

Inteligentni sistemi v avtomatiki

VS-3-AV 2022/2023

Govorna komunikacija "človek - stroj"

Simon Dobrišek

POVZETEK

- » Uvod
- » Gradniki sistema za govorno komunikacijo "človek-stroj"
- » Dialog "človek stroj"
- » Načrtovanje sistema za dialog
- » Preizkušanje sistema za dialog
- » Programska orodja za razvoj sistemov za dialog



Uvod

- » Inteligentni sistemi naj bi se ne le zavedali prisotnosti ljudi v svoji bližini, temveč naj bi njimi tudi komunicirali na njim naraven način.
- » Govor je človeku najbolj naraven način komuniciranja.
- » Eden od ciljev razvoja inteligentnih sistemov je tako tudi podpora človeku naravne komunikacije s strojem.
- » Govorna komunikacija človeka s strojem v veliko primerih poveča učinkovitost in hitrost komunikacije.
- » Pogosto je človeku tudi najbolj prikladen način, saj med komunikacijo ni potrebno ničesar opazovati ali se nečesa dotikati.



Sistem za govorno komunikacijo "človek – stroj"

- » Sistem za govorno komunikacijo "človek stroj" je računalniški sistem, ki uporabniku omogoča, da z govorom dostopa do določenih računalniških aplikacij ali da upravlja/vodi nek stroj.
- » Sistemi za govorno komunikacijo "človek stroj" se gradijo kot računalniški sistemi za dialog, ki se razvijajo širše tudi za negovorno komunikacijo "človek – stroj".
- » Pri interakciji med uporabnikom in strojem se tako poleg govora
 - > za vhod predvideva dotikanje, tipkanje, pisanje, izvajanje gest, menijsko izbiranje ipd,
 - za izhod pa se uporablja zvočne signale, grafične prikazovalnike, animirane grafične in mehanske like, razna svetila ipd.



Sistemi za dialog

- » Sisteme za dialog poskušamo razviti kot inteligentne avtonomne agente, ki izvajajo naloge na področjih, kot so:
 - > podajanje informacij (potovalnih, turističnih, komercialnih, ...)
 - > svetovanje (nakupovalno, potovalno, zdravstveno, ...)
 - > izvajanje storitev (servisnih, prodajnih, finančnih, upravnih, ...)
 - > poučevanje (inštrukcije, tutorstvo, preverjanje znanja, ...)
 - > sodelovanje (reševanje problemov, proizvodni procesi, ...)
 - > gospodinjenje (čiščenje, pospravljanje, kuhanje, pranje, ...)
 - > tekmovanje (igranje raznih iger, šport, ...)
 - > varovanje (nadzor dostopa, varovanje otrok, starejših ...)
 - > konverzacija (spletni klepetalniki, ...)

> ...

Stanje razvoja na področju

- » Predvsem v tujini že obstaja nekaj nameščenih komercialnih sistemov, kot so:
 - > klicni centri za podajanje informacij, preusmerjanje ipd.,
 - > govorno menijsko izbiranje storitev (rezervacije ipd.),
 - > uporabniški vmesniki za dostop do informacij (Apple Siri ipd.),
 - > podpora vozniku avtomobila, ...
- » Pojavljajo se tudi do neke mere že uporabni sistemi:
 - > za vodenje in ukazovanje strojem, napravam, avtomatom,
 - > sistemi za učenje jezika,
 - > govorno vnašanje informacij v podatkovne zbirke, ...



