BITS Pilani

Pilani|Dubai|Goa|Hyderabad

S1-25_DSECLZG530/SSCLZG599
Natural Language Processing
(Lecture #1 - Introduction)

Dr. Naveen Kumar Laskari, WILP, BITS Pilani

DISCLAIMERS

- *The slides presented here are obtained from the authors of the books and from various other contributors. I hereby acknowledge all the contributors for their material and inputs.*
- *I have added and modified a few slides to suit the requirements of the course.*

innovate achieve lead

BITS Pilani

Pilani|Dubai|Goa|Hyderabad

Introduction to NLP

Natural Language Processing

- NLP is the branch of computer science focused on developing systems that allow computers to communicate with people using everyday language.
- Also called Computational Linguistics
  - Also concerns how computational methods can aid the understanding of human language
  - The Association for Computational Linguistics (ACL) is a scientific and professional organization for people working on natural language processing

NLP – Why and What

Why study NLP?

- Text is the largest repository of humans
- News articles, Web pages, Human Knowledge(Philosophy/Science), Patents, Emails/Communications, Administration/Society

What is Goal of NLP?

- Linguistic/Fundamental
  - Deeper understanding of human languages

- Engineering
  - Design, Build, Validate Systems that work with Natural Languages for many practical applications

A few applications of NLP

- Spelling correction
- Grammar checking
- Better search engines
- Information extraction
- Etc.

- Solutions such as:
  - Speech recognition (and text-to-speech)
  - Dialogue systems (or conversational agent (CA), is a computer system intended to converse with a human)
  - Machine translation (Bing/Google)

Related Areas

- Artificial Intelligence
- Formal Language (Automata) Theory
- Machine Learning
- Information Retrieval
- Linguistics
- Cognitive Science
- Philosophy of Language

