

NLP

NATURAL LANGUAGE
PROCESSING

Girish Khanzode

Contents

- Natural Language Understanding
- Text Categorization
- Syntactic Analysis
- Parsing
- Semantic Analysis
- Pragmatic Analysis
- Corpus-based Statistical Approaches
- Measuring Performance
- NLP - Supervised Learning Methods
- Part of Speech Tagging
- Named Entity Recognition
- Simple Context-free Grammars
- N-grams
- References

NLP

- Natural Language Understanding
 - Taking some spoken/typed sentence and working out what it means
- Natural Language Generation
 - Taking some formal representation of what you want to say and working out a way to express it in a natural human language like English
- Fundamental goal: deep understand of broad language
 - Not just string processing or keyword matching
- Target end systems
 - speech recognition, machine translation, question answering...
 - spelling correction, text categorization...

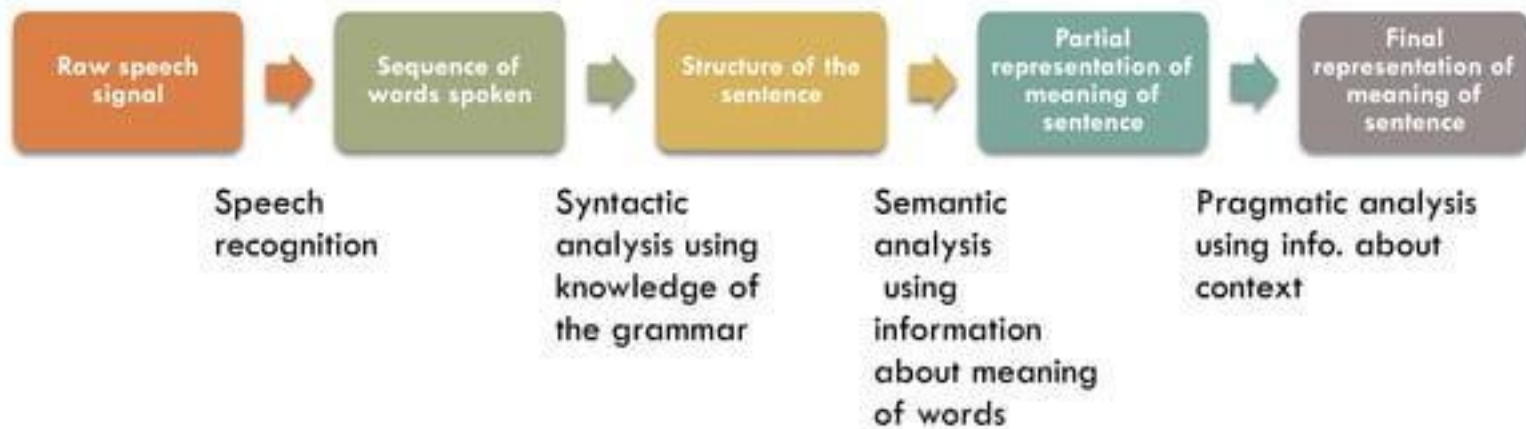
Applications

- Text Categorization - classify documents by topics, language, author, spam filtering, sentiment classification (positive, negative)
- Spelling & Grammar Corrections
- Speech Recognition
- Summarization
- Question Answering
- Better search engines
- Text-to-speech
- Machine aided translation
- Information Retrieval
 - Selecting from a set of documents the ones that are relevant to a query
- Extracting data from text
 - Converting unstructured text into structure data

