati, da za poljuben svet w poljubnega modela $\mathcal{M} = (W, R, V)$ modalne logike velja $w \Vdash_V \Box(A \Rightarrow B) \wedge \Box A \Rightarrow \Box B$. Po definiciji 2 je dovolj pokazati, da iz $w \Vdash_V \Box(A \Rightarrow B) \wedge \Box A$ sledi $w \Vdash_V \Box B$. Predpostavimo torej, da velja $w \Vdash_V \Box(A \Rightarrow B) \wedge \Box A$. To je natanko tedaj, ko $w \Vdash_V \Box(A \Rightarrow B)$ in $w \Vdash_V \Box A$ oz. $\forall v \in W : (iz wRv \text{ sledi } v \Vdash_V (A \Rightarrow B) \text{ in } v \Vdash_V A)$. Sledi $v \Vdash_V B$ in zato $w \Vdash_V \Box B$.

Vrnimo se sedaj k formalizaciji različnih oblik resnice. Tabela 6.1 pri dani interpretaciji operatorja \Box podaja ustrezno interpretacijo modalnega operatorja \diamond , ki je v klasični modalni logiki ekvivalenten $\neg\Box\neg$. Še več, v klasični modalni logiki lahko lastnosti \diamond izpeljemo iz lastnosti \Box in obratno:

$$\Box A \equiv \neg\diamond\neg A \qquad \diamond A \equiv \neg\Box\neg A$$

Množice veljavnih izjav, pripadajočih različnim interpretacijam modalnega operatorja \Box , se med sabo razlikujejo, saj formalizirajo različne oblike resnice. Izjave, ki so veljavne v osnovni modalni logiki (npr. \mathbf{K}), so sicer veljavne tudi v vseh drugih modalnih logikah, ki izvirajo iz različnih interpretacij modalnega operatorja \Box . Na nivoju Kripkejevih modelov pa vsaki interpretaciji pripada ustrezna interpretacija relacije dosegljivosti R , ki je v nekaterih primerih po potrebi lahko npr. tudi reflektivna ali tranzitivna.

$\Box A$	$\diamond A$
nujno velja A	mogoče velja A
vedno bo veljalo A	nekoč bo veljalo A
zapovedano je A	dovoljeno je A
agent verjame A	A je v skladu z agentovim mišljenjem
agent ve, da velja A	agent ne izključuje veljavnosti A
A velja po vsaki izvedbi programa P	A velja po neki izvedbi programa P

Tabela 6.1: Interpretacija modalnega operatorja \diamond pri dani interpretaciji \Box

Za primer vzemimo interpretacijo $\Box A$ kot *nujno velja* A . Pogoj

$$w \Vdash_V \Box A \quad \text{če} \quad \forall v \in W : (iz wRv \text{ sledi } v \Vdash_V A)$$

iz definicije 2 pravi, da je A nujno veljavna v svetu w natanko tedaj, ko je A veljavna v vseh svetovih v , s katerimi je w v relaciji R . Kakšna pa je ta relacija? Intuitivno pravimo, da nujno velja A , če velja A v vseh *možnih* svetovih. Svet v , za katerega

velja wRv , lahko zato interpretiramo kot svet, ki je, glede na znanje sveta w , mogoč oz. dosegljiv.

Interpretaciji $\Box A$ kot *agent ve, da velja A*, ustreza interpretacija wRv , ki pravi, da bi v lahko bil dejanski svet glede na znanje agenta v svetu w . Z drugimi besedami, če je dejanski svet ravno w , potem agent, ki ni vseveden, ne more izključiti možnosti, da je ta svet pravzaprav v . Pomen relacije R dosegljivosti med svetovi pri različnih interpretacijah $\Box A$ je prikazan v tabeli 6.2.

$\Box A$	wRv
nujno velja A	v je možen svet, glede na informacije v w
vedno bo veljalo A	v je v prihodnosti glede na svet w
zapovedano je A	v je dopusten glede na informacije v w
agent verjame A	v bi lahko bil dejanski svet glede na mišljenje agenta v w
agent ve, da velja A	v bi lahko bil dejanski svet glede na znanje agenta v w
A velja po vsaki izvedbi programa P	v je možno stanje po izvedbi programa P v w

Tabela 6.2: Pomen relacije R pri različnih interpretacijah izjave $\Box A$

Lastnosti relacije dosegljivosti R (npr. refleksivnost, simetričnost ali tranzitivnost) so povezane z zakoni dane modalne logike:

- Če je R refleksivna, velja

$$\mathbf{T} : \quad \Box A \Rightarrow A.$$

To dokažemo tako, da za poljuben model osnovne modalne logike $\mathcal{M} = (W, R, V)$ pokažemo $\mathcal{M} \Vdash_V \Box A \Rightarrow A$, kar pomeni, da za vsak $w \in W$ velja $w \Vdash_V \Box A \Rightarrow A$. Naj bo torej w poljuben in naj velja $w \Vdash_V \Box A$. Ker je wRw , iz definicije 2 takoj sledi $w \Vdash_V A$. Torej res velja $w \Vdash \Box A \Rightarrow A$.

- Če je R simetrična, velja

$$\mathbf{B} : \quad A \Rightarrow \Box \Diamond A.$$

Potrebno je pokazati, da za poljuben svet w poljubnega modela $\mathcal{M} = (W, R, V)$ velja $w \Vdash_V A \Rightarrow \Box \Diamond A$. Naj velja $w \Vdash_V A$, in naj bo $v \in W$ takšen, da je wRv . Zaradi simetričnosti je vRw . Za vsak v z lastnostjo wRv torej obstaja svet (to je kar w), s katerim je v v relaciji R in v katerem velja A . Sledi $w \Vdash_V \Box \Diamond A$.

- Če je R tranzitivna, velja

$$\mathbf{4} : \quad \Box A \Rightarrow \Box \Box A.$$

Potrebno je pokazati, da v poljubnem modelu $\mathcal{M} = (W, R, V)$ velja $\mathcal{M} \Vdash_V \Box A \Rightarrow \Box \Box A$. Naj bo $w \in W$. Pokažimo torej, da velja $w \Vdash_V \Box A \Rightarrow \Box \Box A$. Iz $w \Vdash_V \Box A$ po definiciji 2 za vsak v z lastnostjo wRv in vsak u z lastnostjo vRu sledi $u \Vdash A$. Torej res $w \Vdash_V \Box \Box A$.

Zapišimo sedaj formalno definicijo modalne logike z modalnima operatorjema \Box in \Diamond .

Definicija 4 Modalna logika L je podmnožica izjav, sestavljenih iz konstant \perp in \top , iz izjavnih spremenljivk, klasičnih veznikov $\wedge, \vee, \neg, \Rightarrow$ in \Leftrightarrow ter modalnih operatorjev \Box in \Diamond , z naslednjimi lastnostmi:

1. L vsebuje vse izjave, ki jih izpelje klasična izjavna logika.
2. Velja zakon **K**: $\Box(A \Rightarrow B) \wedge \Box A \Rightarrow \Box B$.
3. Če je $A \in L$, potem je tudi $\Box A \in L$.
4. L je zaprta za zamenjave delov izjav.

Primer modalne logike je logika $KT4$, v literaturi pogosto imenovana tudi $S4$. Modeli te modalne logike so natanko Kripkejevi modeli $\mathcal{M} = (W, R, V)$, kjer je relacija R refleksivna (zakon **T**) in tranzitivna (zakon **4**). Za to logiko se izkaže, da velja naslednji izrek.

Izrek 1 V modalni logiki $KT4$ je poljuben niz modalnih operatorjev in negacij ekvivalenten enemu izmed naslednjih nizov: $\neg, \Box, \Diamond, \Box\Diamond, \Diamond\Box, \Box\Diamond\Box, \Diamond\Box\Diamond, \neg, \neg\Box, \neg\Diamond, \neg\Box\Diamond, \neg\Diamond\Box, \neg\Box\Diamond\Box$ in $\neg\Diamond\Box\Diamond$.

Dokaz. Z indukcijo na dolžino niza.

6.2 Intuicionistična modalna logika

V klasični logiki (tudi modalni) je po Aristotelovem zakonu o izključeni tretji možnosti poljubna izjava A ali resnična ali neresnična. V vsakem modelu velja $A \vee \neg A$. Hitro pa pridemo do zaključka, da vsebina tega zakona omejuje naše razmišljanje. Kot primer vzemimo izjavo

V decimalnem zapisu števila π obstaja zaporedje sedmih sedmic.

Trenutno ne poznamo nobene metode, s pomočjo katere bi se lahko prepričali v resničnost ali neresničnost te izjave. Lahko se celo zgodi, da ne bo nikoli nihče ugotovil, ali je ta izjava resnična, čeprav nas klasična logika prisili, da sprejmemo nujnost ene izmed možnosti – resničnost ali neresničnost.

Klasično logiko brez Aristotelovega zakona o izključeni tretji možnosti $A \vee \neg A$, a z zakonom o nasprotovanju $\neg A \wedge A \Rightarrow B$, imenujemo *intuicionistična logika*, zaobsega pa principe logičnega sklepanja, ki jih je leta 1907 začel uporabljati L.E.J. Brouwer, ko je razvijal svojo intuicionistično matematiko [Troelstra-www]. Brouwer je opazil, da je bil Aristotelov zakon o izključeni tretji možnosti oblikovan na podlagi končnih situacij, nato pa neupravičeno razširjen na neskončne. Principi, ki jih je začel uporabljati Brouwer, so podvrženi tudi ruski rekurzivni analizi in konstruktivni analizi E.

Bishopa in njegovih naslednikov, zato lahko na intuicionistično logiko gledamo kot na temelje konstruktivne matematike.

V intuicionistični logiki lahko z gotovostjo trdimo, da objekt z določeno lastnostjo obstaja šele, ko poznamo učinkovito metodo konstrukcije oz. iskanja takega objekta. Znamenita izjava še bolj znamenitega intuicionista Arenda Heytinga pa pravi, da razložiti formulo ne pomeni podati pogoje resničnosti, ampak pogoje dokazljivosti. To se odraža v intuicionistični razlagi izjavnih veznikov \neg in \vee , ki se od razlage klasične logike razlikuje. Izjavi A in $\neg\neg A$ zaradi izključitve omenjenega Aristotelovega zakona nista ekvivalentni, interpretacija izjave $A \vee B$ pa je veliko močnejša.

Vzemimo npr. agenta, ki se znajde na razpotju labirinta in ve, da vodi do izhoda iz labirinta samo ena pot. Naj izjava A pomeni, da vodi do cilja pot 1, izjava B pa, da vodi do cilja pot 2. V klasični logiki izjava $A \vee B$ pove samo, da do izhoda iz labirinta vodi ena izmed obeh poti, kar agentu pri iskanju izhoda ne pomaga. Drugače pa je v intuicionistični logiki, kjer izjava $A \vee B$ pove tudi, katera izmed obeh poti vodi do cilja.

Definicija 5 Intuicionistični izjavni račun (*IPC*) je podteorija intuicionistične logike, ki jo dobimo, ko jezik intuicionistične logike zožimo le na izjave, sestavljene iz izjavnih spremenljivk p_0, p_1, \dots , konstant \top in \perp ter izjavnih veznikov \wedge, \vee in \Rightarrow . V *IPC* izjavo $\neg A$ definiramo kot $A \Rightarrow \perp$, izjavo $A \Leftrightarrow B$ pa kot $(A \Rightarrow B) \wedge (B \Rightarrow A)$. Osnovna sodba intuicionističnega izjavnega računa je logična posledica. To je relacija \vdash med končno množico hipotez $\{A_1, \dots, A_n\}$ in sklepom B :

$$\{A_1, \dots, A_n\} \vdash B.$$

Običajno pišemo $A_1, \dots, A_n \vdash B$, namesto $\{\} \vdash B$ pa kar $\cdot \vdash B$. Relacija \vdash je podana induktivno z naslednjimi pravili sklepanja, pri čemer so A, B in C poljubne izjave, Γ pa poljubna končna množica izjav:

A1. Sklep iz hipoteze

$$\frac{}{\Gamma \vdash A} \text{ če } A \in \Gamma$$

A2. Resnica

$$\frac{}{\Gamma \vdash \top}$$

A3. Laž

$$\frac{\Gamma \vdash \perp}{\Gamma \vdash A}$$

A4. Konjunkcija

$$\frac{\Gamma \vdash A \quad \Gamma \vdash B}{\Gamma \vdash A \wedge B} \quad \frac{\Gamma \vdash A \wedge B}{\Gamma \vdash A} \quad \frac{\Gamma \vdash A \wedge B}{\Gamma \vdash B}$$

A5. Disjunkcija

$$\frac{\Gamma \vdash A}{\Gamma \vdash A \vee B} \quad \frac{\Gamma \vdash B}{\Gamma \vdash A \vee B} \quad \frac{\Gamma \vdash A \vee B \quad \Gamma, A \vdash C \quad \Gamma, B \vdash C}{\Gamma \vdash C}$$

A6. Implikacija

$$\frac{\Gamma, A \vdash B}{\Gamma \vdash A \Rightarrow B} \quad \frac{\Gamma \vdash A \Rightarrow B \quad \Gamma \vdash A}{\Gamma \vdash B}$$

Pravila iz definicije 5 beremo kot veljavne sklepe. Pravilo A1 pravi, da so hipoteze vedno veljavne izjave, pravilo A2 pa, da vedno velja tudi resnica \top . Pravilo A3 pravi, da poljubna izjava sledi kot logična posledica, če le velja laž \perp . Prvo pravilo za konjunkcijo A4 pa npr. pomeni, da iz $\Gamma \vdash A$ in $\Gamma \vdash B$ sledi $\Gamma \vdash A \wedge B$.

Omenimo še, da bi klasični izjavni račun dobili, če bi v definiciji 5 dodali naslednje pravilo sklepanja:

A7. Izključena tretja možnost

$$\frac{\Gamma, A \vdash B \quad \Gamma, A \vdash \neg B}{\Gamma \vdash \neg A}$$

Ljudje svoje znanje gradimo s pomočjo delčkov informacij, ki jih pridobivamo iz različnih zunanjih virov. Znanje zato ni objektivno, proces pridobivanja znanja pa ni končen. Ker svoje znanje gradimo na osnovi več različnih virov, si podatki, ki jih pridobivamo, lahko tudi nasprotujejo (iz enega vira lahko dobimo A , iz drugega pa $\neg A$). Nasprotujoča si podatka A in $\neg A$ lahko v intuicionistični modalni logiki predstavimo kot $\diamond A$ in $\diamond \neg A$, kar pomeni, da dopuščamo obe možnosti.

Vzemimo sedaj model $\mathcal{M} = (W, R, V)$ modalne logike KT4, kjer je relacija R refleksivna in tranzitivna. Modele intuicionistične modalne logike, kjer \square in \diamond beremo kot *nujno* in *mogoče*, dobimo tako, da v $\mathcal{M} = (W, R, V)$ dodamo relacijo \leq delne urejenosti (refleksivna, antisimetrična in tranzitivna) svetov, ki je s pogojem 0^* povezana z obstoječo relacijo R , za funkcijo pripisovanja veljavnosti V pa zahtevamo lastnost *monotonosti*, t.j. iz $w \leq v$ mora slediti $V(w) \subseteq V(v)$.

Definicija 6 Intuicionistični Kripkejevi modeli modalne logike $KT4$ so modeli $\mathcal{M} = (W, \leq, R, V)$ modalne logike $KT4$ z dodatno relacijo \leq delne urejenosti množice svetov W ($w, w', v, v' \in W$),

$$0^*. \quad wRv \text{ in } v \leq v' \quad \text{sledi} \quad \exists w' \in W: (w'Rv' \text{ in } w \leq w'),$$

in s pogojem o monotonosti funkcije V . V teh modelih definiramo relacijo \Vdash_V kot najmanjšo relacijo med svetovi in izjavami, katere pogoji se od pogojev iz definicije 2 razlikujejo le po interpretaciji izjavnih veznikov $\neg, \Rightarrow, \square$ in \diamond . Za $w, v, u \in W$, izjavno spremenljivko p ter izjavi A in B mora veljati:

- 1*. $w \Vdash_V \top$
- 2*. $w \not\Vdash_V \perp$
- 3*. $w \Vdash_V p$ čee $p \in V(w)$
- 4*. $w \Vdash_V \neg A$ čee $\forall v \in W: (iz w \leq v \text{ sledi } v \not\Vdash_V A)$

5*.	$w \Vdash_V A \wedge B$	čee	$w \Vdash_V A$ in $w \Vdash_V B$
6*.	$w \Vdash_V A \vee B$	čee	$w \Vdash_V A$ ali $w \Vdash_V B$
7*.	$w \Vdash_V A \Rightarrow B$	čee	$\forall v \in W : (w \leq v \Rightarrow (iz v \Vdash_V A \text{ sledi } v \Vdash_V B))$
8*.	$w \Vdash_V \Box A$	čee	$\forall v \in W : (w \leq v \Rightarrow \forall u \in W : (iz vRu \text{ sledi } u \Vdash_V A))$
9*.	$w \Vdash_V \Diamond A$	čee	$\forall v \in W : (iz w \leq v \text{ sledi } \exists u \in W : (vRu \text{ in } u \Vdash_V A))$

Izjavne spremenljivke intuicionistične modalne logike lahko pojmujeemo kot *monotone resnice* – ko je v nekem svetu izjavna spremenljivka resnična, je resnična tudi v vseh, iz tega sveta z relacijo \leq dosegljivih svetovih. Naslednja lema pa pravi, da so tudi ostale izjave monotone resnice.

