



Natural Language Processing

Info 159/259

Lecture 1: Introduction (Jan 17, 2023)

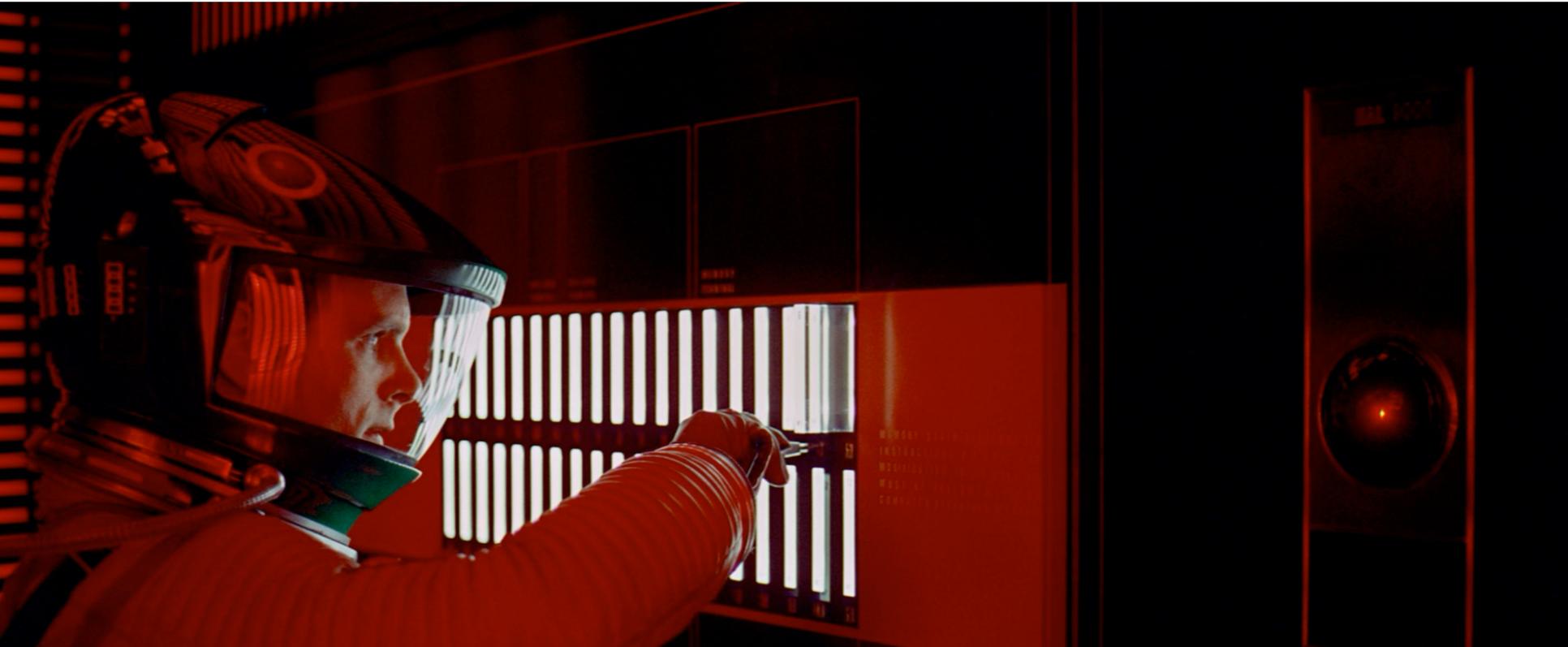
David Bamman, UC Berkeley

NLP is interdisciplinary

- Artificial intelligence
- Machine learning (ca. 2000—today); statistical models, neural networks
- Linguistics (representation of language)
- Social sciences/humanities (models of language at use in culture/society)

NLP = processing^{*} language with computers

processing as “understanding”



JOAQUIN PHOENIX AMY ADAMS ROONEY MARA

OLIVIA WILDE AND SCARLETT JOHANSSON

her

A SPIKE JONZE LOVE STORY

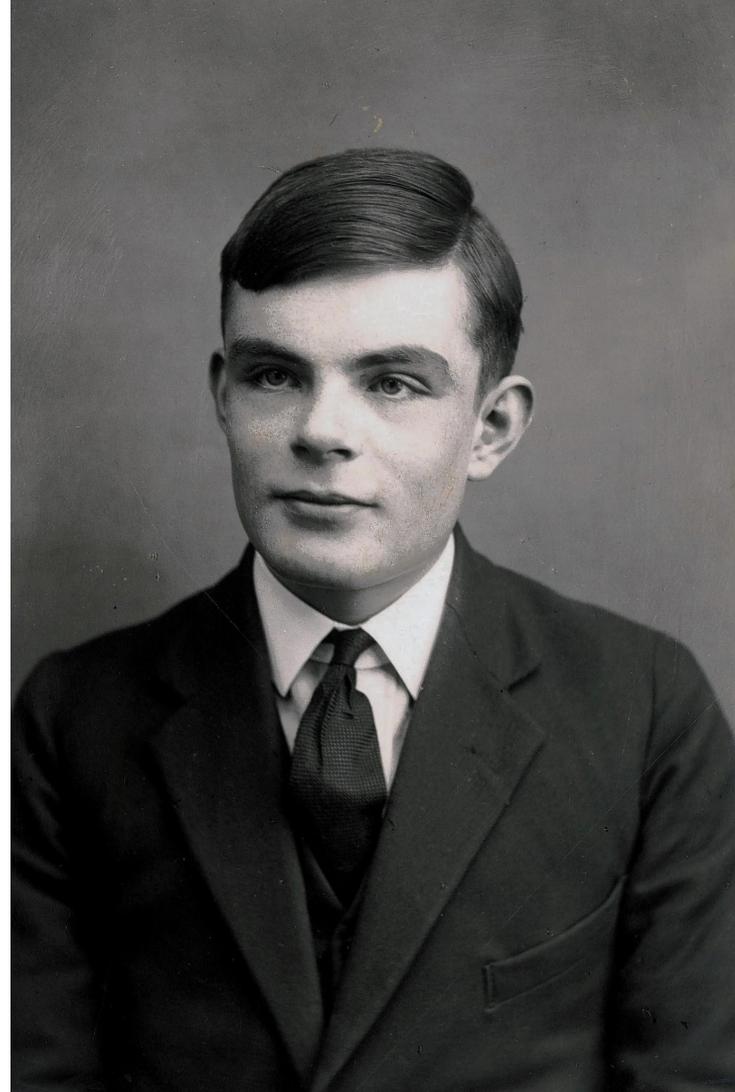
WARNER BROS. PICTURES PRESENTS
AN ANIMAPURNA PICTURES PRODUCTION "HER" JOAQUIN PHOENIX AMY ADAMS ROONEY MARA OLIVIA WILDE AND SCARLETT JOHANSSON
DIRECTED BY SPIKE JONZE
CASTING BY ELLEN LEVINS CASSANDRA KALLIKONDIS MUSIC BY SUPREMACY FEN WU COSTUME DESIGNER ARCADE FIRE PRODUCTION DESIGNER CASEY STORINA EXECUTIVE PRODUCERS ERIC ZANGRUM WEN ALEE JEFF BUCHANAN PRODUCED BY KX BARRETT
WRITTEN BY PHOTOGRAF HOVTE VAN HOYTERA EXECUTIVE PRODUCERS DANIEL LUPI NATALIE PARRYER EMILY CARWARD PRODUCED BY MEGAN ELLISON SPIKE JONZE VINCENT LANDAY DIRECTED BY SPIKE JONZE
COMING SOON herthemovie.com

