Website and Waitlist!

- [https://csci-1460-computational-linguistics.github.io](https://csci-1460-computational-linguistics.github.io)
- Also linked from my webpage (Google “Ellie Pavlick”, go to my cs.brown.edu page, click on “Teaching” at the top)
- Sign up for waitlist if you aren’t registered
  - But, unlikely to let many off the waitlist, unfortunately
NLP Intro: Brief History and Context
Lecture Outline

• What is NLP? Current Progress
• A brief history of the field
• NLP Under-the-Hood: The language processing pipeline
• Course Preview
NLP Intro: Brief History and Context
Lecture Outline

• What is NLP? Current Progress
• A brief history of the field
• NLP Under-the-Hood: The language processing pipeline
• Course Preview
What is NLP?

“The goal of this new field is to get computers to perform useful tasks involving human language, tasks like enabling human-machine communication, improving human-human communication, or simply doing useful processing of text or speech.”

—Jurafsky and Martin, Speech and Language Processing, 2nd Edition
What is NLP?
What is NLP?

Google autocomplete
What is NLP?

Natural language processing (NLP) is the ability of a computer program to understand human language as it is spoken and written — referred to as natural language. It is a component of artificial intelligence (AI). NLP has existed for more than 50 years and has roots in the field of linguistics.

Source: https://www.techtarget.com/searchenterpriseai/definition/What-is-Natural-Language-Processing-An-Introduction-to-NLP
What is Natural Language Processing (NLP)? An Introduction to NLP

People also ask:
- What is NLP and how does it work?
- What is NLP used for?
What is NLP?

Natural language processing (NLP) is the ability of a computer program to understand human language as it is spoken and written — referred to as natural language. It is a component of artificial intelligence (AI). NLP has existed for more than 50 years and has roots in the field of linguistics.

People also ask:
- What is NLP and how does it work?
- What is NLP used for?
What is NLP?

Natural language processing (NLP) is the ability of a computer program to understand human language as it is spoken and written -- referred to as natural language. It is a component of artificial intelligence (AI). NLP has existed for more than 50 years and has roots in the field of linguistics.

https://www.techtarget.com/searchenterpriseai/definition

What is Natural Language Processing? An Introduction to NLP

People also ask:
- What is NLP and how does it work?
- What is NLP used for?

See results about
- Natural language processing
- Neuro-linguistic programming

Ad targeting
- Natural language processing is a subfield of linguistics, ...
- Neuro-linguistic programming is a pseudoscientific approach to ...
What is NLP?
What is NLP?

Please review the latest university announcements and novel coronavirus information.
What is NLP?

Course Name: CSCI 1460
Course Description: This course is an introduction to systems programming and low-level computer architecture. We will explore how to write C/C++ code that interacts with the operating system in interesting ways, and how to write code that is both efficient and secure.

Prerequisites:
CSCI 1200 - Data Structures
CSCI 1201 - Data Structures Laboratory
CSCI 2300 - Computer Science I
CSCI 2301 - Computer Science I Laboratory

Course Length: 16 weeks

Course Materials:

Other Required Materials:
A computer with an Internet connection
Goals for the Semester

• Learn how to build actual NLP systems, and to understand how commercial systems work under the hood

• Understand key models and algorithms from NLP, including both “traditional” and new approaches to solving problems

• Understand the theoretical and intellectual context surrounding major technological advances or design decisions

• Think critically about interface between NLP systems and society

• Roughly: 25% Application, 50% Applied ML and Algorithms, 25% Linguistics/Theory
Who should take this course?

- Prerequisites: Just the intro sequence!
- However, NLP is a bit of everything:
  - Machine Learning
  - Deep Learning
  - Probability, Statistics, Linear Algebra, Calculus
  - Linguistics (Syntax, Semantics)
  - Formal Language Theory/Theory of Computation
- But I believe learning is a undirected, highly cyclic graph ;) so these are not prereqs
- If you have not taken the above, you might need to do extra reading/spend time in office hours some weeks
NLP Intro: Brief History and Context

Lecture Outline

• What is NLP? Current Progress
• A brief history of the field
• NLP Under-the-Hood: The language processing pipeline
• Course Preview
If there were machines which bore a resemblance to our bodies and imitated our actions as closely as possible for all practical purposes, we should still have two very certain means of recognizing that they were not real men. The first is that they could never use words, or put together signs, as we do in order to declare our thoughts to others. For we can certainly conceive of a machine so constructed that it utters words, and even utters words that correspond to bodily actions causing a change in its organs...But it is not conceivable that such a machine should produce different arrangements of words so as to give an appropriately meaningful answer to whatever is said in its presence, as the dullest of men can do...
NLP: A brief history
NLP: A brief history

Logic and Computation: Tarski, Church, Turing
Loosely: A function is “computable” iff it can be computed by a Turing Machine.
Logic and Computation: Tarski, Church, Turing

1936: Church-Turing Thesis

Loosely: A function is “computable” iff it can be computed by a Turing Machine

\( \lambda \)-calculus invented for defining computable functions
“I believe that in about fifty years' time it will be possible to programme computers, with a storage capacity of about $10^9$, to make them play the imitation game so well that an average interrogator will not have more than 70 percent chance of making the right identification after five minutes of questioning. … I believe that at the end of the century the use of words and general educated opinion will have altered so much that one will be able to speak of machines thinking without expecting to be contradicted.”
NLP: A brief history

🤔 Is natural language a “computable function”?

Logic and Computation: Tarski, Church, Turing


1936: Church-Turing Thesis

Turing Test
NLP: A brief history

Logic and Computation: Tarski, Church, Turing

- 1910
- 1930
- 1950
- 1970
- 1990
- 2010
- 2020

Formal Linguistics: Montague, Chomsky

- 1910
- 1930
- 1950
- 1970
- 1990
- 2010
- 2020

1936: Church-Turing Thesis

1965: Chomsky “Aspects of a Theory of Syntax”
“There is in my opinion no important theoretical difference between natural languages and the artificial languages of logicians; indeed I consider it possible to comprehend the syntax and semantics of both kinds of languages with a single natural and mathematically precise theory.”
Career [edit]

At the University of California, Berkeley, Montague earned a B.A. in Philosophy in 1950, an M.A. in Mathematics in 1953, and a Ph.D. in Philosophy in 1957, the latter under the direction of the mathematician and logician Alfred Tarski. Montague, one of Tarski's most accomplished American students, spent his entire career teaching in the UCLA Department of Philosophy, where he supervised the dissertations of Nino Cocchiarella and Hans Kamp.

Montague wrote on the foundations of logic and set theory, as would befit a student of Tarski. His Ph.D. dissertation, titled Contributions to the Axiomatic Foundations of Set Theory,[1] contained the first proof that all possible axiomatizations of the standard axiomatic set theory ZFC must contain infinitely many axioms. In other words, ZFC cannot be finitely axiomatized.

He pioneered a logical approach to natural language semantics that became known as Montague grammar. This approach to language has been especially influential among certain computational linguists—perhaps more so than among more traditional philosophers of language. In particular, Montague's influence lives on in grammar approaches like categorial grammar (such as Unification Categorial Grammar, Left-Associate Grammar, or Combinatory Categorial Grammar), which attempt a derivation of syntactic and semantic representation in tandem and the semantics of quantifiers, scope and discourse (Hans Kamp, a student of Montague's, co-developed Discourse Representation Theory).