Lema 1 (LEMA O MONOTONOSTI) *V intuicionističnem Kripkejevem modelu $\mathcal{M} = (W, \leq, R, V)$ modalne logike $KT4$ velja*

$$iz w \Vdash_V A \text{ sledi } \forall v \in W : (w \leq v \text{ sledi } v \Vdash_V A).$$

Dokaz. Z indukcijo po strukturi izjave.

Pogoj θ^* v definiciji 6 podaja zvezo med obema relacijama na množici svetov W , ki temelji na naslednjem razmisleku. Dejstvo, da je v dosegljiv iz sveta w (wRv), je razumno obravnavati kot neke vrste osnovno resnico. Če je sedaj svet $v' \geq v$, bi moral biti tudi v' dosegljiv iz w , toda medtem smo lahko zbrali že več podatkov o svetu w in je ta lahko prešel v novi svet $w' \geq w$.

Če definicije izjavnih veznikov \Rightarrow , \Box in \Diamond ne bi ustrezno spremenili, ne bi imeli nobenega zagotovila, da bodo izjave oblike $A \Rightarrow B$, $\Box A$ in $\Diamond A$ res zadoščale lemi o monotonosti. Sprememba definicije veznika \neg pa je posledica intuicionizma.

Intuicionistični Kripkejevi modeli modalne logike $KT4$ pripadajo natanko *modalnemu intuicionističnemu izjavnemu računu* $CS4$ [Simpson-1994, Alechina-2001], ki ga dobimo, če IPC razširimo z modalnima operatorjema \Box in \Diamond .

Definicija 7 *Modalni intuicionistični izjavni račun $CS4$ je razširitev jezika IPC (definicija 5) z enomestnima modalnima operatorjema \Box in \Diamond . Logična posledica je najmanjša relacija \vdash_{CS4} , ki poleg pravil $A1$ – $A6$ za IPC zadošča še naslednjim pravilom:*

B1. Nujno

$$\frac{\Gamma \vdash_{CS4} \Box A}{\Gamma \vdash_{CS4} A} \quad \frac{\Gamma \vdash_{CS4} \Box A}{\Gamma \vdash_{CS4} \Box \Box A}$$

B2. Implikacija in nujno

$$\frac{\Gamma \vdash_{CS4} \Box(A \Rightarrow B) \quad \Gamma \vdash_{CS4} \Box A}{\Gamma \vdash_{CS4} \Box B}$$

B3. Logične resnice so znane

$$\frac{\cdot \vdash_{CS4} A}{\cdot \vdash_{CS4} \Box A}$$

B4. Mogoče

$$\frac{\Gamma \vdash_{CS4} A}{\Gamma \vdash_{CS4} \Diamond A} \quad \frac{\Gamma \vdash_{CS4} \Diamond \Diamond A}{\Gamma \vdash_{CS4} \Diamond A}$$

B5. Implikacija, nujno in mogoče

$$\frac{\Gamma \vdash_{CS4} \Box(A \Rightarrow B) \quad \Gamma \vdash_{CS4} \Diamond A}{\Gamma \vdash_{CS4} \Diamond B}$$

Aksiom B1 pravi, da znanje ni zmotno in da se svojega znanja zavedamo. Naslednji aksiom, t.j. aksiom B2, podaja zvezo med implikacijo in \Box ; če nujno velja $A \Rightarrow B$ in A , potem nujno velja B . Aksiom B3 pravi, da nujno velja vse, kar je logično dokazljivo brez hipotez. Aksiom B4 med drugim pravi, da so resnice mogoče. Zadnji aksiom pa podaja zvezo med implikacijo, \Box in \Diamond .

Izrek 2 *Modalni intuicionistični izjavni račun CS4 (definicija 7) je za intuicionistične Kripkejeve modele (definicija 6) poln, t.j.*

$$\Gamma \vdash_{CS4} A \quad \text{če} \quad \Gamma \Vdash_V A.$$

Dokaz. Desno implikacijo (pravilnost semantike) dokažemo z indukcijo na strukturo dokaza formule A iz konteksta Γ . Pokažemo, da so v intuicionističnih Kripkejevih modelih modalne logike KT4 vsi aksiomi veljavni, pravila sklepanja pa ohranjajo veljavnost formul. Z intuicionističnim delom ni težav, zato dokažimo samo veljavnost aksiomov B1–B5. Veljavnost prvega dela aksioma B1 je posledica refleksivnosti relacij R in \leq , veljavnost drugega dela pa tranzitivnosti relacije R , refleksivnosti relacije \leq ter zveze med obema relacijama Kripkejevih modelov. Zaradi refleksivnosti \leq sta veljavna aksioma B2 in B5. Aksiom B3 sledi iz dejstva, da veljavnost formule v vseh modelih pomeni veljavnost formule v vseh svetovih v vseh modelih. Veljavnost prvega dela aksioma B4 je posledica refleksivnosti R , veljavnost drugega dela pa je posledica refleksivnosti \leq in tranzitivnosti R . Dokaz leve implikacije (polnost semantike) je podan v [Alechina-2001]. Konec dokaza.

6.3 Predstavitev vremenskih podatkov

Čeprav so se razvijalci sistemov za dialog pri predstavitvi znanja osredotočili predvsem na tradicionalne logične podatkovne modele kot so hierarhični, mrežni in relacijski, smo se pri razvoju sistema za podajanje informacij o vremenu vprašali, kakšna predstavitev znanja bi bila s stališča sposobnosti usmerjanja k izbiri relevantnih, dosegljivih podatkov najustreznejša. To vprašanje postane smiselno, ko imamo na eni strani opravka s podatkovno zbirko, katere struktura je časovno odvisna in nepredvidljiva (poglavje 2), na drugi strani pa s številnimi pomanjkljivostmi tradicionalnih podatkovnih modelov [Elmasri-2004].

S takšno podatkovno zbirko so imeli opravka tudi razvijalci sistema JUPITER [Zue-2000], ki so vremenske podatke sicer shranili v obliki tradicionalnega relacijskega podatkovnega modela, vendar so dodali še konstantno tabelo geografskih podatkov. Ta dodatna tabela res omogoča odgovore na vprašanja oblike

Katere kraje v Karibih poznaš?

ne pa tudi na vprašanja, ki se nanašajo na časovno odvisno strukturo podatkov, kot na primer

Za katere kraje v Karibih poznaš trenutno temperaturo?

Sistem za dialog je vsekakor veliko zmogljivejši pri usmerjanju uporabnika k izbiri relevantnih, dosegljivih podatkov, če relacije med posameznimi podatki niso fiksno določene. Veliko primerneje je te relacije po vsakem osveževanju podatkovne zbirke (v skladu s trenutno dosegljivimi podatki) ponovno vzpostaviti. Časovno dinamično predstavitev, ki bo zajemala tudi relacije med posameznimi podatki, dosežemo z uporabo teorije intuicionistične modalne logike.

Modalna logika je bila sicer že uporabljena pri vzpostavitvi logičnega podatkovnega modela v primeru razpolaganja z več dokumenti [Nie-1989, Sebastiani-1998] ter za izražanje prepričanj in ciljev v komunikaciji med več razumskimi agenti [Sadek-1994, Smith-1998], vendar imajo ti logični modeli z našo predstavitvijo znanja zelo malo skupnega.

Prednosti predstavljenega podatkovnega modela so predvsem:

- ↪ preprosta pretvorba podatkov,
- ↪ jasne in logične relacije med posameznimi podatki,
- ↪ preprosto ravnanje z nepopolnimi podatki,
- ↪ preprosto povpraševanje oz. iskanje relevantnih, dosegljivih podatkov.

Imejmo torej modul za vodenje dialoga, ki ima opravka z n različnimi predalčki. V primeru našega sistema za podajanje informacij o vremenu je $n = 3$, tipi predalčkov pa so *kraj*, *čas* in *podatek*. Za množico izjavnih spremenljivk vzemimo unijo množic P_1, \dots, P_n , kjer je P_i množica vseh možnih elementov i -tega predalčka.

Zbirko vremenskih podatkov pretvorimo v množico svetov $W = W1 \cup W2$. Prvo podmnožico svetov $W1$ dobimo tako, da vsaki izjavni spremenljivki p_{ij} priredimo svet $W1_{ij} \in W1$, v katerem je p_{ij} , ki jo imenujemo *vodilna atomarna izjava*, veljavna. Drugo podmnožico svetov $W2$ pa dobimo tako, da v svetove iz $W1$ dodamo veljavnost vsaj ene dodatne izjavne spremenljivke, t.j. *omejitve*, pri čemer se noben par *atomarnih izjav* (vodilna atomarna izjava in omejitve) ne nanaša na isti predalček. Te atomarne izjave ne obravnavamo tako kot ostale izjave, saj kljub temu, da natanko določajo zakone danega sveta, ne vsebujejo informacije o dosegljivosti oz. relevantnosti posameznih delčkov informacij. Atomarne izjave se navezujejo na sestavne dele uporabnikovega vprašanja, ostale izjave, ki so sestavljene iz

- izjavnih spremenljivk p_{ij} ,
- konstant \perp in \top ,
- klasičnih izjavnih veznikov \wedge, \vee, \neg in \Rightarrow ter
- modalnih operatorjev *spodnji diamant* ∇ in *zgornji diamant* Δ ,

pa dosegljivost in relevantnost vremenskih informacij opisujejo tako, da je sistem sposoben uporabnika usmerjati k izbiri informacij, ki se glede na vprašanje zdijo smiselne. Vpeljane modalne operatorje interpretiramo takole:

- ∇A pomeni, da je A dosegljivo in relevantno, vendar manj specifično;
- ΔA pomeni, da je A dosegljivo in relevantno, vendar bolj specifično.

Na množici svetov W definiramo reflektivno in tranzitivno relacijo *dosegljivosti* R , kjer

- wRv pomeni, da je svet v bolj specifičen oz. *dosegljiv* iz sveta w , t.j. množici atomarnih izjav svetov w in v se razlikujeta le v vodilni atomarni izjavi, pri čemer je vodilna izjava sveta v bolj specifična od vodilne atomarne izjave sveta w ,

in relacijo delne urejenosti (reflektivna, antisimetrična, tranzitivna) \leq , kjer

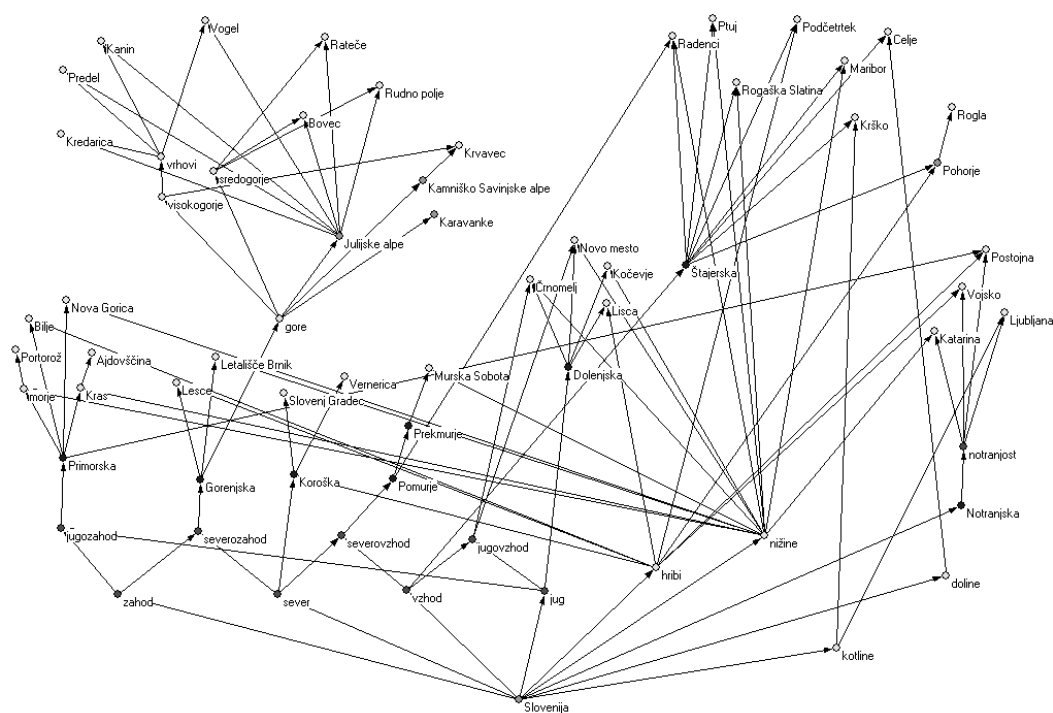
- $w \leq v$ pomeni, da je množica atomarnih izjav sveta w podmnožica množice atomarnih izjav sveta v .

Relacija R temelji na vnaprej definirani reflektivni in tranzitivni relaciji *biti manj specifičen*, ki povezuje manj specifične oz. bolj splošne delčke informacij, navezujoče se na isti predalček, z bolj specifičnimi. Usmerjene Hassejeve diagrame teh relacij, kjer vozlišča ustrezajo posameznim delčkom informacij, usmerjene povezave pa vzpostavljenim relacijam, prikazujejo slike 6.1, 6.2, 6.3 in 6.4. Relacija R tedaj povezuje svetova, katerih množici atomarnih izjav se razlikujeta le v vodilni atomarni izjavi, ti pa sta v relaciji *biti manj specifičen*.

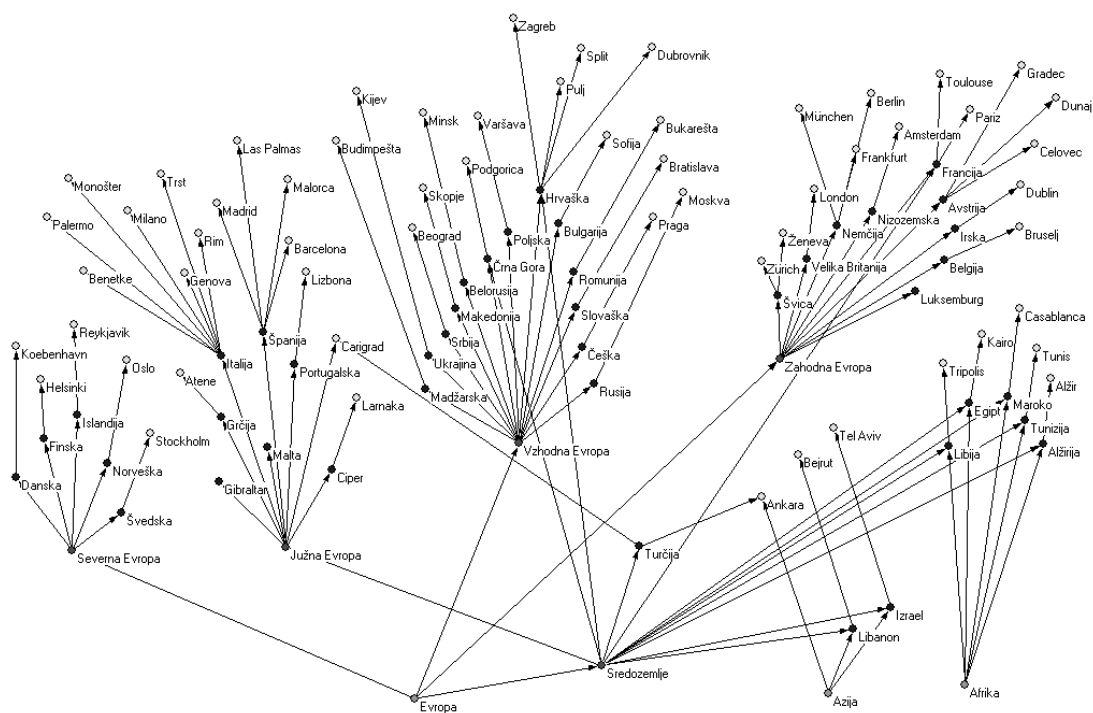
Sledi definicija razreda Kripkejevih modelov, ki mu takšna predstavitev znanja zadošča.