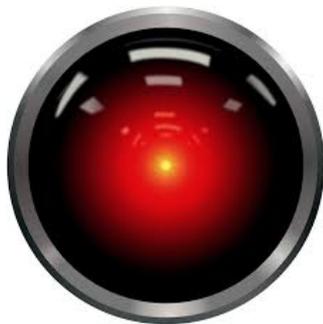


Turing test

Distinguishing human vs.
computer only through
written language

Turing 1950





Dave Bowman: Open the pod bay doors, HAL
HAL: I'm sorry Dave. I'm afraid I can't do that

Agent	Movie	Complex human emotion/goal mediated through language
Hal	2001	Mission execution
Samantha	Her	Love
IG-11	The Mandalorian	"to nurse and protect"

Where we are now

Who is the current President of the [United States](#)?

Joe Biden is the President of the United States.

 KNOWLEDGE



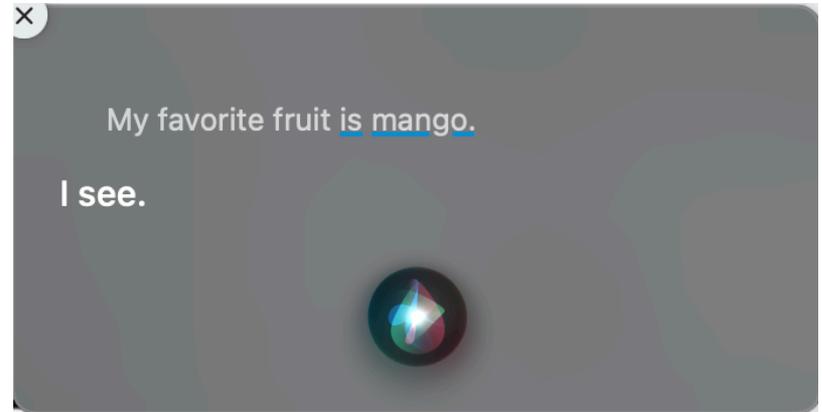
Joe Biden
46th and current President of the United States

Joseph Robinette Biden Jr. is an American politician who is the 46th and current president of the United States. A member of the Democratic Party, he previously served as the 47th vice president from 2009 to 2017 under President Barack Obama, and represented Delaware in the United States Senate from 1973 to 2009. Born in Scranton, Pennsylvania, Biden moved with his family to Delaware in 1953. He studied at the University of De... more

[Wikipedia](#)



Where we are now



Where we are now

My favorite fruit is mango.

I see.

What's my favorite fruit?

I can't read your mind.





OpenAI



My favorite fruit is mango.



That's great! Mangoes are a delicious and nutritious fruit. They are high in vitamins C and A, as well as antioxidants and fiber. Enjoy!



What's my favorite fruit?



You have said that your favorite fruit is mango. ■

<https://chat.openai.com/chat>

What makes language hard?

- Language is a complex social process
- Tremendous ambiguity at every level of representation
- Modeling it is **AI-complete** (requires first solving general AI)

What makes language hard?

- Speech acts (“can you pass the salt?”)
[Austin 1962, Searle 1969]
- Conversational implicature (“The opera singer was amazing; she sang all of the notes”).
[Grice 1975]

What makes language hard?

The icon consists of the letters 'DB' in white, centered within a solid orange square.

If I were to say "The opera singer was amazing; she sang all of the notes", did I think the opera singer was good?



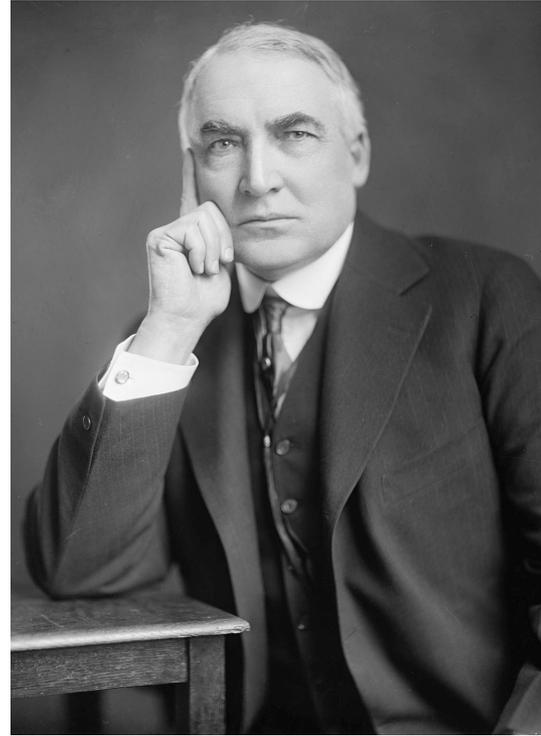
If you said "The opera singer was amazing; she sang all of the notes," it would suggest that you thought the opera singer was good. The statement implies that the singer was able to hit all of the notes correctly and with skill, which is generally considered to be a positive characteristic of a singer.

What makes language hard?

- Speech acts (“can you pass the salt?”)
[Austin 1962, Searle 1969]
- Conversational implicature (“The opera singer was amazing; she sang all of the notes”).
[Grice 1975]
- Shared knowledge (“Warren ran for president”)



Elizabeth Warren
2020



Warren G. Harding
1920

What makes language hard?