Stanje razvoja na področju

- » Izvaja se tudi veliko razvojnih in raziskovalnih projektov, ki vključujejo govorno komunikacijo "človek-stroj".
- » Primeri takšnih raziskovalnih področij so:
 - > sodelovanje ljudi in robotov,
 - > prilagodljivi sistemi za podporo uporabnikom,
 - > ambientalni inteligentni sistemi,
 - > virtualni asistenti,
 - > ne-kooperativni agenti, ki igrajo določeno vlogo, ...



https://www.youtube.com/watch?v=ZmlrkQBjVok



Zgledi sistemov za dialog

- » JUPITER Govorni sistem za podajanje informacij o vremenu
- » TRAINS Govorni sistem za načrtovanje potovanj z železnico
- » Klepec Virtualni klepetalnik https://www.amebis.si/virtualni-asistent
- » SecondEgo Virtualni asistent http://secondego.com/about-us/
- » ALICE Virtualni klepetalnik http://alice.pandorabots.com



GRADNIKI SISTEMA ZA GOVORNO KOMUNIKACIJO "ČLOVEK-STROJ"

- » Glavni gradniki sistema za govorno komunikacijo "človek-stroj" so
 - > razpoznavalnik govora,
 - > sistem za dialog in
 - > sintetizator govora.
- » Sistem za dialog pa sestavljajo naslednje enote:
 - > pomenski analizator,
 - > upravljavec dialoga,
 - > generator sporočil.





DIALOG "ČLOVEK – STROJ"

- » Dialog je vzajemni pogovor dveh ali več oseb oziroma agentov.
- » Obstaja več vrst dialogov in navadno ločimo:
 - > Spontani dialog (pogovor ob kosilu ipd.)
 - > Poučevalni dialog (učenje učenca, inštrukcije, ...)
 - > Usmerjeni dialog s posebno nalogo (rezervacija sedeža ipd.)
- » Dialog "človek-stroj" je navadno usmerjen dialog.
- » Spontani dialog se poskuša simulirati v demonstracijskih sistemih, ki se navadno razvijajo z namenom, da bi prestali popularni t.i. Turingov preizkus umetne inteligence.

Turingov preizkus umetne inteligence

- » Izvedba preizkusa
 - Pogovor človeka s prikritim strojem na eni strani ter človekom na drugi strani.
 - > Preizkus je uspešen, če ne more ugotoviti, kateri od sogovornikov ni človek.
- » Zadržki pri tem preizkusu
 - > Stroj, ki simulira človeške pogovorne navade ni nujno inteligenten.
 - Stroj bi lahko bil nadvse inteligenten, pa vendar ni zmožen debatirati s človekom.
 - > Veliko manj izobraženim ljudem bi spodletelo na tem preizkusu.
- » Nedavno je bila predlagana izboljšava preizkusa z uporabo t.i. Winogradove sheme oblikovanja vprašanj, na primer:

```
"Miha ni mogel dvigniti Janeza, ker je bil pretežek. Kdo je pretežek?"
"Pokal ne gre v rjav kovček, ker je prevelik. Kdo je prevelik?"
"Miha se prepira z Janezom, ker je zelo slabe volje. Kdo je slabe volje?"
```



Modeliranje med-agentne komunikacije

» Pri modeliranju komunikacije med človekom in strojem je sorodno modeliranju komunikacija med agenti v t.i. več-agentnih sistemih.



- » Modeliranje komunikacije med agenti se naslanja na Teorijo govornih dejanj (angl. Speech act theory).
- » Ta teorija obravnava jezik kot sredstvo, s katerim ljudje dosegajo svoje cilje in namere.
- » Izjave obravnava kot svojevrstna fizična dejanja, za katere se zdi, da spremenijo stanje agenta in tudi stanje sveta (npr. napoved vojne).
- » V okviru organizacije FIPA (Foundation for Intelligent Physical Agents) so bili predlagani standardi za med-agentne jezike.

Govorna dejanja

- » Predpostavlja se več vrst osnovnih govornih dejanj, kot so:
 - informacije: informiranje o stanju, ("Zunaj dežuje."),
 - ukazi: poskus pripraviti poslušalca izvesti nekaj, ("Pripravi čaj!"),
 - obveze: obvezovanje govorca k nekemu dejanju, ("Obljubljam, da ..."),
 - izrazi: izražanja govorčevega duševnega stanja ("Hvala!"),
 - deklaracije: deklariranje novega stanja okolja ("Napovedujem ti vojno!")
- » Za govorna dejanja se predpostavlja, da ima dve glavni komponenti performativni glagol (zahtevati, informirati, ...) in vsebino izjave.

