

Robot, ki sliši in govori

JOTA seminar

25. 5. 2015

Vid Podpečan

Inštitut Jožef Stefan
Odsek za tehnologije znanja

Vsebina



1. predstavitev robota NAO
2. govorne zmogljivosti
3. slušne zmogljivosti
4. NAO in slovenski jezik
5. ideje in načrti za prihodnost

Kdo je NAO?



- humanoidni robot
- podjetje Aldebaran robotics (Pariz)
- cilj: inteligentni robotski družabnik

Zgodovina robota NAO

- **2005:** Bruno Maisonnier ustanovi Aldebaran Robotics, pridruži se mu še 5 entuziastov
- **2006:** prvi prototip
- **2007:** ekipa NAO robotov že tekmuje v robotskem nogometu
- **2015:**
 - peta različica robota NAO
 - več kot 7000 prodanih robotov
 - robot Pepper, ki prepoznava čustva
 - v razvoju: robot Romeo
- **cena robota NAO: trenutno ~7000€**



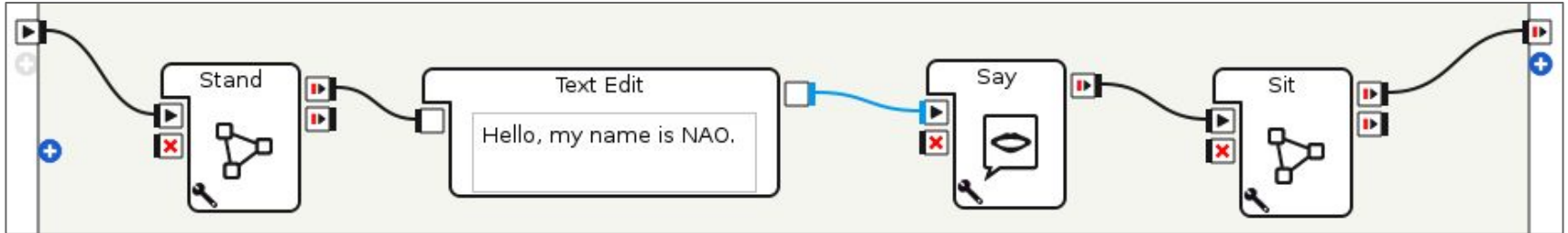
Robotski sistem NAO

- tehnične značilnosti
 - **zaznava sveta:** 2 kameri, 4 mikrofoni, sonar, IR, senzorji dotika in inercije
 - **komunikacija:** 2 zvočnika, brezžična in žična povezava, USB priključek
 - **gibanje:** človeku podobne okončine, 25 prostostnih stopenj (človek ~240)
 - **programska oprema:** temelji na odprti kodi, podpira več programskih jezikov, *vizualno programiranje*, velik nabor vgrajenih sposobnosti in osnovnih gradnikov (npr. hoja, govor, koordinacija gibov, ...)

- osnovna uporaba robota NAO je enostavna tudi za laike!

Primer za laike

- NAO vstane, se predstavi in se usede



- programiranje ni potrebno, le pravilno povežemo škatljice

Govorne zmogljivosti robota NAO

Pregled govornih sposobnosti

- sinteza govora za več jezikov (21)
 - angleščina, francoščina, japonščina, kitajščina, španščina, nemščina,...
 - primarni jezik je angleščina, sekundarnega lahko izberemo
 - dodatni jeziki po dogovoru
 - slovenščina ni podprta
- animirani govor
 - veliko število kretenj
 - dva avtomatska načina izbire kretnje: vsebinski in naključni
- predvajanje poljubne zvočne datoteke
 - lahko vsebuje npr. vnaprej posnet govor v nepodprtem jeziku
- uporaba lastnih sintetizatorjev
 - je možna
 - kasneje: sintentizator za slovenski jezik

Demonstracija: NAO pesnik

- NAO bo sestavil in prebral pesem (v angleščini)
- rimo bomo določili s pritiski na senzorje na robotovi glavi
 - na primer: **AA BB AA CC**
- pesem bo poslal tudi na posebno spletno stran
- kako deluje?
 - zbrano je veliko število pesmi znanih pesnikov
 - slovar izgovorjav za ugotavljanje rime
- NAO naključno izbere vrstice
 - upošteva rimo
 - interpretira s kretnjami



Slušne zmogljivosti

Prepoznavanje govora

- veliko zahtevnejši problem kot sinteza
 - človeški možgani to nalogo opravljajo samodejno in izjemno dobro
 - računalnik ima težave
- do nedavnega:
 - dobro delovanje le za posamezne besede
 - lahko deluje precej dobro, če se nauči posebnosti nekega govorca
 - za poljubnega govorca pa deluje slabše
 - redki programi omogočajo sprejemljivo dobro prepoznavo tekočega govor (npr. Dragon NaturallySpeaking)
- 2010: uporaba večnivojskih nevronske mreže (deep learning)
 - preskok v natančnosti
 - odlična prepoznavna tekočega govora
 - Google speech recognition, Apple Siri, Microsoft Cortana

