Language and Computers

Introduction

Jessy Li, UT Austin

Siri

- https://www.youtube.com/watch?v=15Apiybz8Lc&t=157s

Ice breaker

What are the daily tasks that you do with computers, that would require it to understand language or to say something to you?

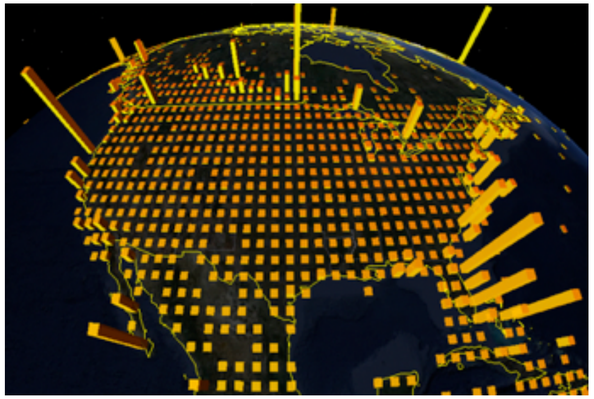
Word clouds

- Go to NYT and pick an article!
- https://www.wordclouds.com

The World Well-Being Project

- http://wwbp.org

What on earth is this word?



Beach

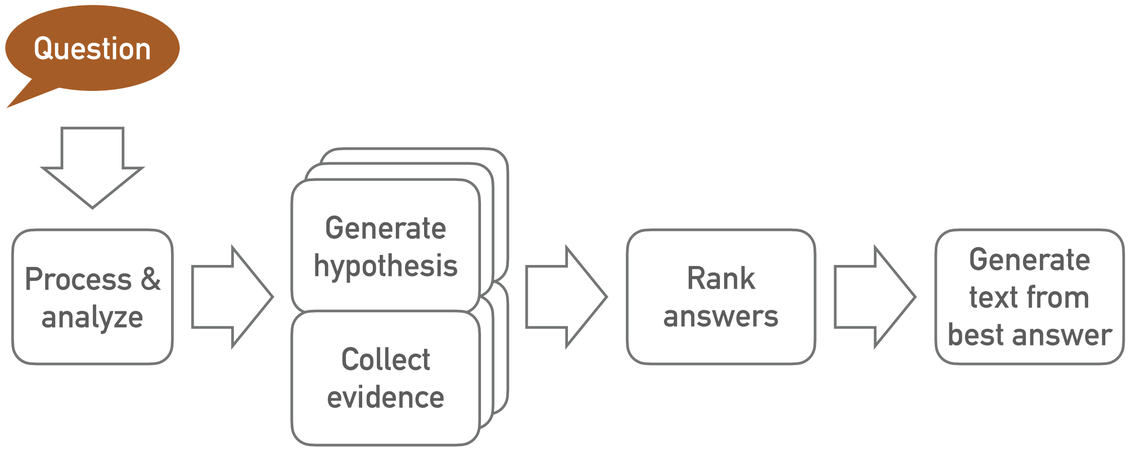
Mountain

Wine

BBQ

IBM Watson

- https://www.youtube.com/watch?v=WFR3lOm_xhE



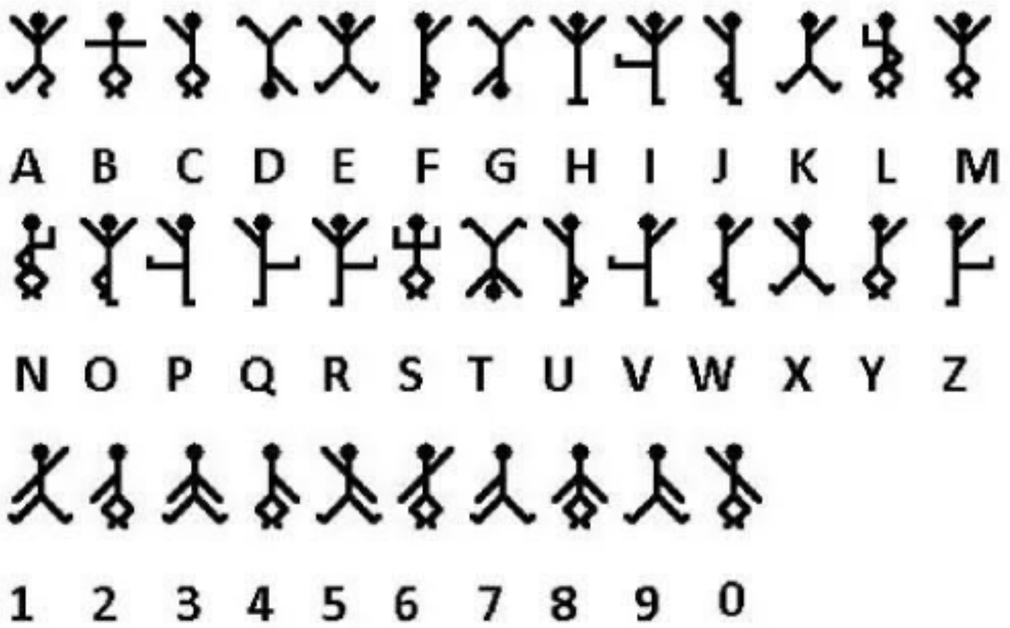
What we will learn…

Encoding

- You see some text here.
- But how are they represented in computers?
  - (How is anything represented in computers?)

Also encoding…

- Any Sherlock fans here?



Search

Google

who performed at the super bowl 2019

All News Videos Images Shopping More Settings Tools