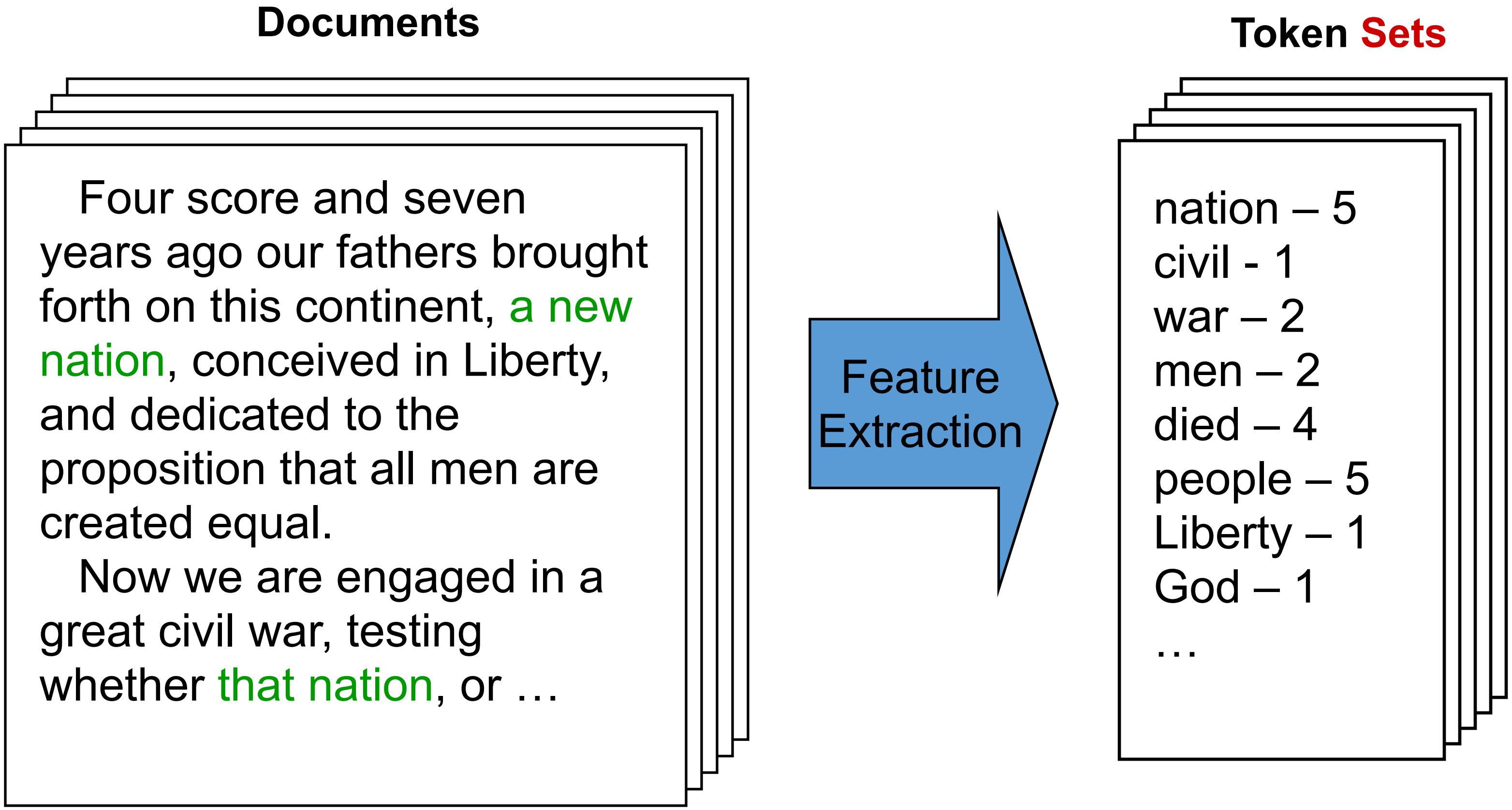
innovate achieve lead

BITS Pilani

Pilani|Dubai|Goa|Hyderabad

NLP Challenges

Bag-of-Tokens Approaches



Documents

Four score and seven years ago our fathers brought forth on this continent, a new nation, conceived in Liberty, and dedicated to the proposition that all men are created equal.