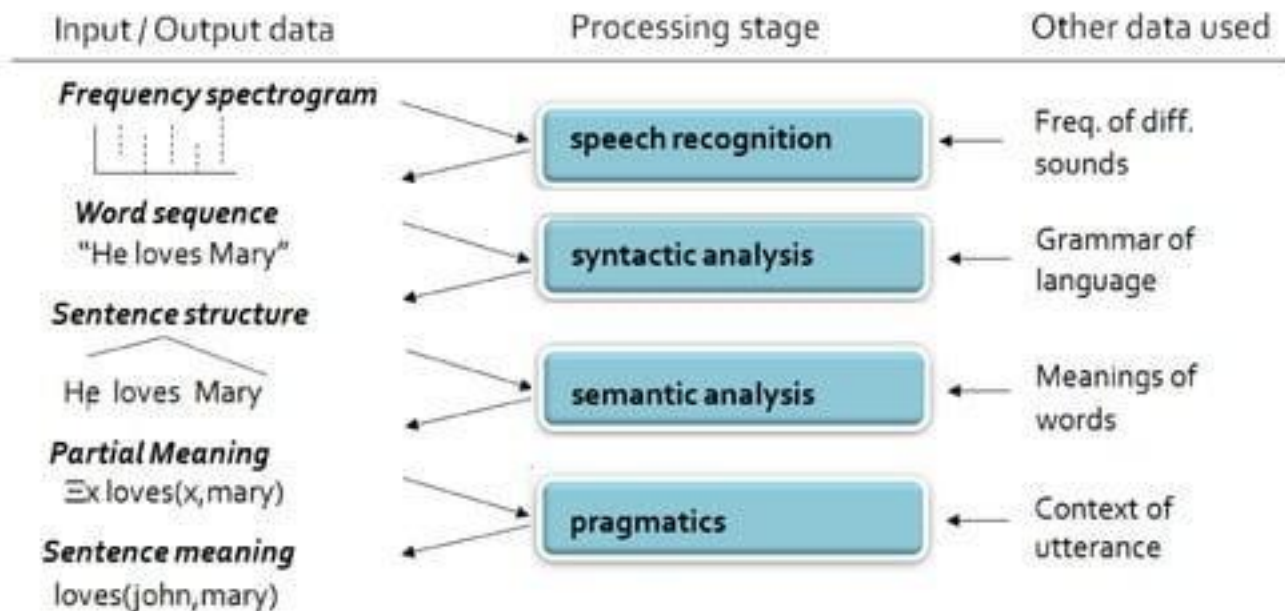
Natural Language Understanding

- Answering an essay question in exam
- Deciding what to order at a restaurant by reading a menu
- Realizing you've been praised
- Appreciating a poem

Natural Language Understanding



Natural Language Understanding



Text Categorization

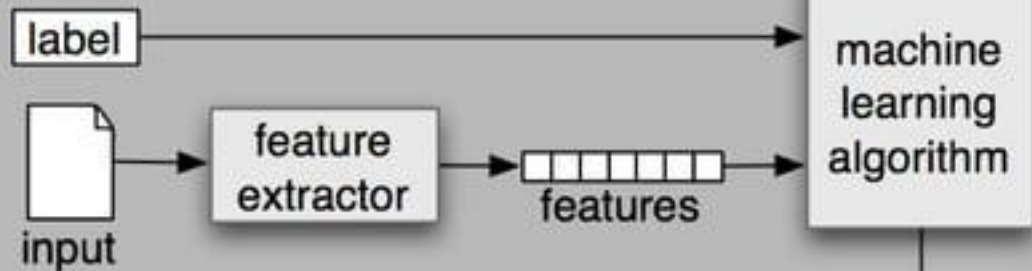
- Text annotation - classify entire document
- Sentiment classification
 - What features of the text could help predict # of likes?
 - How to identify customer opinions?
 - Are the features hard to compute? (syntax? sarcasm?)
- Is it spam?
- What medical billing code for this visit?
- What grade for an answer to this essay question?

Text Categorization

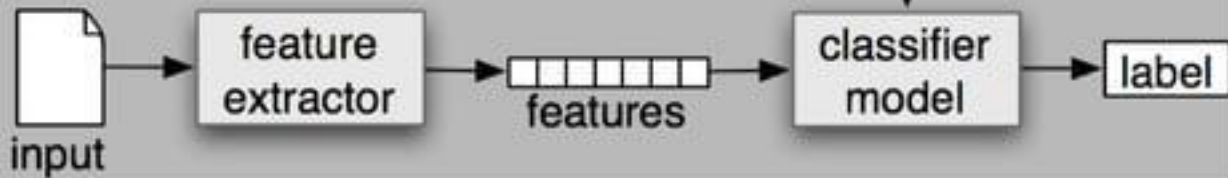
- Is it interesting to this user?
 - News filtering; helpdesk routing
- Is it interesting to this NLP program?
 - If it's Spanish, translate it from Spanish
 - If it's subjective, run the sentiment classifier
 - If it's an appointment, run information extraction
- Where should it be filed?
 - Which mail folder? (work, friends, junk, urgent ...)
 - Yahoo! / Open Directory / digital libraries

Text Categorization

(a) Training



(b) Prediction



Syntactic Analysis

- Rules of syntax (grammar) specify the possible organization of words in sentences and allows us to determine sentence's structure(s)
 - John saw Mary with a telescope
 - John saw (Mary with a telescope)
 - John (saw Mary with a telescope)
- Parsing: given a sentence and a grammar
 - Checks that the sentence is correct according with the grammar and if so returns a parse tree representing the structure of the sentence

Syntactic Analysis

- Syntax mapped into semantics
 - Nouns ↔ things, objects, abstractions.
 - Verbs ↔ situations, events, activities.
 - Adjectives ↔ properties of things, ...
 - Adverbs ↔ properties of situations, ...
- A parser recovers the phrase structure of an utterance, given a grammar (rules of syntax)
- Parser's outcome is the structure (groups of words and respective parts of speech)

