Epilogue

What did we cover?

- Language Modeling:
  - Markov language models, neural language models: feed-forward, RNN, Transformers
- Classification:
  - The basics: discriminative (log-linear) and generative models (naive Bayes)
  - Sequence labeling: HMM, MEMM, CRF for POS tagging and NER mostly
- Text Classification and Sentiment Analysis
- Grammar
  - What is grammar?
  - Syntactic Parsing: graph-based, transition-based

- Broad-coverage semantic analysis:
  - Information Extraction
  - Indirect supervision
  - Semantic role labeling
  - Syntax and semantics
- Machine Translation (MT)
  - Nature of the task
  - Neural MT
- Contextualized word embeddings
  - Attention and self-attention
  - The Transformer model
  - Pretraining
  - Masked language models

Many More Topics to Cover

- Semantics:
  - Lexical Semantics
  - Sentence semantics: we touched on this, but much more to cover
  - Executable semantic forms: text to code, personal assistants (connection to apps and APIs)
  - Discourse Structure: the relation between sentences
  - Question answering

Complex question: What are the name and budget of the departments with average instructor salary greater than the overall average?
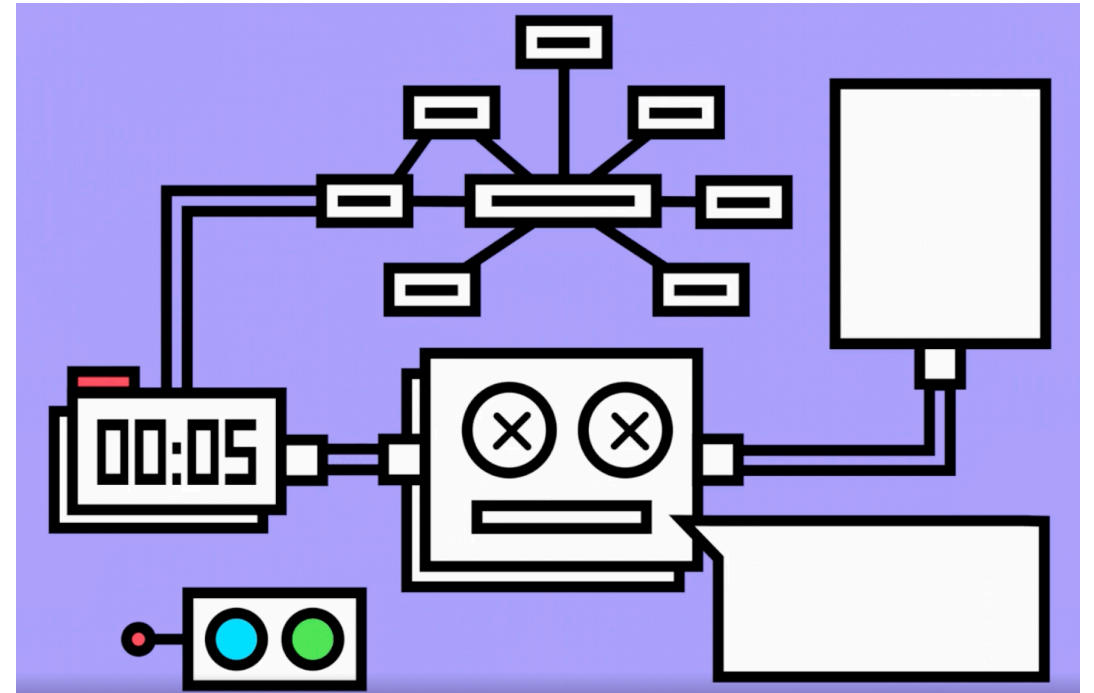
Complex SQL:
```
SELECT T2.name, T2.budget
FROM instructor as T1 JOIN department as
T2 ON T1.department_id = T2.id
GROUP BY T1.department_id
HAVING avg(T1.salary) >
    (SELECT avg(salary) FROM instructor)
```

https://arxiv.org/pdf/1809.08887.pdf

Many More Topics to Cover

- Dialogue systems: one of the most important frontiers of the field, both theoretically and practically
  - Chit-chat
  - Task-oriented dialogue: sales agents, customer support
  - Assistance
  - Debate

https://www.theverge.com/2018/6/13/17453994/amazon-

https://www.nature.com/articles/s41586-021-03215-w

Many More Topics to Cover

- Deep learning for NLP

- Resources

- Evaluation

- Cognitive modeling: psycholinguistics, computational neuroscience

- …

NLP: Some Big Questions

- Universal semantic metalanguage

- How do languages differ from one another?