Montague was an accomplished organist and a successful real estate investor. He died violently in his own home; the case is unsolved to this day. Anita Feferman and Solomon Feferman argue that he usually went to bars "cruising" and bringing people home with him.[2]

On the day that he was murdered, he brought home several people "for some kind of soirée", but they strangled him.[2]
NLP: A brief history

1910

1930

1950

1970

1990

2010

1936: Church-Turing Thesis

Logic and Computation: Tarski, Church, Turing

Formal Linguistics: Montague, Chomsky

1965: Chomsky “Aspects of a Theory of Syntax”
He also became involved in left-wing activism. Chomsky refused to pay half his taxes, publicly supported students who refused the draft, and was arrested while participating an anti-war teach-in outside the Pentagon. During this time, Chomsky co-founded the anti-war collective RESIST with Mitchell Goodman, Denise Levertov, William Sloane Coffin, and Dwight Macdonald. Although he questioned the objectives of the 1968 student protests, Chomsky gave many lectures to student activist groups and, with his colleague Louis Kampf, ran undergraduate courses on politics at MIT independently of the conservative-dominated political science department. When student activists campaigned to stop weapons and counterinsurgency research at MIT, Chomsky was sympathetic but felt that the research should remain under MIT’s oversight and limited to systems of deterrence and defense. In 1970 he visited southeast Asia to lecture at Vietnam’s Hanoi University of Science and Technology and toured war refugee camps in Laos. In 1973 he helped lead a committee commemorating the 50th anniversary of the War Resisters League.

Because of his anti-war activism, Chomsky was arrested on multiple occasions and included on President Richard Nixon's master list of political opponents. Chomsky was aware of the potential repercussions of his civil disobedience, and his wife began studying for her own doctorate in linguistics to support the family in the event of Chomsky's imprisonment or joblessness. Chomsky's scientific reputation insulated him from administrative action based on his beliefs.

His work in linguistics continued to gain international recognition as he received multiple honorary doctorates and delivered public lectures at the University of Cambridge, Columbia University (Woodbridge Lectures), and Stanford University. In his appearance in a 1971 debate with French continental philosopher Michel Foucault positioned Chomsky as a symbolic figure in modern philosophy. He continued to publish extensively on linguistics, producing Studies on Semantics in Generative Grammar (1972), an enlarged edition of Language and Mind (1972), and Reflections on Language (1975). In 1974 Chomsky became a corresponding fellow of the British Academy.

1965: Chomsky “Aspects of a Theory of Syntax”
NLP: A brief history

Logic and Computation: Tarski, Church, Turing

Formal Linguistics: Montague, Chomsky

1910: Church-Turing Thesis
1936: Church-Turing Thesis
1965: Chomsky "Aspects of a Theory of Syntax"

Linguists working on codifying syntax and semantics using precise representations and computable algorithms...
NLP: A brief history

1936: Church-Turing Thesis

1950: Turing Test

1956: Dartmouth Workshop

1965: Chomsky “Aspects of a Theory of Syntax”

Formal Linguistics: Montague, Chomsky

Logic and Computation: Tarski, Church, Turing
NLP: A brief history

1956: Dartmouth Workshop

“We propose that a 2-month, 10-man study of artificial intelligence be carried out during the summer of 1956 at Dartmouth College in Hanover, New Hampshire. The study is to proceed on the basis of the conjecture that every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it. An attempt will be made to find how to make machines use language, form abstractions and concepts, solve kinds of problems now reserved for humans, and improve themselves. We think that a significant advance can be made in one or more of these problems if a carefully selected group of scientists work on it together for a summer.”

McCarthy, Minsky, Shannon
NLP: A brief history

1936: Church-Turing Thesis

1956: Dartmouth Workshop

1962: ACL Founded

1965: Chomsky “Aspects of a Theory of Syntax”

Turing Test

Logic and Computation: Tarski, Church, Turing

Formal Linguistics: Montague, Chomsky

Early NLP: Mix of rule-based and info-theory methods
The Association for Computational Linguistics (ACL) is the premier international scientific and professional society for people working on computational problems involving human language, a field often referred to as either computational linguistics or natural language processing (NLP). The association was founded in 1962, originally named the Association for Machine Translation and Computational Linguistics (AMTCL), and became the ACL in 1968. Activities of the ACL include the holding of an annual meeting each summer and the sponsoring of the journal Computational Linguistics, published by MIT Press; this conference and journal are the leading publications of the field.

https://www.aclweb.org/portal/what-is-cl

1965: Chomsky “Aspects of a Theory of Syntax”
NLP: A brief history

1962: ACL Founded

These are the proceedings of the 17th Annual Meeting of the Association for Computational Linguistics. The field which began as a study of mechanical translation has broadened to include much more. The Program Committee tried to have as many areas represented here as possible. In the conference's sessions are papers on topics as varied as speech perception, the formal definition of knowledge representations, the reality of a cognitive model of spatial knowledge, and computer-assisted dialect adaption. At the same time, we have tried to highlight current trends. So there are several papers on topics like word-driven versus rule-driven parsers, graph representations of knowledge, dialogue models, and natural language access to databases.

1936: Church-Turing Thesis

1965: Chomsky “Aspects of a Theory of Syntax”
SYNTAX AND SEMANTICS

Automatic creation of an AIIN grammar from a transformational grammar. Perry Miller, Massachusetts Institute of Technology

String transformations in the REQUEST system. Warren J. Plath, IBM Research

A computational treatment of coordinate conjunctions. Carol Rase, New York University.

Toward formal solutions to philosophical problems. James Dunn, Princeton University.

An investigation of algorithmic translation procedures from standard semantic feature representation to predicate logic. Robert M. Harnish and Michael Houghtaling, University of Arizona.

Lana’s progress. Ernest von Glasersfeld, University of Georgia.

NATURAL LANGUAGE SYSTEMS

Computer as model and metaphor. Stephan Isard, University of Edinburgh.


Linguistics and graphics: a compatible system. Christina Davis and Jeffrey Eastman, North Carolina State University.

CLET: A computer program that learns arithmetic from an elementary textbook. Sagib Sadre, IBM.


The believer system. Geoffrey Brown, Rutgers University.

Junction grammar as a base for automatic language processing. Eldon Lytle, Brigham Young University.

DISCOURSE AND COMPREHENSION

Simulating processes of verbalization and translation. Wallace Chafe, University of California at Berkeley

Episode understanding and belief guided parsing. Bertram Bruce and C. F. Schmidt, Rutgers University

Computer understanding of metaphorical phrases. Sylvia Russell

Understanding by conceptual inference. Charles Rieger, University of Maryland

Paraphrasing paragraphs. Roger Schank, Institute for Semantics-Cognition, Switzerland

Cybernetic model of conscious behavior. A nawim Vinje-Morpurgo

SPEECH PRODUCTION AND UNDERSTANDING

Simple digital speech synthesis. William Fisher and A. M. Engebretson, Central Institute for the Deaf, St. Louis, Missouri

Fundamental frequency contours of auxiliary phrases in English. Jonathan Allen and O. O’Shaughnessy, Massachusetts Institute of Technology


Computer testing of fast speech rules. Douglas R. Moran, University of Michigan

The role of lexical semantics in automated speech understanding. Tim Diller, System Development Corporation

An experiment in the use of iconic language to present graphic structures via the auditory channel. John B. Eulenberg and Mortiza Amir Rahimi, Michigan State University

Phonological rules for a text-to-speech system. Sharon Sennett, Massachusetts Institute of Technology
NLP: A brief history

1910

1920

1930

1936: Church-Turing Thesis

Logic and Computation: Tarski, Church, Turing

1940s

Formal Linguistics: Montague, Chomsky

1950

1956: Dartmouth Workshop

1962: ACL Founded

Early NLP: Mix of rule-based and information-theory methods

Linguists using computational tools to study theories of syntax/semantics

SYNTAX AND SEMANTICS

Automatic creation of an ANL grammar from a transformational grammar. Perry Miller, Massachusetts Institute of Technology

String transformations in the REQUEST system. Warren J. Plath, IBM Research

A computational treatment of coordinate conjunctions. Carol Rase, New York University

Toward formal solutions to philosophical problems. James Dunn, Princeton University

An investigation of algorithmic translation procedures from standard semantic feature representation to predicate logic. Robert M. Harnish and Michael Noughtaling, University of Arizona