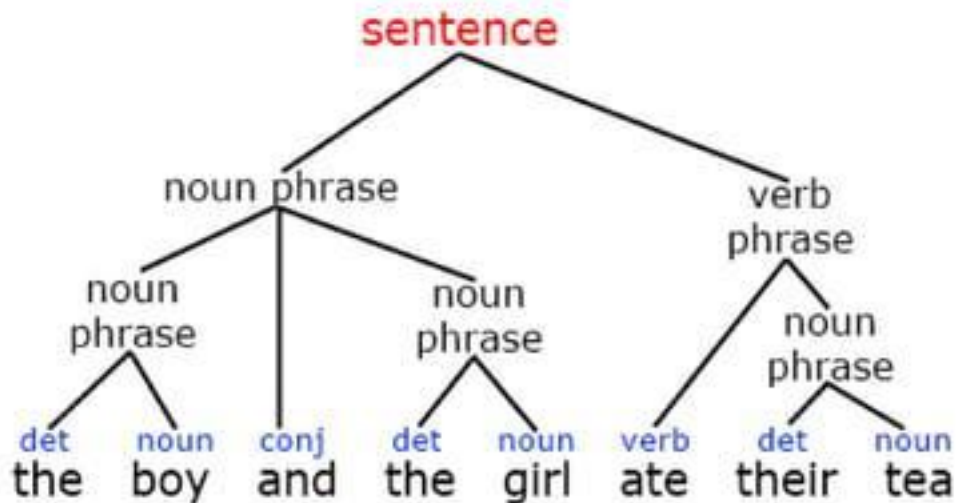
Syntactic Analysis

- Phrase structure is represented in a parse tree
- Parsing is the first step towards determining the meaning of an utterance
- Outcome of the syntactic analysis can still be a series of alternate structures with respective probabilities
- Sometimes grammar rules can disambiguate a sentence
 - John set the set of chairs
- Sometimes they can't.
- ...the next step is semantic analysis

Symbols in Grammar

- | | |
|---|---|
| <ul style="list-style-type: none">• S - Sentence• NP- Noun Phrase• PN- Proper Noun• N-Noun• VP-Verb Phrase• Adv-Adverb• V-Verb• Adj-Adjective• Prep-Preposition• Art-Article | <ul style="list-style-type: none">• Pro-Pronoun• PP-Prepositional Phrase• * Ungrammatical Sentence• → Consists of / rewrites as• () Optional Constituent• { } Only one of these constituents must be selected |
|---|---|

Syntactic Analysis - Grammar



Parsing

- A method to analyze a sentence to determine its structure as per grammar
- Grammar - formal specification of the structures allowable in the language
- Syntax is important - a skeleton on which various linguistic elements, meaning among them depends
- So recognizing syntactic structure is also important

Parsing

- Some researchers deny syntax its central role
- There is a verb-centered analysis that builds on Conceptual Dependency [textbook, section 7.1.3] - a verb determines almost everything in a sentence built around it - Verbs are fundamental in many theories of language
- Another idea is to treat all connections in language as occurring between pairs of words, and to assume no higher-level groupings
 - Structure and meaning are expressed through variously linked networks of words

Syntactic Analysis - Challenges

- Number (singular vs. plural) and gender
 - sentence-> noun_phrase(n), verb_phrase(n)
 - proper_noun(s) -> [mary]
 - noun(p) -> [apples]
- Adjective
 - noun_phrase-> determiner,adjectives,noun
 - adjectives-> adjective, adjectives
 - adjective->[ferocious]

Syntactic Analysis - Challenges

- Adverbs, ...
- Handling ambiguity
 - Syntactic ambiguity - fruit flies like a banana
- Having to parse syntactically incorrect sentences

Semantic Analysis

- Generates meaning/representation of the sentence from its syntactic structures
- Represents the sentence in meaningful parts
- Uses possible syntactic structures and meaning
- Builds a parse tree with associated semantics
- Semantics typically represented with logic

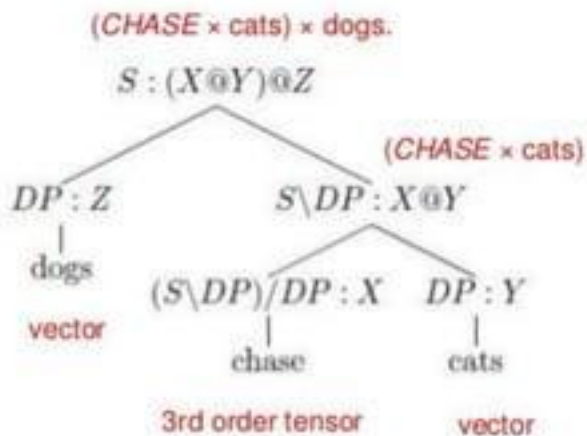
Semantic Analysis

- Compositional semantics: meaning of the sentence from the meaning of its parts
 - Sentence - A tall man likes Mary
 - Representation - $\exists x \text{ man}(x) \ \& \ \text{tall}(x) \ \& \ \text{likes}(x, \text{mary})$
- Grammar + Semantics
 - Sentence (Smeaning) \rightarrow noun_phrase(NPmeaning), verb_phrase(VPmeaning), combine(NPmeaning, VPmeaning, Smeaning)
- Complications - Handling ambiguity
 - Semantic ambiguity - I saw the prudential building flying into Paris

Compositional Semantics

- The semantics of a phrase is a function of the semantics of its sub-phrases
- It does not depend on any other phrase
- So if we know the meaning of sub-phrases, then we know the meaning of the phrases
- A goal of semantic interpretation is to find a way that the meaning of the whole sentence can be put together in a simple way from the meanings of the parts of the sentence - Alison, 1997 p. 112

Compositional Semantics



Pragmatic Analysis

- Uses context
- Uses partial representation
- Includes purpose and performs disambiguation
- Where, when, by whom an utterance was said
- ▣ Uses context of utterance
 - Where, by who, to whom, why, when it was said
 - Intentions - inform, request, promise, criticize, ...