Definicija 8 Dialog mIPC Kripkejevi modeli so strukture $\mathcal{M} = (W, \leq, R, \Vdash)$, kjer je (W, \leq) neprazna, delno urejena množica svetov, R je reflektivna in tranzitivna relacija na W , ki za $w, w', v, v' \in W$ zadošča

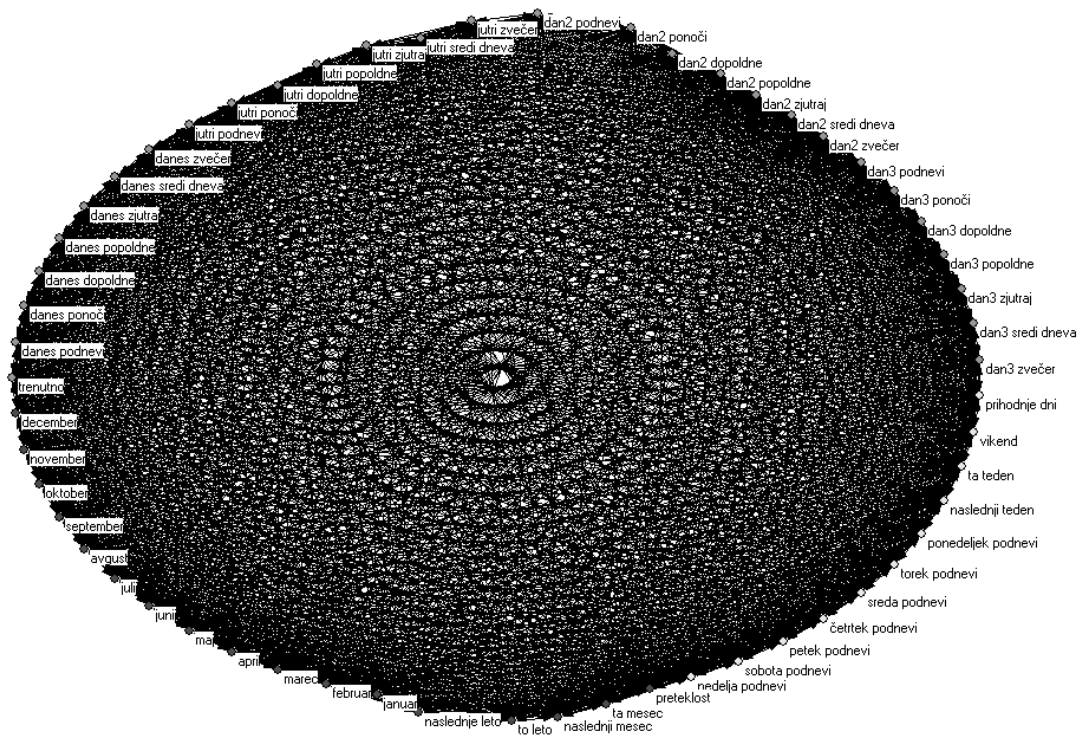
- 0^{**} .
- iz wRv in $v \leq v'$ sledi $\exists w' \in W : (w'Rv' \text{ in } w \leq w')$,
 - iz $w \leq v$ in vRv' sledi $\exists w' \in W : (wRw' \text{ in } w' \leq v')$,
 - iz $w \leq v$ in wRv' sledi $\exists w' \in W : (vRw' \text{ in } v' \leq w')$,



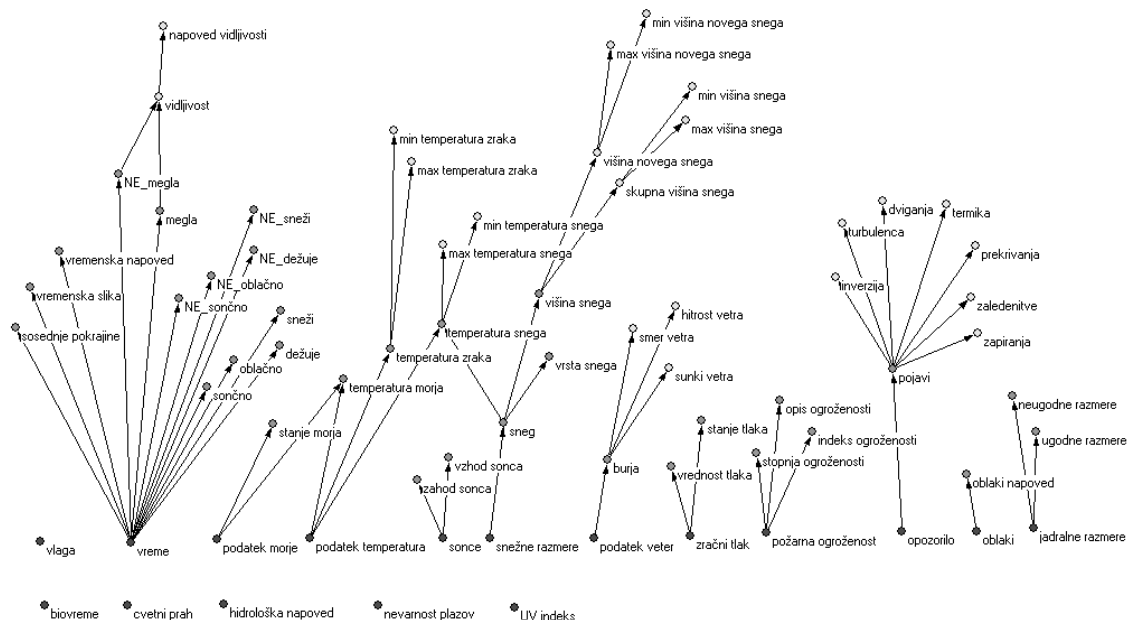
Slika 6.1: Hassejev diagram relacije *biti manj specifičen* med krajevnimi delčki informacij, ki se navezujejo na Slovenijo



Slika 6.2: Hassejev diagram relacije *biti manj specifičen* med krajevnimi delčki informacij, ki se navezujejo na ostalo Evropo



Slika 6.3: Hassejev diagram relacije *biti manj specifičen* med časovnimi delčki informacij. Zaradi majhnega števila možnih dosegljivih časovnih trenutkov, so vzpostavljene vse povezave, zaradi česar relacija *biti manj specifičen* v tem primeru izgubi pomen, ki izvira iz imena.



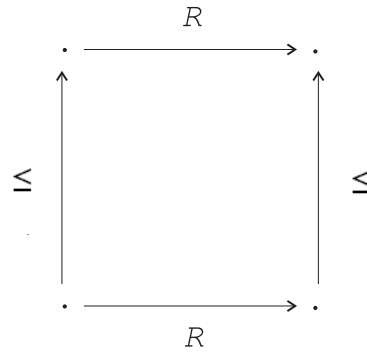
Slika 6.4: Hassejev diagram relacije *biti manj specifičen* med podatkovnimi delčki informacij

iz $v \leq w$ in $v'Rw$ sledi $\exists w' \in W : (w'Rv \text{ in } w' \leq v')$,

\Vdash pa je dvomestna relacija med svetovi in izjavami ($w \Vdash A$ beremo kot izjava A je v svetu w veljavna), da za poljubni izjavi A in B , izjavno spremenljivko p , ter svetova w in v velja:

1**.	$w \Vdash \top$		
2**.	$w \not\Vdash \perp$		
3**.	$w \Vdash p$ in $w \leq v$	sledi	$v \Vdash p$ (MONOTONOST IZJAVNIH SPREMENLJIVK)
4**.	$w \Vdash \neg A$	čee	$\forall v \in W : (iz w \leq v \text{ sledi } v \not\Vdash A)$
5**.	$w \Vdash A \wedge B$	čee	$w \Vdash A$ in $w \Vdash B$
6**.	$w \Vdash A \vee B$	čee	$w \Vdash A$ ali $w \Vdash B$
7**.	$w \Vdash A \Rightarrow B$	čee	$\forall v \in W : (w \leq v \Rightarrow (iz v \Vdash A \text{ sledi } v \Vdash B))$
8**.	$w \Vdash \nabla A$	čee	$\exists v \in W : (vRw \text{ in } v \Vdash A)$
9**.	$w \Vdash \Delta A$	čee	$\exists v \in W : (wRv \text{ in } v \Vdash A)$

Pogoj 0** lahko krajše izrazimo takole – za poljubne tri svetove v kvadratu relacij R in \leq



obstaja še četrti. Omenimo še, da se pravila 1**–7** v definiciji dialog mIPC Kripkejevih modelov ujemajo s pravili iz definicije intuicionističnih Kripkejevih modelov modalne logike KT4 (definicija 6). Tudi v razredu mIPC Kripkejevih modelov pa velja lema o monotonosti.

Lema 2 (LEMA O MONOTONOSTI) *V dialog mIPC Kripkejevih modelih intuicionistične modalne logike $\mathcal{M} = (W, \leq, R, \Vdash)$*

$$iz w \Vdash A \text{ sledi } \forall w' \in W : (w \leq w' \text{ sledi } w' \Vdash A).$$

Dokaz. Z indukcijo na strukturo izjave A . Naj bosta B in C poljubni izjavi, za $w, w' \in W$ pa naj velja $w \Vdash A$ in $w \leq w'$.

- A je izjavna spremenljivka. Tedaj $w' \Vdash A$ sledi iz pravila 3** v definiciji 8 dialog mIPC Kripkejevih modelov.

- A je $\neg B$. Tedaj $w \Vdash \neg B$ pomeni, da za vsak svet $v \geq w$ sledi $v \not\Vdash B$. Zaradi tranzitivnosti relacije \leq za $v' \geq w'$ velja $v' \geq w$ in zato $v' \not\Vdash B$. Sledi $w' \Vdash \neg B$.
- A je $B \wedge C$. Tedaj $w \Vdash B \wedge C$ pomeni, da $w \Vdash B$ in $w \Vdash C$. Po indukcijski predpostavki velja $w' \Vdash B$ in $w' \Vdash C$, torej $w' \Vdash B \wedge C$.
- A je $B \vee C$. Tedaj $w \Vdash B \vee C$ pomeni, da velja $w \Vdash B$ ali $w \Vdash C$. Če velja $w \Vdash B$, po indukcijski predpostavki sledi $w' \Vdash B$ in zato tudi $w' \Vdash B \vee C$. Za $w \Vdash C$ podobno.
- A je $B \Rightarrow C$. Tedaj $w \Vdash B \Rightarrow C$ pomeni, da za vse $v \geq w$, iz $v \Vdash B$ sledi $v \Vdash C$. Naj bo $v' \geq w'$. Zaradi tranzitivnosti \leq sledi $w \leq v'$, zaradi $w \Vdash B \Rightarrow C$ pa iz $v' \Vdash B$ sledi $v' \Vdash C$. Torej res $w' \Vdash B \Rightarrow C$.
- A je ∇B . Tedaj $w \Vdash \nabla B$ pomeni, da obstaja $v \in W$, da je vRw in $v \Vdash B$. Pravilo 0^{**} v definiciji 8 zaradi vRw in $w \leq w'$ pravi, da obstaja svet v' , da velja $v \leq v'$ in $v'Rw'$. Po indukcijski predpostavki sledi $v' \Vdash B$, torej $w' \Vdash \nabla B$.
- A je ΔB . Tedaj $w \Vdash \Delta B$ pomeni, da obstaja $v \in W$, da je wRv in $v \Vdash B$. Pravilo 0^{**} v definiciji 8 zaradi wRv in $w \leq w'$ pravi, da obstaja svet v' , da velja $v \leq v'$ in $w'Rv'$. Po indukcijski predpostavki sledi $v' \Vdash B$, torej $w' \Vdash \Delta B$.

Konec dokaza.

Zaradi leme o monotonosti lahko o svetovih Kripkejevih modelov razmišljamo kot o možnih stopnjah znanja; svet, ki je v delno urejeni množici (W, \leq) uvrščen višje, predstavlja razširitev našega znanja. Ko se pomikamo po delni urejenosti navzgor, se naše znanje kopiči – ničesar ne pozabimo, vemo kvečjemu več.

Predpostavljajmo sedaj, da smo v svetu v . Dejstvo, da je ta svet dosegljiv iz sveta w lahko spet obravnavamo kot neke vrste osnovno resnico. Zaradi tega bi moral iz w biti dosegljiv tudi vsak svet $v' \geq v$, vendar je smiselno pričakovati, da smo medtem o svetu w zbrali že več podatkov, zaradi česar je w prešel v svet $w' \geq w$. Takšen razmislek nas pripelje do prvega dela pravila 0^{**} iz definicije 8. Ostale dele tega pravila pojasnimo na podoben način.

Primer 1 *Oglejmo si primere svetov v predstavitvi vremenskih podatkov. Predpostavimo, da vir znanja sistema vsebuje podatke o trenutni temperaturi v Ljubljani, o temperaturi v Ljubljani danes zvečer ter o pričakovani jutrišnji dnevni temperaturi v Sloveniji. Naj bo Ljubljana svet z vodilno atomarno izjavo Ljubljana, Slovenija pa svet z vodilno atomarno izjavo Slovenija, oba brez dodatnih omejitev. Ker je izjavna spremenljivka Slovenija v relaciji biti manj specifičen z izjavno spremenljivko Ljubljana (slika 6.1), svet Ljubljana vsebuje naslednje izjave:*

Ljubljana

temperatura \Rightarrow trenutno (15)

trenutno \Rightarrow temperatura (15)

temperatura \Rightarrow danes zvečer (10)

danes zvečer \Rightarrow temperatura (10)

$\nabla^{\text{Slovenija}}$ (temperatura \Rightarrow jutri podnevi) (od 13 do 17)

$\nabla^{\text{Slovenija}}$ (jutri podnevi \Rightarrow temperatura) (od 13 do 17)

\vdots

Svet Slovenija pa vsebuje izjave:

Slovenija

temperatura \Rightarrow jutri podnevi (od 13 do 17)

jutri podnevi \Rightarrow temperatura (od 13 do 17)

$\Delta^{\text{Ljubljana}}$ (temperatura \Rightarrow trenutno) (15)

$\Delta^{\text{Ljubljana}}$ (trenutno \Rightarrow temperatura) (15)

$\Delta^{\text{Ljubljana}}$ (temperatura \Rightarrow danes zvečer) (10)

$\Delta^{\text{Ljubljana}}$ (danes zvečer \Rightarrow temperatura) (10)

\vdots

Pri tem smo modalnima operatorjema ∇ in Δ dodali indeks atomarnih izjav ($\nabla^{\text{Slovenija}}$ oz. $\Delta^{\text{Ljubljana}}$), ki pove, na kateri relevanten svet se izjava nanaša, dejansko informacijo (npr. opis temperature) pa smo zapisali v oklepajih na koncu izjave.

Naslednji izrek pravi, da dialog mIPC Kripkejevi modeli (definicija 8) zadoščajo vsem pravilom sklepanja IPC (definicija 5) ter navaja dodatna pravila sklepanja, ki zajemajo spodnji ali zgornji diamant. Tako dobljene sklepe, sestavljene le iz ene izjavne spremenljivke in enomestnih izjavnih veznikov, imenujemo *končne izjave* in jih iz nadaljnjih sklepov izločimo.

Izrek 3 *Dialog mIPC Kripkejevi modeli zadoščajo vsem pravilom sklepanja intuicionističnega izjavnega računa IPC, t.j. za poljubno množico izjav Γ in nekončno izjavo A*

iz $\Gamma \vdash A$ sledi $\Gamma \Vdash A$.

Za svetova w in v , nekončni izjavi A in B , ter kontekst Γ v dialog mIPC Kripkejevih modelih veljajo tudi naslednja pravila sklepanja:

B1. Spodnji diamant

$$\frac{\Gamma \vdash A}{\Gamma \vdash \nabla A} \quad \frac{\Gamma \vdash \nabla \nabla A}{\Gamma \vdash \nabla A}$$

B2. Zgornji diamant

$$\frac{\Gamma \vdash A}{\Gamma \vdash \Delta A} \quad \frac{\Gamma \vdash \Delta \Delta A}{\Gamma \vdash \Delta A}$$

B3. Konjunkcija in spodnji/zgornji diamant

$$\frac{\Gamma \vdash \nabla(A \wedge B)}{\Gamma \vdash \nabla A \wedge \nabla B} \quad \frac{\Gamma \vdash \Delta(A \wedge B)}{\Gamma \vdash \Delta A \wedge \Delta B}$$

B4. Disjunkcija in spodnji/zgornji diamant

$$\frac{\Gamma \vdash \nabla A \vee \nabla B}{\Gamma \vdash \nabla(A \vee B)} \quad \frac{\Gamma \vdash \Delta A \vee \Delta B}{\Gamma \vdash \Delta(A \vee B)}$$

B5. Konjunkcija, implikacija in spodnji/zgornji diamant

$$\frac{\Gamma \vdash \nabla(A \wedge (A \Rightarrow B))}{\Gamma \vdash \nabla B} \quad \frac{\Gamma \vdash \Delta(A \wedge (A \Rightarrow B))}{\Gamma \vdash \Delta B}$$

Pravila sklepanja iz izreka 3 z dvojno črto beremo kot veljavne sklepe v obe smeri. Na primer, drugi del pravila B1 pravi, da iz $\Gamma \vdash \nabla \nabla A$ sledi $\Gamma \vdash \nabla A$ in obratno.

Dokaz. Z indukcijo na strukturo izpeljave A iz konteksta Γ . Pokažemo, da so vsi aksiomi veljavni, pravila sklepanja pa veljavnost ohranjajo. Z intuicionističnim delom ni težav, zato dokažimo le veljavnost pravil, ki vsebujejo katerega izmed modalnih operatorjev. Naj v svetu w velja množica izjav Γ .