Now we are engaged in a great civil war, testing whether that nation, or …

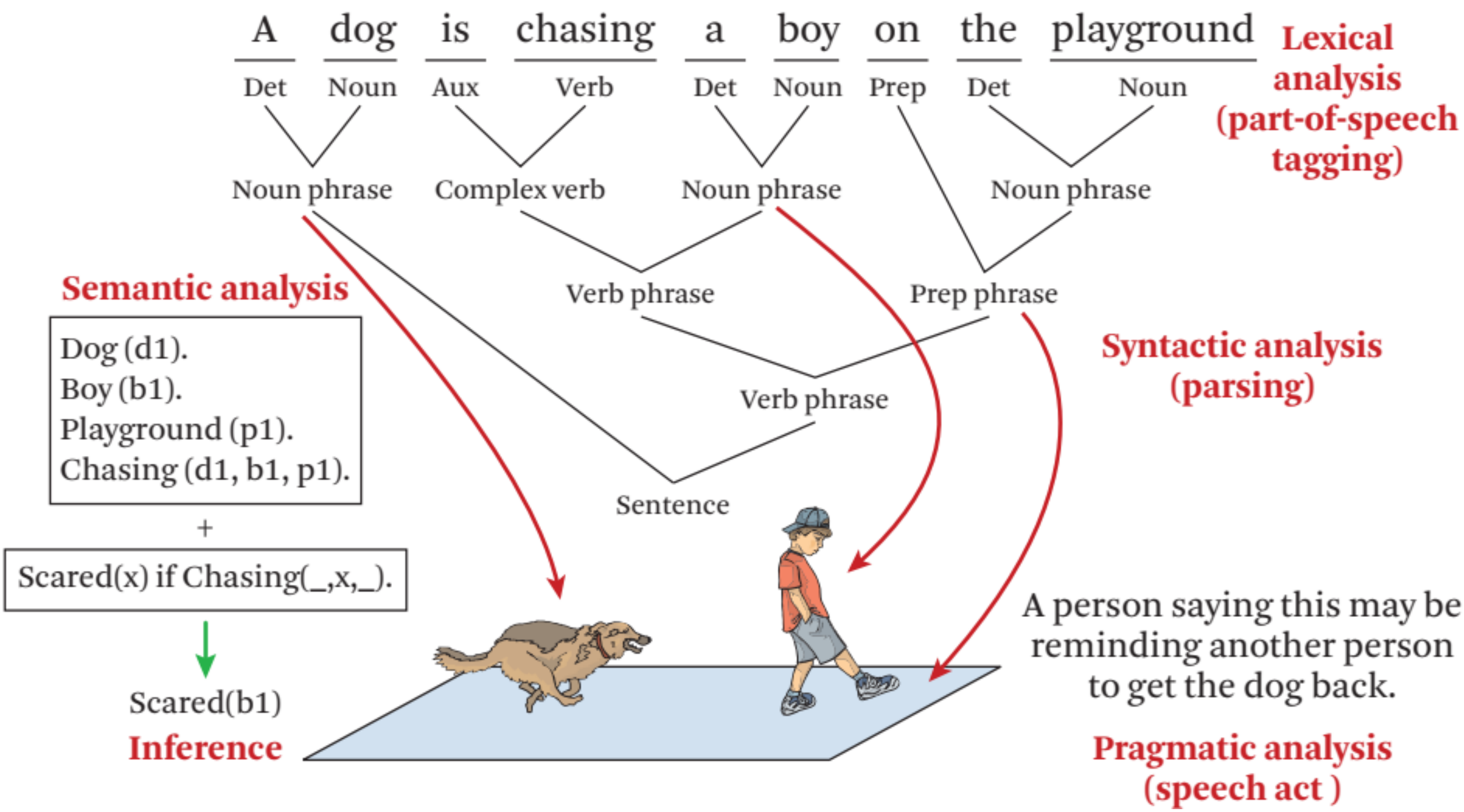
Feature Extraction

Token Sets

nation – 5
civil - 1
war – 2
men – 2
died – 4
people – 5
Liberty – 1
God – 1
…

Loses all order-specific information!
Severely limits context!

Natural Language Processing



A dog is chasing a boy on the playground

Det Noun Aux Verb Det Noun Prep Det Noun

Lexical analysis (part-of-speech tagging)

Noun phrase

Complex verb

Noun phrase

Noun phrase

Verb phrase

Prep phrase

Verb phrase

Sentence

Syntactic analysis (parsing)

Semantic analysis

Dog (d1).
Boy (b1).
Playground (p1).
Chasing (d1, b1, p1).

+

Scared(x) if Chasing(_,x,_).

Scared(b1)

Inference

A person saying this may be reminding another person to get the dog back.

Pragmatic analysis (speech act)

Text Data Management and Analysis: A Practical Introduction to Information Retrieval and Text Mining ChengXiang Zhai, Sean Massung, ACM Books 2016

Structural Knowledge

Beyond individual words, NLP Software must use structural knowledge to string together the words that constitute its response to a question

Consider a question-answering system dealing with the following question:

How much Chinese silk was exported to Western Europe by the end of the 18th century?

System need to know

Lexical semantics, the meaning of all the words (export or silk)

Compositional semantics

- what exactly constitutes Western Europe as opposed to Eastern or Southern Europe,
- What does end mean when combined with the 18th century
- we need to know that by the end of the 18th century is a temporal end-point and not a description of the agent (e.g. How much Chinese silk was exported to Western Europe by southern merchants?)

General NLP—Too Difficult!

Word-level ambiguity

- **"design" can be a noun or a verb** (Ambiguous POS)
- **"root" has multiple meanings** (Ambiguous sense)

Syntactic ambiguity

- **"natural language processing"** (Modification)
- **"A man saw a boy *with a telescope*."** (PP Attachment)

Anaphora resolution

- **"John persuaded Bill to buy a TV for *himself*."**

  (*himself* = John or Bill?)

Presupposition

- **"He has quit smoking." implies that he smoked before.**

**Humans rely on context to interpret (when possible).
This context may extend beyond a given document!**

Knowledge in Language Processing

What distinguishes *language processing applications* from other data processing systems is their use of **knowledge of language**.

- Some simple NLP tasks require limited knowledge of language.
- Big NLP tasks such as conversational agents, machine translation systems, robust question-answering systems, require much broader and deeper knowledge of language.
- **Phonology**–concerns how words are related to the sounds that realize them.
- **Morphology**–concerns how words are constructed from more basic meaning units called morphemes. A ***morpheme*** is the primitive unit of meaning in a language.
- **Syntax**–concerns how can be put together to form correct sentences and determines what structural role each word plays in the sentence and what phrases are subparts of other phrases.