Pragmatic Analysis

- Handling Pronouns

- Mary eats apples. She likes them

- She = Mary, them = apples

- Handling ambiguity

- Pragmatic ambiguity - you're late - What's the speaker's intention - informing or criticizing?



NLP Challenges

- ❑ NLP systems needs to answer the question “who did what to whom”
- ❑ MANY hidden variables
 - ❑ Knowledge about the world
 - ❑ Knowledge about the context
 - ❑ Knowledge about human communication techniques
 - Can you tell me the time?
- ❑ Problem of scale
 - ❑ Many (infinite?) possible words, meanings, context

NLP Challenges

- ❑ Problem of sparsity
 - ❑ Very difficult to do statistical analysis, most things (words, concepts) are never seen before
- ❑ Long range correlations
- ❑ Key problems
 - ❑ Representation of meaning
 - ❑ Language presupposes knowledge about the world
 - ❑ Language only reflects the surface of meaning
 - ❑ Language presupposes communication between people

NLP Challenges

- ❑ Different ways of Parsing a sentence
- ❑ Word category ambiguity
- ❑ Word sense ambiguity
- ❑ Words can mean more than their sum of parts - The Times of India
- ❑ Imparting world knowledge is difficult - the blue pen ate the ice-cream

NLP Challenges

- ❑ Fictitious worlds - people on mars can fly
- ❑ Defining scope
 - ❑ people like ice-cream - does this mean all people like ice cream?
- ❑ Language is changing and evolving
- ❑ Complex ways of interaction between the kinds of knowledge
- ❑ Exponential complexity at each point in using the knowledge

NLP Challenges - Ambiguity

- At all levels - lexical, phrase, semantic
- Institute head seeks aid
 - Word sense is ambiguous (head)
- Stolen Painting Found by Tree
 - Thematic role is ambiguous: tree is agent or location?
- Ban on Dancing on Governor's Desk
 - Syntactic structure (attachment) is ambiguous - is the ban or the dancing on the desk?
- Hospitals Are Sued by 7 Foot Doctors
 - Semantics is ambiguous : what is 7 foot?

Meaning

- From NLP viewpoint, meaning is a mapping from linguistic forms to some kind of representation of knowledge of the world
 - Physical referent in the real world
 - Semantic concepts, characterized also by relations.
- It is interpreted within the framework of some sort of action to be taken

Meaning – Representation and Usage

- I am Italian
 - From lexical database (WordNet)
 - Italian = a native or inhabitant of Italy → Italy = republic in southern Europe [..]
- I am Italian
 - Who is "I"?
- I know she is Italian/I think she is Italian
 - How do we represent I know and I think
 - Does this mean that I is Italian?
 - What does it say about the I and about the person speaking?
- I thought she was Italian
 - How do we represent tenses?

Corpus-based Statistical Approaches

- How can a machine understand these differences?
 - Decorate the cake with the frosting
 - Decorate the cake with the kids
- Rules based approaches
 - Hand coded syntactic constraints and preference rules
 - The verb decorate require an animate being as agent
 - The object cake is formed by any of the following, inanimate entities
 - cream, dough, frosting.....

Corpus-based Statistical Approaches

- ❑ These approaches are time consuming to build, do not scale up well and are very brittle to new, unusual, metaphorical use of language
- ❑ To swallow requires an animate being as agent/subject and a physical object as object
 - I swallowed his story
 - The supernova swallowed the planet
- ❑ A Statistical NLP approach seeks to solve these problems by automatically learning lexical and structural preferences from text collections (corpora)

Corpus-based Statistical Approaches

- Statistical models are robust, generalize well and behave gracefully in the presence of errors and new data
- Steps
 - ▣ Get large text collections
 - ▣ Compute statistics over those collections
 - The bigger the collections, the better the statistics

Corpus-based Statistical Approaches

- Decorate the cake with the frosting
- Decorate the cake with the kids
- From (labeled) corpora we can learn that
 - $\#(\text{kids are subject/agent of decorate}) > \#(\text{frosting is subject/agent of decorate})$
- From (unlabelled) corpora we can learn that
 - $\#(\text{"the kids decorate the cake"}) \gg \#(\text{"the frosting decorates the cake"})$
 - $\#(\text{"cake with frosting"}) \gg \#(\text{"cake with kids"})$
- Given these facts, we need a statistical model for the attachment decision

Corpus-based Statistical Approaches

- Topic categorization: classify the document into semantics topics

Document 1

The nation swept into the Davis Cup final on Saturday when twins Bob and Mike Bryan defeated Belarus's Max Mirnyi and Vladimir Voltchkov to give the Americans an unsurmountable 3-0 lead in the best-of-five semi-final tie.