- B1. Predpostavimo $w \Vdash A$. Zaradi refleksivnosti R (wRw) velja $w \Vdash \nabla A$. Za drugi del predpostavimo $w \Vdash \nabla \nabla A$. Obstaja torej svet $w'Rw$, v katerem velja ∇A in zato obstaja še en svet vRw' , da velja $v \Vdash A$. Zaradi tranzitivnosti relacije dosegljivosti (vRw) tedaj sledi $w \Vdash \nabla A$. Obratni sklep je posledica prvega dela pravila B1.
- B2. Predpostavimo $w \Vdash A$. Zaradi refleksivnosti R (wRw) velja $w \Vdash \Delta A$. Za drugi del spet predpostavimo $w \Vdash \Delta \Delta A$. Obstaja torej svet w' , da je wRw' in $w' \Vdash \Delta A$. Tedaj obstaja tudi svet v , da je $w'Rv$ in $v \Vdash A$. Zaradi tranzitivnosti relacije dosegljivosti (wRv) in veljavnosti $v \Vdash A$ sledi $w \Vdash \Delta A$. Obratni sklep je posledica prvega dela pravila B2.
- B3. Predpostavimo $w \Vdash \nabla(A \wedge B)$. To pomeni, da $\exists v : (vRw \text{ in } v \Vdash A \wedge B)$. Torej $v \Vdash A$ in $v \Vdash B$, iz česar sledi $w \Vdash \nabla A \wedge \nabla B$. Za zgornji diamant podobno.
- B4. Predpostavimo $w \Vdash \nabla A \vee \nabla B$. To pomeni, da $\exists v : (vRw \text{ in } v \Vdash A)$ ali $\exists v' : (v'Rw \text{ in } v' \Vdash B)$. V obeh primerih velja $w \Vdash \nabla(A \vee B)$. Za obratni sklep predpostavimo $w \Vdash \nabla(A \vee B)$, kar pomeni, da $\exists v : (vRw \text{ in } v \Vdash A \vee B)$. Sledi $v \Vdash A$ ali $v \Vdash B$, torej res $w \Vdash \nabla A \vee \nabla B$. Za zgornji diamant podobno.

B5. Predpostavimo $w \Vdash \nabla(A \wedge (A \Rightarrow B))$. To pomeni, da $\exists v : (vRw \text{ in } v \Vdash A \text{ in } v \Vdash A \Rightarrow B)$. Zaradi refleksivnosti \leq sledi $v \Vdash B$ in zato $w \Vdash \nabla B$. Za drugi del predpostavimo $w \Vdash \Delta(A \wedge (A \Rightarrow B))$. To pomeni, da $\exists v : (wRv \text{ in } v \Vdash A \text{ in } v \Vdash A \Rightarrow B)$. Zaradi refleksivnosti \leq spet sledi $v \Vdash B$ in zato $w \Vdash \Delta B$.

Konec dokaza.

Na primerih si oglejmo, kako se modul za vodenje dialoga obnaša, ko uporabnik naredi poteze tipa QUERY-YN, QUERY-WR, QUERY-WI, QUERY-WL ali QUERY-WT (poglavje 5), t.j. poteze, katerih namen je pridobiti določeno vremensko informacijo. Obnašanje sistema v primeru potez tipa QUERY-YN, QUERY-WR in QUERY-WT ponazarja primer 2 – sistem se premakne v svet z atomarnima izjavama zahtevanega kraja in vrste podatka ali pa v svet z atomarnima izjavama zahtevanega časovnega trenutka in kraja. Pri potezah tipa QUERY-WI (primer 3) se sistem premakne v svet z atomarnima izjavama zahtevanega kraja in časovnega trenutka, pri potezah tipa QUERY-WL (primer 4) pa v svet z atomarnima izjavama zahtevane vrste vremenskega podatka in časovnega trenutka.

Primer 2 *V mislih imejmo primer 1 in predpostavimo, da uporabnika zanima jutrišnja temperatura v Ljubljani. Sistem lahko (ko nima zahtevane informacije) uporabnika usmerja k izbiri dosegljivih, relevantnih podatkov tako, da ponudi relevantne kraje ali relevantne časovne trenutke, za katere je sposoben podati zahtevano informacijo. Relevantne kraje izbere tako, da se iz sveta Ljubljana premakne v svet, katerega dodatna omejitvev je temperatura. Svetova Ljubljana in Slovenija iz primera 1 tako preideta v svetova Ljubljana-temperatura in Slovenija-temperatura, ki ju dobimo tako, da množici prvotnih izjav dopolnimo z omejitvijo temperatura in uporabimo pravila sklepanja iz izreka 3. Pri tem je Ljubljana \leq Ljubljana-temperatura in Slovenija \leq Slovenija-temperatura. V svetu Ljubljana-temperatura so veljavne naslednje izjave:*

Ljubljana, temperatura

temperatura \Rightarrow trenutno (15)

trenutno \Rightarrow temperatura (15)

temperatura \Rightarrow danes zvečer (10)

danes zvečer \Rightarrow temperatura (10)

$\nabla^{\text{Slovenija-temperatura}}$ (temperatura \Rightarrow jutri podnevi) (od 13 do 17)

$\nabla^{\text{Slovenija-temperatura}}$ (jutri podnevi \Rightarrow temperatura) (od 13 do 17)

⋮

Naslednje končne izjave dobimo s pomočjo pravil sklepanja iz izreka 3.

trenutno (15)

danés zvečer (10)

$\nabla^{\text{Slovenija-temperatura}}$ jutri podnevi (od 13 do 17)

⋮

Dane izjave nazorno kažejo, da ima sistem podatke o trenutni temperaturi v Ljubljani in o temperaturi v Ljubljani danés zvečer, dosegljiv pa je tudi podatek o jutrišnji temperaturi v Sloveniji, kar je zagotovo relevantno, saj sta izjavni spremenljivki oz. vodilni atomarni izjavi Slovenija in Ljubljana v relaciji biti manj specifičen (slika 6.1).

Relevantne časovne trenutke pa sistem najde tako, da se premakne v svet jutri podnevi-Ljubljana, katerega vodilna atomarna izjava je **jutri podnevi**, dodatna omejitev pa Ljubljana. Svetovi jutri podnevi, trenutno in danés zvečer tedaj preidejo v svetove jutri podnevi-Ljubljana, trenutno-Ljubljana in danés zvečer-Ljubljana. Pri tem je jutri podnevi \leq jutri podnevi-Ljubljana, trenutno \leq trenutno-Ljubljana in danés zvečer \leq danés zvečer-Ljubljana, v svetu jutri podnevi-Ljubljana pa veljajo naslednje izjave:

jutri podnevi, Ljubljana

temperatura \Rightarrow Slovenija (od 13 do 17)

Slovenija \Rightarrow temperatura (od 13 do 17)

$\Delta^{\text{trenutno-Ljubljana}}$ (temperatura \Rightarrow Ljubljana) (15)

$\Delta^{\text{trenutno-Ljubljana}}$ (Ljubljana \Rightarrow temperatura) (15)

$\nabla^{\text{trenutno-Ljubljana}}$ (temperatura \Rightarrow Ljubljana) (15)

$\nabla^{\text{trenutno-Ljubljana}}$ (Ljubljana \Rightarrow temperatura) (15)

$\Delta^{\text{danés zvečer-Ljubljana}}$ (temperatura \Rightarrow Ljubljana) (10)

$\Delta^{\text{danés zvečer-Ljubljana}}$ (Ljubljana \Rightarrow temperatura) (10)

$\nabla^{\text{danés zvečer-Ljubljana}}$ (temperatura \Rightarrow Ljubljana) (10)

$\nabla^{\text{danés zvečer-Ljubljana}}$ (Ljubljana \Rightarrow temperatura) (10)

⋮

Naslednje končne izjave dobimo s pomočjo pravil sklepanja iz izreka 3.

$\Delta^{\text{trenutno-Ljubljana}}$ temperatura (15)

$\nabla^{\text{trenutno-Ljubljana}}$ temperatura (15)

$\Delta^{\text{danés zvečer-Ljubljana}}$ temperatura (10)

$\nabla^{\text{danés zvečer-Ljubljana}}$ temperatura (10)

⋮

Vidimo, da so podatki o trenutni temperaturi v Ljubljani in o temperaturi v Ljubljani danes zvečer relevantni, saj sta izjavni spremenljivki jutri podnevi in trenutno ter izjavni spremenljivki jutri podnevi in danes zvečer v obojesmerni relaciji biti manj specifičen (slika 6.3).

Primer 3 *Predpostavimo, da uporabnika zanimajo vse dosegljive vrste podatkov za Mursko Soboto (QUERY-WI poteza). Sistem najde odgovor na to vprašanje v svetu Murska Sobota-trenutno, katerega atomarni izjavi sta Murska Sobota (vodilna atomarna izjava) in trenutno. Vse izjave tega sveta dobimo tako, da vzamemo svet Murska Sobota in mu dodamo omejitev trenutno ter uporabimo pravila sklepanja iz izreka 3. Primeri izjav sveta Murska Sobota-trenutno so:*

Murska Sobota, trenutno

vreme \Rightarrow trenutno (*pretežno jasno*)

trenutno \Rightarrow vreme (*pretežno jasno*)

temperatura zraka \Rightarrow trenutno (6)

trenutno \Rightarrow temperatura zraka (6)

hitrost vetra \Rightarrow trenutno (1)

trenutno \Rightarrow hitrost vetra (1)

smer vetra \Rightarrow trenutno (*severozahodni veter*)

trenutno \Rightarrow smer vetra (*severozahodni veter*)

$\nabla^{\text{Slovenija-trenutno}}$ (max temperatura zraka \Rightarrow trenutno) (9)

$\nabla^{\text{Slovenija-trenutno}}$ (trenutno \Rightarrow max temperatura zraka) (9)

$\nabla^{\text{Slovenija-trenutno}}$ (min temperatura zraka \Rightarrow trenutno) (-4)

$\nabla^{\text{Slovenija-trenutno}}$ (trenutno \Rightarrow min temperatura zraka) (-4)

⋮

Naslednje končne izjave dobimo s pomočjo pravil sklepanja iz izreka 3.

vreme (*pretežno jasno*)

temperatura zraka (6)

hitrost vetra (1)

smer vetra (*severozahodni veter*)

$\nabla^{\text{Slovenija-trenutno}}$ max temperatura zraka (9)

$\nabla^{\text{Slovenija-trenutno}}$ min temperatura zraka (-4)

⋮

Izjave pravijo, da ima sistem podatke o vremenu, temperaturi zraka, hitrosti vetra in smeri vetra v Murski Soboti, pozna pa tudi trenutno najvišjo in najnižjo temperaturo zraka v Sloveniji.

Primer 4 *Predpostavimo, da uporabnika zanimajo vsi kraji, za katere sistem dostopa do podatkov o vremenu jutri popoldan (QUERY-WL poteza). Odgovor na to vprašanje sistem najde v svetu vreme-jutri popoldan, katerega atomarni izjavi sta vreme (vodilna atomarna izjava) in jutri popoldan. Vse izjave tega sveta dobimo tako, da vzamemo svet vreme in mu dodamo omejitev jutri popoldan ter spet uporabimo pravila sklepanja iz izreka 3. Primeri izjav sveta vreme-jutri popoldan so:*

vreme, jutri popoldan

jutri popoldan \Rightarrow morje (pretežno jasno)

morje \Rightarrow jutri popoldan (pretežno jasno)

jutri popoldan \Rightarrow gore (krajevne plohe)

gore \Rightarrow jutri popoldan (krajevne plohe)

jutri popoldan \Rightarrow sever (posamezne plohe)

sever \Rightarrow jutri popoldan (posamezne plohe)

jutri popoldan \Rightarrow zahod (delno jasno, občasno bolj oblačno)

zahod \Rightarrow jutri popoldan (delno jasno, občasno bolj oblačno)

$\triangle^{\text{NE-dežuje-jutri popoldan}}$ (jutri popoldan \Rightarrow morje) (pretežno jasno)

$\triangle^{\text{NE-dežuje-jutri popoldan}}$ (morje \Rightarrow jutri popoldan) (pretežno jasno)

$\triangle^{\text{vidljivost-jutri popoldan}}$ (jutri popoldan \Rightarrow Štajerska) (do 10 km)

$\triangle^{\text{vidljivost-jutri popoldan}}$ (Štajerska \Rightarrow jutri popoldan) (do 10 km)

⋮

Naslednje končne izjave dobimo s pomočjo pravil sklepanja iz izreka 3.

morje (pretežno jasno)

gore (krajevne plohe)

sever (*posamezne plohe*)

zahod (*delno jasno, občasno bolj oblačno*)

$\Delta^{\text{NE_de\u017euje-jutri popoldan}}$ morje (*prete\u017eno jasno*)

$\Delta^{\text{vidljivost-jutri popoldan}}$ \u0160tajerska (*do 10 km*)

⋮

Dobljene izjave pravijo, da sistem razpolaga s podatki o vremenu jutri popoldan za slovensko obalo, za gore ter za severni in zahodni del Slovenije. Poleg tega ve \u0161e, da ob slovenski obali jutri popoldan ne be de\u017eevalo, in zna povedati, kak\u0161na bo takrat vidljivost na \u0160tajerskem (vreme in NE_de\u017euje oz. vidljivost sta v relaciji biti manj specifi\u010den (slika 6.4)).

Omenimo \u0161e, kdaj svet w vsebuje izjavo $\neg A$ in kako se sistem v tem primeru odziva. Izjava $\neg A$ pomeni, da izjava A ne velja v nobenem svetu v , ki je z relacijo \leq dosegljiv iz sveta w . Tak\u0161ni svetovi $v \geq w$ vsebujejo vse atomarne izjave sveta w in vsaj eno dodatno omejitev. Izjava $\neg A$ zato v svetu w velja, \u010de za nobeno izbiro dodatnih omejitev vir znanja ne vsebuje informacije A . Recimo, da svet Ljubljana z edino atomarno izjavo Ljubljana vsebuje izjavo $\neg vreme$. To pomeni, da sistem nima podatka o vremenu za nobeno izbiro \u010dasovnega trenutka, kar sistem ve, \u0161e preden se premakne v svet Ljubljana-vreme, da bi uporabniku ponudil vreme v dosegljivih \u010dasovnih trenutkih. Svet Ljubljana-vreme v tem primeru vsebuje atomarno izjavo $vreme$ in neatomarno izjavo $\neg vreme$ (ekvivalentna $vreme \Rightarrow \perp$), iz \u010desar po aksiomu A6 iz definicije 5 izpeljemo la\u017e \perp .

7 Evalvacija modula za vodenje dialoga

7.1 Ogrodje PARADISE

7.2 Rezultati evalvacije dialogov iz prvega eksperimenta Čarovnik iz Oza

7.3 Rezultati evalvacije dialogov iz drugega eksperimenta Čarovnik iz Oza

Opisali bomo potencialno splošno metodologijo evalvacije sistemov za dialog, namreč ogrodje PARADISE (PARAdigm for DIAlogue System Evaluation), ki omogoča izpeljavo ocene učinkovitosti sistema kot uteženo linearno kombinacijo od domene odvisnih *parametrov uspešnosti naloge in cen dialoga*.

Ogrodje PARADISE bomo uporabili za evalvacijo sistemov iz obeh eksperimentov Čarovnik iz Oza in prišli do zaključka, da lahko učinkovitost modula za vodenje dialoga opazno izboljšamo, če zagotovimo čim večjo fleksibilnost sistema pri iskanju dosegljivih, relevantnih podatkov.

To fleksibilnost sicer v precejšnji meri zagotavlja že *dialog mIPC Kripkejev model* znanja, zaradi velikega vpliva na zadovoljstvo uporabnikov pa jo bomo v poglavju 8 poskušali še izboljšati.

7.1 Ogradje PARADISE

Z razvojem sistemov za dialog se pojavljajo tudi potrebe po evalvaciji in primerjavi teh sistemov. V zadnjih dveh desetletjih so bili zato predlagani številni *objektivni parametri dialoga* [Price-1992, Danieli-1995, Smith-1997],

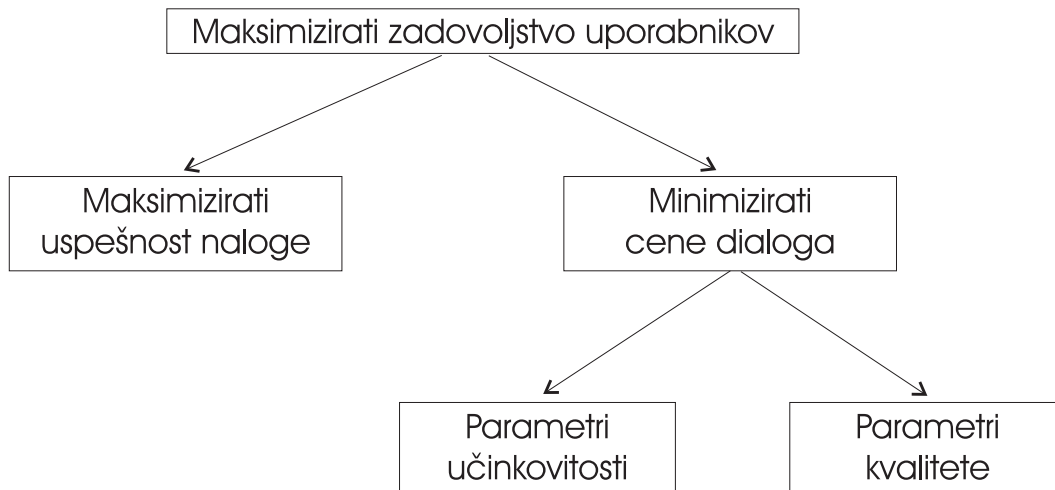
- število izjav,
- čas dialoga,
- povprečni čas odziva uporabnika,
- povprečni čas odziva sistema,
- delež izjav, sestavljenih iz več kot ene besede,
- povprečna dolžina izjav, sestavljenih iz več kot ene besede,

ki jih lahko določimo brez mnenja človeka, in parametri, ki temeljijo na mnenju človeka, namreč *subjektivni parametri dialoga* [Shriberg-1992, Danieli-1995, Boyce-1996],

- delež izjav, s katerimi sistem popravlja napake,
- delež kontekstno primernih izjav sistema,
- hevristična evalvacija stopnje sodelovanja sistema na podlagi Griceovih maksim [Grice-1975], opisanih v podpoglavju 4.2,
- delež pravih in delno pravih odgovorov,
- delež primernih in neprimernih izjav, s katerimi sistem usmerja uporabnika,
- zadovoljstvo uporabnika.