Knowledge in Language Processing

- Semantics–concerns what words mean and how these meaning combine in sentences to form sentence meaning. The study of context-independent meaning.
- Pragmatics–concerns how sentences are used in different situations and how use affects the interpretation of the sentence.
- Discourse–concerns how the immediately preceding sentences affect the interpretation of the next sentence.
    - For example, interpreting pronouns and interpreting the temporal aspects of the information.
- World Knowledge–includes general knowledge about the world.
    - What each language user must know about the other’s beliefs and goals.

Consider a dialogue system

S: How may I help you?

U: When is Saving Private Ryan playing?

S: For what theater?

U: The Paramount theater.

S: Saving Private Ryan is not playing at the Paramount theater, but it's playing at the Madison theater at 3:00, 5:30, 8:00, and 10:30.

https://www.nltk.org/book/ch01.html

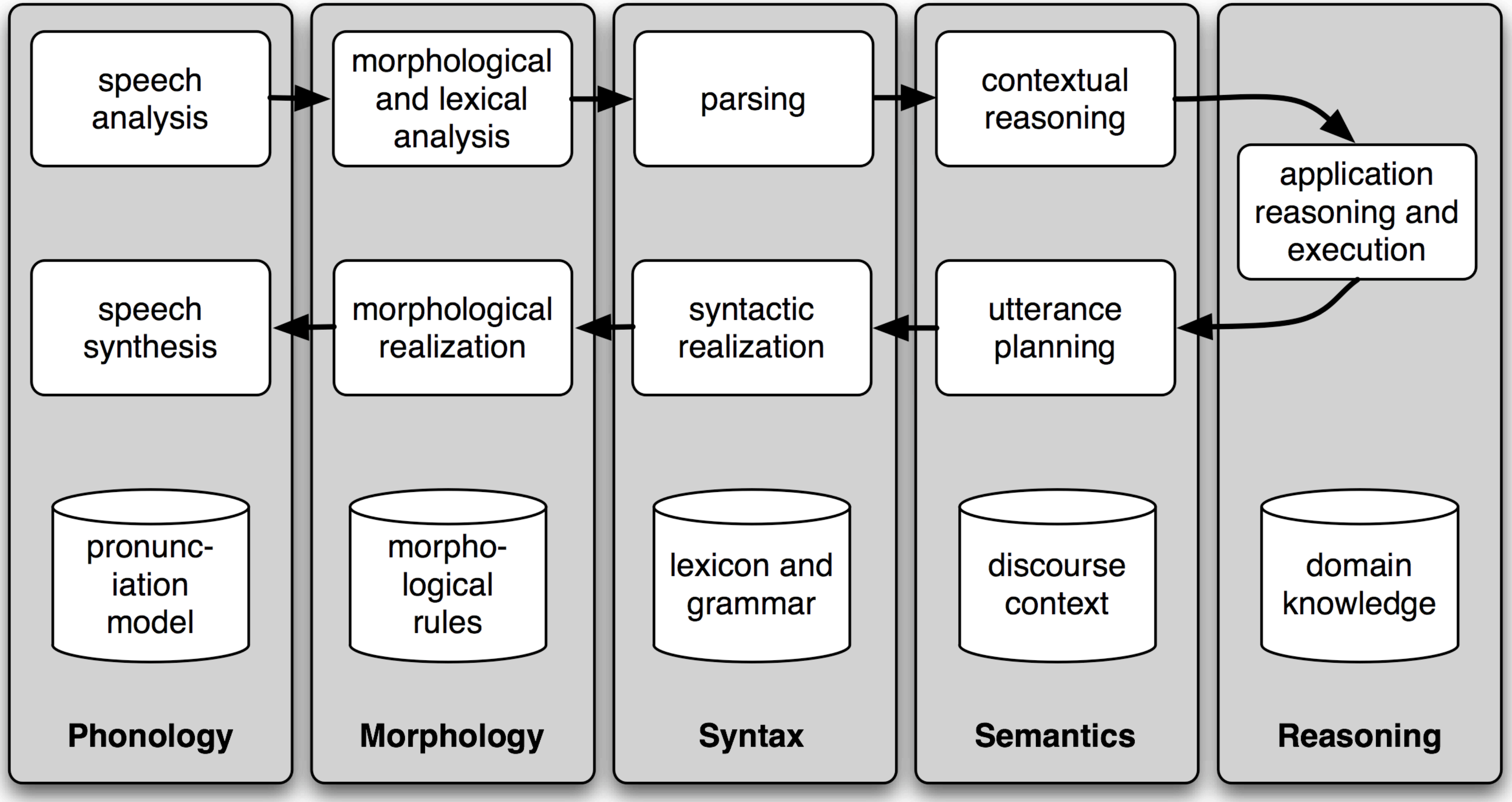
Consider a dialogue system

Observe that this system seems to understand the user's goals:

- The user asks when a movie is showing and the system correctly determines from this that the user wants to see the movie.
  - This inference seems so obvious that you probably didn't notice it was made, yet a natural language system needs to be endowed with this capability in order to interact naturally

- Without it, when asked Do you know when Saving Private Ryan is playing?, a system might unhelpfully respond with a cold Yes.

Consider a dialogue system



https://www.nltk.org/book/ch01.html

Ambiguity

Why NLP is hard?

- Natural language is extremely rich in form and structure, and very ambiguous.
    - How to represent meaning,
    - Which structures map to which meaning structures.
- One input can mean many different things and Ambiguity can be at different levels.
    - Lexical (word level) ambiguity --different meanings of words
    - Syntactic ambiguity --different ways to parse the sentence
    - Interpreting partial information --how to interpret pronouns
    - Contextual information --context of the sentence may affect the meaning of that sentence.
- Many input can mean the same thing.
- Interaction among components of the input is not clear.

Ambiguity

I made her duck.

- How many different interpretations does this sentence have?
- What are the reasons for the ambiguity?
- The categories of knowledge of language can be thought of as ambiguity resolving components.
- How can each ambiguous piece be resolved?
- Does speech input make the sentence even more ambiguous?
    - Yes –deciding word boundaries

Ambiguity (cont.)

Some interpretations of : **I made her duck.**

1. I cooked *duck* for her.
2. I cooked *duck* belonging to her.
3. I created a toy duck which she owns.
4. I caused her to quickly lower her head or body.
5. I used magic and turned her into a duck.

- duck–morphologically and syntactically ambiguous: noun or verb.
- her–syntactically ambiguous: dative or possessive.
- make–semantically ambiguous: cook or create.
- make–syntactically ambiguous:
  - –Transitive –takes a direct object. => 2
  - –Di-transitive –takes two objects. => 5
  - –Takes a direct object and a verb. => 4

Extreme Ambiguity

"Buffalo buffalo Buffalo buffalo buffalo buffalo Buffalo buffalo" is a grammatically correct sentence in English

It is an example of how homonyms and homophones can be used to create complicated linguistic constructs through lexical ambiguity

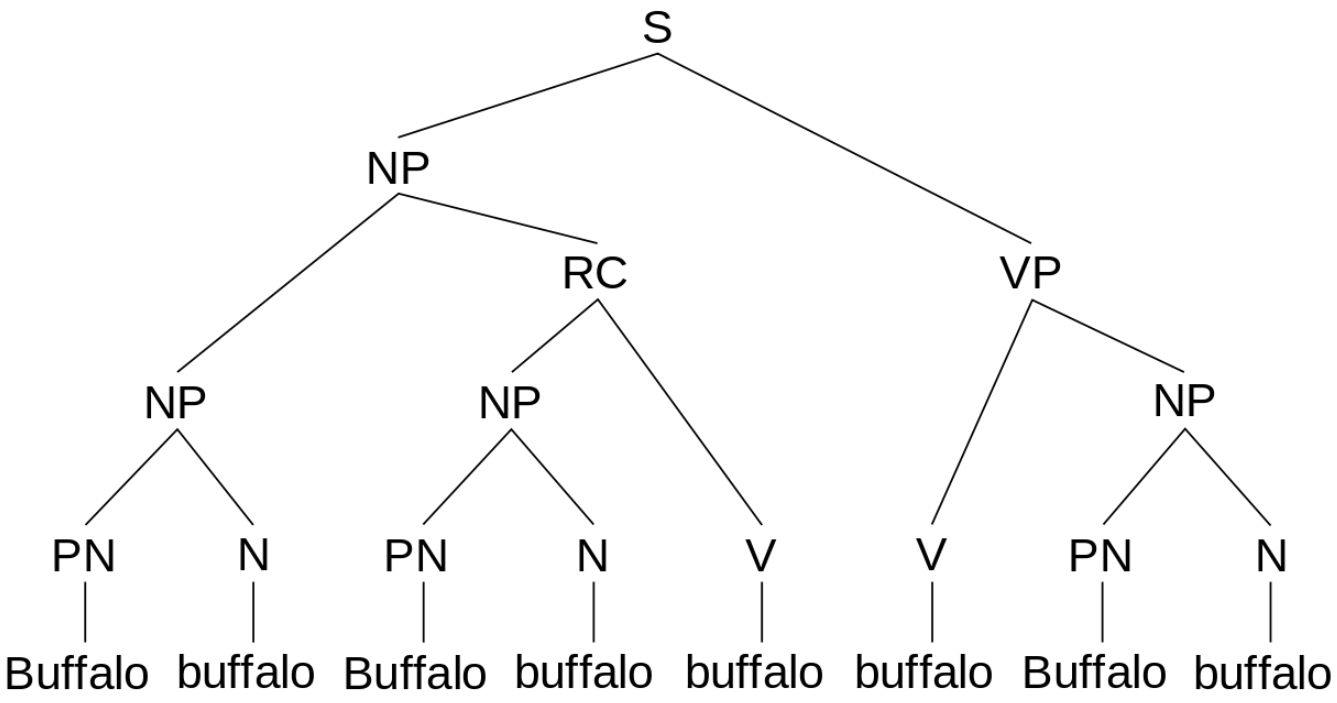
from Dmitri Borgmann's Beyond Language: Adventures in Word and Thought