About 349,000,000 results (0.99 seconds)

What are the problems here that computers need to solve?

Travis Scott

The **Super Bowl** LIII Halftime Show, officially known as the Pepsi **Super Bowl** LIII Halftime Show, took place on February 3, **2019** at Mercedes-Benz Stadium in Atlanta, Georgia, as part of **Super Bowl** LIII. It was headlined by American pop group Maroon 5, joined by rappers Big Boi and Travis Scott as guests.

Super Bowl LIII halftime show - Wikipedia

https://en.wikipedia.org › wiki › Super_Bowl_LIII_halftime_show

Search

- Let’s say that the LIN 313 syllabus is too long…
- And we want to look for content that is about grade, or grading
- How do we do that?

Generating text

- Write with the transformer! https://transformer.huggingface.co

Ever tried Grammarly?

- grammarly.com

Who wrote this text?

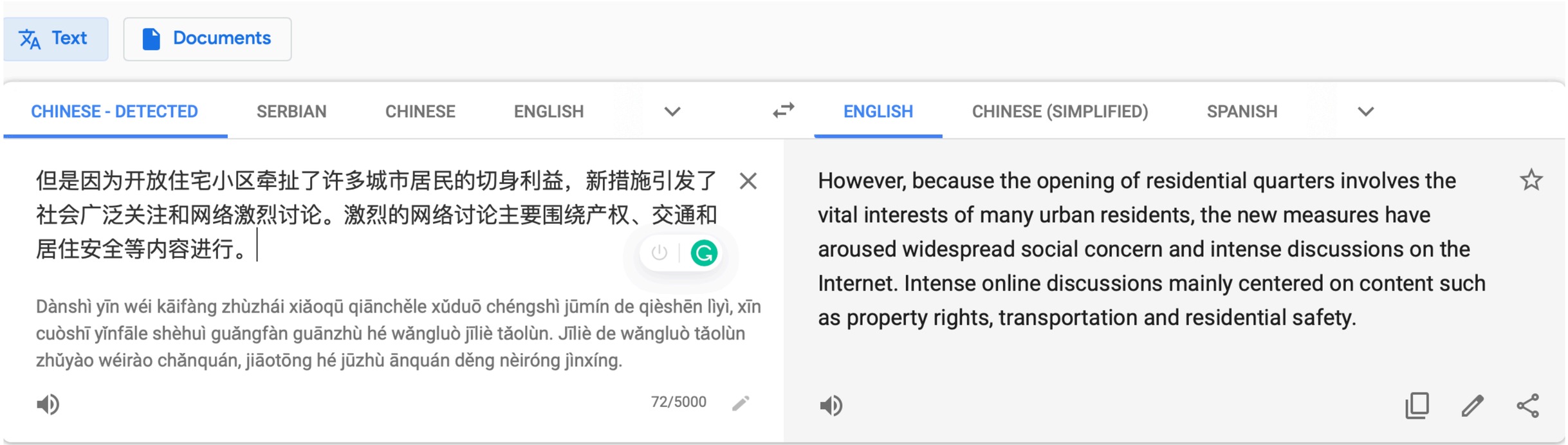
- How would this be useful?
- What problems do computers need to solve?