- Speech acts (“can you pass the salt?”)
[Austin 1962, Searle 1969]
- Conversational implicature (“The opera singer was amazing; she sang all of the notes”).
[Grice 1975]
- Shared knowledge (“Warren ran for president”)
- Variation/Indexicality (“This homework is wicked hard”)
[Labov 1966, Eckert 2008]

Ambiguity

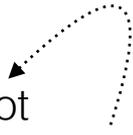
“One morning I shot
an elephant in my pajamas”



Animal Crackers

Ambiguity

“One morning I shot
an elephant in my pajamas”



Animal Crackers

Ambiguity



“One morning I shot
an elephant in my pajamas”



Ambiguity

verb noun



“One morning I shot
an elephant in my pajamas”

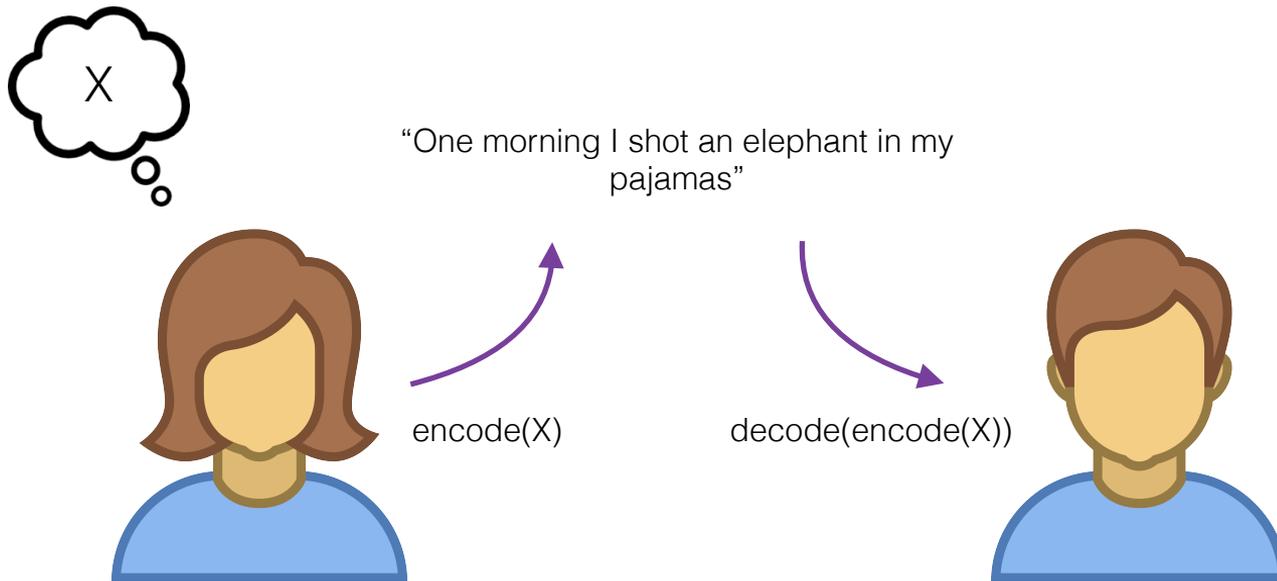


Animal Crackers

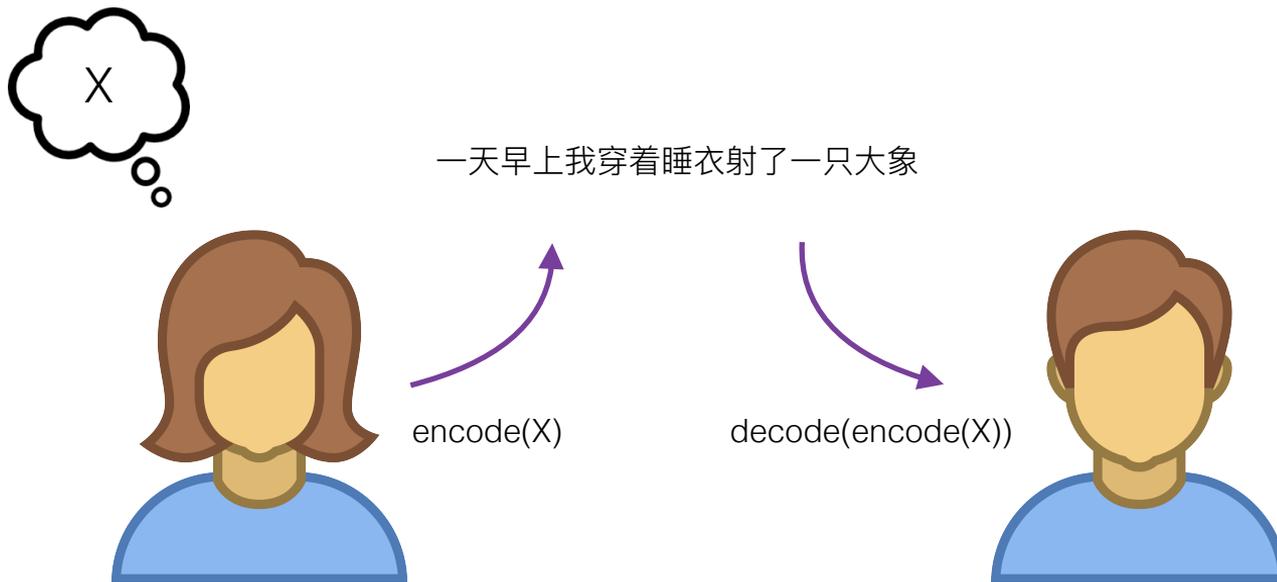
processing as representation

- NLP generally involves **representing language** for some end, e.g.:
 - dialogue
 - translation
 - speech recognition
 - text analysis

Information theoretic view



Information theoretic view



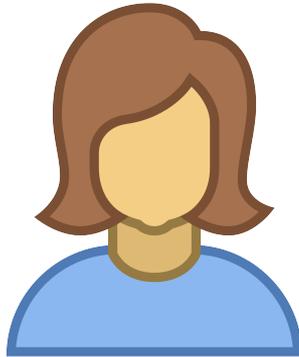
When I look at an article in Russian, I say: 'This is really written in English, but it has been coded in some strange symbols. I will now proceed to decode.'

Weaver 1955

Rational speech act view



“One morning I shot an elephant in my pajamas”



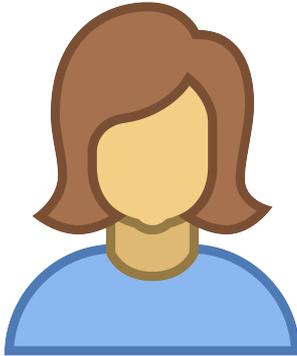
Communication involves **recursive reasoning**: how can X choose words to maximize understanding by Y?