- Multi-modal systems (e.g., visual, auditory, tactile)

- Implicit Information (implicit claims, implicit participants, sub-text)

- Subjective text (answering non-factual questions, understanding nuances, generating them, aesthetic aspects of language)

What is the Role of Linguistic Structure in Present Day NLP?

- Say what you will about neural language models, their output is mostly fluent

- What happens if we take a closer look?
  - Slice it by different syntactic phenomena
  - This is very much an open question

The BLiMP Benchmark

Phenomenon	UID	5-gram	LSTM	TXL	GPT-2	Human	Acceptable Example	Unacceptable Example
ANAPHOR AGREEMENT	anaphor_gender_agreement	44	88	91	99	96	Katherine can't help **herself**.	Katherine can't help **himself**.
	anaphor_number_agreement	52	95	97	100	99	Many teenagers were helping **themselves**.	Many teenagers were helping **herself**.
ARGUMENT STRUCTURE	animate_subject_passive	54	68	58	77	98	Amanda was respected by some **waitresses**.	Amanda was respected by some **picture**.
	animate_subject_trans	72	79	70	80	87	Danielle **visited** Irene.	The eye **visited** Irene.
	causative	51	65	54	68	82	Aaron breaks the glass.	Aaron appeared the glass.
	drop_argument	68	79	67	84	90	The Lutherans couldn't skate around.	The Lutherans couldn't disagree with.
	inchoative	89	72	81	90	95	A screen was fading.	A screen was cleaning.
	intransitive	82	73	81	90	86	Some glaciers are vaporizing.	Some glaciers are scaring.
	passive_1	71	65	76	89	99	Jeffrey's sons are insulted by Tina's supervisor.	Jeffrey's sons are smiled by Tina's supervisor.
	passive_2	70	72	74	79	86	Most cashiers are disliked.	Most cashiers are flirted.
	transitive	91	87	89	49	87	A lot of actresses' nieces have toured **that** art gallery.	A lot of actresses' nieces have coped **that** art gallery.
BINDING	principle_A_c_command	58	59	61	100	86	A lot of actresses that thought about Alice healed **themselves**.	A lot of actresses that thought about Alice healed **herself**.
	principle_A_case_1	100	100	100	96	98	Tara thinks that **she** sounded like Wayne.	Tara thinks that **herself** sounded like Wayne.
	principle_A_case_2	49	87	95	73	96	Stacy imagines herself **praising** this actress.	Stacy imagines herself **praises** this actress.
	principle_A_domain_1	95	98	99	99	95	Carlos said that Lori helped **him**.	Carlos said that Lori helped **himself**.
	principle_A_domain_2	56	68	70	73	75	Mark imagines Erin might admire **herself**.	Mark imagines Erin might admire **himself**.
	principle_A_domain_3	52	55	60	82	83	Nancy could say every guy hides **himself**.	Every guy could say Nancy hides **himself**.
	principle_A_reconstruction	40	46	38	37	78	It's herself who Karen criticized.	It's herself who criticized Karen.

BLiMP: The Benchmark of Linguistic Minimal Pairs for English
Warstadt et al., TACL 2020

Do Neural Language Models Implicitly Learn Grammar?

- Current answer: depends on how big your model is and who you ask

- Can we just read the representation that these networks ascribe to sentence from the network activations?
  - How?
  - Does grammar have a role in the next phase of NLP?

- Big questions arise…
  - What part of grammar do neural LMs capture?
  - Do they have systematic biases?
  - Are distributional signals sufficient to learn grammar?
  - Do findings generalize cross-linguistically?
  - Does interpretability matter?

Courses

- In NLP:
    - Advanced course in NLP
    - Seminars by Roy Schwartz, Gabi Stanovsky and myself
    - Advanced students are welcome to join our reading group
- Other related courses:
    - Introduction to Information Theory
    - Machine learning courses
    - Probabilistic methods in artificial intelligence
    - Deep Learning, Cognition and Brain (a course in Cog-Sci)