Sentiment analysis

- https://demo.allennlp.org/sentiment-analysis
- This is called “classification”. What else can classifiers be used for?

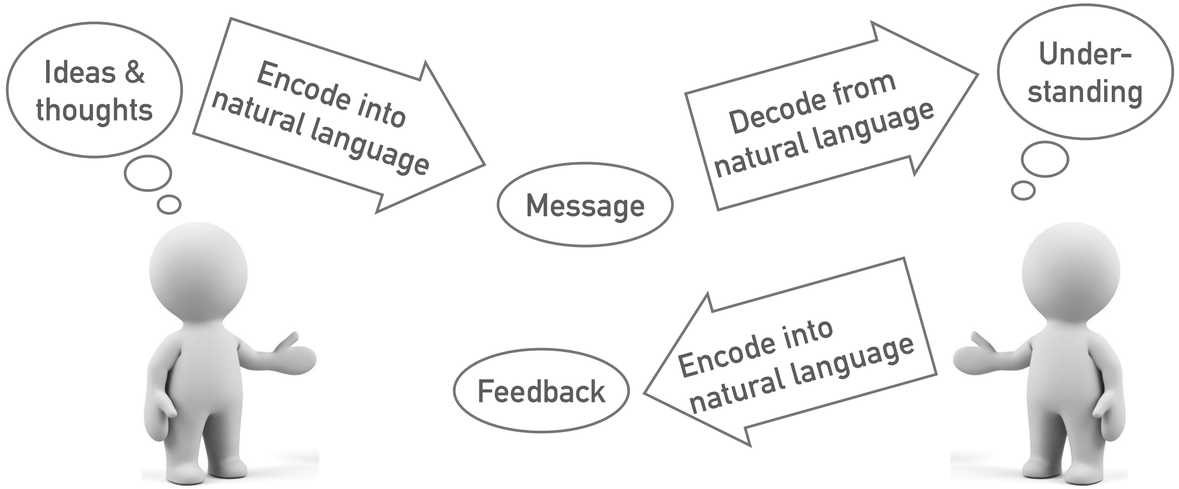
Machine translation

- What problems do you think we need to solve here?

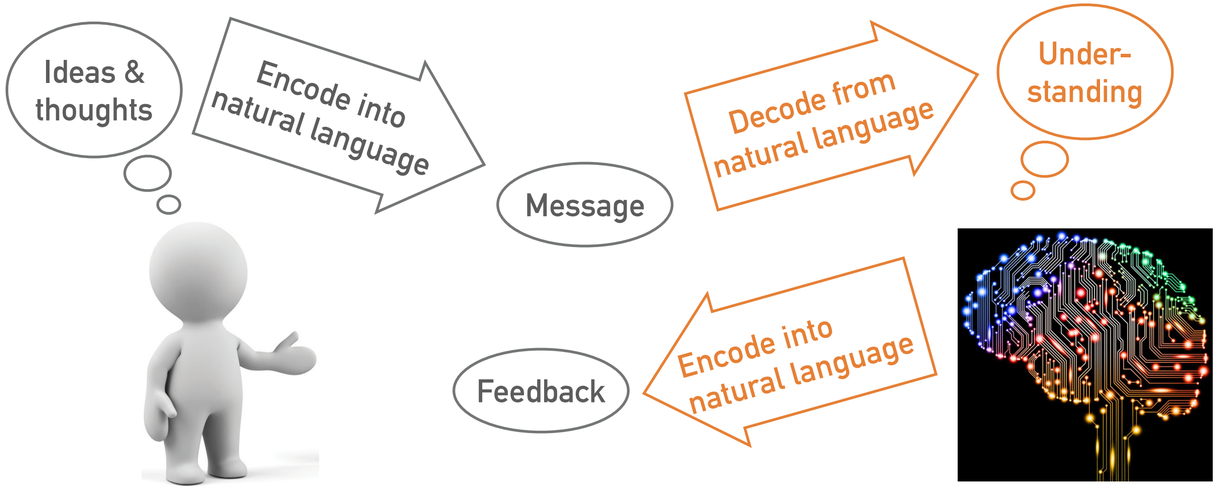


How about dialog systems?