Upoštevani performativni glagoli po FIPA standardu

performative	passing	requesting	negotiation	performing	error		
	info	info		actions	handling		
accept-proposal			Х				
agree				X			
cancel		X		X			
cfp			X				
confirm	X						
disconfirm	X						
failure					X		
inform	X					7gled	govornega dejanja
inform-if	X				(inform	_Bica	- Sovornega dejanja
inform-ref	X				:sender	r	agent1
not-understood					:recei	ver	agent5
propose			X		:conter	nt	(price good200 150)
query-if		X			:langua	age	sl
query-ref		X			:ontolo	ogy	hpl-auction
refuse				X)		
reject-proposal			X				,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,
request				X			
request-when				X			
request-whenever				X			
subscribe		X					

Dejanja v usmerjenem dialogu

- » V usmerjenih sistemih za dialog predstavimo dialog kot zaporedje specifičnih sporočil (govornih dejanj) med sistemom in uporabnikom.
- » Posamezna sporočila sistema za dialog ali uporabnika običajno imenujemo dejanja v dialogu (angl. dialogue acts (DA)).
- » Razlikujemo med:
 - > dejanji sistema (angl. system dialogue acts (SDA)) in
 - > dejanji uporabnika (angl. user dialogue acts (UDA)).
- » Dejanja uporabnika v dialogu (UDA) delimo naprej na informativna in na kontrolna dejanja:
 - informativna dejanja uporabnika (IUDA) vsebujejo konkretne podatke vezane na aplikacijo,
 - kontrolna dejanja uporabnika (CUDA) pa so splošna sporočila, katerih namen je krmiljenje dialoga.



Primeri zapisa dejanj v usmerjenem dialogu

CUDA	Razlaga
Accept	Potrditev
Reject	Zavrnitev
Open	Uvodni pozdrav
Close	Slovo
Unknown	"Ne vem."
Repeat	Prošnja za ponovitev
Please	Prošnja
Thanks	Zahvala
Fail	Neuspešno razpoznane besede

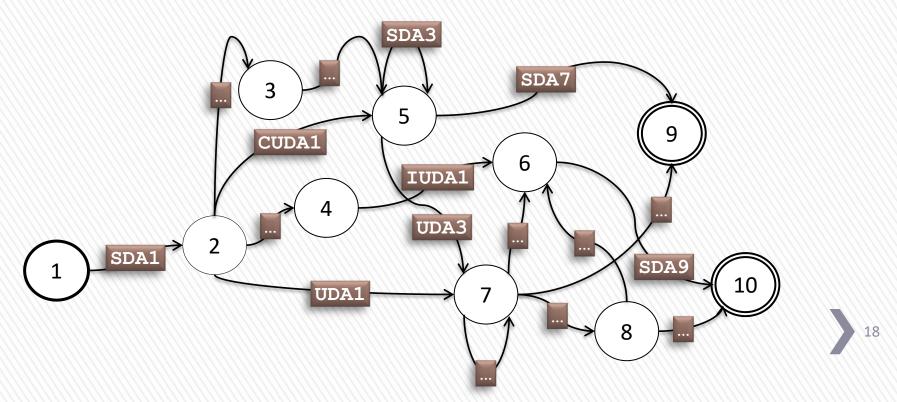
IUDA	Razlaga
Request	Zahteva po podatku
Provide	Posredovanje podatkov

SDA	Razlaga
Open	Predstavitev
Request	Zahteva po podatku
Confirm	Potrditveno vprašanje o podatku
Inform	Posredovanje podatkov iz zbirke
Repeat	Prošnja za ponovitev
Explain	Prošnja za dodatno razlago
Further_info	Prošnja za dodatno informacijo
Close	Slovo