Lana's progress. Ernst von Glasersfeld, University of Georgia

NATURAL LANGUAGE SYSTEMS

Computer as model and metaphor. Stephan Iard, University of Edinburgh

SOPHIE: A knowledge based CAI system using an analogical inference mechanism. Richard K. Burton, Bolt Beranek and Newman Inc

Linguistics and graphics: a compatible system. Christina Davis and Jeffrey Eastman, North Carolina State University

CLET: A computer program that learns arithmetic from an elementary textbook. Sagib Sadro, IBM

Construct. Robert Smith, F. L. Rawson, and Nancy Smith, Stanford University

The believer system. Geoffrey Brown, Rutgers University

Junction grammar as a base for automatic language processing. Eldon Lytle, Brigham Young University

DISCOURSE AND COMPREHENSION

Simulating processes of verbalization and translation. Wallace Chafe, University of California at Berkeley

Episode understanding and belief guided parsing. Bertram Bruce and C. F. Schmidt, Rutgers University

Computer understanding of metaphorical phrases. Sylvia Russell

Understanding by conceptual inference. Charles Rieser, University of Maryland

Paraphrasing paragraphs. Roger Schank, Institute for Semantics-Cognition, Switzerland

Cybernetic model of conscious behavior. Anawim Vinje-Morpurgo

SPEECH PRODUCTION AND UNDERSTANDING

Simple digital speech synthesis. William Fisher and A. M. Engebretson, Central Institute for the Deaf, St. Louis, Missouri

Fundamental frequency contours of auxiliary phrases in English. Jonathan Allen and O. O'Shaughnessy, Massachusetts Institute of Technology


Computer testing of fast speech rules. Douglas B. Moran, University of Michigan

The role of lexical semantics in automated speech understanding. Tim Diller, System Development Corporation

An experiment in the use of iconic language to present graphic structures via the auditory channel. John B. Eulenberg and Mortaza Amir Rahimi, Michigan State University

Phonological rules for a text-to-speech system. Sharon Nunnicutt, Massachusetts Institute of Technology
Linguists using computational tools to study theories of syntax/semantics

SYNTAX AND SEMANTICS
Automatic creation of an ANL grammar from a transformational grammar. Perry Miller, Massachusetts Institute of Technology
String transformations in the REQUEST system. Warren J. Flath, IBM Research
A computational treatment of coordinate conjunctions. Carol Kase, New York University.
Toward formal solutions to philosophical problems. James Dunn, Princeton University.
Luna's progress. Ernst von Glasersfeld, University of Georgia.

NATURAL LANGUAGE SYSTEMS
Computer as model and metaphor. Stephan Isard, University of Edinburgh.
Linguistics and graphics: a compatible system. Christina Davis and Jeffrey Eastman, North Carolina State University.
CLET: A computer program that learns arithmetic from an elementary textbook. Nagib Sadro, IBM.
The believer system. Geoffrey Brown, Rutgers University.
Junction grammar as a base for automatic language processing. Eldon Lytle, Brigham Young University.

1974
Program

DISCOURSE AND COMPREHENSION
Simulating processes of verbalization and translation. Wallace Chafe, University of California at Berkeley
Episode understanding and belief guided parsing. Bertram Bruce and C. F. Schmidt, Rutgers University
Computer understanding of metaphorical phrases. Sylvia Russell
Understanding by conceptual inference. Charles Rieger, University of Maryland
Paraphrasing paragraphs. Roger Schank, Institute for Semantics Cognition, Switzerland
Cybernetic model of conscious behavior. Anawim Vinje-Morpugo

SPEECH PRODUCTION AND UNDERSTANDING
Simple digital speech synthesis. William Fisher and A. N. Engebretson, Central Institute for the Deaf, St. Louis, Missouri
Fundamental frequency contours of auxiliary phrases in English. Jonathan Allen and D. O'Shaughnessy, Massachusetts Institute of Technology
Computer testing of fast speech rules. Douglas B. Moran, University of Michigan
The role of lexical semantics in automated speech understanding. Tim Diller, System Development Corporation
An experiment in the use of iconic language to present graphic structures via the auditory channel. John B. Eulenberg and Mortaza Amir Rahimi, Michigan State University
Giles for a text-to-speech system. Sharon Kannicutt, Institute of Technology

Early NLP: Mix of rule-based and info-theory methods
Linguists using computational tools to study theories of syntax/semantics

Engineers building systems to “do stuff with language”
Typically (not always):
Deterministic algorithms or “rule-based” approaches
Typically (not always):
Deterministic algorithms or "rule-based" approaches
NLP: A brief history

1910: Logic and Computation: Tarski, Church, Turing

1930: Formal Linguistics: Montague, Chomsky

1936: Church-Turing Thesis

1950: Turing Test

1956: Dartmouth Workshop

1962: ACL Founded

1965: Chomsky “Aspects of a Theory of Syntax”

1988: IBM Model 1

1990: Statistical NLP: “Traditional” ML, standardized evals

Early NLP: Mix of rule-based and info-theory methods
A STATISTICAL APPROACH TO LANGUAGE TRANSLATION

P. BROWN, J. COCKE, S. DELLA PIETRA, V. DELLA PIETRA, F. JELINEK, R. MERCER, and P. ROOSSIN

IBM Research Division
T.J. Watson Research Center
Department of Computer Science
P.O. Box 218
Yorktown Heights, N.Y. 10598

1936: Church-Turing Thesis
1965: Chomsky “Aspects of a Theory of Syntax”

Early NLP: Mix of rule-based and info-theory methods

Statistical NLP: “Traditional” ML, standardized evals

IBM Model 1

Logic and Computation: Tarski, Church, Turing

1962: ACL Founded

1988: IBM Model 1

1956: Dartmouth Workshop

1990

2010

2020
A STATISTICAL APPROACH TO LANGUAGE TRANSLATION

P. BROWN, J. COCKE, S. DELLA PIETRA, V. DELLA PIETRA
F. JELINEK, R. MERCER, and P. ROOSSIN

IBM Research Division
T.J. Watson Research Center
Department of Computer Science
P.O. Box 218
Yorktown Heights, N.Y. 10598

1936: Church-Turing Thesis

Turing Test

1956: Dartmouth Workshop

1962: ACL Founded

1965: Chomsky “Aspects of a Theory of Syntax”

1988: IBM Model 1

Early NLP: Mix of rule-based and info-theory methods

Statistical NLP: “Traditional” ML, standardized evals

1910

Logic and Computation: Tarski, Church, Turing

Turing Test

1970

1990

2010

2020
NLP: A brief history

• IBM Model 1 — Statistical alignment model that replaced hand-written dictionaries
• Kicked off the “statistical revolution” in NLP
  • Focus on more data rather than theory/linguistic expertise
  • Prioritization on empirical evaluation
  • Benchmarking: standardized metrics and test sets to enable stable tracking of progress over time
• 1996: First EMNLP ("Empirical Methods for Natural Language Processing")
NLP: A brief history

- 1910: Logic and Computation: Tarski, Church, Turing
- 1930: Logic and Computation: Tarski, Church, Turing
- 1936: Church-Turing Thesis
- 1950: Logic and Computation: Tarski, Church, Turing
- 1956: Dartmouth Workshop
- 1962: ACL Founded
- 1965: Chomsky “Aspects of a Theory of Syntax”
- 1966: Turing Test
- 1988: IBM Model 1
- 2012: AlexNet
- 2010: End-to-end deep learning
- 2020: Statistical NLP: “Traditional” ML, standardized evals
NLP: A brief history

1910: Logic and Computation: Tarski, Church, Turing
1930: Turing Test
1936: Church-Turing Thesis
1950: Formal Linguistics: Montague, Chomsky
1956: Dartmouth Workshop
1962: ACL Founded
1965: Chomsky “Aspects of a Theory of Syntax”
1968: Chomsky “Aspects of a Theory of Syntax”
1970: Early NLP: Mix of rule-based and info-theory methods
1988: IBM Model 1
1990: Statistical NLP: “Traditional” ML, standardized evals
2010: End-to-end deep learning
2012: AlexNet