Topic = sport

Document 2

One of the strangest, most relentless hurricane seasons on record reached new bizarre heights yesterday as the plodding approach of Hurricane Jeanne prompted evacuation orders for hundreds of thousands of Floridians and high wind warnings that stretched 350 miles from the swamp towns south of Miami to the historic city of St. Augustine.

Topic = Natural Event

Corpus-based Statistical Approaches

- Topic categorization: classify the document into semantics topics

Document 1 (sport)

The nation swept into the Davis Cup final on Saturday when twins Bob and Mike Bryan ...

Document 2 (Natural Event)

One of the strangest, most relentless hurricane seasons on record reached new bizarre heights yesterday as....

- From (labeled) corpora we can learn that
 - $\#(\text{sport documents containing word cup}) > \#(\text{nature event documents containing word cup})$ - feature
- We then need a statistical model for the topic assignment

Corpus-based Statistical Approaches

- Feature extractions
 - Usually linguistics motivated
- Statistical models
- Data
 - Corpora, labels, linguistic resources

Measuring Performance

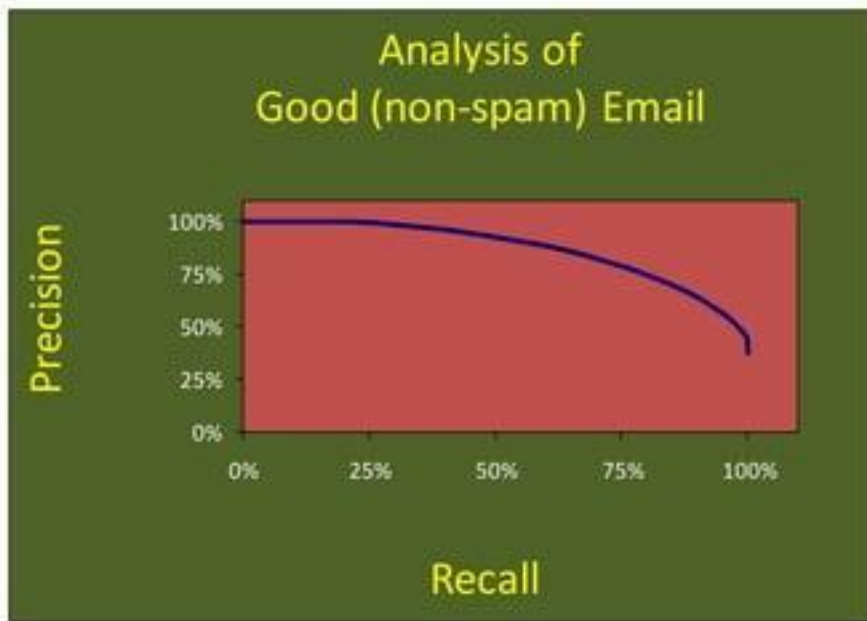
- Classification accuracy
 - What % of messages were classified correctly?
- Is this what we care about?
- Which system is better?

	Overall accuracy	Accuracy on spam	Accuracy on gen
System 1	95%	99.99%	90%
System 2	95%	90%	99.99%

Measuring Performance

- Precision = $\frac{\text{good messages kept}}{\text{all messages kept}}$
- Recall = $\frac{\text{good messages kept}}{\text{all good messages}}$
- Move from high precision to high recall by deleting fewer messages
 - delete only if spam content > high threshold

Measuring Performance



NLP - Supervised Learning Methods

- Conditional log-linear models
 - Feature engineering - Throw in enough features to fix most errors
 - Training - Learn weights θ such that in training data, the true answer tends to have a high probability
 - Test - Output the highest-probability answer
 - If the evaluation metric allows for partial credit, can do fancier things
 - minimum-risk training and decoding

NLP - Supervised Learning Methods

- The most popular alternatives are roughly similar
 - Perceptrons, SVM, MIRA, neural network, ...
 - These also learn a usually linear scoring function
 - However, the score is not interpreted as a log-probability
 - Learner just seeks weights θ such that in training data, the desired answer has a higher score than the wrong answers

Features

- ❑ Example - word to left
- ❑ Spelling correction using an n-gram language model
($n \geq 2$) would use words to left and right to help predict the true word
- ❑ Similarly, an HMM (Hidden Markov Model) would predict a word's class using classes to left and right
- ❑ But we'd like to throw in all kinds of other features too

Features

Word to left	Frequency as Aid	Frequency as Aide
foreign	718	1
federal	297	0
western	146	0
provide	88	0
covert	26	0
oppose	13	0
future	9	0
similar	6	0
presidential	0	63
chief	0	40
longtime	0	26
aids-infected	0	2
sleepy	0	1
disaffected	0	1
indispensable	2	1
practical	2	0
squander	1	0

Part of Speech Tagging

- Treat tagging as a token classification problem
 - Tag each word independently given
 - features of context
 - features of the word's spelling (suffixes, capitalization)
- Use an HMM
 - the tag of one word might depend on the tags of adjacent words
- Combination of both
 - Need rich features (in a log-linear model), but also want feature functions to depend on adjacent tags
 - So, the problem is to predict **all** tags together.