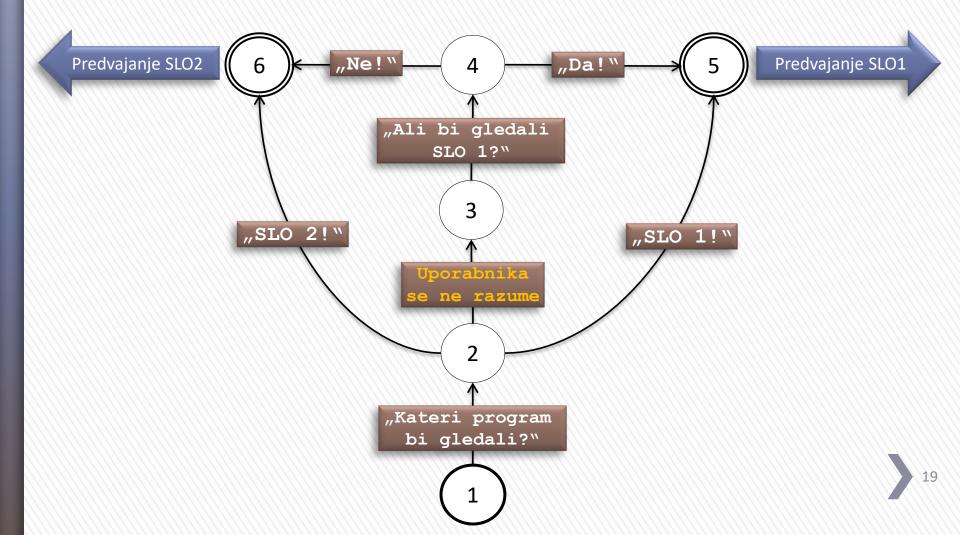


Usmerjen dialog

- » Usmerjen dialog navadno modeliramo kot stroj stanj (končni avtomat), ki ima neko začetno stanje in eno ali več končnih stanj, k doseganju katerih teži upravljavec dialoga.
- » Dejanja dialoga SDA, UDA, CUDA, in IUDA spremenijo/preklopijo stanje avtomata (dialoga).



Preprost primer končnega avtomata usmerjenega dialoga



Načini vodenja usmerjenega dialoga

» Uporablja se več načinov vodenja usmerjenega dialoga, kot so enostavne menijske izbire, potrjevanje posameznih podatkov, potrjevanje izvedenih poizvedb, iniciativno vodenje, iniciativno vodenje in potrjevanje ipd.

Menijske izbire

```
"Na voljo so prispevki med
naslednjimi tematikami:"
"Aktualno?"
"Naprej!"
"Gospodarstvo?"
"Preberi!"
"..."
```

Iniciativno vodenje

```
"Rad bi letel v London."
"Kdaj želite leteti?"
"..."
```

Potrjevanje posameznih podatkov

```
"Rad bi letel v London."
"Želite leteli iz Ljubljane?"
"Da!"
"Želite leteli v London?"
"Da!"
"Kdaj želite leteti?"
```

Iniciativno vodenje s potrjevanjem

```
"Rad bi letel v London."
"Kdaj želite leteti iz Ljubljane v London?"
"…"
```

Potrjevanje izpeljanih poizvedb

```
"Rad bi letel v London."
"Razumem, da želite leteli iz
Ljubljane v London."
"Kdaj želite leteti?"
"..."
```

20

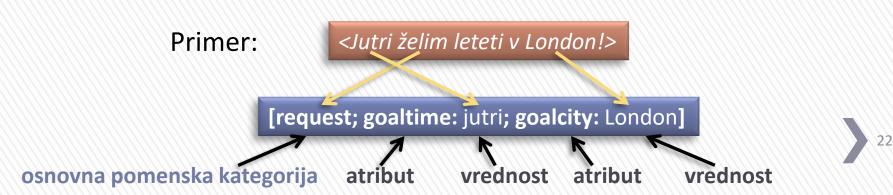
Upravljavec dialoga

- » Jedro sistema za dialog je upravljavec dialoga (angl. dialogue manager), ki mora delovati smiselno in "inteligentno".
- » Upravljavec dialoga "upravlja" sporazumevanje z uporabnikom sistema za komunikacijo "človek-stroj".
- » Njegove naloge so:
 - > sprejemanje pomensko obdelanih izjav uporabnika,
 - > ugotavljanje in spreminjanja stanja v dialogu,
 - > generiranje izjav sistema,
 - > napovedovanje naslednjih uporabnikovih izjav .