Neural networks for cogsci and AI
NLP: A brief history

- Deep CNNs take over Computer Vision with AlexNet in 2012
- Resource improvements—GPUs, big data, auto-differentiation—make NNs viable
- Enters NLP a bit later…
  - 2013: word2vec for representing words
  - 2015: “end-to-end” MT with seq2seq models
NLP: A brief history

1910

Logic and Computation: Tarski, Church, Turing

1936: Church-Turing Thesis

1936: Church-Turing Thesis

1950

Turing Test

1956: Dartmouth Workshop

Formal Linguistics: Montague, Chomsky

1965: Chomsky “Aspects of a Theory of Syntax”

1962: ACL Founded

1988: IBM Model 1

1990

Statistical NLP: “Traditional” ML, standardized evals

Early NLP: Mix of rule-based and info-theory methods

2010

End-to-end deep learning

2018: ELMo

Pretraining and Transfer Learning

2012: AlexNet

Neural networks for cogsci and AI
NLP: A brief history

• Very large models trained without a specific task in mind
• Later adapted (ideally using a small amount of data) for specific tasks
NLP: A brief history

- **1936:** Church-Turing Thesis
- **1950:** Logic and Computation: Tarski, Church, Turing
- **1956:** Dartmouth Workshop
- **1962:** ACL Founded
- **1965:** Chomsky “Aspects of a Theory of Syntax”
- **1988:** IBM Model 1
- **1990:** Early NLP: Mix of rule-based and info-theory methods
- **2010:** You are here!
- **2012:** AlexNet
- **2018:** ELMo

**Key Events:**
- **1930s:** Logic and Computation
- **1950s:** Formal Linguistics
- **1960s:** Dartmouth Workshop, Chomsky’s seminal work
- **1970s:** Turing Test, ACL Founding
- **1980s:** IBM Model 1, Statistical NLP
- **1990s:** Early NLP advancements
- **2000s:** End-to-end deep learning
- **2010s:** Pretraining and Transfer Learning
NLP Intro: Brief History and Context

Lecture Outline

• What is NLP? Current Progress
• A brief history of the field
• **NLP Under-the-Hood: The language processing pipeline**
• Course Preview
The Language Processing Pipeline

phonology

lexicon

syntax

semantics

pragmatics

discourse
The Language Processing Pipeline

- phonology
- lexicon
- syntax
- semantics
- pragmatics
- discourse
The Language Processing Pipeline

- phonology
- lexicon
- syntax
- semantics
- pragmatics
- discourse
The Language Processing Pipeline

- phonology
- lexicon
- syntax
- semantics
- pragmatics
- discourse

Alexa! um can you get me up at uh six
The Language Processing Pipeline

- **phonology**
- **lexicon**
- **syntax**
- **semantics**
- **discourse**
- **pragmatics**

Automatic Speech Recognition (ASR)

(A&M 3e Ch. 26)

Alexa! um can you get me up at uh six
The Language Processing Pipeline

- **phonology**
- **lexicon**
- **syntax**
- **semantics**
- **discourse**
- **pragmatics**

Tokenization/preprocessing

```
['get', 'me', 'up', 'at', '6']
```

Alexa! um can you get me up at uh six
The Language Processing Pipeline

- **Phonology**
- **Lexicon**
- **Syntax**
- **Semantics**
- **Discourse**
- **Pragmatics**

**Tokenization/Preprocessing**

- Alexa! um can you get me up at uh six
  - ['get', 'me', 'up', 'at', '6']

  Sometimes combine words instead!
  - ['get', 'me', 'up', 'at', '6']
  - ['hot', 'dogs', 'and', 'impossible', 'burgers'] -> ['hot_dogs', 'and', 'impossible_burgers']
The Language Processing Pipeline

[Image of a diagram with the following stages: phonology, lexicon, syntax, semantics, pragmatics, discourse.]

The input sentence is: 

Alexa! um can you get me up at uh six

This is mapped to the list: ['get', 'me', 'up', 'at', '6']

This list is then mapped to "meanings".
The Language Processing Pipeline

The processing pipeline includes:
- **Phonology**
- **Lexicon**
- **Syntax**
- **Semantics**
- **Discourse**
- **Pragmatics**

### Lexicon

[Alexa!] “um can you get me up at uh six”

- **Map to “meanings”**
  - Lookup in a database (e.g., WordNet, PropBank, custom “intent” database)

### Example Sentences

**WordNet Search - 3.1**

- **Nouns**
  - get (a return on a shot that seemed impossible to reach and would normally have resulted in a point for the opponent)

- **Verbs**
  - get (v), acquire (come into the possession of something concrete or abstract): “She got a lot of paintings from her uncle”, “They acquired a new pet”; “Get your results the next day”; “Get permission to take a few days off from work”
  - become (v), get (enter or assume a certain state or condition): “He became annoyed when he heard the bad news”; “It must be getting more serious”; “her face went red with anger”; “She went into ecstasy”
  - get (v), have (cause to move; cause to be in a certain condition): “He got his squad on the ball”; “This let me in for trouble”; “He got a girl into trouble”
  - get (v), receive, get, find, obtain, incur (receive a specified thing): “These aspects of civilization do not find expression or read interpretation”; “His movie received a good review”; “I got into trouble for my good intentions”
  - get (v), arrive, get, come (reach a destination; arrive by moving)

- **Sentences**
  - get_music(song)
  - set_time(time)
  - set_temp(temperature)
  - check_weather(day, location)
The Language Processing Pipeline

- phonology
- lexicon
- syntax
- semantics
- discourse
- pragmatics

[Image: WordNet Search 3.1]

Word to search for: get
Search WordNet

Display Options: [Select option to change]  Change

Key: "S." = Show Synset (semantic) relations, "W." = Show Word (lexical) relations
Display options for sense: (gloss) "an example sentence"

Noun
- S: (n) get (a return on a shot that seemed impossible to reach and would normally have resulted in a point for)

Verb
- S: (v) get, acquire (come into the possession abstract) "She got a lot of paintings out of it"; "Get your results the next day from work"
- S: (v) become, go, get (enter or assimilate) "She became more serious"; "Her face went red with an expression"
- S: (v) get, let, have (cause to move; condition) "He got his squad on the move"
- S: (v) receive, get, find, obtain, get into "These aspects of civilization do not get into interpretation"; "His movie received the highest praise for good intentions"
- S: (v) arrive, get, come (reach a destination) "She arrived home at 7 o'clock"; "She arrived at midnight"

[Image: Vector representation of word embeddings]

[Text: ['get', 'me', 'up', 'at', '6']]

map to "meanings"

map onto vectors
The Language Processing Pipeline

phonology

lexicon

syntax

semantics

pragmatics
discourse

Alexa! um can you get me up at uh six

[‘get’, ‘me’, ‘up’, ‘at’, ‘6’]

<table>
<thead>
<tr>
<th>Action</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>play</td>
<td>play_music(song)</td>
</tr>
<tr>
<td>set</td>
<td>set_time(time)</td>
</tr>
<tr>
<td>set</td>
<td>set_temp(temperature)</td>
</tr>
<tr>
<td>weather</td>
<td>check_weather(day, location)</td>
</tr>
</tbody>
</table>
The Language Processing Pipeline

phonology

lexicon

syntax

semantics

discourse

pragmatics

['get', 'me', 'up', 'at', '6']

play

play_music(song)

set

set_time(time)

set

set_temp(temperature)

weather

check_weather(day, location)
The Language Processing Pipeline

- **Phonology**
- **Lexicon**
- **Syntax**
- **Semantics**
- **Pragmatics**
- **Discourse**

```
Alexa! um can you get me up at uh six
```

- **Play**
  - `play_music(song)`
- **Set**
  - `set_time(time)`
  - `set_temp(temperature)`
- **Weather**
  - `check_weather(day, location)`

Diagram showing the processing steps and the breakdown of the spoken command into its constituent parts.
Alexa! um can you get me up at uh six

set_time(6)