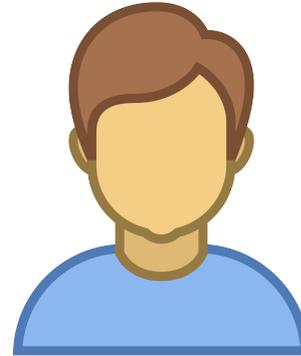
Pragmatic view



“One morning I shot an elephant in my pajamas”



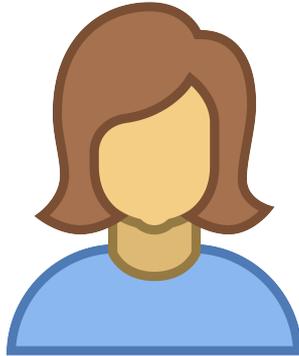
Meaning is co-constructed by the interlocutors and the **context** of the utterance



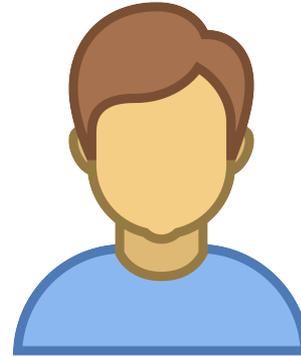
Whorfian view



“One morning I shot an elephant in my pajamas”



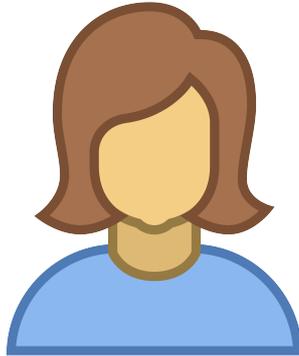
Weak relativism: structure of language influences thought



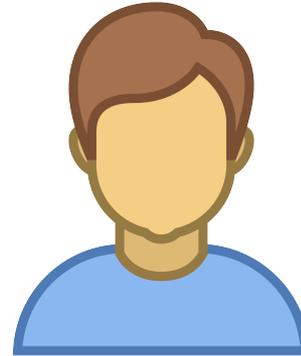
Whorfian view



一天早上我穿着睡衣射了一
只大象

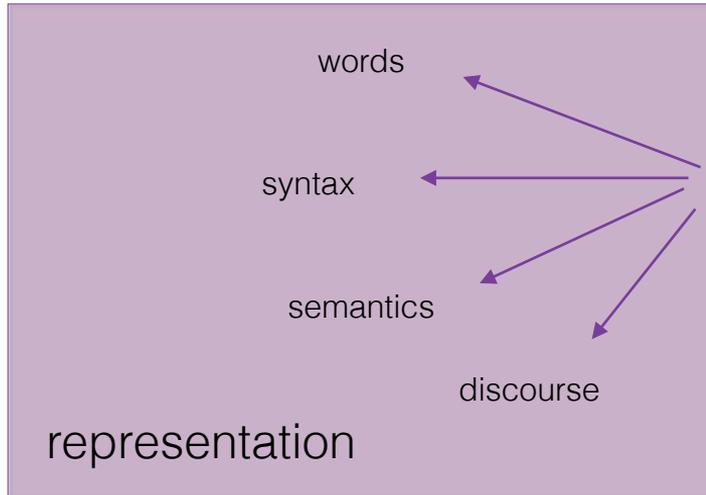


Weak relativism: structure of
language influences thought

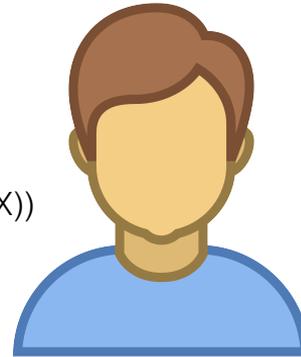


Decoding

“One morning I shot an elephant in my pajamas”



$\text{decode}(\text{encode}(X))$



discourse

semantics

syntax

morphology

words

Words

- One morning I shot an elephant in my pajamas
- I didn't shoot an elephant
- **Imma** let you finish but Beyonce had one of the best videos of all time
- 一天早上我穿着睡衣射了一只大象



How many words are in the sentence "I didn't shoot an elephant"?



The sentence "I didn't shoot an elephant" contains 7 words.



What are they?



The words in the sentence "I didn't shoot an elephant" are : "I", "didn't", "shoot", "an", "elephant"

Parts of speech

noun

verb

noun

noun

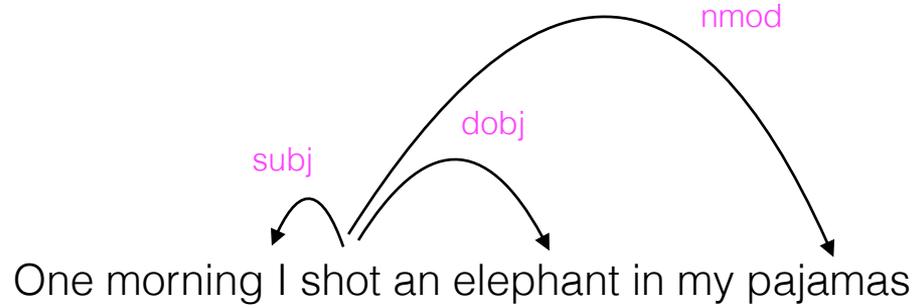
One morning I shot an elephant in my pajamas

Named entities

person

Imma let you finish but Beyonce had one of the best videos of all time

Syntax



Sentiment analysis



"Unfortunately I already had this exact picture tattooed on my chest, but **this shirt** is very useful in colder weather."