Enota za pomensko analizo sporočil

- » Pomensko obdelana sporočila iz vhodnega besedila oblikuje enota za pomensko analizo sporočil (angl. natural language understanding unit).
- » V usmerjenih sistemih za dialog pomenska analiza sporočil pomeni iskanje besede ali fraz v vhodnem (razpoznanem) besedilu, ki predstavljajo eno izmed podprtih pomenskih kategorij.
- » Pomensko obdelan stavek predstavimo s ključnimi besedami pomenskimi kategorijami, ki predstavljajo govorno dejanje, in vsebino sporočila, navadno oblikovano kot seznam atributov (prilastkov) in vrednosti atributov.
- » Število pomenskih kategorij je običajno omejeno na nekaj deset.



Primeri pomenskih kategorij

» Primeri pomenskih kategorij, ki označujejo govorno dejanje oziroma atribute vsebine in so se uporabljale pri sistemu za podajanju informacij o letalskem prometu so naslednji:

Govorna dejanja

- > pozdrav
- > poslovitev
- > zahteva
- > potrditev
- > zanikanje

Atributi (prilastki)

- > časovna kategorije (absolutni čas, relativni čas, čas odhoda, čas prihoda, trajanje, ...)
- > letalski prevoznik
- > mesto odhoda
- > mesto prihoda
- > ime letališča
- > ...



Enota za generiranje sporočil

- » Izhodna sporočila oz. izhodno besedilo oblikuje enota za generiranje sporočil (angl. natural language generation unit).
- » Enota za generiranje sporočil preslika sistemska dejanja v dialogu v besedilna sporočila v naravnem (npr. slovenskem) jeziku.
- » Število različnih sporočil je omejeno in navadno temelji na uporabi stavčnih predlog (šablon).

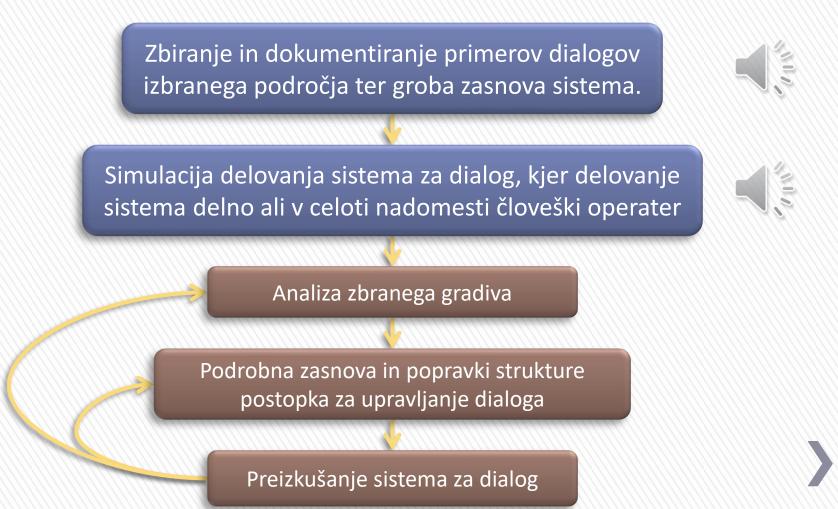
Primeri iz sistema za podajanje letalskih informacij:

- 1. Lepo pozdravljeni, poklicali ste samodejni sistem za podajanje informacij o letalskem prometu. Kaj vas zanima?
- 2. Na katero letališče želite prispeti?
- 3. Želite prispeti v <*cilj poleta*>?
- 4. Prvi polet v <*cilj poleta*> <*datum*> je ob <*ura*> iz letališča <*ime letališča*>.
- 5. Naj ponovim?
- 6. Želite še kakšen podatek?
- 7. Žal vas ne razumem, prosim pokličite na <telefon v sili>.
- 8. Lep pozdrav!