[‘get’, ‘me’, ‘up’, ‘at’, ‘6’]

play
play_music(song)

set
set_time(time)

set
set_temp(temperature)

weather
check_weather(day, location)
The Language Processing Pipeline

Alexa! um can you get me up at uh six

set_time(6am)

depends on info about the speaker (time, location) and requires some “common sense”

<table>
<thead>
<tr>
<th>play</th>
<th>play_music(song)</th>
</tr>
</thead>
<tbody>
<tr>
<td>set</td>
<td>set_time(time)</td>
</tr>
<tr>
<td>set</td>
<td>set_temp(temperature)</td>
</tr>
<tr>
<td>weather</td>
<td>check_weather(day, location)</td>
</tr>
</tbody>
</table>
Alexa! When is Cora’s doctors appointment? Can you get her up an hour before?

set_time(6am)

[‘get’, ‘me’, ‘up’, ‘at’, ‘6’]

<table>
<thead>
<tr>
<th>play</th>
<th>play_music(song)</th>
</tr>
</thead>
<tbody>
<tr>
<td>set</td>
<td>set_time(time)</td>
</tr>
<tr>
<td>set</td>
<td>set_temp(temperature)</td>
</tr>
<tr>
<td>weather</td>
<td>check_weather(day, location)</td>
</tr>
</tbody>
</table>
NLP Intro: Brief History and Context
Lecture Outline

• What is NLP? Current Progress
• A brief history of the field
• NLP Under-the-Hood: The language processing pipeline
• Course Preview
<table>
<thead>
<tr>
<th>Lecture #</th>
<th>Date</th>
<th>Topic</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>9 Sept</td>
<td>Intro Lecture</td>
</tr>
<tr>
<td>1</td>
<td>13 Sept</td>
<td>Tasks and Metrics</td>
</tr>
<tr>
<td>2</td>
<td>15 Sept</td>
<td>Text Classifiers Part 1: Models</td>
</tr>
<tr>
<td>3</td>
<td>20 Sept</td>
<td>Text Classifiers Part 2: Features</td>
</tr>
<tr>
<td>4</td>
<td>22 Sept</td>
<td>Tokenization</td>
</tr>
<tr>
<td>5</td>
<td>29 Sept</td>
<td>No Class</td>
</tr>
<tr>
<td>6</td>
<td>27 Sept</td>
<td>Distributional Hypothesis</td>
</tr>
<tr>
<td>7</td>
<td>4 Oct</td>
<td>Topic Modeling</td>
</tr>
<tr>
<td>8</td>
<td>6 Oct</td>
<td>Deep Learning 101</td>
</tr>
<tr>
<td>9</td>
<td>11 Oct</td>
<td>Language Modeling Part 1: Ngram Models</td>
</tr>
<tr>
<td>10</td>
<td>13 Oct</td>
<td>Language Modeling Part 2: Neural Architectures</td>
</tr>
<tr>
<td>12</td>
<td>20 Oct</td>
<td>Guest Lecture (Julia Netter): Risks of LM Misuse</td>
</tr>
<tr>
<td>13</td>
<td>25 Oct</td>
<td>Statistical Machine Translation</td>
</tr>
<tr>
<td>14</td>
<td>27 Oct</td>
<td>Neural Machine Translation</td>
</tr>
<tr>
<td>15</td>
<td>1 Nov</td>
<td>Guest Lecture (Dan Garrette): Technical and Societal Challenges of Multilingual NLP</td>
</tr>
<tr>
<td>16</td>
<td>3 Nov</td>
<td>Part of Speech Tagging</td>
</tr>
<tr>
<td>17</td>
<td>8 Nov</td>
<td>No Class: Election Day!</td>
</tr>
<tr>
<td>18</td>
<td>10 Nov</td>
<td>Formal Grammars and Constituency Parsing</td>
</tr>
<tr>
<td>19</td>
<td>15 Nov</td>
<td>Dependency Grammars and Parsing</td>
</tr>
<tr>
<td>20</td>
<td>17 Nov</td>
<td>Semantic Representations and Semantic Parsing</td>
</tr>
<tr>
<td>21</td>
<td>22 Nov</td>
<td>No Class: Thanksgiving</td>
</tr>
<tr>
<td>22</td>
<td>24 Nov</td>
<td>No Class: Thanksgiving</td>
</tr>
<tr>
<td>19</td>
<td>20 Dec</td>
<td>Research Trends 1: Linguistic Structure in the Age of Deep Learning</td>
</tr>
<tr>
<td>20</td>
<td>1 Dec</td>
<td>Research Trends 2: Grounding and Multimodal NLP</td>
</tr>
<tr>
<td>21</td>
<td>6 Dec</td>
<td>Last Lecture: State-of-the-art and Open Questions</td>
</tr>
<tr>
<td>Lecture #</td>
<td>Date</td>
<td>Topic</td>
</tr>
<tr>
<td>-----------</td>
<td>----------</td>
<td>-------------------------------------------------</td>
</tr>
<tr>
<td>0</td>
<td>9 Sept</td>
<td>Intro Lecture</td>
</tr>
<tr>
<td>1</td>
<td>13 Sept</td>
<td>Tasks and Metrics</td>
</tr>
<tr>
<td>2</td>
<td>15 Sept</td>
<td>Text Classifiers Part 1: Models</td>
</tr>
<tr>
<td>3</td>
<td>20 Sept</td>
<td>Text Classifiers Part 2: Features</td>
</tr>
<tr>
<td>4</td>
<td>22 Sept</td>
<td>Tokenization</td>
</tr>
<tr>
<td>5</td>
<td>27 Sept</td>
<td>No Class</td>
</tr>
<tr>
<td>6</td>
<td>4 Oct</td>
<td>No Class</td>
</tr>
<tr>
<td>7</td>
<td>6 Oct</td>
<td>Distributional Hypothesis</td>
</tr>
<tr>
<td>8</td>
<td>11 Oct</td>
<td>Deep Learning 101</td>
</tr>
<tr>
<td>9</td>
<td>13 Oct</td>
<td>Language Modeling Part 1: Ngram Models</td>
</tr>
<tr>
<td>10</td>
<td>15 Oct</td>
<td>Language Modeling Part 2: Neural Architectures</td>
</tr>
<tr>
<td>11</td>
<td>16 Oct</td>
<td>Language Modeling Part 3: Pretraining and Finetuning</td>
</tr>
<tr>
<td>12</td>
<td>20 Oct</td>
<td>Guest Lecture (Julia Netter): Risks of LM Misuse</td>
</tr>
<tr>
<td>13</td>
<td>25 Oct</td>
<td>Statistical Machine Translation</td>
</tr>
<tr>
<td>14</td>
<td>27 Oct</td>
<td>NEURAL MACHINE TRANSLATION</td>
</tr>
<tr>
<td>15</td>
<td>3 Nov</td>
<td>Guest Lecture (Dan Garrette): Technical and Societal Challenges of Multilingual NLP</td>
</tr>
<tr>
<td>16</td>
<td>8 Nov</td>
<td>Part of Speech Tagging</td>
</tr>
<tr>
<td>17</td>
<td>10 Nov</td>
<td>No Class: Election Day!</td>
</tr>
<tr>
<td>18</td>
<td>15 Nov</td>
<td>Formal Grammars and Constituency Parsing</td>
</tr>
<tr>
<td>19</td>
<td>17 Nov</td>
<td>Dependency Grammars and Parsing</td>
</tr>
<tr>
<td>20</td>
<td>22 Nov</td>
<td>No Class: Thanksgiving</td>
</tr>
<tr>
<td>21</td>
<td>24 Nov</td>
<td>No Class: Thanksgiving</td>
</tr>
<tr>
<td>22</td>
<td>29 Nov</td>
<td>No Class: Thanksgiving</td>
</tr>
<tr>
<td>23</td>
<td>1 Dec</td>
<td>No Class: Thanksgiving</td>
</tr>
<tr>
<td>24</td>
<td>6 Dec</td>
<td>No Class: Thanksgiving</td>
</tr>
<tr>
<td>25</td>
<td>9 Dec</td>
<td>No Class: Thanksgiving</td>
</tr>
<tr>
<td>26</td>
<td>12 Dec</td>
<td>No Class: Thanksgiving</td>
</tr>
<tr>
<td>27</td>
<td>15 Dec</td>
<td>No Class: Thanksgiving</td>
</tr>
<tr>
<td>28</td>
<td>18 Dec</td>
<td>No Class: Thanksgiving</td>
</tr>
<tr>
<td>29</td>
<td>21 Dec</td>
<td>No Class: Thanksgiving</td>
</tr>
<tr>
<td>30</td>
<td>24 Dec</td>
<td>No Class: Thanksgiving</td>
</tr>
<tr>
<td>31</td>
<td>27 Dec</td>
<td>No Class: Thanksgiving</td>
</tr>
<tr>
<td>32</td>
<td>30 Dec</td>
<td>No Class: Thanksgiving</td>
</tr>
<tr>
<td>Lecture #</td>
<td>Date</td>
<td>Topic</td>
</tr>
<tr>
<td>----------</td>
<td>---------</td>
<td>-------------------------------------------------</td>
</tr>
<tr>
<td>0</td>
<td>9 Sept</td>
<td>Intro Lecture</td>
</tr>
<tr>
<td>1</td>
<td>13 Sept</td>
<td>Tasks and Metrics</td>
</tr>
<tr>
<td>2</td>
<td>15 Sept</td>
<td>Text Classifiers Part 1: Models</td>
</tr>
<tr>
<td>3</td>
<td>20 Sept</td>
<td>Text Classifiers Part 2: Features</td>
</tr>
<tr>
<td>4</td>
<td>22 Sept</td>
<td>Tokenization</td>
</tr>
<tr>
<td>5</td>
<td>29 Sept</td>
<td>No Class</td>
</tr>
<tr>
<td>6</td>
<td>27 Sept</td>
<td>Distributional Hypothesis</td>
</tr>
<tr>
<td>7</td>
<td>4 Oct</td>
<td>Topic Modeling</td>
</tr>
<tr>
<td>8</td>
<td>6 Oct</td>
<td>Deep Learning 101</td>
</tr>
<tr>
<td>9</td>
<td>11 Oct</td>
<td>Language Modeling Part 1: Ngram Models</td>
</tr>
<tr>
<td>10</td>
<td>13 Oct</td>
<td>Language Modeling Part 2: Neural Architectures</td>
</tr>
<tr>
<td>11</td>
<td>18 Oct</td>
<td>Language Modeling Part 3: Pretraining and Finetuning</td>
</tr>
<tr>
<td>12</td>
<td>20 Oct</td>
<td>Guest Lecture (Julia Netter): Risks of LM Misuse</td>
</tr>
<tr>
<td>13</td>
<td>25 Oct</td>
<td>Statistical Machine Translation</td>
</tr>
<tr>
<td>14</td>
<td>27 Oct</td>
<td>Neural Machine Translation</td>
</tr>
<tr>
<td>15</td>
<td>1 Nov</td>
<td>Guest Lecture (Dan Garrette): Technical and Societal Challenges of Multilingual NLP</td>
</tr>
<tr>
<td>16</td>
<td>3 Nov</td>
<td>Part of Speech Tagging</td>
</tr>
<tr>
<td>17</td>
<td>8 Nov</td>
<td>No Class: Election Day</td>
</tr>
<tr>
<td>18</td>
<td>10 Nov</td>
<td>Formal Grammars and Constituency Parsing</td>
</tr>
<tr>
<td>19</td>
<td>15 Nov</td>
<td>Dependency Grammars and Parsing</td>
</tr>
<tr>
<td>20</td>
<td>17 Nov</td>
<td>Semantic Representations and Semantic Parsing</td>
</tr>
<tr>
<td>21</td>
<td>22 Nov</td>
<td>No Class: Thanksgiving</td>
</tr>
<tr>
<td>22</td>
<td>24 Nov</td>
<td>No Class: Thanksgiving</td>
</tr>
<tr>
<td>23</td>
<td>29 Nov</td>
<td>Research Trends 1: Linguistic Structure in the Age of Deep Learning</td>
</tr>
<tr>
<td>24</td>
<td>1 Dec</td>
<td>Research Trends 2: Grounding and Multimodal NLP</td>
</tr>
<tr>
<td>25</td>
<td>6 Dec</td>
<td>Last Lecture: State-of-the-art and Open Questions</td>
</tr>
</tbody>
</table>
# Lectures