Izkaže pa se, da kombinacija različnih parametrov evalvacije dialoga lahko pripelje do težav. V [Danieli-1995] podana primerjava dveh sistemov za podajanje informacij o železniških povezavah na primer kaže, da lahko en sistem sicer omogoča več uspešno zaključenih dialogov in povzroča manj neustreznih in popravilnih izjav, dialogi z drugim sistemom pa so zato lahko za polovico krajši. V tem primeru ni bilo mogoče ugotoviti, kateri od sistemov je bolj uporaben.

Z namenom omogočiti primerjavo sistemov z različnimi domenami, kjer je pomembno vedeti, v kolikšni meri posamezni parametri vplivajo na učinkovitost in kako strategija vodenja dialoga vpliva na zadovoljstvo uporabnikov, je bilo leta 1997 kot potencialna splošna metodologija evalvacije sistemov za dialog predlagano ogradje PARADISE (PARAdigm for DIalogue System Evaluation) [Walker-1997]. Ogradje PARADISE omogoča izpeljavo ocene učinkovitosti sistema kot uteženo linearno kombinacijo od domene odvisnih *parametrov uspešnosti naloge* in *cen dialoga*, zajema pa model učinkovitosti sistema (slika 7.1), ki za osnovni cilj postavlja maksimizirati zadovoljstvo uporabnikov [Shriberg-1992].



Slika 7.1: Model učinkovitosti sistema za dialog, ki ga zajema ogrodje PARADISE

Cene dialoga lahko razdelimo na dva tipa: *parametri učinkovitosti* in *parametri kvalitete*. Parametri učinkovitosti dialoga (npr. število izjav, ki jih uporabnik potrebuje, da izpolni svojo namero, ali čas dialoga) merijo, kako učinkovito sistem uporabniku pomaga pri doseganju njegove namere. Parametri kvalitete dialoga (npr. kolikokrat mora uporabnik ponoviti svojo izjavo, da ga sistem razume, ali kakšen je čas čakanja na odziv sistema) pa zajemajo ostale vidike, ki lahko na zadovoljstvo uporabnika prav tako močno vplivajo.

Če želimo sistem za dialog evalvirati z ogrodjem PARADISE, moramo podatke zbrati v eksperimentu, v katerem bodo uporabniki ocenili svoje zadovoljstvo, t.j. podali stopnjo strinjanja z izjavami o obnašanju oz. učinkovitosti sistema. Ostale parametre modela (parametri uspešnosti naloge, cene dialoga) pa je treba določati samodejno ali jih ročno označiti. Model učinkovitosti sistema, ki ga zajema ogrodje PARADISE, trdi, da lahko funkcijo učinkovitosti sistema določimo s pomočjo *multiple linearne regresije* (MLR) z zadovoljstvom uporabnikov kot neodvisno spremenljivko ter parametri uspešnosti naloge, parametri učinkovitosti dialoga in parametri kvalitete dialoga kot neodvisnimi spremenljivkami:

$$Učinkovitost = (\alpha * \mathcal{N}(\kappa)) - \sum_{i=1}^n w_i * \mathcal{N}(c_i)$$

Pri tem je α utež Kappa koeficienta κ [Carletta-1996b], ki ga izračunamo iz kontingenčne tabele uspešnosti razumevanja uporabnikovih govorjenih izjav (podpoglavje 7.2), w_i so uteži cen c_i , \mathcal{N} pa je funkcija normalizacije,

$$\mathcal{N}(x) = \frac{x - \bar{x}_0}{\sigma_{x_0}},$$

kjer sta \bar{x}_0 in σ_{x_0} srednja vrednost in standardni odklon spremenljivke x_0 v učni množici. Srednja vrednost s funkcijo normalizacije \mathcal{N} preslikanih parametrov učne množice je 0, standardni odklon pa 1. Tako se znebimo težav, ki se pojavijo, če primerjamo vrednosti parametrov, ki se raztezajo na različnih intervalih in/ali so njihove vrednosti različno razpršene. Z normalizacijo parametrov κ in c_i dosežemo relevantnost prispevkov

uteži preslikanih parametrov $\mathcal{N}(\kappa)$ in $\mathcal{N}(c_i)$ k učinkovitosti sistema za dialog. Vrednosti funkcije učinkovitosti sistema lahko uporabimo za napovedovanje zadovoljstva uporabnikov, pri čemer za učno množico vrednosti odvisne spremenljivke uporabimo vrednosti zadovoljstva uporabnikov, pridobljene v ustreznem eksperimentu. Rezultat multiple linearne regresije na učni množici parametrov, ki praviloma tvorijo predoločen sistem, je množica uteži, ki pomenijo relativni prispevek teh parametrov k učinkovitosti sistema.

MLR temelji na metodi najmanjših kvadratov, saj minimizira vsoto kvadratov razlik med v eksperimentu pridobljenimi vrednostmi in napovedanimi vrednostmi zadovoljstva uporabnikov. Torej

$$\mathcal{N}(US) = \widehat{\mathcal{N}(US)} + \epsilon,$$

kjer je $\mathcal{N}(US)$ normalizirana pridobljena, $\widehat{\mathcal{N}(US)}$ pa napovedana normalizirana vrednost zadovoljstva uporabnikov. Napako, t.j. razliko med obema vrednostma, smo označili z ϵ . Nenormalizirano vrednost zadovoljstva uporabnikov lahko tedaj ocenimo kot

$$US = \widehat{\mathcal{N}(US)}\sigma_{US_0} + \overline{US_0} + \epsilon\sigma_{US_0} = \widehat{US} + \epsilon\sigma_{US_0},$$

kjer sta $\overline{US_0}$ in σ_{US_0} srednja vrednost in standardna deviacija vrednosti zadovoljstva uporabnikov v učni množici MLR modela učinkovitosti sistema.

Kako dobro \widehat{US} napoveduje US , lahko vidimo s pomočjo razmerja absolutnih vrednosti njune razlike in pridobljene vrednosti zadovoljstva US danega uporabnika,

$$q(US, \widehat{US}) = \frac{|US - \widehat{US}|}{|US|}.$$

Naslednje razmerje pa kaže, da ocena normalizirane vrednosti zadovoljstva uporabnika ni tako dobra kot ocena nenormalizirane vrednosti, če je le $US > \frac{\overline{US_0}}{2}$:

$$\frac{q(\mathcal{N}(US), \widehat{\mathcal{N}(US)})}{q(US, \widehat{US})} = \frac{\frac{|\mathcal{N}(US) - \widehat{\mathcal{N}(US)}|}{|\mathcal{N}(US)|}}{\frac{|US - \widehat{US}|}{|US|}} = \frac{\frac{|US - \overline{US_0} - \widehat{\mathcal{N}(US)}\sigma_{US_0}|}{|US - \overline{US_0}|}}{\frac{|US - \widehat{\mathcal{N}(US)}\sigma_{US_0} - \overline{US_0}|}{|US|}} = \frac{|US|}{|US - \overline{US_0}|}$$

Oceno nenormalizirane vrednosti zadovoljstva uporabnika \widehat{US} , ki je v večini primerov veliko boljše od ocene normalizirane vrednosti $\widehat{\mathcal{N}(US)}$, dobimo takole:

$$\widehat{US} = \widehat{\mathcal{N}(US)}\sigma_{US_0} + \overline{US_0}$$

Ne samo, da v literaturi [Walker-1997, Walker-1998, Kamm-1999, Litman-2002] temu ni posvečena pozornost, ampak tudi ni omenjeno, da je treba vrednosti zadovoljstva uporabnikov v učni množici, preden začnemo z izpeljavo takšnega MLR modela učinkovitosti, normalizirati, če želimo preprečiti prevelike napake ocen.

Uporaba te metode pri reševanju predoločenega sistema pa zahteva izpolnitev vrste pogojev [Johnson-2002]. Najbolj očitna zahteva je, da mora obstajati približno linear-na zveza med odvisno spremenljivko na eni in neodvisnimi spremenljivkami na drugi

strani. Naslednja zahteva ne dovoljuje visoke koreliranosti neodvisnih spremenljivk (po absolutni vrednosti okrog 0.5), kar lahko preverimo z izračunom korelacij med spremenljivkami. Odvečne neodvisne spremenljivke je zato potrebno odstraniti iz predločenega sistema. Pokazati pa se da tudi, da neodvisne spremenljivke, ki so z odvisno spremenljivko v zelo nizki korelaciji (po absolutni vrednosti pod 0.1), povečajo varianco napovedi, zaradi česar lahko napako napovedi zmanjšamo, če takšne spremenljivke iz sistema odstranimo.

Potem ko smo zgradili modul za vodenje dialoga v sistemu za podajanje informacij o vremenu, smo želeli ovrednotiti njegovo učinkovitost oz. videti, kako ga sprejemajo uporabniki. V ta namen smo vzeli podatke iz obeh eksperimentov Čarovnik iz Oza (poglavje 3), jih ovrednotili z ogrođjem PARADISE in rezultate med sabo primerjali.

7.2 Rezultati evalvacije dialogov iz prvega eksperimenta Čarovnik iz Oza

Sistem Čarovnik iz Oza, ki smo ga uporabili v prvem eksperimentu Čarovnik iz Oza (podpoglavje 3.3), bomo evalvirali z ogrođjem PARADISE in rezultate primerjali z rezultati evalvacije podatkov iz drugega eksperimenta (podpoglavje 3.4). Edino razliko sistemov, uporabljenih v teh eksperimentih, predstavlja način vodenja dialoga. Prvič je to nalogo s pomočjo grafičnega vmesnika opravljal čarovnik, drugič pa modul za vodenje dialoga, opisan v poglavju 5.

V obeh eksperimentih so uporabniki ocenili svoje zadovoljstvo tako, da so podali stopnjo strinjanja z izjavami o obnašanju oz. učinkovitosti sistema (tabela 3.2). Splošno **zadovoljstvo uporabnika** smo dobili tako, da smo vse točke te ankete sešteli. Vrednosti teh parametrov, ki jih bomo uporabljali kot vrednosti odvisne spremenljivke v MLR modelu učinkovitosti sistema, zato ležijo med 8 in 40. Del parametrov uspešnosti naloge, parametrov učinkovitosti dialoga in parametrov kvalitete dialoga, ki v MLR modelu učinkovitosti predstavljajo neodvisne spremenljivke, smo določili samodejno, del pa smo ročno označili. V prvem eksperimentu smo tako določili parametra uspešnosti naloge

- **Kappa koeficient** (κ) in
- **izpolnitev naloge** (Comp), mnenje uporabnika o izpolnitvi prve naloge, ki smo mu jo v eksperimentu zastavili,

parameter učinkovitosti dialoga

- **povprečni čas dialoga** (MET), povprečni čas trajanja dialoga brez vštetega trajanja odzivov sistema,

in parametre kvalitete dialoga

- **povprečni čas čakanja na odziv** (MRT), povprečni čas, ki ga sistem porabi, preden se odzove,
- **delež nepodanih informacij** (NPR), delež začetnih potez uporabnika, na katere sistem ne poda relevantnega odgovora,
- **delež zavrnitev** (RR), delež izjav sistema, s katerimi uporabnika prosi, naj ponovi zadnjo izjavo, t.j. *Pardon* potez,
- **delež nudenja pomoči** (HMR), delež izjav sistema, s katerimi uporabniku pomaga nadaljevati dialog, t.j. *Reply-help* in *Reply-timeout* potez,
- **delež negativnih odgovorov** (NDR), delež potez, s katerimi sistem sporoča, da nima zahtevanega podatka in pri tem uporabnika ne usmerja k izbiri relevantnih, dosegljivih podatkov,
- **delež izbire relevantnih podatkov** (RDR), delež potez sistema, ki uporabnika usmerjajo k izbiri relevantnih, dosegljivih podatkov, in
- **delež neprimernih iniciativ** (UIR), delež začetnih potez uporabnika, katerih vsebina ne ustreza domeni sistema.

Kappa koeficient (κ) izračunamo s pomočjo kontingenčne tabele, ki podaja ujemanje med vsebino zahtev uporabnika (*ključ*) in dojetanjem te s strani sistema (*odgovor*). Primer kontingenčne tabele prikazuje slika 7.2.

ODGOVOR	KLJUČ	kraj				čas				podatek			
		Krivavec	Primorje	obala	Rogla	trenutno	vikend	sobota	danes	temp.	viš. snega	temp. močja	smer vetra
kraj	Krivavec	2											
	Primorje		1										
	obala			1									
	Rogla				1								
čas	trenutno					3							
	vikend						0						
	sobota						1	0					
	danes						1		0				
podatek	temp.									2			
	viš. snega										1		
	temp. močja											1	
	smer vetra												1
	15	2	1	1	1	3	2	0	0	2	1	1	1

Slika 7.2: Primer kontingenčne tabele dialoga

Vsako ujemanje ključa z odgovorom poveča število v ustrezni diagonalni celici kontingenčne tabele za 1. Zunajdiagonalne celice pripadajo neujemanjem, ki znotraj tekoče konverzijske igre niso bila odpravljena. Začasni nesporazumi, ki jih sogovornika znotraj tekoče konverzijske igre odpravita, se namreč odražajo v cenah dialoga. **Kappa koeficient**, ki meri ujemanje opisov različnih označevalcev, izračunamo kot

$$\kappa = \frac{P(A) - P(E)}{1 - P(E)},$$

kjer

$$P(A) = \frac{\sum_{i=1}^n M(i, i)}{T}$$

izraža ujemanje ključa in odgovora v kontingenčni tabeli M z vsoto vseh števil enako T , ki opisuje dialog z n konverzacijskimi igrami,

$$P(E) = \sum_{i=1}^n \left(\frac{t_i}{T}\right)^2$$

pa je verjetnost pričakovanega ujemanja dveh različnih označevalcev, ki dane kategorije tabele izbirata z enako verjetnostjo, pri čemer je t_i vsota števil i -tega stolpca tabele. Če je $P(A)$ enako verjetnosti pričakovanega ujemanja $P(E)$, je $\kappa = 0$, pri popolnem ujemanju ($P(A) = 1$), pa je $\kappa = 1$.

Srednje vrednosti vseh parametrov dialogov, pridobljenih v obeh eksperimentih Čarovnik iz Oza, so podane v tabeli 7.1.

	WOZ1	WOZ2
Kappa koeficient (κ)	0.94	0.98
izpolnitev naloge (Comp)	0.97	0.96
povprečni čas dialoga (MET)	15.41 s	22.03 s
povprečni čas čakanja na odziv (MRT)	5.57 s	6.61 s
delež nepodanih informacij (NPR)	0.31	0.28
delež zavrnitev (RR)	0.01	0.03
delež nudenja pomoči (HMR)	0.01	0.06
delež negativnih odgovorov (NDR)	0.08	0.07
delež izbire relevantnih podatkov (RDR)	0.06	0.20
delež neprimernih iniciativ (UIR)	0.13	0.04
delež neiniciativnih potez (NIR)	-	0.50
delež preverjanj (CR)	-	0.20
zadovoljstvo uporabnika (US)	34.08	31.96

Tabela 7.1: Srednje vrednosti parametrov v prvem (WOZ1) in v drugem (WOZ2) eksperimentu Čarovnik iz Oza

Zanimivo je, da se cene kvalitete dialoga **delež nepodanih informacij**, **delež negativnih odgovorov** in **delež izbire relevantnih podatkov** v literaturi ne pojavljajo. Razlog je verjetno ta, da imajo razvijalci sistemov za dialog redko na razpolago podatkovno zbirko, katere struktura bi bila tako zelo časovno odvisna in skopa kot je naša.

Izračun korelacijskih koeficientov posameznih parametrov (tabela 7.2) pokaže, da je iz MLR modela učinkovitosti sistema zaradi zelo nizke korelacije (-0.01) z odvisno

spremenljivko **zadovoljstvo uporabnika** (US) potrebno odstraniti **delež izbire relevantnih podatkov** (RDR). Razen tega je zaradi visoke korelacije (0.47 oz. 0.48) s parametrom **delež nepodanih informacij** (NPR) potrebno odstraniti tudi **delež negativnih odgovorov** (NDR) in **delež neprimernih iniciativ** (UIR). Pri tem omenimo, da sta mejni vrednosti 90% intervala zaupanja za korelacijski koeficient 0.1 (meja izločanja zaradi prenizke korelacije) enaki 0.08 in 0.13, mejni vrednosti 90% intervala zaupanja za korelacijski koeficient 0.5 (meja izločanja zaradi previsoke korelacije) pa 0.39 in 0.67.