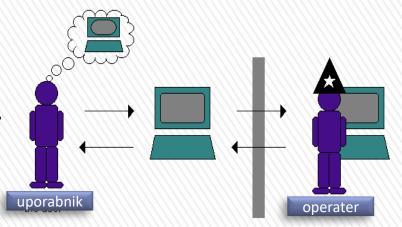
NAČRTOVANJE SISTEMA ZA DIALOG

» Načrtovanje sistema za dialog poteka v več fazah.



Simulacija "Čarovnik iz Oza"

- » Simulacijo delovanja sistema za dialog izvedemo po zgledu "Čarovnika iz Oza".
- » Uporabniki so prepričani, da se pogovarjajo z računalniškim sistemom.



- » Sistem je delno pod kontrolo človeškega operaterja "čarovnika".
- » Dialogi izražajo jezik komunikacij "človek-stroj".
- » Pri uporabnikih se navadno opazi prilagajanje na jezikovne sposobnosti sistema; uporabo bolj artikuliranega, počasnejšega in/ali glasnejšega govora; smeh, šale in duhovite opazke.
- » Človeški operater izbira in prehaja med vnaprej izbranimi odzivi sistema ter ročno spreminja parametrov simulacijskega modela.



Zgled simulacije sistema za dialog

Vremenske informacije odziv sistema osnovni osveži Da, v Karavankah jutri podnevi - povečini jasno. Imate še kakšno vprašanje? naravno besedilo ODATKI meni Del dneva podnevi Dan Slovenija jutri časovni interval + krajevna informacija Evropa Deli Pokrajine Karavanke · lege ... Slovenije Lavinske postaje slikovna polja (image fields) -Požarna ogroženost Stanje voda sončno, oblačno, dežuje, sneži APOVEDI Slovensko Primorje Slovenija Juliiske Alpe vreme vrsta podatka -Vodostaji Letalstvo vreme, temperatura, veter, sneg temperatura

Analiza gradiva, zbranega s simulacijo dialoga

- » Z analizo zbranega gradiva s simulacijo dialoga določimo:
 - > omejitev področja uporabe in izberemo cilje sistema,
 - > slovar in leksikon uporabljenega besedišča,
 - > sestavine pomenske analize uporabnikovih sporočil,
 - > uporabnikova in sistemska dejanja v dialogu,
 - > Seznam šablon in pravil za tvorjenje sporočil sistema.
- » Zasnovo strukture upravljanja dialoga navadno izvedemo s končnim avtomatom, obstajajo pa tudi drugi pristopi, kot je "polnjenje predalčkov" za vpogled v podatkovno zbirko ipd.
- » Preizkušanje sistema za dialog izvedemo z beleženjem dejanskih dialogov in oceno uspešnosti delovanja.



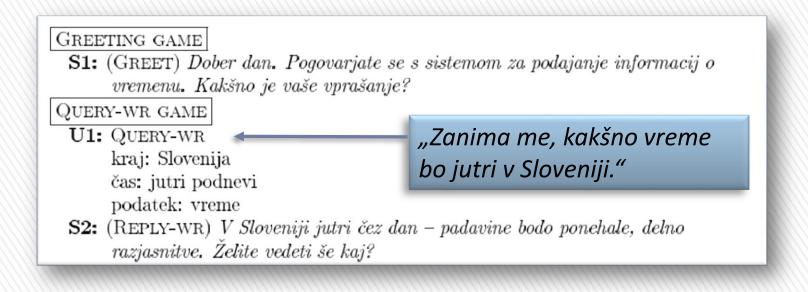
PREIZKUŠANJE SISTEMA ZA DIALOG

- » Pri vrednotenju uspešnosti sistema za dialog navadno beležimo in ovrednotimo naslednje parametre:
 - kapa κ koeficient,
 (ujemanje med vsebino zahtev uporabnika in "dojemanjem" sistema)
 - > delež izpolnjenih nalog,
 - > povprečni čas dialoga,
 - > povprečni čas čakanja na odziv,
 - > delež nepodanih informacij,
 - > delež zavrnitev,
 - > delež nudenja pomoči,
 - > delež negativnih odgovorov,
 - > delež izbire relevantnih podatkov,
 - > delež preverjanj,
 - > zadovoljstvo uporabnika,
 - > ...