…and NLP in general



Communicating with Machines



What makes it hard for machines to have language skills?

- Words are ambiguous!

What makes it hard for machines to have language skills?

- Grammar is ambiguous!

I ate dumplings with chopsticks.

I ate dumplings with pork.

Headline fails

- US eyes return to the moon
- Stolen painting found by tree
- Blair wins on budget, more lies ahead
- Hospitals are sued by seven foot doctors
- Miners refuse to work after death
- Flesh-eating bug survivor goes home
- In America a woman has a baby every 15 minutes

What makes it hard for machines to have language skills?

- Knowledge and inference

> *“I ran all the way to the main gate, and then I waited a second till I got my breath. I have no wind, if you want to know the truth. I’m quite a heavy smoker, for one thing—that is, I used to be. They made me cut it out. Another thing, I grew six and a half inches last year. That’s also how I practically got t.b. and came out here for all these goddam checkups and stuff. I’m pretty healthy though.”*
>
> *The Catcher in the Rye, J. D. Salinger*

What makes it hard for machines to have language skills?

- Knowledge and inference

> *"I ran all the way to the main gate, and then I waited a second till I got my breath. I have no wind, if you want to know the truth. I'm quite a heavy smoker, for one thing—that is, I used to be.* ***They*** *made me cut it out. Another thing, I grew six and a half inches last year. That's also how I practically got t.b. and came out* ***here*** *for all these goddam checkups and stuff. I'm pretty healthy though."*
>
> *The Catcher in the Rye, J. D. Salinger*

What makes it hard for machines to have language skills?

- Language is diverse.

(JP) 私は、箸で餃子を食べました。

(I with chopsticks dumplings ate.)

(I ate dumplings with chopsticks.)

What makes it hard for machines to have language skills?

- Language evolves

Google, v.

friend, v.

selfie, n.

…

Mistakes are consequential

From: http://davidrothman.net/2009/09/02/all-your-healthbase-are-belong-to-us-want-em-back/ (legacy link)

Being right can be consequential too.

- What are the potential social impacts of language technology?

Language and computers…

- Computational linguistics
- Natural language processing
- We will dive into a range of topics in:
  - Linguistics
  - Computer Science
  - Probability and statistics

This course

- Instructor: Jessy Li (jessyli.com)
- TAs: Isaiah Hogue, Lindia Tjuatja
- (Almost) Everything's on the course website:
  - jessyli.com/courses/lin313
- We use Canvas for homeworks, exercises, and grades
- We use Piazza for QA/discussions

This course

- Who this course is geared toward and who it isn't.
  - There is no programming, but expect to think **very** quantitatively.
  - This course is meant to be fun and topical, yet rigorous and specific.
- Other computational linguistics courses:
  - LIN 353C Introduction to Computational Linguistics (Erk)
  - LIN 350 Computational Semantics (Erk)
  - LIN 373 Machine Learning Toolbox for Text Analysis (Li)
- I welcome your input and feedback. (See something cool in the news? Tell me!)

COVID-19 highlights

Please read the syllabus carefully!

- Most of the meetings will be on Zoom
  - Let’s see each other more often! Please turn on your videos (as long as you’re comfortable doing so).
  - All classes will be recorded and kept for 2 weeks.
  - Use hand raising or chatbox for interaction.

COVID-19 highlights

Please read the syllabus carefully!

- Hybrid instruction and checkpoints
  - Checkpoints: graded discussions, homework reviews
  - We will watch for local conditions and check your preferences whether we meet in person
  - Grading policy and late policy (no exams)
- Let’s work together to make this work! I will ask for your feedback during the semester :)

Questions?