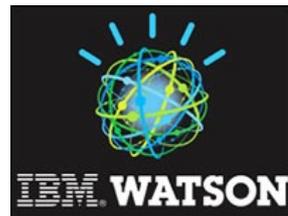
[overlook1977]

NLP

- Machine translation
- Question answering
- Information extraction
- Conversational agents
- Summarization



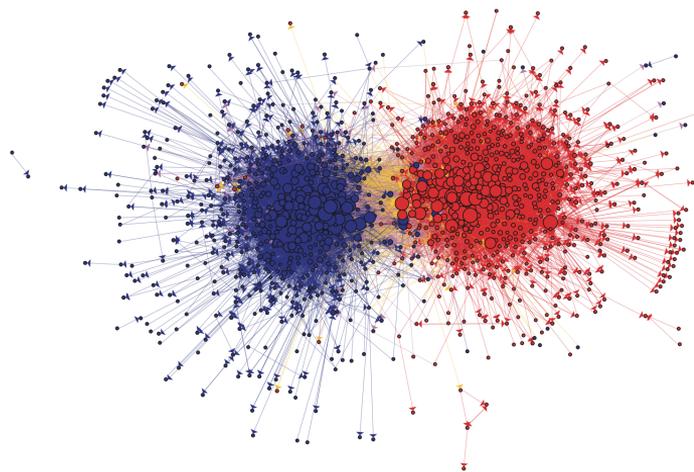
Google



NLP + X

Computational Social Science

- Inferring ideal points of politicians based on voting behavior, speeches
- Detecting the triggers of censorship in blogs/social media
- Inferring power differentials in language use



Link structure in political blogs
Adamic and Glance 2005

Computational Journalism

What do Journalists do with Documents? Field Notes for Natural Language Processing Researchers

Jonathan Stray
Columbia Journalism School
jms2361@columbia.edu

- Robust import
- Robust analysis
- Search, not exploration
- Quantitative summaries
- Interactive methods
- Clarity and Accuracy

Computational Humanities

Ted Underwood (2018), “Why Literary **Time** is Measured in Minutes”

Ryan Heuser, Franco Moretti, Erik Steiner (2016), The **Emotions** of London

Richard Jean So and **Hoyt Long** (2015), “Literary Pattern Recognition”

Ted Underwood, David Bamman and Sabrina Lee, The Transformation of **Gender** in English-Language Fiction (2018)

Franco Moretti (2005), Graphs, Maps, Trees

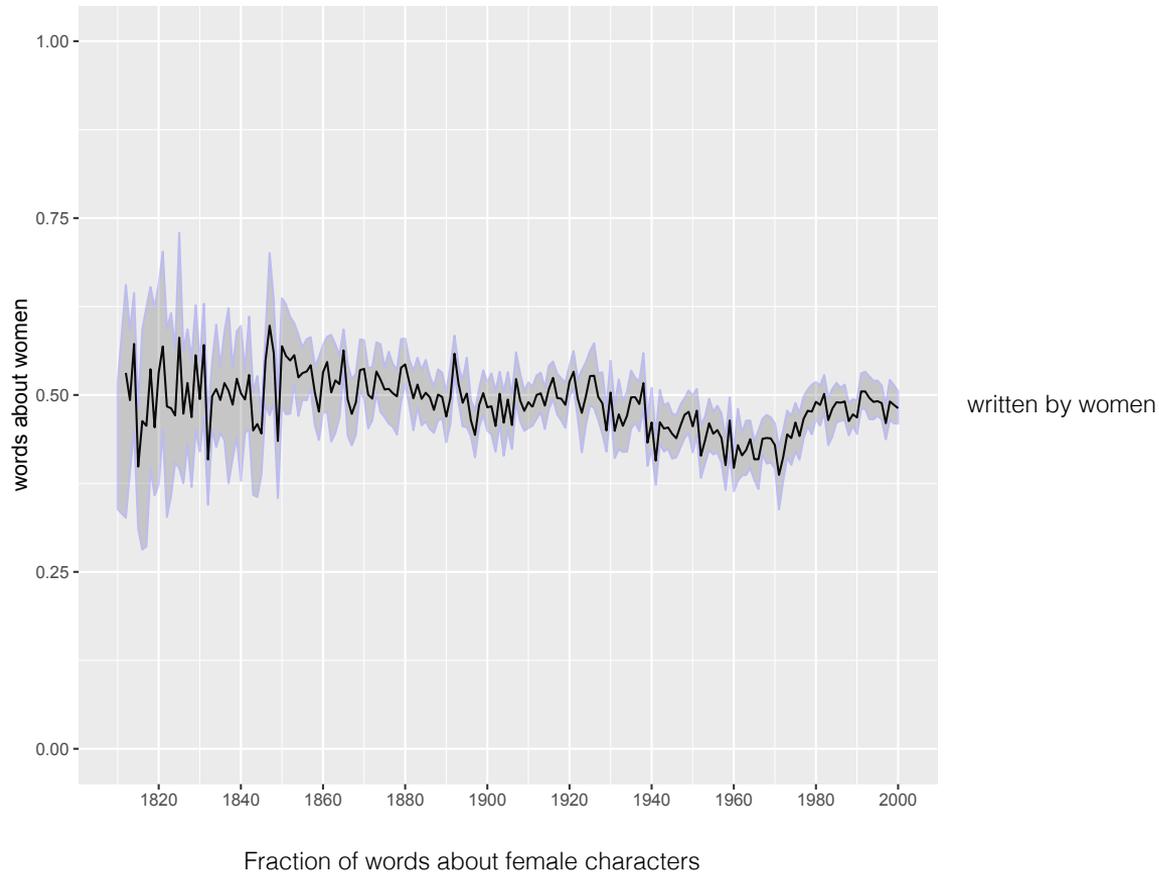
Holst Katsma (2014), **Loudness** in the Novel

So et al (2014), “**Cents** and Sensibility”

