Intelligent User Interfaces (IUI)

Block 1: Voice User interfaces
Session 1: Introduction, Terminology, Concepts

LMU München, Medieninformatik
Prof. Dr. Andreas Butz
WS2018/19
Voice UI: Introduction, Terminology, Concepts

• Organization of the lecture

• A brief history of NLP and VUI

• Basic concepts in NLP
  – syntax
  – semantics
  – pragmatics
  – prosody

• Modern Voice Assistants

• Voice UI design
Organization

- **Lecture**: Thursday 10-12h, Luisenstr. 37 (C), C 106
- **Practical projects**: one practical project for each topic, to be presented in the end of the block
- **Blog posts**: students have to document 3 appointments in a 1,000 word blog post, adding *additional resources* they found on the topic.

- **Website**:
  - [http://mimuc.de/iui](http://mimuc.de/iui) (Slides, Exercises, additional material)
  - User: iui, Password: iui_ws1819
Structure of the class

- 4 theme blocks
  - introductory lecture (to kick off the block)
  - practical project accompanying the class
  - guest lecture (industry or academic expert)
  - student lecture (results of project, additional resources)

- Block 1: Voice UI
- Block 2: TBD
- Block 3: Text Input
- Block 4: Adaptive UI
Exam and bonus points

• 100 Points for exam across everything we teach and from (projects or reading material).
  – some questions will be on the practical projects or alternatively on the reading material
  – if you have read the material or worked on the projects, you can answer them
  – If you haven’t, you can still pass the exam with the remaining questions

• 9 bonus points for 3 blog posts of 1000 words each (3 per post. Max 3 posts)
Voice UI: Introduction, Terminology, Concepts

• Organization of the lecture

• A brief history of NLP and VUI

• Basic concepts in NLP
  – syntax
  – semantics
  – pragmatics
  – prosody

• Modern Voice Assistants

• Voice UI design
Eliza (Weizenbaum, 1964)

Eliza is a mock Rogerian psychotherapist. The original program was described by Joseph Weizenbaum in 1966. This implementation by Norbert Landsteiner 2005.

ELIZA: Is something troubling you?
YOU: Men are all alike.
ELIZA: What is the connection, do you suppose?
YOU: They're always bugging us about something or other.
ELIZA: Can you think of a specific example?
YOU: Well, my boyfriend made me come here.
ELIZA: Is it important to you that your boyfriend made you come here?
YOU: He says I'm depressed much of the time.
ELIZA: I am sorry to hear that you are depressed.
YOU: It's true. I am unhappy.
ELIZA: Can you explain what made you unhappy?
YOU: 
SHRDLU (Winograd, 1970)

- Domain: simple blocks world
- can execute commands in this world
- keeps a memory of world state
- keeps a memory of previous operations
- can learn names of structures
- was considered extremely successful AI
Parry (Colby, 1972)

- simulated a person with paranoid schizophrenia
- passed a modified version of the Turing test
- was repeatedly connected to doctor (ELIZA)
PARRY Encounters the DOCTOR

Good Evening.

Tell me your problems. Please terminate input with a period or a question mark.

*People get on my nerves sometimes.
   I am not sure I understand you fully.

*You should pay more attention.
   Suppose you should pay more attention.

*You're entitled to your own opinion.
   What makes you think I am entitled to my own opinion?

*It bothers me just to be around people in general.
   Please go on.

*I went to the track at Bay Meadows a while back.
   What does that suggest to you?

*I went to the races.
   What else would you like to discuss?

*Horseracing interests me.
   Why do you say that just now?
KL-One (Sondheimer et al., 1974)

- Knowledge representation system
- Semantic network, frames

http://www.jfsowa.com/figs/kl1truck.gif
Fast forward…

• 1971: US Department of Defense (DARPA) “Speech Understanding Research” (SUR) program. An outcome was “Harpy”, a speech-understanding system which could understand more than 1000 words.
• 1984: “Speechworks”, Automated Speech Recognition over IVR (Interactive Voice Response) on telephone
• 1996: BellSouth launched VAL, first voice portal based on dial-in interactive voice recognition system
• 1997: “Dragon Dictate”, first software to enable recognition of continuous speech
• 2007: “Siri” was founded to enable natural human to machine interaction through voice interface
• 2008: Google launched Voice search app
• 2011: Siri launched with iPhone 4S integrated with iOS.
• 2011: IBM Watson wins in Jeopardy
• 2014: Amazon Echo launched.
• 2016: Google Home launched.