Part of Speech Tagging

- Idea 1 - classify tags one at a time from left to right
 - $p(\text{tag} \mid \text{wordseq}, \text{prevtags}) = (1/Z) \exp \text{score}(\text{tag}, \text{wordseq}, \text{prevtags})$
 - where Z sums up $\exp \text{score}(\text{tag}', \text{wordseq}, \text{prevtags})$ over all possible tags
- Idea 2 - maximum entropy Markov model (MEMM)
 - Same model, but don't commit to a tag before we predict the next tag
 - Instead, evaluate probability of every tag sequence

Part of Speech Tagging

- Idea 3 - linear-chain conditional random field (CRF)
 - Symmetric and very popular
 - Score each tag sequence as a whole, using arbitrary features
 - $p(\text{tagseq} \mid \text{wordseq}) = (1/Z) \exp \text{score}(\text{tagseq}, \text{wordseq})$
 - where Z sums up $\exp \text{score}(\text{tagseq}', \text{wordseq})$ over competing tagseqs
 - Can still compute Z and best path using dynamic programming
 - Dynamic programming works if each feature $f(\text{tagseq}, \text{wordseq})$ considers at most an n -gram of tags
 - Score a $(\text{tagseq}, \text{wordseq})$ pair with a WFST whose state remembers the previous $(n-1)$ tags
 - As in 2, arc weight can consider the current tag n -gram and all words
 - But unlike 2, arc weight isn't a probability - only normalize at the end

Named Entity Recognition

- Deals with the *detection* and *categorization* of proper names
- Labeling all occurrences of named entities in a text
 - ▣ Named Entity = People, organizations, lakes, bridges, hospitals, mountains...
- Well-understood technology, readily available and works well
- Uses a combination of enumerated lists (often called gazetteers) and regular expressions

Named Entity Recognition

Commonly Used Named Entity Types

NE Type	Examples
ORGANIZATION	<i>Georgia-Pacific Corp., WHO</i>
PERSON	<i>Eddy Bonte, President Obama</i>
LOCATION	<i>Murray River, Mount Everest</i>
DATE	<i>June, 2008-06-29</i>
TIME	<i>two fifty a m, 1:30 p.m.</i>
MONEY	<i>175 million Canadian Dollars, GBP 10.40</i>
PERCENT	<i>twenty pct, 18.75 %</i>
FACILITY	<i>Washington Monument, Stonehenge</i>
GPE	<i>South East Asia, Midlothian</i>

Complex Entities and Relationships

- Uses Named Entities as components
- Pattern-matching rules specific to a given domain
- May be multi-pass - One pass creates entities which are used as part of later passes
 - First pass locates CEOs, second pass locates dates for CEOs being replaced...
- May involve syntactic analysis as well, especially for things like negatives and reference resolution

Complex Entities and Relationships

KEEP UP **ON** YOUR **READING** WITH AUDIO **BOOKS**
Vietnam UK Louisiana, USA

Audio **books** are highly **popular** with **library** patrons in the **town**
Louisiana, USA S. Carolina, USA Pennsylvania, USA Mass., USA

of **Springfield,** **Greene** County, **MO.** "People are **mobile**
Turkey Virginia, USA Maine, USA Norway Alabama, USA

and busier, and audio **books** fit into that lifestyle" says **Gary**
Louisiana, USA Indiana, USA

Sanchez, who oversees the **library's** **\$2** **million** budget...
Dominican Republic Pennsylvania, USA Kentucky, USA

Why Does It Work?

- The problem is very constrained
 - Only looking for a small set of items that can appear in a small set of roles
- We can ignore stuff we don't care about
- BUT creating the knowledge bases can be very complex and time-consuming

Simple Context-free Grammars

- Consider the simplest Context-Free Grammars - without and with parameters
 - parameters allows to express more interesting facts
- sentence \rightarrow noun_phrase verb_phrase
- noun_phrase \rightarrow proper_name
- noun_phrase \rightarrow article noun
- verb_phrase \rightarrow verb
- verb_phrase \rightarrow verb noun_phrase
- verb_phrase \rightarrow verb noun_phrase
prep_phrase
- verb_phrase \rightarrow verb prep_phrase
- prep_phrase \rightarrow preposition noun_phrase

Simple CF Grammars

- The still-undefined syntactic units are pre-terminals
- They correspond to parts of speech
- We can define them by adding lexical productions to the grammar
- `article` \rightarrow `the` | `a` | `an`
- `noun` \rightarrow `pizza` | `bus` | `boys` | ...
- `preposition` \rightarrow `to` | `on` | ...
- `proper_name` \rightarrow `Jim` | `Dan` | ...
- `verb` \rightarrow `ate` | `yawns` | ...
- This is not practical on a large scale
- Normally, we have a lexicon (dictionary) stored in a database, that can be interfaced with the grammar