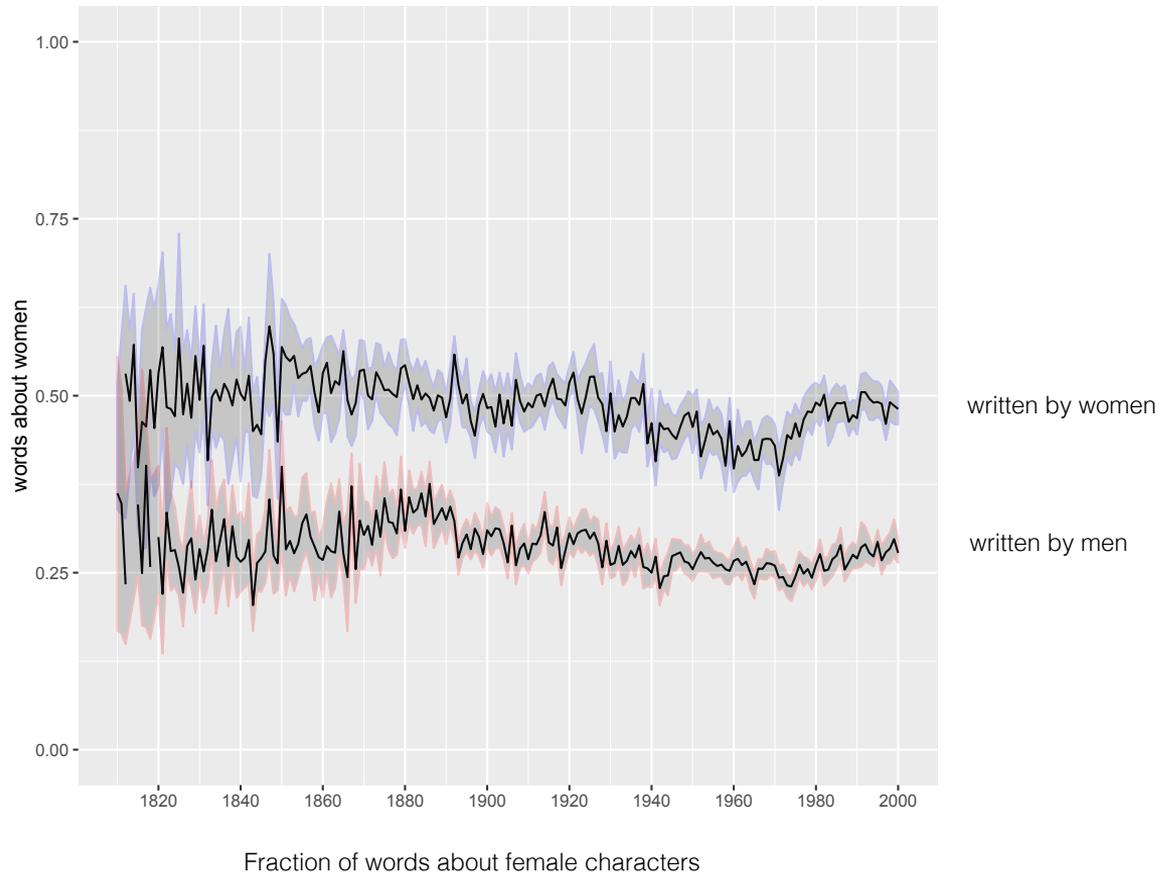
Matt Wilkens (2013), “The **Geographic** Imagination of Civil War Era American Fiction”

Jockers and Mimno (2013), “Significant **Themes** in 19th-Century Literature,”

Ted Underwood and Jordan Sellers (2012). “The Emergence of **Literary Diction**.” JDH



Ted Underwood, David Bamman, and Sabrina Lee (2018), "The Transformation of Gender in English-Language Fiction," *Cultural Analytics*



Ted Underwood, David Bamman, and Sabrina Lee (2018), "The Transformation of Gender in English-Language Fiction," *Cultural Analytics*

Methods

- Finite state automata/transducers (tokenization, morphological analysis)
- Rule-based systems

Methods

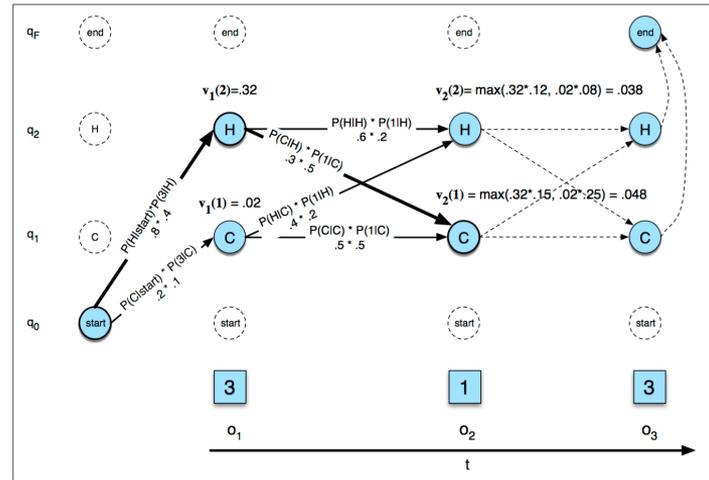
- Probabilistic models
- Naive Bayes, Logistic regression, HMM, MEMM, CRF, language models

$$P(Y = y|X = x) = \frac{P(Y = y)P(X = x|Y = y)}{\sum_y P(Y = y)P(X = x|Y = y)}$$

Methods

- Dynamic programming (combining solutions to subproblems)

Viterbi algorithm,
CKY



Methods

- Latent variable models (specifying probabilistic structure between variables and inferring likely latent values)

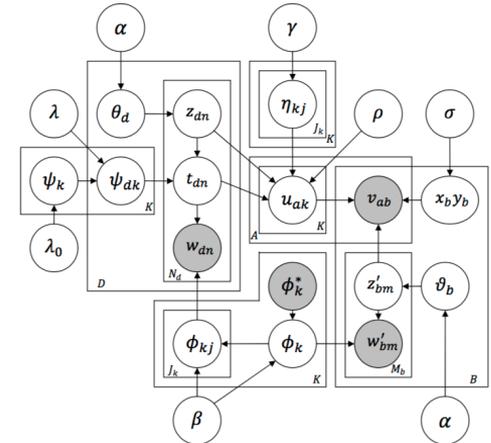
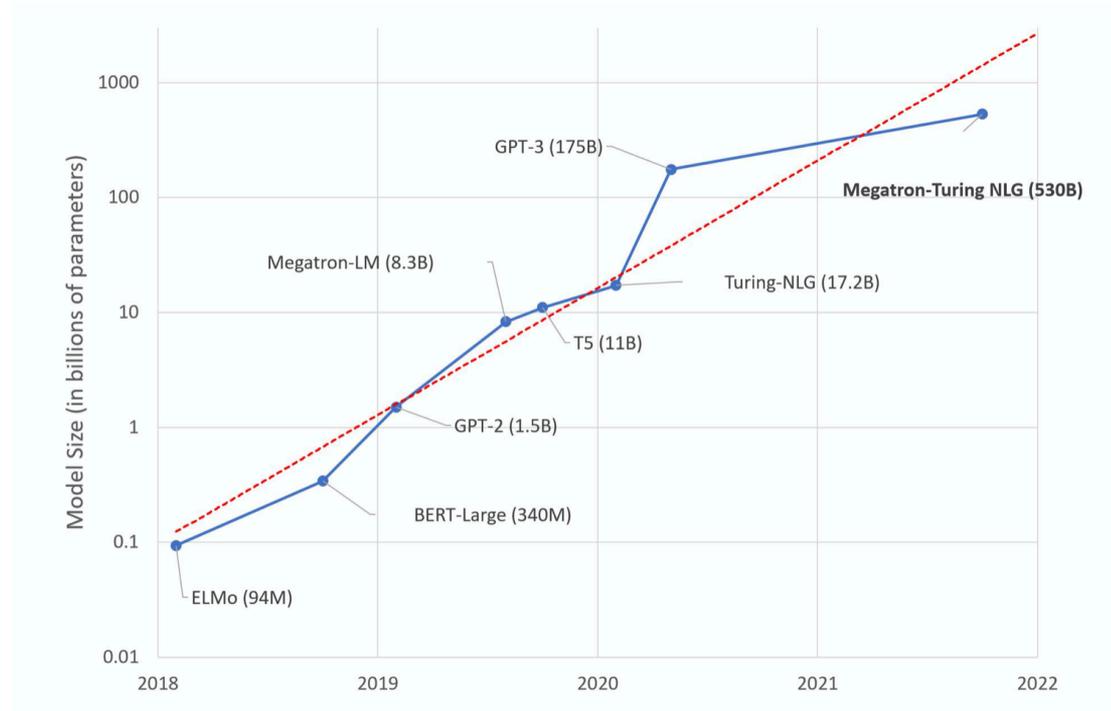


Figure 1: Plate notation diagram of HIPTM.

