

0. Introduction

NLP for CogSci Research

Marlene Staib

September 3, 2018

Tasks and Toolkits

e.g.: Tokenisation, POS-tagging, Parsing, (ASR, MT), ...
nltk,...

Modelling Approach

A set of modelling assumptions; building predictive classifiers
Basically ML for language data...

Algorithms

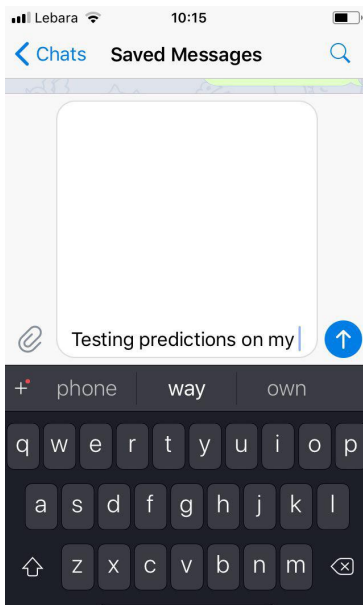
Dynamic Programming, Viterbi, Lambda Calculus, Optimisation
Algorithms, ...

Summary of Contents

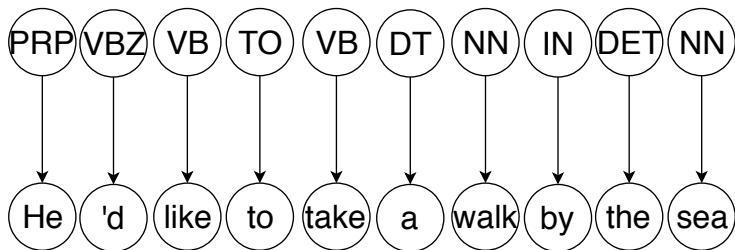
1	Introduction.....	9
2	Regular Expressions, Text Normalization, Edit Distance.....	10
3	Finite State Transducers.....	34
4	Language Modeling with N-grams.....	35
5	Spelling Correction and the Noisy Channel.....	61
6	Naive Bayes and Sentiment Classification.....	74
7	Logistic Regression.....	92
8	Neural Networks and Neural Language Models.....	102
9	Hidden Markov Models.....	122
10	Part-of-Speech Tagging.....	142
11	Formal Grammars of English.....	168
12	Syntactic Parsing.....	197
13	Statistical Parsing.....	212
14	Dependency Parsing.....	245
15	Vector Semantics.....	270
16	Semantics with Dense Vectors.....	286
17	Computing with Word Senses.....	300
18	Lexicons for Sentiment and Affect Extraction.....	326
19	The Representation of Sentence Meaning.....	346
20	Computational Semantics.....	347
21	Information Extraction.....	348
22	Semantic Role Labeling.....	377
23	Coreference Resolution and Entity Linking.....	396
24	Discourse Coherence.....	397
25	Machine Translation and Seq2Seq Models.....	398
26	Summarization.....	399
27	Question Answering.....	400
28	Dialog Systems and Chatbots.....	418
29	Advanced Dialog Systems.....	441
30	Speech Recognition.....	459
31	Speech Synthesis.....	460
	Bibliography.....	461
	Author Index.....	485