Watson (IBM, 2006)

- Question-answering system
- IBM DeepQA project
- won Jeopardy in 2011

Watson wins Jeopardy

https://www.youtube.com/watch?v=P18EdAKuC1U
Voice UI: Introduction, Terminology, Concepts

• Organization of the lecture
• A brief history of NLP and VUI

• Basic concepts in NLP
  – syntax
  – sematics
  – pragmatics
  – prosody

• Modern Voice Assistants
• Voice UI design
Phonology, Syntax

Syntax

Phonology, Syntax

Phonology, Syntax
Semantics and ambiguity

- Time flies like an arrow, fruit flies like a peach.
- Time flies like an arrow, time horses like a bow.
- …many more…
Pragmatics

• Describes how we actually use language

• A: „Have you got coffee to go?”
• B: „Milk and Sugar?”
• A: „black as my soul…”

• …now try to teach a computer
  – to understand this (in general)
  – to react adequately
  – to generate such things
Prosody

• The „melody“ (volume, pitch, rate) of speech
• Used differently in different languages
  – prosody in English can modify the meaning of sentences
    • question or statement?
    • angry or friendly?
    • emphasis on words
  – tones in Chinese change the meaning of words
  – prosody generally can convey additional meaning
Speech Synthesis Markup Language (SSML)

- **Do**
  - Great job, you are correct
  - `<prosody rate="x-slow">Great</prosody> job, you are correct.`
  - `<emphasis level="strong">Great</emphasis> job, you are correct.`

- **Don't**
  - Grrrrreeaaat job, you are correct.

Voice UI: Introduction, Terminology, Concepts

- Organization of the lecture
- A brief history of NLP and VUI
- Basic concepts in NLP
  - syntax
  - semantics
  - pragmatics
  - prosody
- Modern Voice Assistants
- Voice UI design
Modern Voice assistants (Siri, Google, Alexa)

- Use simple front end devices (basically elaborate networked microphones)
- All „intelligence“ happens in the backend/cloud
- Can integrate your own services („skills“)
Which Voice Assistant is best?

https://www.youtube.com/watch?v=GZ2LiyisSiE
Voice UI: Introduction, Terminology, Concepts

- Organization of the lecture
- A brief history of NLP and VUI
- Basic concepts in NLP
  - syntax
  - semantics
  - pragmatics
  - prosody
- Modern Voice Assistants
- Voice UI design
No Affordances in Voice!

• Many principles for designing understandable GUI

• Some continue to work
  – Consistency (especially lexical)
  – Feedback (visual, acoustic or spoken)
  – Metaphors (when expressed in language)

• Some work differently in VUI
  – Constraints (only logical and cultural, no physical)
  – Error tolerance (homophones, ambiguities, …)

• Some are absent in VUI:
  – Spatial mappings
  – Visual affordances
Some Voice UI guidelines

• Tell users what they can do
  – e.g. „You can ask for today’s weather or a weekly forecast.“

• Tell users where they are
  – user: „What’s the weather today?“
    – system: „Today’s weather forecast is mostly sunny and dry“ rather than just „sunny and dry.“

• Give examples rather than instructions
  – e.g. in the help function or in the greeting

• Limit the amount of information
  – e.g. not more than three different options for an interaction
  – tradeoff between efficiency and short term memory!

• Use visual feedback (if possible)

https://www.interaction-design.org/literature/topics/voice-user-interfaces
Soul Machines: A glimpse into the uncanny valley…

https://www.youtube.com/watch?v=khr-eWGHtSI
Managing User Expectations

• Users tend to overestimate machine capabilities
  – normally we use voice to communicate with humans
  – humans possess true intelligence
  – …hence, the thing we talk to must be truly intelligent ;-) 

• Some guidelines for managing limitations:
  – Don’t assume your product can fully understand the user’s context.
  – If in doubt, go for limited scenarios
  – Make options and suggestions easily available to users
  – …but never interrupt to provide them.
VUI design process

• Design a dialog structure

• Think of alternatives
  – structure
  – wording

• Try out your dialog
  – wizard of Oz technique!
  – use outside people

• Refine, Revise, Repeat
(Res)Sources

- https://www.interaction-design.org/literature/topics/voice-user-interfaces
- https://developer.amazon.com/designing-for-voice/
- https://www.w3.org/TR/speech-synthesis11/