Prepoznavanje govora in robot NAO

- prepoznavanje za vseh 21 podprtih jezikov
 - prepoznavna je omejena na **posamezne besede**
- uporabnik določi seznam besed za prepoznavo
 - npr. [up, down, left, right]
 - NAO reagira, če izgovorjeno besedo najde v seznamu
- lovljenje besed iz tekočega govora
 - t.i. **word spotting**
 - NAO reagira, če v tekočem govoru prepozna besedo iz seznama

Demonstracija: prepoznavanje besed

- seznam besed: različno sadje
 - [apple, orange, strawberry, blackberry, papaya, lemon, watermelon, grapes]
- NAO bo prepoznal besedo in jo izgovoril




NAO - inteligentni sogovornik

- ali lahko NAO robota naučimo inteligentnega pogovora?
- kaj potrebujemo
 - prepoznavo tekočega govora
 - program, ki je sposoben voditi pogovor
- tehnologije so na voljo
 - avtomatski klepetalniki (chatterbot)
 - ChatScript, ALICE, Mitsuko
 - vsebujejo veliko število odgovorov, prepoznavajo vzorce vprašanj, pametno izbirajo smiselne odgovore
 - Loebner Prize: tekmovanje klepetalnikov
 - prepoznavna govora
 - Google speech recognition
- kako dobro to deluje v praksi?

Demonstracija:

Na klepetu z robotom



- NAO posluša, razume vprašanja ter smiselno odgovarja nanje
 - pogovor bo prikazan na spletni strani

- kako deluje?
 - naš govor se pošlje v Googlov prepoznavalnik
 - prepoznan govor je vhod v avtomatski klepetalnik
 - NAO pove odgovor avtomatskega klepetalnika in ga opremi s kretnjami

NAO in slovenski jezik

Programska oprema za slovenski jezik

- sintetizatorji
 - Govorec, e-Govorec (IJS, Amebis)
 - Proteus TTS (podjetje Alpineon)
 - kvaliteta govora ni na ravni sintetizatorja angleščine na NAO robotu
- domačega odprtokodnega sintetizatorja ni!

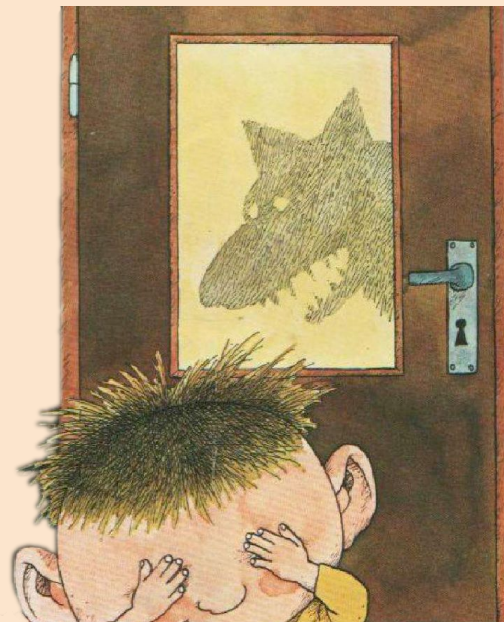
- prepoznavalniki slovenskega jezika
 - domači produkti: jih ni
 - na univerzah je nastalo nekaj prototipov, projekt VoiceTRAN
 - **Googlov prepoznavalnik podpira slovenščino**
 - deluje solidno, postaja vedno boljši

Odprtokodni sintetizatorji in NAO

- NAO lahko uporablja naš lasten sintetizator
 - če je tehnično kompatibilen
- odprtokodnih sintetizatorjev je veliko, za nas sta zanimiva dva
 - **eSpeak**
 - **MaryTTS**
- **eSpeak**
 - izjemno kompakten, dodajanje novih jezikov je relativno preprosto
 - trenutno podpira ~30 jezikov, ene boljše, druge slabše
- **MaryTTS**
 - napreden, odlična orodja in dokumentacija za dodajanje novih jezikov

Demonstracija: NAO govori slovensko

- NAO bo recitiral pesmico *Pedenjped* Nika Grafenauerja
- uporabili bomo sintetizator **eSpeak**
 - ker slovenščine eSpeak še ne podpira, bomo uporabili modul za hrvaški jezik



Zaključek

Pogled v prihodnost

- razvoj robotskega sogovornika
 - za splošen klepet
 - za posebne domene
 - strokovne konference
 - promocija znanosti in robotike
- podpora za slovenski jezik v sintetizatorjih eSpeak in MaryTTS
- kreativno računalništvo z robotom NAO
 - ples, glasba, poezija



DEPARTMENT OF
**KNOWLEDGE
TECHNOLOGIES**

Jožef Stefan Institute