Simple CF Grammars

- sentence →
- noun_phrase verb_phrase →
- proper_name verb_phrase →
- Jim verb_phrase →
- Jim verb noun_phrase prep_phrase →
- Jim ate noun_phrase prep_phrase →
- Jim ate article noun prep_phrase →
- Jim ate a noun prep_phrase →
- Jim ate a pizza prep_phrase →
- Jim ate a pizza preposition noun_phrase →
- Jim ate a pizza on noun_phrase →
- Jim ate a pizza on article noun →
- Jim ate a pizza on the noun →
- Jim ate a pizza on the bus

Simple CF Grammars

- Other examples of sentences generated by this grammar:

- Jim ate a pizza
- Dan yawns on the bus

- These wrong data will also be recognized:

- Jim ate an pizza
- Jim yawns a pizza
- Jim ate to the bus
- the boys yawns
- the bus yawns

- ... but not these, obviously correct:

- the pizza was eaten by Jim
- Jim ate a hot pizza

- and so on, and so forth.

Simple CF Grammars

- This simple grammar can be improved in many interesting ways
 - Add productions, for example to allow adjectives
 - Add words (in lexical productions, or in a more realistic lexicon)
 - Check agreement (noun-verb, noun-adjective, and so on)
 - rabbitspl runpl ↔ a rabbitsg runsg
 - le bureau blancm ↔ la tablef blanchef
- An obvious, but naïve, method of enforcing agreement is to duplicate the productions and the lexical data.

Simple CF Grammars

- sentence → noun_phr_sg verb_phr_sg
- sentence → noun_phr_pl verb_phr_pl
- noun_phr_sg → art_sg noun_sg
- noun_phr_sg → proper_name_sg
- noun_phr_pl → art_pl noun_pl
- noun_phr_pl → proper_name_pl
- art_sg → the | a | an
- art_pl → the
- noun_sg → pizza | bus | ...
- noun_pl → boys | ... and so on.

Simple CF Grammars

- A much better method is to add parameters, and to parameterize words as well as productions
- `sentence` \rightarrow `noun_phr`(Num) `verb_phr`(Num)
- `noun_phr`(Num) \rightarrow `art`(Num) `noun`(Num)
- `noun_phr`(Num) \rightarrow `proper_name`(Num)
- `art`(sg) \rightarrow `the` | `a` | `an`
- `art`(pl) \rightarrow `the`
- `noun`(sg) \rightarrow `pizza` | `bus` | ...
- `noun`(sg) \rightarrow `boys` | ... and so on.
- This notations slightly extends the basic Context-Free Grammar formalism.

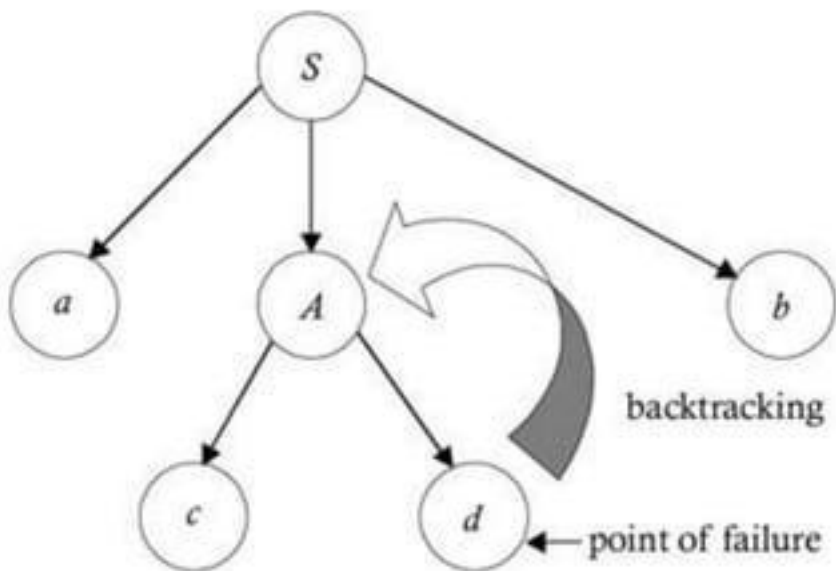
Simple CF Grammars

- Another use of parameters in productions: represent transitivity
- We want to exclude such sentences as
 - Jim yawns a pizza
 - Jim ate to the bus
- `verb_phr(Num) → verb(intrans, Num)`
- `verb_phr(Num) →`
 - `verb(trans, Num) noun_phr(Num1)`
- `verb(intrans, sg) → yawns | ...`
- `verb(trans, sg) → ate | ...`
- `verb(trans, pl) → ate | ...`

Direction of Parsing – Top Down

- ❑ Starts from root and moves to leaves
- ❑ Top-down, hypothesis-driven - assume that we have a sentence, keep rewriting, aim to derive a sequence of terminal symbols, backtrack if data tell us to reject a hypothesis
- ❑ For example, we had assumed a noun phrase that begins with an article, but there is no article
- ❑ Problem - wrong guesses, wasted computation