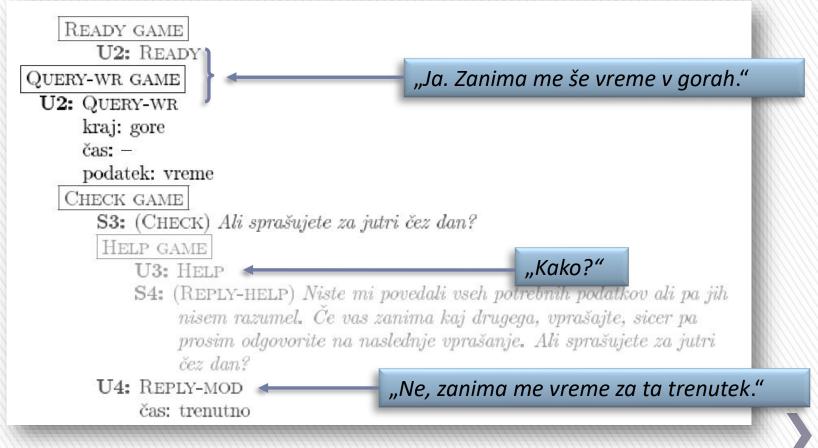
Primer dialoga v sistemu LUKS za podajanje informacij o vremenu (1/4)

» Vhod v enoto za upravljanje dialoga je pomenska predstavitev uporabnikovega sporočila (izjave).





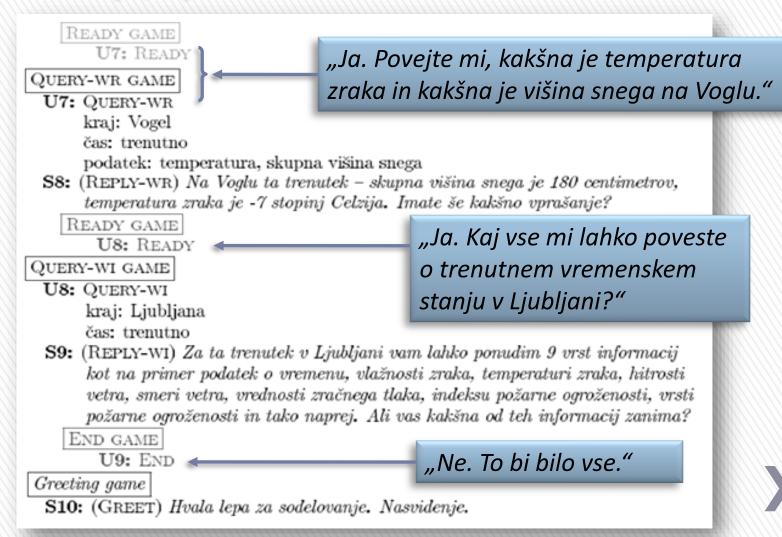
Primer dialoga v sistemu za podajanje informacij o vremenu (2/4)



Primer dialoga v sistemu za podajanje informacij o vremenu (3/4)

```
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
                                   "Povejte mi prosim za Vogel."
      U5: Query-wr. 		■
           kraj: Vogel
           čas: -
           podatek: -
       Check game
          S6: (Check) Ali sprašujete za ta trenutek o vremenu?
                                                                      "Ja."
          U6: Reply-y
      S7: (Reply-wr) Na Voglu ta trenutek – megla. Vam lahko še kako
           pomagam?
```

Primer dialoga v sistemu za podajanje informacij o vremenu (4/4)



PROGRAMSKA ORODJA ZA RAZVOJ SISTEMOV ZA DIALOG

- » Razvitih je bilo več programskih orodij in jezikov za razvoj govornih sistemov za dialog.
- » Svetovni spletni konzorcij (World Wide Web Consortium W3C) je predlagal več standardov na osnovi razširljivega označevalnega jezika (Extended Markup Language - XML), ki so omogočili razvoj programskih orodij za razvoj govornih sistemov za dialog.