Methods

- Pretraining highly parametrized large language models.



Info 159/259

- This is a class about **models**.
 - You'll learn and implement algorithms to solve NLP tasks efficiently and understand the fundamentals to innovate new methods.
- This is a class about the **linguistic representation** of text.
 - You'll annotate texts for a variety of representations so you'll understand the phenomena you'll be modeling

Prerequisites

- Strong programming skills
 - Translate pseudocode into code (Python)
 - Analysis of algorithms (big-O notation)
- Basic probability/statistics
- Calculus

function VITERBI(*observations* of len T , *state-graph* of len N) **returns** *best-path*

create a path probability matrix $viterbi[N+2, T]$

for each state s **from** 1 **to** N **do** ; initialization step

$viterbi[s, 1] \leftarrow a_{0,s} * b_s(o_1)$

$backpointer[s, 1] \leftarrow 0$

for each time step t **from** 2 **to** T **do** ; recursion step

for each state s **from** 1 **to** N **do**

$viterbi[s, t] \leftarrow \max_{s'=1}^N viterbi[s', t-1] * a_{s',s} * b_s(o_t)$

$backpointer[s, t] \leftarrow \operatorname{argmax}_{s'=1}^N viterbi[s', t-1] * a_{s',s}$

$viterbi[q_F, T] \leftarrow \max_{s=1}^N viterbi[s, T] * a_{s,q_F}$; termination step

$backpointer[q_F, T] \leftarrow \operatorname{argmax}_{s=1}^N viterbi[s, T] * a_{s,q_F}$; termination step

return the backtrace path by following backpointers to states back in time from $backpointer[q_F, T]$

Grading

- Info 159:
 - Homeworks (25%)
 - Annotation project (25%)
 - Weekly quizzes (10%)
 - Midterm (20%)
 - NLP subfield survey (20%)

Annotation project

- This course covers many of the methods and existing tasks in NLP →
- But the most exciting applications of NLP have yet to be invented.
- Design a new NLP task and annotate data to support it, working in groups of **exactly 3 students**.

Existing tasks
Question answering
Named entity recognition
Sentiment analysis
Machine translation
Syntactic parsing
Coreference resolution
Text generation
Word sense disambiguation
...

Respect

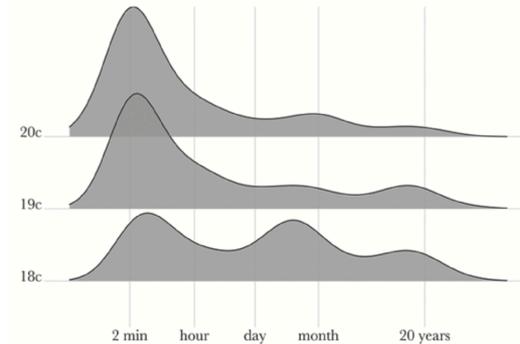
- Present one dialogue turn (police/driver) to be rated by people for respect (4-point Likert scale). High IAA.
- Build a predictive model mapping **text** to **respect**.

Voigt et al. 2017, "Language from police body camera footage shows racial disparities in officer respect"

EXAMPLE	RESPECT SCORE
<p>FIRST NAME ASK FOR AGENCY QUESTIONS</p> <p>[name], can I see that driver's license again? It- it's showing suspended. Is that- that's you?</p> <p>DISFLUENCY NEGATIVE WORD DISFLUENCY</p>	-1.07
<p>INFORMAL TITLE ASK FOR AGENCY ADVERBIAL "JUST"</p> <p>All right, my man. Do me a favor. Just keep your hands on the steering wheel real quick.</p> <p>"HANDS ON THE WHEEL"</p>	-0.51
<p>APOLOGY INTRODUCTION LAST NAME</p> <p>Sorry to stop you. My name's Officer [name] with the Police Department.</p>	0.84
<p>FORMAL TITLE SAFETY PLEASE</p> <p>There you go, ma'am. Drive safe, please.</p>	1.21
<p>ADVERBIAL "JUST" FILLED PAUSE REASSURANCE</p> <p>It just says that, uh, you've fixed it. No problem. Thank you very much, sir.</p> <p>GRATITUDE FORMAL TITLE</p>	2.07

Time

Underwood (2018), “Why Literary Time is Measured in Minutes”. Measuring how much time has passed in 250-word chunks of text.



Underwood (2018), “Why Literary Time Is Measured in Minutes”

Passage	Mins
I fear then, Emma, Sewell is a knave, and joined in mean collusion with his brother, to distress your husband, who looks upon him as his friend. You are deceived, Charles, I am sure he is Sir James's friend, and mine, by his perpetually dissuading him from play. It may be so; but tell me, Emma, all you know, and all you think of Lady Juliana's sudden departure, what can it mean? ...	5.0
At length we reached the gates of this noble edifice, and had the pleasure to find the family not retired to rest, by perceiving lights in the hall. ... In a few minutes all was hushed, and a man, whom I believed to be an upper servant, was sent to reconnoitre my person, and enquire my name and business. I told him I should not reveal either, but to his master. ...	15.0

Dogmatism

“I’m supposed to trust the opinion of a MS minion? The people that produced Windows ME, Vista and 8? They don’t even understand people, yet they think they can predict the behavior of new, self-guiding AI?” *–anonymous*

“I think an AI would make it easier for Patients to confide their information because by nature, a robot cannot judge them. Win-win? :D” *–anonymous*

Fast and Horvitz (2016), “Identifying Dogmatism in Social Media: Signals and Models”

Given a comment, imagine you hold a well-informed, different opinion from the commenter in question. **We’d like you to tell us how likely that commenter would be to engage you in a constructive conversation about your disagreement, where you each are able to explore the other’s beliefs.** The options are:

(5): It’s unlikely you’ll be able to engage in any substantive conversation. When you respectfully express your disagreement, they are likely to ignore you or insult you or otherwise lower the level of discourse.