<table>
<thead>
<tr>
<th>Lecture #</th>
<th>Date</th>
<th>Topic</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>9 Sept</td>
<td>Intro Lecture</td>
</tr>
<tr>
<td>1</td>
<td>13 Sept</td>
<td>Tasks and Metrics</td>
</tr>
<tr>
<td>2</td>
<td>15 Sept</td>
<td>Text Classifiers Part 1: Models</td>
</tr>
<tr>
<td>3</td>
<td>20 Sept</td>
<td>Text Classifiers Part 2: Features</td>
</tr>
<tr>
<td>4</td>
<td>22 Sept</td>
<td>Tokenization</td>
</tr>
<tr>
<td>5</td>
<td>27 Sept</td>
<td>No Class</td>
</tr>
<tr>
<td>6</td>
<td>4 Oct</td>
<td>Distributional Hypothesis</td>
</tr>
<tr>
<td>7</td>
<td>6 Oct</td>
<td>Topic Modeling</td>
</tr>
<tr>
<td>8</td>
<td>11 Oct</td>
<td>Deep Learning 101</td>
</tr>
<tr>
<td>9</td>
<td>13 Oct</td>
<td>Language Modeling Part 1: Ngram Models</td>
</tr>
<tr>
<td>10</td>
<td>18 Oct</td>
<td>Language Modeling Part 2: Neural Architectures</td>
</tr>
<tr>
<td>11</td>
<td>20 Oct</td>
<td>Language Modeling Part 3: Pretraining and Finetuning</td>
</tr>
<tr>
<td>13</td>
<td>27 Oct</td>
<td>Statistical Machine Translation</td>
</tr>
<tr>
<td>14</td>
<td>1 Nov</td>
<td>Neural Machine Translation</td>
</tr>
<tr>
<td>15</td>
<td>3 Nov</td>
<td>Guest Lecture (Dan Garrette): Technical and Societal Challenges of Multilingual NLP</td>
</tr>
<tr>
<td>16</td>
<td>8 Nov</td>
<td>No Class: Election Day!</td>
</tr>
<tr>
<td>17</td>
<td>10 Nov</td>
<td>Formal Grammars and Constituency Parsing</td>
</tr>
<tr>
<td>18</td>
<td>15 Nov</td>
<td>Dependency Grammars and Parsing</td>
</tr>
<tr>
<td>19</td>
<td>17 Nov</td>
<td>Semantic Representations and Semantic Parsing</td>
</tr>
<tr>
<td>20</td>
<td>22 Nov</td>
<td>No Class: Thanksgiving</td>
</tr>
<tr>
<td>21</td>
<td>24 Nov</td>
<td>No Class: Thanksgiving</td>
</tr>
<tr>
<td>22</td>
<td>29 Nov</td>
<td>No Class: Thanksgiving</td>
</tr>
<tr>
<td>23</td>
<td>1 Dec</td>
<td>Research Trends 1: Linguistic Structure in the Age of Deep Learning</td>
</tr>
<tr>
<td>24</td>
<td>6 Dec</td>
<td>Research Trends 2: Grounding and Multimodal NLP</td>
</tr>
<tr>
<td>25</td>
<td>6 Dec</td>
<td>Last Lecture: State-of-the-art and Open Questions</td>
</tr>
<tr>
<td>Lecture #</td>
<td>Date</td>
<td>Topic</td>
</tr>
<tr>
<td>----------</td>
<td>----------</td>
<td>----------------------------------------------------------------------</td>
</tr>
<tr>
<td>0</td>
<td>9 Sept</td>
<td>Intro Lecture</td>
</tr>
<tr>
<td>1</td>
<td>13 Sept</td>
<td>Tasks and Metrics</td>
</tr>
<tr>
<td>2</td>
<td>15 Sept</td>
<td>Text Classifiers Part 1: Models</td>
</tr>
<tr>
<td>3</td>
<td>20 Sept</td>
<td>Text Classifiers Part 2: Features</td>
</tr>
<tr>
<td>4</td>
<td>22 Sept</td>
<td>Tokenization</td>
</tr>
<tr>
<td>5</td>
<td>27 Sept</td>
<td>Distributional Hypothesis</td>
</tr>
<tr>
<td>6</td>
<td>4 Oct</td>
<td>Topic Modeling</td>
</tr>
<tr>
<td>7</td>
<td>6 Oct</td>
<td>Deep Learning 101</td>
</tr>
<tr>
<td>8</td>
<td>11 Oct</td>
<td>Language Modeling Part 1: Ngram Models</td>
</tr>
<tr>
<td>9</td>
<td>13 Oct</td>
<td>Language Modeling Part 2: Neural Architectures</td>
</tr>
<tr>
<td>10</td>
<td>18 Oct</td>
<td>Language Modeling Part 3: Pretraining and Finetuning</td>
</tr>
<tr>
<td>11</td>
<td>20 Oct</td>
<td>Guest Lecture (Jula Netter): Risks of LM Misuse</td>
</tr>
<tr>
<td>12</td>
<td>25 Oct</td>
<td>Statistical Machine Translation</td>
</tr>
<tr>
<td>13</td>
<td>27 Oct</td>
<td>Neural Machine Translation</td>
</tr>
<tr>
<td>14</td>
<td>1 Nov</td>
<td>Guest Lecture (Dan Garrette): Technical and Societal Challenges of Multilingual NLP</td>
</tr>
<tr>
<td>15</td>
<td>3 Nov</td>
<td>Part of Speech Tagging</td>
</tr>
<tr>
<td>16</td>
<td>8 Nov</td>
<td>No Class: Election Day!</td>
</tr>
<tr>
<td>17</td>
<td>10 Nov</td>
<td>Formal Grammars and Constituency Parsing</td>
</tr>
<tr>
<td>18</td>
<td>15 Nov</td>
<td>Dependency Grammars and Parsing</td>
</tr>
<tr>
<td>19</td>
<td>17 Nov</td>
<td>Semantic Representations and Semantic Parsing</td>
</tr>
<tr>
<td>20</td>
<td>22 Nov</td>
<td>No Class: Thanksgiving</td>
</tr>
<tr>
<td>21</td>
<td>24 Nov</td>
<td>No Class: Thanksgiving</td>
</tr>
<tr>
<td>22</td>
<td>29 Nov</td>
<td>Research Trends 1: Linguistic Structure in the Age of Deep Learning</td>
</tr>
<tr>
<td>23</td>
<td>1 Dec</td>
<td>Research Trends 2: Grounding and Multimodal NLP</td>
</tr>
<tr>
<td>24</td>
<td>6 Dec</td>
<td>Last Lecture: State-of-the-art and Open Questions</td>
