Language and Computers

Can computers talk?

- No . . . Yes?
- It isn't currently possible to give machines the full conversational skills of a human being; programming a computer to work with language to some degree is possible.
- Both computers and people are information processing systems.
- The processing tasks of human language can be automated to some degree on a computer:
  - Recognizing words in speech
  - Pronouncing words
  - Translating from one language to another

Speech Synthesis

- Speech synthesis: the use of a machine, a computer, to produce human-like speech.
  - Canned speech: playing prerecorded utterances and phrases
  - Piecing together smaller recorded units of speech into new utterances
  - Synthesized speech: creating speech "from scratch"
- Applications: cars, elevators, automatic lottery machines, answering machines, telephone customer service centers, automated grocery checkout lanes
- Intelligibility & Naturalness
- The early synthesis machines
  - The Voder https://www.youtube.com/watch?v=0rAyrmm7vv0
- Articulatory synthesis: generates speech "from scratch" based on computational models of the shape of the human vocal tract and the articulation processes. Not have been very successful.
- Concatenative Synthesis: the most commonly used speech synthesis technology today because it generates very natural sounding speech; use recorded speech and manipulate speech samples
- Text to Speech Synthesis
  - Stephen Hawking—How is He Talking https://www.youtube.com/watch?v=UErbwiJH1dI

Automatic Speech Recognition

- Automatic speech recognition
  - The conversion of an acoustic speech waveform into text
  - A computer takes sounds of speech and converts them into words of some particular language
- Applications
  - Telephone applications that allow customers to enter or request verbally
    - Calling to find the status of an airplane flight
    - Call routing "Operator, please"
  - Dictation: producing a transcription of a spoken document
  - Desirable in environments that require a user's hands to control equipment or demand that visual attention be focused elsewhere, e.g. driving, flying fighter jet in combat

- The noisy channel model
  - Treats speech input as if it has been passed through a communication channel that garbles the speech waveform, producing a "noisy" or distorted version of the original words spoken
  - Noises: variations in pronunciation that distort words' canonical form
    - "Did you go yet?" [did] [ju] vs [did3ju]
    - Acoustic variation introduced by the microphone or telephone network
- Components of an automatic speech recognition system
  - Signal processing
  - Acoustic modeling: mapping acoustic waveforms into symbols for phones
  - Pronunciation modeling: knowledge of which sequences of sounds are mostly likely. **ŋasal/nasal*
  - Language modeling: calculating the probability of word sequences. *drive a car/drive a call*
  - Putting it all together
- Types of speech recognition systems
  - Speaking mode: isolated word input vs continuous speech input
  - Vocabulary size: small (<20) vs large (>20,000)
  - Speaker enrollment: speaker-dependent vs speaker independent

Siri—the Apple speech recognition system

- A part of Apple Inc.'s iOS which works as an intelligent personal assistant and knowledge navigator. The feature uses a natural language user interface to answer questions, make recommendations, and perform actions by delegating requests to a set of Web services such as Wolfram Alpha. The software, both in its original version and as an iOS feature, adapts to the user's individual language usage and individual searches (preferences) with continuing use, and returns results that are individualized.
- Talking with Siri
  - https://www.youtube.com/watch?v=P8zG_gGmfmo
  - https://www.youtube.com/watch?v=mztUL9HG_Lc

Communicating with Computers

- As for humans, natural language is the most familiar, convenient, and efficient method of communication in most situations, it is desirable for people to be able to use natural language to interact with computers.
- Systems allowing effective natural language communication with them are relatively new.
- Jules, a Conversational Character Robot designed and built by David Hanson. https://www.youtube.com/watch?v=ysU56JzBjTY
- "Her" and Artificial Intelligence https://www.youtube.com/watch?v=0ANXEJrtJMg
- How far have computers emulate human?

Interactive Text-Based Systems

- ELIZA (a computer psychiatrist). 1966. Joseph Weizenbaum.
  - Text-based

User: I am stressed.

ELIZA: How long have you been stressed?

User: Since I started writing this chapter.

ELIZA: Can you elaborate on that?

User: I can't stop thinking about human-computer
nteraction.

ELIZA: Have you tried?

User: No I guess not really.

ELIZA: You are being a bit negative.

User: Well what can I do to change my attitude?

ELIZA: Why do you say your attitude?

User: Because you said I was being negative.

(2) User: I was once the captain of a ship.
ELIZA: Why do you tell me you were once the captain of
ship just now?
User: I was never good at soccer.
ELIZA: Why do you tell me you were never good at socce
just now?

- Word-spotting strategy:
  - I was [. . .] → Why do you tell me you were [. . .]
  - . . . No. . . Not. . . → You are being a bit negative
- http://www.manifestation.com/neurotoys/eliza.php
- A user can't get any useful behavior or information by talking to her. ㅠㅠ

MUD (Multi-User Dungeon) "bots"

- An internet role-playing game
- Can give informative replies to questions about who is playing and how long they have been playing, how to use the game, etc.
- Use the simple pattern-matching strategies
- Perform better when their use I restricted to a limited domain.
  - specific virtual world, flight booking system
- Julia. Michael Mauldin
  - Can recognize words she knows, "Where are we?", retrieve appropriate information from a database, and report
  - Productive, informative communication

(3) User: Where are we?

Julia: We are in "The Inter Nexus"

User: Who is here?

Julia: Well Lenny I see 8 players awake here: guest byte MadraRua Leda Lenny Mandi TheHighMage Samuel

User: Who is asleep here?

Julia: Well Lenny I see 1 player asleep here: Jacques

Spoken-language Dialogue Systems

- Handle isolated speech, continuous speech, or barge-in (where the impatient user interrupts and talks over the computer)
- Commercial use: cheaper for a company/organization to use a computer than to pay people
- Combined with telephone technology. Since 1980s
  - To place and accept or reject calls, use a calling card, and perform other services
  - To place orders, request or cancel a service, or check an account status
  - Inquire about train/flight timetables, reserve flight/train tickets
- CMU Darpa Communicator, an automated telephone-based dialogue system for booking flight information
- http://www.speech.cs.cmu.edu/Communicator/sample2.wav

- Components of a Spoken-Language Dialogue System
  - Automatic Speech Recognition
  - Language Processing and Understanding
  - Dialogue Management
  - Text Generation
  - Speech Synthesis
- Evaluation of interactive systems
  - Wizard of Oz simulations: the user thinks they are interacting with the actual computer system but in fact, a hidden human controller simulates some aspects of the system

Machine Translation (MT)

- Translation: source language → target language
- The use of computer to carry out translation
- A viable alternative to human translators
  - With increasing globalization, the volume of business-oriented translation has increased so much in recent years that often there aren't enough translator to meet the demand
  - Human translators can be extremely expensive
- Not enough is known about language and the process of translation to enable a computer to duplicate the efforts of a human being.
- Canadian METEO system: translating English/French weather reports
- Google translate https://translate.google.com/

Corpus Linguistics

- Using computer programs to help us analyze language
- Corpus: a collected body of text
- Corpus linguistics: the design and the annotation of corpus materials that are required for specific purposes
  - Frequency
  - Collocation
- Kinds of Corpora
  - Source of the corpus
  - Levels of annotation
  - BYU Corpora http://corpus.byu.edu/
  - BNC(British National Corpus) http://corpus.byu.edu/bnc/
  - COCA(Corpus of Contemporary American English) http://corpus.byu.edu/coca/
  - Sejong Corpus (세종말뭉치) https://ithub.Korean.go.kr/user/main.do

Machine learning

- http://www.ted.com/talks/deb_roy_the_birth_of_a_word