(4): They are deeply rooted in their opinion, but you are able to exchange your views with- out the conversation degenerating too much.

(3): It’s not likely you’ll be able to change their mind, but you’re easily able to talk and understand each other’s point of view.

(2): They may have a clear opinion about the subject, but would likely be open to discussing alternative viewpoints.

(1): They are not set in their opinion, and it’s possible you might change their mind. If the comment does not convey an opinion of any kind, you may also select this option.

AP deliverables

- **AP1**. Design a new document classification task and gather data to support it (must be shareable with the public — nothing private or in copyright).
- **AP2**. Annotate the data, creating at least 1000 labeled examples + robust set of annotation guidelines, reporting inter-annotator agreement rates.
- **AP3**. In a separate assignment, a different group will annotate your data only using your annotation guidelines (are your guidelines comprehensive enough that an independent third party could reproduce your judgments?).
- **AP4**. Build a classifier to automatically predict the labels using the data you've annotated.

NLP subfield survey

- 4-page survey for a specific NLP subfield of your choice (e.g., coreference resolution, question answering, interpretability, narrative generation, etc.), synthesizing at least 25 papers published at ACL, EMNLP, NAACL, EACL, AACL, *Transactions of the ACL* or *Computational Linguistics*.
- This survey should be able to provide a newcomer (such as yourself at the start of the semester) a sense of the current state of the art in that subfield in 2023, the major historical papers that have defined that area, and the different schools of thought within it.

Grading

- Info 259:
 - Homeworks (20%)
 - Annotation project (20%)
 - Weekly quizzes (10%)
 - Midterm (20%)
 - Project (30%)

259 Project

- Semester-long project (involving 1-3 students) involving natural language processing -- either focusing on core NLP methods or using NLP in support of an empirical research question
 - Project proposal/literature review
 - Midterm report
 - 6-page final report, *workshop quality*
 - Poster presentation

ACL 2023 workshops

- The 18th Workshop on Innovative Use of NLP for Building Educational Applications (BEA)
- The 7th Workshop on Online Abuse and Harms
- The 17th Workshop on Linguistic Annotation (LAW)
- The 22nd Workshop on Biomedical Natural Language Processing (BioNLP)
- The 5th Workshop on NLP for Conversational AI
- The 3rd Workshop on Trustworthy NLP
- The 5th Clinical Natural Language Processing Workshop (Clinical NLP)
- The 3rd Workshop on NLP for Indigenous Languages of the Americas
- The 5th Workshop on Narrative Understanding

Exams

- We'll have two exams:
 - Midterm 1 (2/23, 2-3:30pm, [remote](#)).
 - Midterm 2 (3/23, 2-3:30pm, [remote](#)).
- We will [not](#) be offering alternative exam dates, so if you anticipate a conflict, don't take this class!
- Your midterm grade will be the $\max(\text{midterm 1}, \text{midterm 2})$ — you will drop your lowest-scoring midterm grade.

Late submissions

- All homeworks and quizzes are due on the date/time specified.
- You have 3 late days total over the semester to use when turning in homeworks/quizzes (not group annotation project deliverables or 259 project deliverables); each day extends the deadline by 24 hours. If all late days have been used up, homeworks/quizzes can be turned in up to 48 hours late for 50% credit; anything submitted after 48 hours late = 0 credit.
- Late days are assessed immediately once homeworks or quizzes are submitted late and can't be retroactively changed (if you submit 2 homeworks and 2 quizzes late, for example, you can't decide after the fact which ones to apply your 3 slip days to -- they apply to whichever homeworks or quizzes use them up first).

Academic integrity

- We'll follow the UC Berkeley code of conduct
<http://sa.berkeley.edu/code-of-conduct>
- You may discuss homeworks at **a high level** with your classmates (if you do, include their names on the submission), but each homework deliverable must be completed **independently** -- all writing and code must be your own; and all quizzes and exams must be completed independently.

Academic integrity

- If you mention the work of others, you must be clear in citing the appropriate source:
<http://gsi.berkeley.edu/gsi-guide-contents/academic-misconduct-intro/plagiarism/>
- This holds for source code as well: if you use others' code (e.g., from StackOverflow), you must cite its source.
- We have zero tolerance policy for cheating and plagiarism; violations will be referred to the Center for Student Conduct and will likely result in failing the class.

AI Assistants

- If you use the output of automatic writing assistants (e.g. ChatGPT) or code suggestions (e.g. Copilot) **you must cite the source and be clear what code/text came from it** — just as you would with anything else you did not create yourself.
- You retain responsibility for anything you submit and should be prepared to demonstrate your understanding of it
- Be honest about your use of these tools; if multiple students submit the same text/code, easy to find plagiarism.
- The work you submit should largely be your own, representing your own ideas, code and words; submissions that rely too heavily on AI tools will not be graded favorably.

Curve

Grades in this class will **not** be curved.

Lectures

- Recordings of lectures will be available on bCourses through course capture.
- Attendance is not required for lectures but is highly encouraged!

Ed Discussion

- We'll use Ed Discussion as a platform for asking and answering questions about the course material, including homeworks.
- Students are encouraged to actively participate on this forum and help others by answering questions that arise (helpful students can see a grade bump across a threshold (e.g., B+ to A-) for this participation).
- When helping with homework questions, keep the discussion to the high-level concepts; don't post answers to homeworks or quiz/exam questions.

TAs

- Kent Chang
- Jingshu Rui
- Tim Schott
- Aryia Dattamajumdar
- Rachel McCarty
- Divya Tadimeti
- Nancy Xu
- Visit TA office hours for help with homeworks/quizzes/exams or just to chat about NLP.
- Office hours start next week.

TAs

- Keep academic integrity in mind during TA office hours: you may discuss homework questions at a high level with others present, but don't discuss specific answers or share screens with code solutions. Neither the TA office hours nor Ed Discussion should be used for pre-grading (asking if a specific answer to a homework or quiz question is correct before the assignment is due).

DB office hours

- DB office hours Wednesday:
 - 10-11am, in person (314 South Hall)
 - 11am-noon on Zoom (link on bCourses)
- Come talk to me to discuss concepts from class and NLP more generally — I'm happy to chat!



Next time:

Lexical semantics/static word embeddings