	κ	Comp	MET	MRT	NPR	RR	HMR	NDR	RDR	UIR	US
κ	1.0										
Comp	0.0	1.0									
MET	-0.2	0.2	1.0								
MRT	-0.2	0.0	0.3	1.0							
NPR	0.1	-0.2	-0.3	-0.1	1.0						
RR	-0.1	-0.2	-0.1	0.3	0.3	1.0					
HMR	0.1	-0.1	0.3	-0.0	0.1	-0.0	1.0				
NDR	-0.1	-0.1	-0.1	0.1	0.5	0.2	0.1	1.0			
RDR	0.1	-0.0	0.0	-0.1	0.1	0.0	0.1	-0.2	1.0		
UIR	0.3	-0.1	-0.1	-0.3	0.5	0.1	0.4	-0.1	-0.1	1.0	
US	0.1	0.3	0.1	-0.1	-0.3	-0.3	-0.4	-0.3	-0.0	-0.2	1.0

Tabela 7.2: Korelacijski koeficienti parametrov v prvem eksperimentu

Pri izpeljavi funkcije učinkovitosti sistema z metodo multiple linearne regresije smo ugotovili, da na zadovoljstvo uporabnikov najbolj vplivajo **izpolnitev naloge** (Comp), **povprečni čas čakanja na odziv** (MRT), **delež nepodanih informacij** (NPR), **delež zavrnitev** (RR) in **delež nudenja pomoči** (HMR):

$$\widehat{\mathcal{N}}(US) = 0.21\mathcal{N}(Comp) - 0.13\mathcal{N}(MRT) - 0.24\mathcal{N}(NPR) - 0.11\mathcal{N}(RR) - 0.35\mathcal{N}(HMR)$$

Napovedane ocene $\widehat{\mathcal{N}}(US)$ preslikamo nazaj na prvotno skalo. Dejanske vrednosti **zadovoljstva uporabnikov** v prvem eksperimentu Čarovnik iz Oza so bližje tem preslikanim ocenam kot pa normalizirane vrednosti ugotovljenega **zadovoljstva uporabnikov** nepreslikanim ocenam. Razmerje dolžin razlike vektorja pridobljenih vrednosti US_1 in vektorja preslikanih napovedanih vrednosti \widehat{US}_1 ter dolžine prvega,

$$Q(US_1, \widehat{US}_1) = \frac{\|US_1 - \widehat{US}_1\|_2}{\|US_1\|_2},$$

je bilo 0.12. Razmerje dolžin vektorjev US_1 in \widehat{US}_1 je bilo 0.99, kot med njima pa je meril 6.92° . Za primerjavo povejmo, da je bilo razmerje dolžin $Q(\mathcal{N}(US_1), \widehat{\mathcal{N}}(US_1))$ enako 6.85, razmerje dolžin obeh vektorjev 0.57, kot med njima pa je meril 55.56° .

7.3 Rezultati evalvacije dialogov iz drugega eksperimenta Čarovnik iz Oza

V drugem eksperimentu Čarovnik iz Oza smo definirali dva dodatna parametra kvalitete dialoga, namreč

- **delež neinicijativnih potez** (NIR), delež potez uporabnika, ki ne spadajo med začetne poteze, in
- **delež preverjanj** (CR), delež *Check* potez sistema.

Srednje vrednosti vseh parametrov dialogov, pridobljenih v drugem eksperimentu Čarovnik iz Oza, so, skupaj z vrednostmi iz prvega eksperimenta, podane v tabeli 7.1.

Izračun korelacijskih koeficientov posameznih parametrov (tabela 7.3) pokaže, da je iz MLR modela učinkovitosti sistema zaradi zelo nizke korelacije (-0.05 oz. 0.07) z odvisno spremenljivko **zadovoljstvo uporabnika** (US) potrebno odstraniti **delež zavrnitev** (RR) in **delež negativnih odgovorov** (NDR). Razen tega je zaradi visoke korelacije (0.69) s parametrom **povprečni čas dialoga** (MET) potrebno odstraniti **delež neinicijativnih potez** (NIR), zaradi visoke korelacije (0.90 oz. 0.86) s parametrom **delež preverjanj** (CR) pa še **delež nudenja pomoči** (HMR) in **delež izbire relevantnih podatkov** (RDR). Spet omenimo, da sta mejni vrednosti 90% intervala zaupanja za korelacijski koeficient 0.1 (meja izločanja zaradi prenizke korelacije) enaki 0.08 in 0.14, mejni vrednosti 90% intervala zaupanja za korelacijski koeficient 0.5 (meja izločanja zaradi previsoke korelacije) pa 0.38 in 0.68.

	κ	Comp	MET	MRT	NIR	NPR	RR	HMR	CR	NDR	RDR	UIR	US
κ	1.0												
Comp	0.0	1.0											
MET	-0.4	-0.3	1.0										
MRT	-0.2	-0.0	0.4	1.0									
NIR	-0.2	-0.2	0.7	-0.2	1.0								
NPR	0.1	-0.2	-0.1	0.1	-0.3	1.0							
RR	-0.1	-0.3	0.3	0.1	0.3	0.0	1.0						
HMR	0.1	0.1	-0.0	-0.1	0.0	0.3	-0.0	1.0					
CR	-0.0	0.1	0.1	-0.1	0.2	0.2	-0.1	0.9	1.0				
NDR	0.2	0.1	-0.2	0.1	-0.4	0.6	-0.1	-0.1	-0.2	1.0			
RDR	0.1	0.0	-0.0	-0.1	0.0	0.3	-0.1	0.9	0.9	-0.1	1.0		
UIR	-0.4	-0.1	0.3	0.3	0.1	0.2	0.2	-0.0	0.0	-0.0	-0.1	1.0	
US	0.3	0.2	-0.3	-0.2	-0.2	-0.2	-0.0	-0.1	-0.3	0.1	-0.1	-0.2	1.0

Tabela 7.3: *Korelacijski koeficienti parametrov v drugem eksperimentu*

Pri izpeljavi funkcije učinkovitosti sistema z metodo multiple linearne regresije smo ugotovili, da so na zadovoljstvo uporabnikov v drugem eksperimentu Čarovnik iz Oza

najbolj vplivali **Kappa koeficient** (κ), **izpolnitev naloge** (Comp), **povprečni čas dialoga** (MET), **delež nepodanih informacij** (NPR) in **delež preverjanj** (CR):

$$\widehat{\mathcal{N}}(US) = 0.20\mathcal{N}(\kappa) + 0.13\mathcal{N}(Comp) - 0.17\mathcal{N}(MET) - 0.14\mathcal{N}(NPR) - 0.25\mathcal{N}(CR)$$

Napovedane vrednosti preslikamo nazaj na prvotno skalo. Razmerje dolžin $Q(US_2, \widehat{US}_2)$ je bilo 0.13, razmerje dolžin vektorjev US_2 in \widehat{US}_2 pa 0.99. Kot med njima je meril 7.70° . Podobno kot prej, za primerjavo povejmo, da je bilo razmerje $Q(\mathcal{N}(US_2), \widehat{\mathcal{N}}(US_2))$ enako 6.53, razmerje dolžin obeh vektorjev 0.48, kot med njima pa je meril 61.00° .

V evalvaciji sistemov za branje elektronske pošte [Walker-1998] je bilo ugotovljeno, da **izpolnitev naloge** močnejše vpliva na **zadovoljstvo uporabnika** kot pa **Kappa koeficient**. Razlog, ki ga navajajo, je ta, da naj bi uporabniki velikokrat drugače dojemali delovanje sistema, kot ga podaja **Kappa koeficient**. V naših eksperimentih je bil **Kappa koeficient** takorekoč odvisen le od čarovnika, ki je nadomeščal razumevanje govora, parameter **izpolnitev naloge** pa se je nanašal le na prvo nalogo, ki smo jo uporabniku zastavili, kar je najverjetneje razlog, zakaj sami nismo prišli do podobnega zaključka. Po eni strani sta bila v naših eksperimentih **Kappa koeficient** in **izpolnitev naloge** nekorelirana (0.00 v prvem in 0.05 v drugem eksperimentu), po drugi strani pa je v drugem eksperimentu **Kappa koeficient** celo močnejše vplival na **zadovoljstvo uporabnikov**.

Parametri, ki so pomembno vplivali na učinkovitost sistema, pri tem pa niso bili podvrženi čarovniku, so **delež nepodanih informacij** in **delež nudenja pomoči** v prvem ter **delež nepodanih informacij** in **delež preverjanj** v drugem eksperimentu. Vrednost parametra **delež nudenja pomoči** je odvisna od tega, kako se uporabnik v dialogu obnaša, to pa je spet odvisno od stopnje prijaznosti in sodelovanja, ki jo sistem pri vodenju dialoga nudi. Vsekakor, edini cilj sistema za dialog ne sme biti le uspešen zaključek naloge, ampak tudi sposobnost prevzeti iniciativo in nuditi pomoč, ko jo uporabnik potrebuje. Ker nekateri novi uporabniki sistema, ki se niso sposobni hitro prilagajati, pri vodenju dialoga pogosto potrebujejo pomoč, vpliva parametra **delež nudenja pomoči** na učinkovitost sistema ni mogoče odpraviti. Tudi vpliva parametra **delež preverjanj** ni mogoče odpraviti, saj razumevanje govora ponavadi predstavlja najtežavnejši del, če ne celo oviro do učinkovitosti sistema za dialog. Vidimo, da lahko zadovoljstvo uporabnikov pomembno povečamo le z zmanjšanjem vpliva parametra **delež nepodanih informacij**, t.j. deleža začetnih potez uporabnika, na katere sistem ne poda relevantnega odgovora. Zmanjšanje vrednosti tega parametra lahko dosežemo tako, da sistemu ne dovolimo odziva, preden se ni zares prepričal, da ne dostopa do nobenega relevantnega podatka. Sistem naj torej uporabnika čim bolj fleksibilno usmerja k izbiri dosegljivih, relevantnih podatkov. Čeprav fleksibilnost sistema za podajanje informacij o vremenu v veliki meri podpira v poglavju 6 opisan *dialog mIPC Kripkejev model* znanja, ki je bil uporabljen v drugem eksperimentu Čarovnik iz Oza, bomo v poglavju 8 to predstavitev znanja poskušali še nekoliko izboljšati.

Na koncu smo želeli videti tudi, kakšen je vpliv posameznih parametrov na določene vidike učinkovitosti sistemov iz obeh eksperimentov Čarovnik iz Oza. V ta namen točke vprašalnikov (tabela 3.2), v katerih so uporabniki ocenjevali obnašanje in učinkovitost

sistemov, nismo sešteli, ampak vsako izmed njih vzeli za odvisno spremenljivko in ji določili MLR model. Ker je bil naš cilj v prvi vrsti evalvirati modul za vodenje dialoga, smo se pri tem osredotočili na *pridobivanje informacij* (Ali ste brez težav prišli do odgovora na vaša vprašanja?), *ustreznost odzivov* (Ali se je sistem na vaše izjave odzival hitro (brez pojasnilnih vprašanj)?) in *pričakovano obnašanje* (Ali se je sistem obnašal tako, kot ste med dialogom od njega pričakovali?). Ugotovili smo, da so na *pridobivanje informacij* najbolj vplivali **Kappa koeficient, povprečni čas dialoga, delež nepodanih informacij, delež zavrnitev in delež preverjanj**; na *ustreznost odzivov* so najbolj vplivali **Kappa koeficient, povprečni čas dialoga in delež preverjanj**; na *pričakovano obnašanje* pa **Kappa koeficient, izpolnitev naloge, delež nepodanih informacij in delež preverjanj**. Ti rezultati naše sklepe vsekakor potrjujejo.

8 Zaključek

8.1 Razširitev predstavitve znanja

8.2 Nadaljnje delo

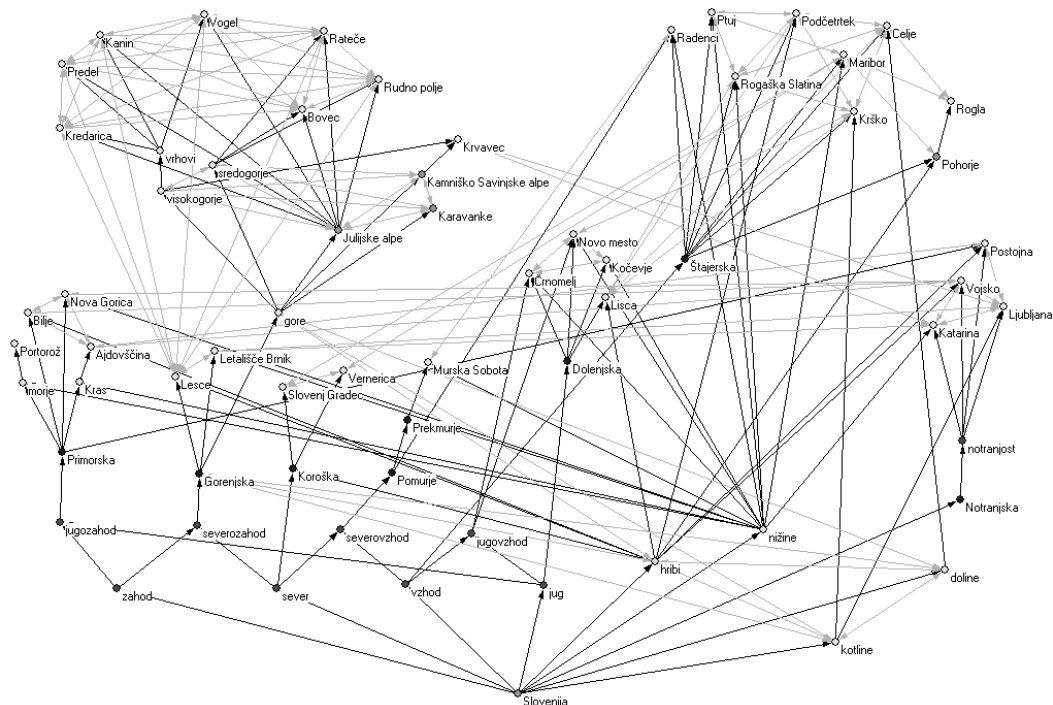
Zaključki, do katerih smo prišli na podlagi rezultatov obeh eksperimentov Čarovnik iz Oza, vodijo do razmišljanja o morebitnih izboljšavah predstavitve znanja, predstavljene v poglavju 6. V skladu s tem bomo predlagali manjšo spremembo modela, ki je trenutno udejanjen v sistemu za podajanje informacij o vremenu.

Magistrsko delo bomo zaključili z opisom nadaljnjega dela. Predstavili bomo zastavljene cilje in ideje na poti do dokončne izgradnje dvojezičnega sistema za podajanje informacij o vremenu.

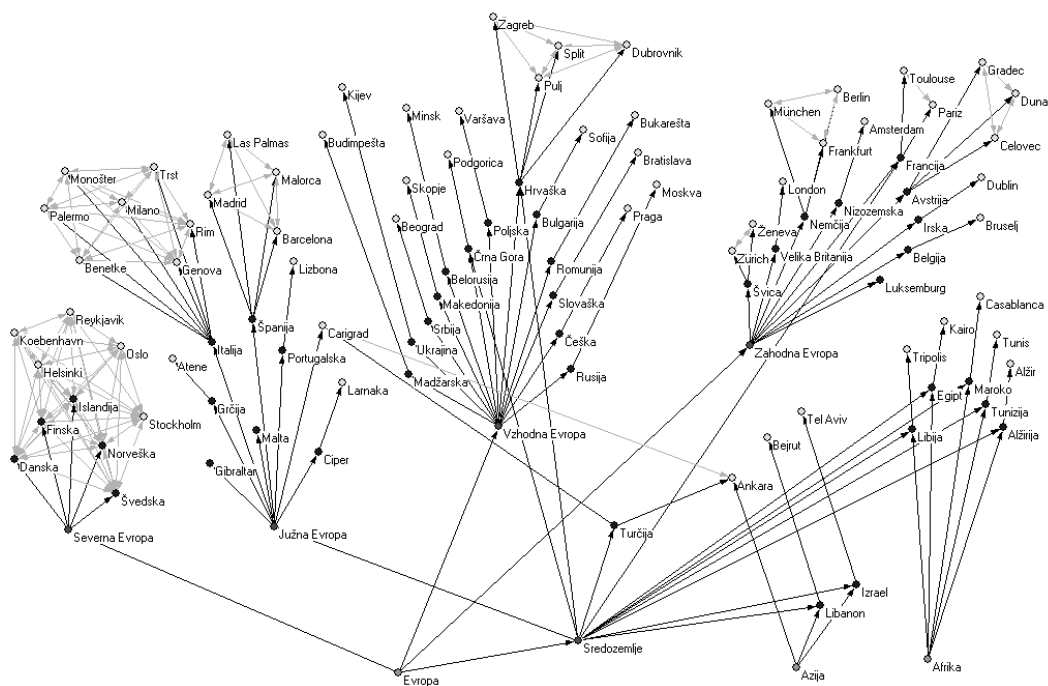
8.1 Razširitev predstavitve znanja

Oba eksperimenta Čarovnik iz Oza sta pokazala, da je predstavitev znanja v sistemu za dialog izrednega pomena. Čeprav smo v drugem eksperimentu že uporabili predstavitev znanja (poglavje 6), ki temelji na intuicionistični modalni logiki in v veliki meri omogoča ponujanje relevantnih, dosegljivih podatkov, je za izboljšanje zadovoljstva uporabnikov takorekoč nujno razmišljati o morebitnih dodatnih razširitvah oz. izboljšavah.