</tr>
<tr>
<td>Lecture #</td>
<td>Date</td>
<td>Topic</td>
</tr>
<tr>
<td>----------</td>
<td>--------</td>
<td>-------------------------------------------------</td>
</tr>
<tr>
<td></td>
<td>9 Sept</td>
<td>Intro Lecture</td>
</tr>
<tr>
<td>1</td>
<td>13 Sept</td>
<td>Tasks and Metrics</td>
</tr>
<tr>
<td>2</td>
<td>15 Sept</td>
<td>Text Classifiers Part 1: Models</td>
</tr>
<tr>
<td>3</td>
<td>20 Sept</td>
<td>Text Classifiers Part 2: Features</td>
</tr>
<tr>
<td>4</td>
<td>22 Sept</td>
<td>Tokenization</td>
</tr>
<tr>
<td>5</td>
<td>27 Sept</td>
<td>No Class</td>
</tr>
<tr>
<td>6</td>
<td>4 Oct</td>
<td>Distributional Hypothesis</td>
</tr>
<tr>
<td>7</td>
<td>6 Oct</td>
<td>Topic Modelling</td>
</tr>
<tr>
<td>8</td>
<td>11 Oct</td>
<td>Deep Learning 101</td>
</tr>
<tr>
<td>9</td>
<td>13 Oct</td>
<td>Language Modeling Part 1: Ngram Models</td>
</tr>
<tr>
<td>10</td>
<td>18 Oct</td>
<td>Language Modeling Part 2: Neural Architectures</td>
</tr>
<tr>
<td>11</td>
<td>20 Oct</td>
<td>Language Modeling Part 3: Pretraining and Finetuning</td>
</tr>
<tr>
<td>13</td>
<td>27 Oct</td>
<td>Statistical Machine Translation</td>
</tr>
<tr>
<td>14</td>
<td>1 Nov</td>
<td>Neural Machine Translation</td>
</tr>
<tr>
<td>15</td>
<td>3 Nov</td>
<td>Guest Lecture (Dan Garrette): Technical and Societal Challenges of Multilingual NLP</td>
</tr>
<tr>
<td>16</td>
<td>8 Nov</td>
<td>Part of Speech Tagging</td>
</tr>
<tr>
<td>17</td>
<td>10 Nov</td>
<td>No Class: Election Day!</td>
</tr>
<tr>
<td>18</td>
<td>15 Nov</td>
<td>Formal Grammars and Constituency Parsing</td>
</tr>
<tr>
<td>19</td>
<td>17 Nov</td>
<td>Dependency Grammars and Parsing</td>
</tr>
<tr>
<td>20</td>
<td>22 Nov</td>
<td>Semantic Representations and Semantic Parsing</td>
</tr>
<tr>
<td>21</td>
<td>24 Nov</td>
<td>No Class: Thanksgiving</td>
</tr>
<tr>
<td>22</td>
<td>29 Nov</td>
<td>Research Trends 1: Linguistic Structure in the Age of Deep Learning</td>
</tr>
<tr>
<td>23</td>
<td>1 Dec</td>
<td>Research Trends 2: Grounding and Multimodal NLP</td>
</tr>
<tr>
<td>24</td>
<td>6 Dec</td>
<td>Last Lecture: State-of-the-art and Open Questions</td>
</tr>
<tr>
<td>Lecture #</td>
<td>Date</td>
<td>Topic</td>
</tr>
<tr>
<td>-----------</td>
<td>-----------</td>
<td>-------------------------------------------------</td>
</tr>
<tr>
<td>0</td>
<td>9 Sept</td>
<td>Intro Lecture</td>
</tr>
<tr>
<td>1</td>
<td>13 Sept</td>
<td>Tasks and Metrics</td>
</tr>
<tr>
<td>2</td>
<td>15 Sept</td>
<td>Text Classifiers Part 1: Models</td>
</tr>
<tr>
<td>3</td>
<td>20 Sept</td>
<td>Text Classifiers Part 2: Features</td>
</tr>
<tr>
<td>4</td>
<td>22 Sept</td>
<td>Tokenization</td>
</tr>
<tr>
<td>5</td>
<td>27 Sept</td>
<td>No Class</td>
</tr>
<tr>
<td>6</td>
<td>4 Oct</td>
<td>Distributional Hypothesis</td>
</tr>
<tr>
<td>7</td>
<td>6 Oct</td>
<td>Topic Modeling</td>
</tr>
<tr>
<td>8</td>
<td>11 Oct</td>
<td>Deep Learning 101</td>
</tr>
<tr>
<td>9</td>
<td>13 Oct</td>
<td>Language Modeling Part 1: Ngram Models</td>
</tr>
<tr>
<td>10</td>
<td>18 Oct</td>
<td>Language Modeling Part 2: Neural Architectures</td>
</tr>
<tr>
<td>11</td>
<td>20 Oct</td>
<td>Language Modeling Part 3: Pretraining and Finetuning</td>
</tr>
<tr>
<td>13</td>
<td>27 Oct</td>
<td>Statistical Machine Translation</td>
</tr>
<tr>
<td>14</td>
<td>1 Nov</td>
<td>Neural Machine Translation</td>
</tr>
<tr>
<td>15</td>
<td>3 Nov</td>
<td>Guest Lecture (Dan Garrette): Technical and Societal Challenges of Multilingual NLP</td>
</tr>
<tr>
<td>16</td>
<td>8 Nov</td>
<td>Part of Speech Tagging</td>
</tr>
<tr>
<td>17</td>
<td>10 Nov</td>
<td>No Class: Election Day!</td>
</tr>
<tr>
<td>18</td>
<td>15 Nov</td>
<td>Formal Grammars and Constituency Parsing</td>
</tr>
<tr>
<td>19</td>
<td>17 Nov</td>
<td>Dependency Grammars and Parsing</td>
</tr>
<tr>
<td>20</td>
<td>22 Nov</td>
<td>Semantic Representations and Semantic Parsing</td>
</tr>
<tr>
<td>21</td>
<td>24 Nov</td>
<td>No Class: Thanksgiving</td>
</tr>
<tr>
<td>22</td>
<td>29 Nov</td>
<td>No Class: Thanksgiving</td>
</tr>
<tr>
<td>23</td>
<td>29 Nov</td>
<td>Research Trends 1: Linguistic Structure in the Age of Deep Learning</td>
</tr>
<tr>
<td>24</td>
<td>1 Dec</td>
<td>Research Trends 2: Grounding and Multimodal NLP</td>
</tr>
<tr>
<td>25</td>
<td>6 Dec</td>
<td>Last Lecture: State-of-the-art and Open Questions</td>
</tr>
</tbody>
</table>
Assignments