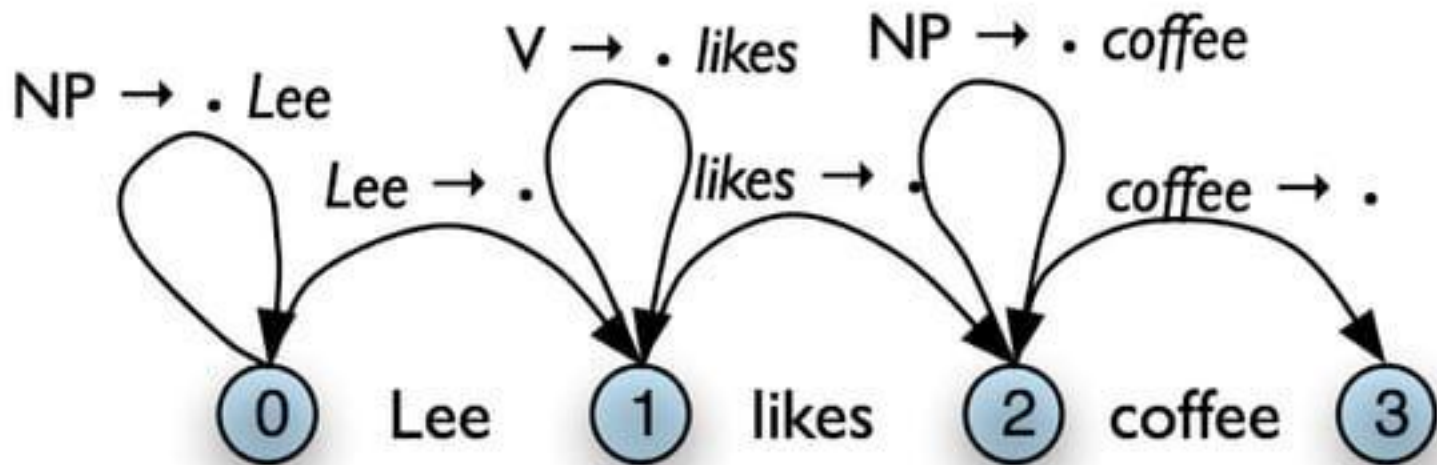
Direction of Parsing – Top Down



Direction of Parsing – Bottom Up

- Data-driven - look for complete right-hand sides of productions, keep rewriting, aim to derive the goal symbol
 - Problem: lexical ambiguity that may lead to many unfinished partial analyses
 - Lexical ambiguity is generally troublesome. For example, in the sentence "Johnny runs the show", both runs and show can be a verb or a noun, but only one of $2*2$ possibilities is correct

Direction of Parsing – Bottom Up



Direction of Parsing

- In practice, parsing is never pure
- Top-down, enriched - check data early to discard wrong hypotheses (somewhat like recursive-descent parsing in compiler construction)
- Bottom-up, enriched - use productions, suggested by data, to limit choices (somewhat like LR parsing in compiler construction)

Direction of Parsing

- A popular bottom-up analysis method - chart parsing
- Popular top-down analysis methods
 - transition networks (used with Lisp)
 - logic grammars (used with Prolog)

N-grams

- Letter or word frequencies - 1-grams - unigrams
 - useful in solving cryptograms - ETAOINSHRDLU...
- If you know the previous letter - 2-grams - bigrams
 - "h" is rare in English (4%)
 - but "h" is common after "t" (20%)
- If you know the previous two letters - 3-grams - trigrams
 - "h" is really common after "(space) t"

References

1. Ruslan Mitkov, *The Oxford Handbook Of Computational Linguistics*, Oxford University Press, 2003.
2. Robert Dale, Herman Mohl, Harold Somers, *Handbook Of Natural Language Processing*, Marcel Dekker Inc.
3. James Allen, *Natural Language Processing*, Pearson Education, 2003.
4. Manning C, Raghavan P, Schuetze H. *Introduction to Information Retrieval*. Cambridge, UK: Cambridge University Press, 2008.
5. Douglas Biber, Susan Conrad, Randi Reppen, *Corpus Linguistics – Investigating Language Structure And Use*, Cambridge University Press, 2000.
6. David Singleton, *Language And The Lexicon: An Introduction*, Arnold Publishers, 2000.
7. Allen, James, *Natural Language Understanding*, second edition (Redwood City: Benjamin/Cummings, 1995).
8. Ginsberg, Matt, *Essentials of Artificial Intelligence* (San Mateo: Morgan Kaufmann, 1993).
9. Hutchins W. *The First Public Demonstration of Machine Translation: the Georgetown-IBM System, 7th January 1954*. 2005.
10. Chomsky N. Three models for the description of language. *IRE Trans Inf Theory* 1956;2:113–24.
11. Aho AV, Sethi R, Ullman JD. *Compilers: Principles, Techniques, Tools*. Reading, MA: Addison-Wesley, 1988.
12. Chomsky N. On certain formal properties of grammars. *Inform Contr* 1959;2:137–67.
13. Friedl JEF. *Mastering Regular Expressions*. Sebastopol, CA: O'Reilly & Associates, Inc., 1997.

Thank You



Check Out My LinkedIn Profile at

<https://in.linkedin.com/in/girishkhanzode>