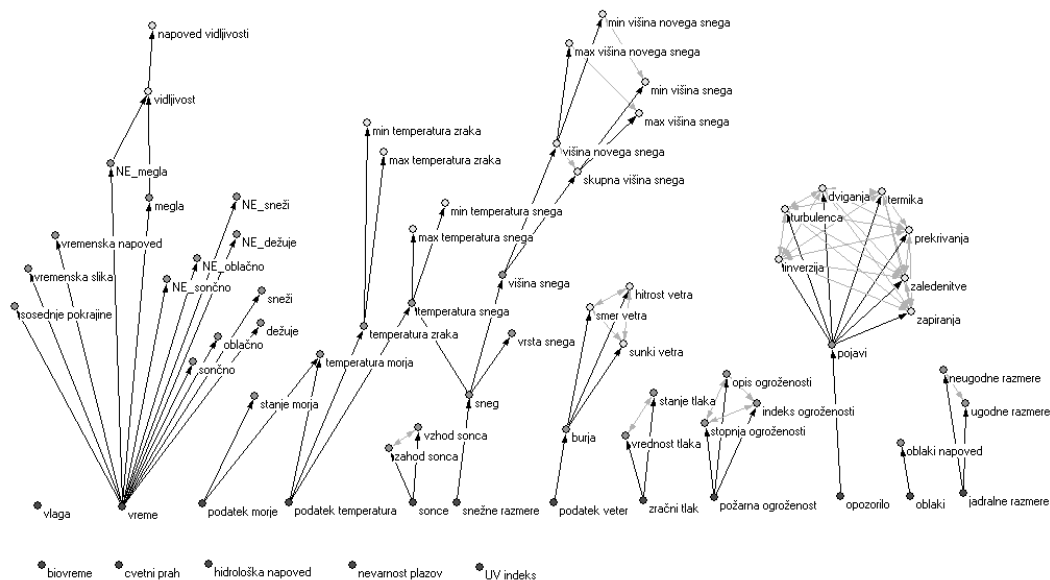
Slabost prvega podatkovnega modela predstavlja tranzitivnost relacije dosegljivosti R , definirane med svetovi dialog mIPC Kripkejevih modelov (definicija 8), oz. tranzitivnost relacije *biti manj specifičen*, definirane med delčki informacij (slike 6.1, 6.2, 6.3 in 6.4). Če bi namreč z relacijo *biti manj specifičen* povezali dva bližnja kraja, ki ležita v različnih pokrajinah (npr. Postojna in Ajdovščina), bi zaradi tranzitivnosti lahko dobili odvečne, nerelevantne povezave (npr. da je *Notranjska* manj specifična od *Ajdovščine*, kar bi pomenilo, da kraj Ajdovščina leži na Notranjskem). To težavo odstranimo z uvedbo nove relacije med delčki informacij, imenovane relacija *sosebnosti* \mathcal{N} , ki delčka informacij p in q ($p\mathcal{N}q$) povezuje natanko tedaj, ko je q na podlagi svoje bližine (bodisi krajevne bodisi podatkovne) delčku p sosedni oz. relevantni. Na slikah 8.1, 8.2 in 8.3 so prikazani Hassejevi diagrami obeh relacij. S črnimi povezavami je podana relacija *biti manj specifičen*, s sivimi pa je med delčki informacij definirana relacija *sosebnosti* \mathcal{N} . Na množici časovnih trenutkov je relacija *sosebnosti* prazna, saj že relacija *biti manj specifičen* povezuje vse možne časovne trenutke (slika 6.3).



Slika 8.1: Hassejev diagram relacije *biti manj specifičen* (povezave črne barve) in relacije *sosebnosti* (povezave sive barve) med krajevnimi delčki informacij, ki se navezujejo na Slovenijo



Slika 8.2: Hassejev diagram relacije *biti manj specifičen* (povezave črne barve) in relacije sosednosti (povezave sive barve) med krajevnimi delčki informacij, ki se navezujejo na ostalo Evropo



Slika 8.3: Hassejev diagram relacije *biti manj specifičen* (povezave črne barve) in relacije sosednosti (povezave sive barve) med podatkovnimi delčki informacij

Relacija sosednosti je sicer pomembna predvsem pri podatkovnem rudarjenju prostorskih informacij [Ester-2001], kjer so lastnosti danega prostorskega objekta pogosto pod vplivom sosednjih objektov.

Na podlagi relacije sosednosti \mathcal{N} lahko definiramo modalni operator *sosed* N , kjer

- NA pomeni, da je A dosegljivo in relevantno, vendar sosedno.

V definicijo dialog mIPC Kripkejevih modelov, katerim bo zadoščala predstavitev znanja, ki bo zajemala tudi modalni operator N in relacijo \mathcal{N} , dodamo naslednje pravilo:

$$10^{**}. \quad w \Vdash NA \quad \text{čee} \quad \exists v \in W : (wNv \text{ in } v \Vdash A)$$

Za svetova w in v , nekončni izjavi A in B , ter kontekst Γ v tako razširjenih dialog mIPC Kripkejevih modelih velja tudi naslednje pravilo sklepanja:

B6. Konjunkcija, implikacija in sosed

$$\frac{\Gamma \vdash N(A \wedge (A \Rightarrow B))}{\Gamma \vdash NB}$$

Razširitev predstavitve znanja s pomočjo modalnega operatorja sosed je le ena od idej, ki se porodijo. Druga ideja bi bila spremeniti definicijo relacije dosegljivosti oz. relacije *biti manj specifičen* (npr. opustiti lastnost tranzitivnosti) in se tako izogniti najbolj izraziti omejitvi dialog mIPC Kripkejevih modelov.

8.2 Nadaljnje delo

Dvojezični sistem za podajanje informacij o vremenu (slika 1.2) [Žibert-2003], ki je predmet tega magistrskega dela, je trenutno še v fazi razvoja. Poleg vira vremenskih informacij in modula za vodenje dialoga, katerih razvoj je v tem delu podrobneje opisan, je veliko dela narejenega tudi na področju razpoznavanja slovenskega in hrvaškega govora, omejenega na tematsko področje vremenskih informacij [Martinčić-2003].

Modul za tvorjenje slovenskega jezika, ki je bil že uporabljen v eksperimentih Čarovnik iz Oza, je takorekoč dograjen. Dograjen je tudi modul za tvorjenje slovenskega govora [Vesnicer-2003], posebej primeren prav za komunikacijsko področje vremenskih informacij. Razvoj modulov za tvorjenje hrvaškega jezika in hrvaškega govora bo potekal po vodilih razvoja modulov za slovenski jezik oz. slovenski govor.

Trenutno razvijamo tudi modul za razumevanje naravnega slovenskega jezika, ki je omejen na komunikacijsko področje vremenskih informacij. Pri tem se zgledujemo po [Wutiw WATCHAI-2003], postopek pa združuje izločanje pomenskih podvzorcev, ki napolnijo predalčke, ter ravrščanje teh s pomočjo nevronske mreže. Učno množico predstavlja zbirka spontanih dialogov SSSQ in del zbirke SSSQ2 (poglavje 3), katerih vsebina zrcali komunikacijo človek – računalnik. Izhod iz modula za razumevanje naravnega slovenskega jezika bo poleg v določeni meri napolnjenih predalčkov še konverzacijska poteza. Razvoj modula za razumevanje hrvaškega jezika bo sledil vodilom razvoja modula za razumevanje slovenskega jezika.

Največ pozornosti smo v magistrskem delu posvetili vodenju dialoga in s tem povezani predstavitvi znanja sistema. Tako modul za vodenje dialoga kakor tudi predstavitev znanja sta od jezika neodvisna in zato predstavljata osrednji podsistem dvojezičnega sistema za podajanje informacij o vremenu. Nekatere ideje glede razširitve udejanjene predstavitve znanja (poglavje 6) smo predstavili v prejšnjem razdelku. Zanimivo pa bi bilo poiskati (če seveda obstaja) različico intuicionističnega modalnega izjavnega računa, ki bi bila za razred dialog mIPC Kripkejevih modelov (definicija 8) ali kakšne njegove razširitve polna.

V modulu za vodenje dialoga smo poleg že znanih konverzacijskih iger [Larsson-2000] definirali in udejanili dodatne igre in poteze, za katere trdimo, da povečajo prenosljivost modula na druge domene. Eden izmed ciljev je zato modul za vodenje dialoga dejansko prirediti za druge domene in na podlagi rezultatov takšnih prirejanj oblikovati množico konverzacijskih iger in potez, ki bi se izkazale za dovolj reprezentativne in bi prenosljivost sistema kar najbolj povečale.

Viri in literatura

- [Alechina-2001] N. Alechina, M. Mendler, V. de Paiva in E. Ritter. *Categorical and Kripke Semantics for Constructive Modal Logics*. V Proceedings Computer Science Logic, Pariz, Francija. 2001.
- [Allen-1995] J. F. Allen, L. K. Schubert, G. Ferguson, P. Heeman, C.-H. Hwang, T. Kato, M. Light, N. G. Martin, B. W. Miller, M. Poesio in D. R. Traum. *The TRAINS Project: A case study in building a conversational planning agent*. Journal of Experimental and Theoretical AI. Zv. 7, str. 7–48. 1995.
- [Allen-1997] J. F. Allen in M. Core. *Draft of DAMSL: Dialog Act Markup in Several Layers*. Multiparty Discourse Group at the Discourse Research Initiative Meeting, Schloss Dagstuhl, Nemčija. Dosegljivo na <http://www.cs.rochester.edu/research/cisd/resources/damsl/RevisedManual/RevisedManual.html>.
- [Appelt-1985] D. Appelt. *Planning English Referring Expressions*. Cambridge University Press, New York. 1985.
- [Barras-2001] C. Barras, E. Geoffrois, Z. Wu in M. Liberman. *Transcriber: use of a tool for assisting speech corpora production*. Speech Communication: Special issue on Speech Annotation and Corpus Tools. Zv. 33(1–2), str. 5–22. 2001.
- [Boyce-1996] A. L. Boyce in A. L. Gorin. *User Interface Issues for Natural Spoken Dialogue Systems*. V Proceedings of the 1996 International Symposium on Spoken Dialogue, Philadelphia, ZDA. str. 65–68. 1996.
- [Bratman-1988] M. Bratman, D. Israel in M. Pollack. *Plans and Resource-bounded Practical Reasoning*. Computational Intelligence. Zv. 4, str. 349–355. 1988.
- [Carletta-1996a] J. Carletta, A. Isard, S. Isard, J. Kowtko, G. Doherty-Sneddon in A. Anderson. *HCRC Dialogue Structure Coding Manual*. Research paper 82, Human Communication Research Centre, University of Edinburgh, Edinburgh, Velika Britanija. 1996.
- [Carletta-1996b] J. C. Carletta. *Assessing the Reliability of Subjective Codings*. Computational Linguistics. Zv. 22(2), str. 249–254. 1996.

- [Codd-1974] E. F. Codd. *Seven Steps to RENDEZVOUS with the Casual User*. V Data Base Management. Uredila J. Kimbie in K. Koffeman. North-Holland Publishers, Amsterdam. 1974.
- [Cohen-1990] P. Cohen in H. Levesque. *Rational Interaction as the Basis for Communication*. str. 221–255. V Intentions in Communication. Uredili P. Cohen, J. Morgan in M. Pollack. MIT Press, Cambridge. 1990.
- [Cohen-1995] P. R. Cohen in S. L. Oviatt. *The Role of Voice Input for Human-Machine Communication*. V Proceedings of the National Academy of Sciences, ZDA. Zv. 92(22), str. 9921–9927. 1995.
- [Cosky-1995] M. Cosky. *Talking to Machines Today and Tomorrow: Designing for the User*. AT&T Technical Journal. str. 81–90. 1995.
- [Dahlbäck-1993] N. Dahlbäck, A. Jönsson in L. Ahrenberg. *Wizard of Oz studies: why and how*. V Proceedings of the international workshop on Intelligent user interfaces, Orlando, ZDA. str. 193–200. 1993.
- [Danieli-1995] M. Danieli in E. Gerbino. *Metrics for Evaluating Dialogue Strategies in a Spoken Language System*. V Proceedings of the 1995 AAAI Spring Symposium on Empirical Methods in Discourse Interpretation and Generation, Stanford, ZDA. str. 34–39. 1995.
- [Danlos-1984] L. Danlos. *Conceptual and Linguistic Decisions in Generation*. V Proceedings of the 22nd Meeting of the Association for Computational Linguistics and the 10th International Conference on Computational Linguistics, Kyoto, Japonska. str. 501–504. 1984.
- [Elmasri-2004] R. Elmasri in S. B. Navathe. *Fundamentals of Database Systems, Fourth Edition*. Pearson Education, Inc., Boston. 2004.
- [Eskenazi-1999] M. Eskenazi, A. Rudnicky, K. Gregory, P. Constantinides, R. Brennan, C. Bennett in J. Allen. *Data Collection and Processing in the Carnegie Mellon Communicator*. V Proceedings of the 6th European Conference on Speech Communication and Technology, Budimpešta, Madžarska. Zv. 6, str. 2695–2698. 1999.
- [Ester-2001] M. Ester, H.-P. Kriegel in J. Sander. *Geographic Data Mining and Knowledge Discovery, Research Monographs in GIS*. str. 160–187. V Algorithms and Applications for Spatial Data Mining. Uredila H. J. Miller in J. Han. Taylor and Francis, London. 2001.
- [Fraser-1991] N. M. Fraser in G. N. Gilbert. *Simulating Speech Systems*. Computer Speech and Language. Zv. 5(1), str. 81–99. 1991.
- [Gardner-1999] C. Gardner-Bonneau. *Guidelines for Speech-enabled IVR Application Design*. str. 147–162. V Human Factors and Voice Interactive Systems. Uredil D. Gardner-Bonneau. Kluwer Academic Publishers, Boston. 1999.

- [Goddeau-1996] D. Goddeau, H. Meng, J. Polifroni, S. Seneff in S. Busayapongchai. *A Form-Based Dialogue Manager for Spoken Language Applications*. V Proceedings of the International Conference on Spoken Language Processing, Philadelphia, ZDA. str. 701–704. 1996.
- [Grice-1975] H. Grice. *Logic and Conversation*. V Syntax and Semantics, Zv. 3, Speech Acts. str. 41–58. Uredila P. Cole in J. Morgan. Academic Press, New York. 1975.
- [Gros-1997a] J. Gros. *Samodejno tvorjenje govora iz besedil*. Doktorska disertacija. Fakulteta za elektrotehniko, Univerza v Ljubljani. Ljubljana. 1997.
- [Gros-1997b] J. Gros, N. Pavešić in F. Mihelič. *Text-to-Speech synthesis: a complete system for the Slovenian language*. Journal of Computing and Information Technology (CIT). Zv. 5(1), str. 11–19. 1997.
- [Hajdinjak-2002a] M. Hajdinjak in F. Mihelič. *Podatkovna baza nastajajočega sistema za podajanje vremenskih napovedi*. V Zbornik 11. mednarodne Elektrotehniške in računalniške konference ERK 2002, Portorož, Slovenija. Zv. B, str. 299–302. 2002.
- [Hajdinjak-2002b] M. Hajdinjak in F. Mihelič. *Semantična analiza vremenskih napovedi*. V Zbornik B 5. mednarodne multikonference Informacijska družba IS'2002, Ljubljana, Slovenija. str. 10–13. 2002.
- [Hajdinjak-2003a] M. Hajdinjak in F. Mihelič. *The wizard of Oz system for weather information retrieval*. V Lecture notes in computer science, Lecture notes in artificial intelligence, 2807, Text, speech and dialogue : 6th International Conference, TSD 2003, České Budějovice, Češka. str. 400–405. Uredila V. Matoušek in P. Mautner. Springer, Berlin. 2003.
- [Hajdinjak-2003b] M. Hajdinjak in F. Mihelič. *Wizard of Oz experiments*. V Proceedings of the IEEE Region 8 EUROCON 2003 : computer as a tool, Ljubljana, Slovenija. Zv. 2, str. 112–116. 2003.
- [Hapeshi-1993] K. Hapeshi. *Design Guidelines for Using Speech in Interactive Multimedia Systems*. str. 177–188. V Interactive Speech Technology: Human Factors Issues in the Application of Speech Input/Output to Computers. Uredila C. Barber in J. Noyes. Taylor & Francis, London. 1993.
- [Hendrix-1978] G. Hendrix, E. Sacerdoti, D. Sagalowicz in J. Slocum. *Developing a Natural Language Interface to Complex Data*. ACM Transactions on Database Systems. Zv. 3(2), str. 105 – 147. 1978.
- [Hintikka-1975] J. Hintikka. *The Intentions of Intentionality and Other New Models for Modalities*. D. Reidel Publishing Company, Dordrecht. 1975.