- Machine Learning: Basic Text Classifier
- Data Science: Topic Modeling
- Deep Learning: Fine-Tuned BERT Classifier
- Deep Learning: Machine Translation
- Algorithms: Dependency Parsing
Final Project

• Reimplement a recent/classic research result!
• I will choose a set of papers you can select from
• You can propose your own idea, but need to clear it with me
• More details later in the semester!
Grading

• 65% Assignments:
  • 5 Technical Assignments (12% each)
  • 1 SRC Writing Assignment (5%)
• 15% Project
• 20% Quizzes
  • 20 lectures, each equally weighted
  • Graded on correctness
  • Lowest 3 scores dropped
Late Days

- Assignments are “due” at 11:59pm on the listed due date
- No late days, no extensions, but hear me out…
  - Grading dates will be posted on the website. If your assignment is in before we start grading, I don’t care when you turned it in.
  - If you turn in your assignment after we start grading, we will most likely still grade it. But I cap TA’s work hours every week, and on-time assignments take precedence, so it’s not 100% guaranteed it will be graded.
- So, the only sure thing is to turn it in before the posted “grading” date.
- I believe that this policy is logically equivalent to what happens when I use a more standard late days+extensions policy. But, I think it is more transparent and more fair. If you disagree, let’s chat! Perhaps you can convince me otherwise.
Should you take the course? Let’s ask GPT3…
Brown Critical Review.
Course: 1460 Natural Language Processing
Instructor: Ellie Pavlick
Review:
This course was extremely difficult, and I struggled a lot throughout the semester. The assignments were often very convoluted and required a lot of outside research. The exams were even worse, and I felt like I needed to spend an unreasonable amount of time studying for them. Overall, I would not recommend this course to anyone.

Brown Critical Review.
Course: 1460 Natural Language Processing
Instructor: Ellie Pavlick
Review:
The course was very well organized and the lectures were clear and informative. The assignments were challenging and required a good understanding of the material. The instructor was always available to answer questions. Overall, the course was very enjoyable and I learned a lot.

Brown Critical Review.
Course: 1460 Natural Language Processing
Instructor: Ellie Pavlick
Review:
Brown is a great NLP course that covers a lot of material in a very short amount of time. The course is very densely packed, but it is still possible to understand the concepts if you are willing to put in the effort. The assignments are very challenging, but they are also very rewarding. I would definitely recommend this course to anyone interested in NLP.

Brown Critical Review.
Course: 1460 Natural Language Processing
Instructor: Ellie Pavlick
Review:
This course was difficult for me. I struggled with the readings and the assignments. I found the TAs to be unhelpful and the instructor was often unresponsive to questions. Overall, I did not enjoy this course.

Brown Critical Review.
Course: 1460 Natural Language Processing
Instructor: Ellie Pavlick
Review:
I found this book to be an interesting and insightful read. It was very informative and provided a lot of detail on the topic of natural language processing. The author did a great job of explaining the various concepts and algorithms in a way that was easy to understand. However, I did find the book to be somewhat dense and heavy on the technical details. Overall, I would recommend this book to anyone interested in learning more about natural language processing.
Should you take the course? Let's ask GPT3…

⚠️ Warning! SRC Violation! ⚠️
Fake news! Do not trust!

NLP is incredible. These models know everything.

Go home GPT you drunk…

NLP is incredible. These models know everything.

⚠️ Warning! SRC Violation! ⚠️
Fake news! Do not trust!

NLP is incredible. These models know everything.

⚠️ Warning! SRC Violation! ⚠️
Fake news! Do not trust!

Brown Critical Review.
Course: 1460 Natural Language Processing
Instructor: Ellie Pavlick
Review:
This course was extremely difficult, and I struggled a lot throughout the semester. The assignments were often very convoluted and required a lot of outside research. The exams were even worse, and I felt like I needed to spend an unreasonable amount of time studying for them. Overall, I would not recommend this course to anyone.

Brown Critical Review.
Course: 1460 Natural Language Processing
Instructor: Ellie Pavlick
Review:
I found this book to be an interesting and insightful read. It was very informative and provided a lot of detail on the topic of natural language processing. The author did a great job of explaining the various concepts and algorithms in a way that was easy to understand. However, I did find the book to be somewhat dense and heavy on the technical details. Overall, I would recommend this book to anyone interested in learning more about natural language processing.

Brown Critical Review.
Course: 1460 Natural Language Processing
Instructor: Ellie Pavlick
Review:
The course was very well organized and the lectures were clear and informative. The assignments were challenging and required a good understanding of the material. The instructor was always available to answer questions. Overall, the course was very enjoyable and I learned a lot.

Brown Critical Review.
Course: 1460 Natural Language Processing
Instructor: Ellie Pavlick
Review:
The Brown Critical Review is a well-written and informative overview of current research in natural language processing. The author provides a clear and concise explanation of the state of the field, and offers thoughtful critiques of recent work. The review is clearly structured and easy to follow, and the author does a good job of highlighting the key points of each article. The only downside of the review is that it is somewhat brief, and thus may not provide enough detail for readers who are new to the field. However, overall, the Brown Critical Review is an excellent resource for anyone interested in keeping up with the latest developments in natural language processing.

Brown Critical Review.
Course: 1460 Natural Language Processing
Instructor: Ellie Pavlick
Review:
Brown is a great NLP course that covers a lot of material in a very short amount of time. The course is very densely packed, but it is still possible to understand the concepts if you are willing to put in the effort. The assignments are very challenging, but they are also very rewarding. I would definitely recommend this course to anyone interested in NLP.

Brown Critical Review.
Course: 1460 Natural Language Processing
Instructor: Ellie Pavlick
Review:
This course was difficult for me. I struggled with the readings and the assignments. I found the TAs to be unhelpful and the instructor was often unresponsive to questions. Overall, I did not enjoy this course.
Okay! Let’s do it! 🏃‍♀️🏃‍♂️🏃‍♀️🏃‍♂️