http://www.w3.org/Voice

» Eden najbolj pomembnih standardov na tem področju je označevalni jezik VoiceXML in programski tolmači za ta jezik.

https://www.w3.org/TR/voicexml30/

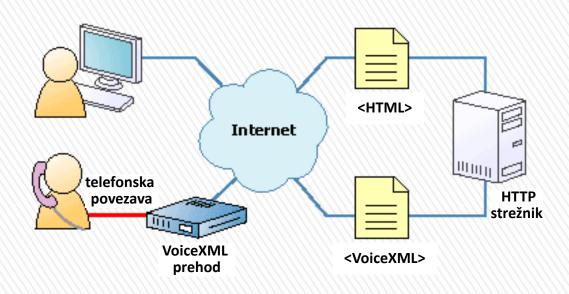


VoiceXML

- » VoiceXML je deklarativni programski jezik za udejanjanje govornih sistemov za dialog, ki vključuje razpoznavalnike in sintetizatorje govora, predvajanje audio posnetkov, razpoznavanje DTMF tonov, telefonijo in dialoge z mešano iniciativo.
- » Poleg osnovnega jezika VoiceXML je predlaganih še več pomožnih jezikov, kot so:
 - > specifikacija za zapis govorne slovnice (angl. Speech Grammar Recognition Specification SRGS),
 - > specifikacija za pomensko tolmačenje razpoznanega govora (angl. Semantic Interpretation for Speech Recognition SISR),
 - > specifikacija za zapis slovarja izgovarjav (angl. Pronunciation Lexicon Specification PLS) in
 - > označevalni jezik za sintetizatorje govora (angl. Speech Synthesis Markup Language SSML)

VoiceXML

- » VoiceXML prehod (angl. gateway) omogoča uporabnikom dostop do spletnih vsebin, do katerih sicer dostopajo preko spletnih brskalnikov, tudi z uporabo govornega sistema za dialog.
- » Iste vsebine tako lahko do neke mere sooblikujemo s spletnim označevalnim jezikom HTML ali z označevalnim jeziko VoiceXML.



VoiceXML – programski zgled

```
<?xml version="1.0" encoding="UTF-8"?>
<vxml version="2.1"
 xmlns="http://www.w3.org/2001/vxml"
 xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
 xsi:schemaLocation="http://www.w3.org/2001/vxml
http://www.w3.org/TR/2007/REC-voicexml21-20070619/vxml.xsd">
 <form>
  <blook>Hello World!</block>
 </form>
<form>
 <field name="drink">
  prompt>Would you like coffee, tea, milk, or nothing?
  <grammar src="drink.grxml" type="application/srgs+xml"/>
 </field>
 <filled>
  <submit next="http://www.drink.example.com/drink2.asp"/>
 </filled>
</form>
</vxml>
```

Programski tolmači za VoiceXML

- » Obstaja več odprtokodnih in predvsem komercialnih arhitektur in tolmačev za VoiceXML, kot so na primer:
 - > Odprtokodna javanska arhitektura JVoiceXML http://jvoicexml.sourceforge.net
 - > VoiceXML tolmač VoiceGlue https://github.com/voiceglue/voiceglue/wiki
 - > Komercialni VoiceXML tolmač BladeWareVXi http://www.commetrex.com/products/bladewarevxi-vxmlinterpreter



VPRAŠANJA

- » Kaj vse se uporablja za vhod in kaj za izhod v sistemih za komunikacijo "človek-stroj"?
- » V katerih primerih in zakaj je govor človeku najbolj prikladen način komunikacije s stroji?
- » Kateri so osnovni gradniki sistema za govorno komunikacijo "človek-stroj"?
- » Katere vrste dejanj v dialogu navadno obravnavamo?
- » Kako navadno modeliramo upravljavca dialoga?
- » Opišite glavne faze načrtovanja sistema za dialog.
- » Katera programska orodja obstajajo za razvoj sistemov za govorno komunikacijo "človek-stroj"?