- [Hoeven-1995] G. van der Hoeven, J. Andernach, S. van der Burgt, G. J. Kruijff, A. Nijholt, J. Schaake in F. de Jong. *SCHISMA: A natural language accessible theatre information and booking system*. V Proceedings of the 1st International Workshop on Applications of Natural Language to Data Bases, Versailles, Francija. str. 271–285. 1995.
- [Houghton-1987] G. Houghton in S. D. Isard. *Why to speak, what to say and how to say it: Modelling language production in discourse*. str. 249–267. V Modelling Cognition. Uredil P. Morris. John Wiley and Sons, London. 1987.
- [Huth-2000] M. Huth in M. Ryan. *Modal logics and agents*. str. 261–315. V Logic in Computer Science, Modelling and reasoning about systems. Uredila M. Huth in M. Ryan. Cambridge University Press, Cambridge. 2000.
- [Ipšić-1999] I. Ipšić, F. Mihelič, S. Dobrišek, J. Gros in N. Pavešić. *A Slovenian Spoken Dialog System for Air Flight Inquires*. V Proceedings of the 6th European Conference on Speech Communication and Technology, Budimpešta, Madžarska. str. 2659–2662. 1999.
- [Johnson-2002] R. A. Johnson in D. W. Wichern. *Applied multivariate statistical analysis*. Prentice-Hall, Upper Saddle River (NJ). 2002.
- [Jurafsky-1994] D. Jurafsky, C. Wooters, G. Tajchman, J. Segal, A. Stolcke, E. Fosler in N. Morgan. *The Berkeley Restaurant Project*. V Proceedings of the International Conference on Spoken Language Processing, Yokohama, Japonska. str. 2139–2142. 1994.
- [Kamm-1999] C. Kamm, M. Walker in D. Litman. *Evaluating Spoken Language Systems*. V Proceedings of the American Voice Input/Output Society, San Jose, ZDA. 1999.
- [Kaplan-1982] S. J. Kaplan. *Cooperative Responses from a Portable Natural Language Query System*. Artificial Intelligence. Zv. 19(2), str. 165–187. 1982.
- [Karat-1999] J. Karat, J. Lai, C. Danis in C. Wolf. *Speech User Interface Evolution*. str. 1–35. V Human Factors and Voice Interactive Systems. Uredil D. Gardner-Bonneau. Kluwer Academic Publishers, Boston. 1999.
- [Kim-1997] W. Kim in M.-W. Koo. *A Korean Speech Corpus for Train Ticket Reservation Aid system Based on Speech Recognition*. V Proceedings of the 5th European Conference on Speech Communication and Technology, Rodos, Grčija. Zv. 4, str. 1723–1726. 1997.
- [Kowtko-1992] J. Kowtko in S. Isard. *Conversational Games Within Dialogue*. Research paper 31, Human Communication Research Centre, University of Edinburgh, Edinburgh, Velika Britanija. 1992.

- [Krahmer-2001] E. J. Krahmer. *The Science and Art of Voice Interfaces*. Philips Research Report, Philips, Eindhoven, Nizozemska. 2001.
- [Larson-1992] J. Larson. *Interactive Software: Tools for Building Interactive User-interfaces*. Prentice Hall, New Jersey. 1992.
- [Larsson-2000] S. Larsson in D. Traum. *Information state and dialogue management in the TRINDI Dialogue Move Engine Toolkit*. Natural Language Engineering. Zv. 6, str. 323–340. 2000.
- [Lea-1994] W. Lea. *Developing Usable Voice Interfaces*. Journal of the American Voice Input/Output Society. Zv. 16, str. 1–48. 1994.
- [Leiser-1993] R. Leiser. *Driver-vehicle Interface: Dialogue Design for Voice Input*. str. 275–293. V *Driving Future Vehicles*. Uredila A. Parkes in S. Franzen. Taylor & Francis, Washington. 1993.
- [Levelt-1989] W. Levelt. *Speaking: From Intention to Articulation*. MIT Press, Cambridge. 1989.
- [Lewin-2000] I. Lewin. *A formal model of Conversational Game Theory*. V Göta-log: Proceedings of the 4th Workshop on the Semantics and Pragmatics of Dialogue, Göteborg, Švedska. 2000.
- [Litman-2002] D. J. Litman in P. Shimei. *Designing and Evaluating an Adaptive Spoken Dialogue System*. User Modeling and User-Adapted Interaction. Zv. 12, str. 111–137. 2002.
- [Marr-1976] D. Marr. *Early Processing of Visual Information*. Philosophical Transactions of the Royal Society. Zv. B275, str. 483–524. 1976.
- [Martinčić-2003] S. Martinčić-Ipšić, J. Žibert, I. Ipšić, F. Mihelič in N. Pavešić. *Bilingual speech recognition for a weather information retrieval dialogue system*. V *Lecture notes in computer science, Lecture notes in artificial intelligence, 2807, Text, speech and dialogue : 6th International Conference, TSD 2003, České Budějovice, Česka*. str. 380–387. Uredila V. Matoušek in P. Mautner. Springer, Berlin. 2003.
- [Montague-1960] R. Montague. *Logical Necessity, Physical Necessity, Ethics and Quantifiers*. Inguiry. Zv. 4, str. 259–269. 1960.
- [Nie-1989] J. Y. Nie. *An Information Retrieval model based on Modal Logic*. Information Processing & Management. Zv. 25(5), str. 477–491. 1989.
- [Nielsen-1993] J. Nielsen. *Usability Engineering*. Morgan Kaufmann, San Diego. 1993.
- [Norman-1988] D. Norman. *The Psychology of Everyday Things*. Basic Books, New York. 1988.

- [Pepelnjak-1996a] K. Pepelnjak. *Pomenska analiza stavkov v sistemu za razumevanje tekočega govora*. Magistrsko delo. Fakulteta za elektrotehniko, Univerza v Ljubljani. Ljubljana. 1996.
- [Pepelnjak-1996b] K. Pepelnjak, F. Mihelič in N. Pavešič. *Semantic decomposition of sentences in the system supporting flight services*. Journal of Computing and Information Technology. Zv. 4(1), str. 17–24. 1996.
- [Perry-www] J. Perry. *Semantics, Possible-Worlds*. <http://www-csli.stanford.edu/~john/PHILPAPERS/posswld.pdf>.
- [Pirker-1999] H. Pirker, G. Loderer in H. Trost. *Thus Spoke the User to the Wizard*. V Proceedings of the 6th European Conference on Speech Communication and Technology, Budimpešta, Madžarska. Zv. 3, str. 1171–1174. 1999.
- [Poesio-1997] M. Poesio in D. R. Traum. *Conversational actions and discourse situations*. Computational Intelligence. Zv. 13(3), str. 309–349. 1997.
- [Power-1979] R. Power. *The organisation of purposeful dialogues*. Linguistics. Zv. 17, str. 107–152. 1979.
- [Price-1992] P. Price, L. Hirschman, E. Shriberg in E. Wade. *Subject-based Evaluation Measures for Interactive Spoken Language Systems*. V Proceedings of the DARPA Speech and NL Workshop, Harriman, ZDA. str. 34–39. 1992.
- [Reiter-1994] E. Reiter. *Has a Consensus NL Generation Architecture Appeared, and is it Psycholinguistically Plausible?*. V Proceedings of the 7th International Workshop on Natural Language Generation, Kennebunkport, ZDA. str. 163–170. 1994.
- [Resnik-1989] P. Resnik. *Access to Multiple Underlying Systems in JANUS*. BBN Report 7142, Bolt Beranek and Newman Inc., Cambridge, Massachusetts. 1989.
- [Sadek-1994] D. Sadek. *Toward a Theory of Belief Reconstruction: Application to Communication*. Speech Communication. Zv. 15, str. 251–263. 1994.
- [Searle-1969] J. Searle. *Speech Acts*. Cambridge University Press, Cambridge. 1969.
- [Sebastiani-1998] F. Sebastiani. *On the role of logic in information retrieval*. Information Processing & Management. Zv. 34(1), str. 1–18. 1998.
- [Seneff-1992] S. Seneff. *TINA: A Natural Languages System for Spoken Language Applications*. Computational Linguistics. Zv. 18(1), str. 61–86. 1992.

- [Shneiderman-1986] B. Shneiderman. *Designing the User Interface: Strategies for Effective Human-Computer Interaction, 3rd Edition*. Addison Wesley, Massachusetts. 1986.
- [Shriberg-1992] E. Shriberg, E. Wade in P. Price. *Human-Machine Problem Solving Using Spoken Language Systems (SLS): Factors Affecting Performance and User Satisfaction*. V Proceedings of the DARPA Speech and NL Workshop, Harriman, ZDA. str. 49–54. 1992.
- [Simpson-1994] A. K. Simpson. *The Proof Theory and Semantics of Intuitionistic Modal Logic*. Doktorska disertacija. Oddelek za računalništvo, Univerza v Edinburghu. Edinburgh. 1994.
- [Smith-1997] R. W. Smith in S. A. Gordon. *Effects of Variable Initiative on Linguistic Behavior in Human-Computer Spoken Natural Language Dialogue*. Computational Linguistics. Zv. 23(1), str. 141–168. 1997.
- [Smith-1998] I. Smith, P. Cohen, J. Bradshaw, M. Greaves in H. Holmback. *Designing conversation policies using joint intention theory*. V Proceedings of the 3rd International Conference on Multi Agent Systems, Pariz, Francija. str. 269–276. 1998.
- [Stallard-2000] D. Stallard. *Talk'n'Travel: A Conversational System for Air Travel Planning*. V Proceedings of the Association for Computational Linguistics 6th Applied Natural Language Processing Conference, Seattle, ZDA. str. 68–75. 2000.
- [Stolcke-2000] A. Stolcke, N. Coccaro, R. Bates, P. Taylor, C. Van Ess-Dykema, K. Ries, E. Shriberg, D. Jurafsky, R. Martin in M. Meteer. *Dialogue Act Modelling for Automatic Tagging and Recognition of Conversational Speech*. Computational Linguistics. Zv. 26(3), str. 339–373. 2000.
- [Sturm-1999] J. Sturm, E. den Os in L. Boves. *Dialogue management in the Dutch ARISE train timetable information system*. V Proceedings of the 6th European Conference on Speech Communication and Technology, Budimpešta, Madžarska. str. 1419–1422. 1999.
- [Swerts-2000] M. Swerts in E. Krahmer. *On the Use of Prosody for On-line Evaluation of Spoken Dialogue Systems*. V Proceedings of the 2nd International Conference on Language Resources and Evaluation, Atene, Grčija. Zv. 3, str. 1571-1578. 2000.
- [Thompson-1983] B. H. Thompson in F. B. Thompson. *Introducing ASK, A Simple Knowledgeable system*. V Proceedings of the 1st Conference on Applied Natural Language Processing, Santa Monica, ZDA. str. 17–24. 1983.
- [Troelstra-www] A. S. Troelstra. *History of constructivism in the 20th century*. Dosegljivo na <http://citeseer.nj.nec.com/96765.html>.

- [Turing-1950] A. M. Turing. *Computing machinery and intelligence*. Mind. Zv. 59, str. 433 – 460. 1950.
- [Vesnicer-2003] B. Vesnicer. *Umetno tvorjenje govora z uporabo prikritih Markovovih modelov*. Magistrsko delo. Fakulteta za elektrotehniko, Univerza v Ljubljani. Ljubljana. 2003.
- [Walker-1997] M. A. Walker, D. Litman, C. A. Kamm in A. Abella. *PARADISE: A General Framework for Evaluating Spoken Dialogue Agents*. V Proceedings of the 35th Annual Meeting of the Association of Computational Linguistics, Madrid, Španija. str. 271–280. 1997.
- [Walker-1998] M. A. Walker, D. J. Litman, C. A. Kamm in A. Abella. *Evaluating Spoken Dialogue Agents with PARADISE: Two Case Studies*. Computer Speech and Language. Zv. 12(3), str. 317–347. 1998.
- [Walker-2000] M. A. Walker. *An Application of Reinforcement Learning to Dialogue Strategy Selection in a Spoken Dialogue System for Email*. Journal of Artificial Intelligence Research. Zv. 12, str. 387–416. 2000.
- [Warren-1982] D. Warren in F. Pereira. *An Efficient Easily Adaptable System for Interpreting Natural Language Queries*. Computational Linguistics. Zv. 8(3–4), str. 110–122. 1982.
- [Weizenbaum-1966] J. Weizenbaum. *ELIZA*. Communications of the ACM. Zv. 9, str. 36–45. 1950.
- [Whittaker-1989] S. Whittaker in P. Stenton. *User Studies and the Design of Natural Language Systems*. V Proceedings of the 4th Conference of the European Chapter of the Association for Computational Linguistics, Manchester, Anglija. str. 116–123. 1989.
- [Woods-1972] W. A. Woods, R. M. Kaplan in B. N. Webber. *The Lunar sciences Natural Language Information System: Final Report*. BBN Report 2378, Bolt Beranek and Newman Inc., Cambridge, Massachusetts. 1972.
- [Wutiwiewatchai-2003] C. Wutiwiewatchai in S. Furui. *Combination of Finite State Automata and Neural Network for Spoken Language Understanding*. V Proceedings of the 8th European Conference on Speech Communication and Technology, Ženeva, Švica. str. 2761–2764. 2003.
- [Xu-2000] W. Xu in A. Rudnicky. *Task-based Dialog Management Using an Agenda*. V Proceedings of the ANLP/NAACL Workshop on Conversational Systems, Seattle, ZDA. str. 42–47. 2000.
- [Zoltan-1991] E. Zoltan-Ford. *How to Get People to Say and Type What Computers Can Understand*. Journal of Man-Machine Studies. Zv. 34, str. 527–547. 1991.

- [Zue-2000] V. Zue, S. Seneff, J. Glass, J. Polifroni, C. Pao, T. J. Hazen in L. Hetherington. *JUPITER: A Telephone Based Conversational Interface for Weather Information*. IEEE Transactions on Speech and Audio Processing. Zv. 8(1), str. 85–96. 2000.
- [Žibert-2003] J. Žibert, S. Martinčić-Ipšić, M. Hajdinjak, I. Ipšić in F. Mihelič. *Development of a Bilingual Spoken Dialog System for Weather Information Retrieval*. V Proceedings of the 8th European Conference on Speech Communication and Technology, Ženeva, Švica. str. 1917–1920. 2003.

Terminološki slovar

V magistrskem delu smo uporabili precej novih terminov. Za boljše razumevanje in lažji pregled podajmo slovar slovenskih in ustreznih angleških izrazov.

atomarna izjava	atomic statement
cene dialoga	dialogue costs
delež izbire relevantnih podatkov	relevant-data ratio
delež negativnih odgovorov	no-data ratio
delež neinicijativnih potez	non-initiating ratio
delež nepodanih informacij	non-provided-information ratio
delež neprimernih iniciativ	unsuitable-initiative ratio
delež nudenja pomoči	help-messages ratio
delež preverjanj	check ratio
delež zavrnitev	rejection ratio
dialog mIPC Kripkejev model	dialogue mIPC Kripke model
dogovor	transaction
eksperiment Čarovnik iz Oza	Wizard-of-Oz experiment
enosmerna cevovodna struktura	one-way pipeline architecture
funkcija učinkovitosti sistema	performance function
intuicionistična logika	intuitionistic logic
izpolnitev naloge	task completion
končna izjava	final statement
konverzacijske igre	conversational games
konverzacijske poteze	conversational moves
konverzacijska teorija iger	conversational game theory
Kripkejev model	Kripke model
Kripkejeva semantika možnih svetov	Kripke possible-worlds semantics

modalna logika	modal logic
modalni operator	modality
modul za vodenje dialoga	dialogue manager
mogoče velja	possibly true
multipla linearna regresija	multivariate linear regression
nujno velja	necessarily true
objektivni parametri dialoga	objective dialogue metrics
odzivne poteze	response moves
ogrodje PARADISE	PARADISE framework
omejitev	limitation
operator diamant	diamond operator
operator sosed	neighbor operator
operator spodnji diamant	diamond down operator
operator škatla	box operator
operator zgornji diamant	diamond up operator
parametri kvalitete	qualitative measures
parametri učinkovitosti	efficiency measures
parametri uspešnosti naloge	task success measure
podmodul za določanje vsebine	content determination submodule
podmodul za oblikovanje stavkov	sentence planning submodule
podmodul za slovnične popravke	surface generation submodule
pojmovni model	conceptual model
poteze pripravljenosti	ready moves
povprečni čas čakanja na odziv	mean system response time
povprečni čas dialoga	mean elapsed time
princip sodelovanja	cooperativity principle
rekurzivna mreža prehodov	recursive transition network
relacija biti manj specifičen	being-less-specific relation
relacija dosegljivosti	accessibility relation
relacija sosednosti	neighborhood relation

razumevanje naravnega jezika	natural language understanding
razumska enota	rational unit
semantični okvir	semantic frame
sistem za govorno upravljanje naprav	command-and-control voice application
sistem za interaktivne govorne odzive	interactive voice response system
sistem za podajanje informacij	information-providing dialogue system
sistem za reševanje problemov	problem-solving dialogue system
sodelujoč odgovor	cooperative answer
subjektivni parametri dialoga	subjective dialogue metrics
tehnika končnih avtomatov	finite-state technique
tehnika množice domen	sets-of-contexts technique
tehnika polnjenja predalčkov	slot-filling technique
tehnika razumskih agentov	rational-conversational-agents technique
tehnika usmerjanja k cilju	goal-directed technique
teorija govornih dejanj	speech act theory
tvorjenje naravnega jezika	natural language generation
vodilna atomarna izjava	main atomic statement
začetne poteze	initiating moves
zadovoljstvo uporabnika	user satisfaction

Izjava

Podpisana Melita Hajdinjak izjavljam, da sem avtorica magistrskega dela z naslovom *Vodenje dialoga med človekom in računalnikom v naravnem jeziku*.

Melita Hajdinjak