NLP Tasks - a *short* intro

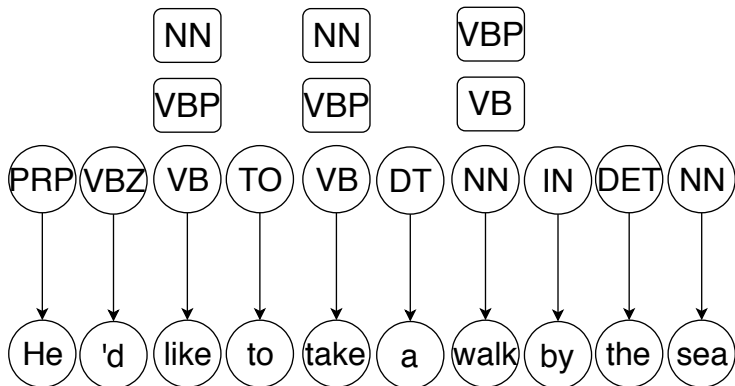
Language Modelling



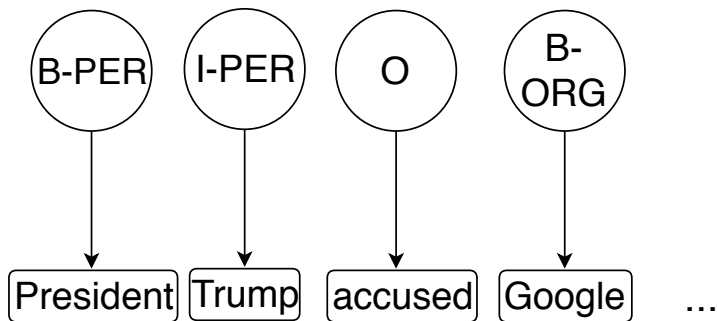
Part-of-Speech (POS) tagging



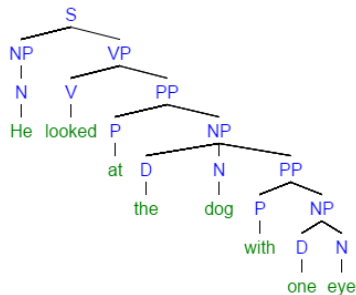
Part-of-Speech (POS) tagging



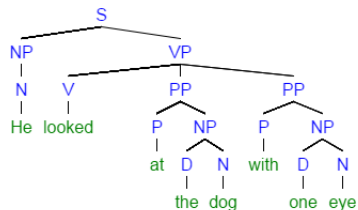
Named Entity Recognition



(Syntactic) Parsing



(a)



(b)

Text Classification



Sentiment Classification

★☆☆☆ **Don't WASTE your money!!!**, June 22, 2010

By [Tommy Lee](#)

Verified Purchase ([What's this?](#))

This review is from: Crush It!: Why NOW Is the Time to Cash In on Your Passion (Hardcover)

Here's the straight to the point short answer for this book: Use facebook, and other social media sites to promote yourself! There, I just gave you all this book is going to give you. I simply gave it to you in one sentence for FREE instead of the \$10 I paid for it on Amazon, and 2-hours of your time wasted reading it. It's really nothing more than a salesman's pitch book. And the pitch here is to buy his book, and read about him telling you THAT he used social media to grow his business, NOT HOW HE USED SOCIAL MEDIA TO GROW HIS BUSINESS.

*** PLEASE NOTE ***

THIS IS NOT A HOW TO BOOK BY ANY MEANS. If that is what you are looking for, then look elsewhere. This book is simply about Gary saying, I used social media to grow my business. And really folks, that's it!



"Awesome dresses!!. Ordered a few dresses for a ball gown event, and they were all lovely. They looked just like the images and fit perfectly, really happy with my purchases.

- Rose D.

More tasks...

- Semantic Parsing
- Relation Extraction
- Semantic Role Labelling
- Topic Modelling
- ...

- Spell Checking, Autocorrect
- Information Retrieval (search engines)
- Question Answering, Chatbots
- Text Analysis (e.g., customer review sentiments, terrorism detection,...)
- Machine Translation
- Automatic Speech Recognition
- Speech Synthesis
- ...

- Corpora: Penn Treebank, Brown, Gutenberg, Europarl, ...
- Lists: stopwords, ...
- Dictionaries: cmudict, ...
- Tools: Tokenizer, Tagger, Parser, ...
- Useful Data Structures: `nltk.tree.Tree`, ...

What now?

- Fill in the pre-course survey:
<https://goo.gl/forms/4okS2qOWKOqbj5os2>
- Install the software (see syllabus)
- Software help: **Monday, 17th of September, 16.00**
- Optional: check out the linked materials in the syllabus ;)