The sentence employs three distinct meanings of the word buffalo:

- As an adjectival proper noun to refer to a specific place named Buffalo, the city of Buffalo, New York, being the most notable;
- As a verb to buffalo, meaning (in American English[1]) "to bully, harass, or intimidate" or "to baffle"; and
- As a noun to refer to the animal, buffalo (often called bison outside US)

https://en.wikipedia.org/wiki/Buffalo_buffalo_Buffalo_buffalo_buffalo_buffalo_Buffalo_buffalo

Ambiguity (cont.)



Simplified parse tree

S = sentence

NP = noun phrase

RC = relative clause

VP = verb phrase

PN = proper noun

N = noun

V = verb

A semantically equivalent form preserving the original word order is: "Buffalo bison that other Buffalo bison bully also bully Buffalo bison."

https://en.wikipedia.org/wiki/Buffalo_buffalo_Buffalo_buffalo_buffalo_buffalo_Buffalo_buffalo

Word Sense Inventory

What is a "sense" of a word?

- Homonyms (disconnected meanings)
  - bank: financial institution
  - bank: sloping land next to a river
- Polysemes (related meanings with joint etymology)
  - bank: financial institution as corporation
  - bank: a building housing such an institution

Sources of sense inventories

- Dictionaries
- Lexical databases

High Street

High Street is a common street name for the primary business street of a city, town, or village, in many English countries. It implies that it is the focal point for business, especially shopping.

- The word "high" denoted superior rank ("high sheriff", "Lord High Chancellor", "high society")
- "High" also applied to roads as they improved: "highway“
- In Britain, the term 'High Street' has both a generic and a specific meaning: people refer to 'shopping on the high street' both when they mean the main retail area, as well as the specific street of that name
- Canada, Australia, New Zealand and New England (especially Massachusetts), adopted the term to refer to retail shopping areas
- High Street is used less commonly in the Republic of Ireland

Machine Translation

Translation of literature, or poetry, is an intense human endeavor, as rich as any other area of human creativity

Machine translation presently focuses on a number of very practical tasks

Common current uses of machine translation is

- for information access.
- translate some instructions on the web,
- Access perhaps the recipe for a favorite dish, or
- the steps for putting together some furniture or
- user might want to read an article in a newspaper, or
- get information from an online resource like Wikipedia or
- translate a government webpage in a foreign language

Machine Translation

English: He wrote a letter to a friend

Japanese: tomodachi ni tegami-o kaita

Friend to letter Wrote

- The elements of the sentences are in very different places in the different languages.
- In English, the verb is in the middle of the sentence, while in Japanese, the verb kaita comes at the end.
- The Japanese sentence doesn't require the pronoun he, while English does.

Ambiguity Resolution for Translation

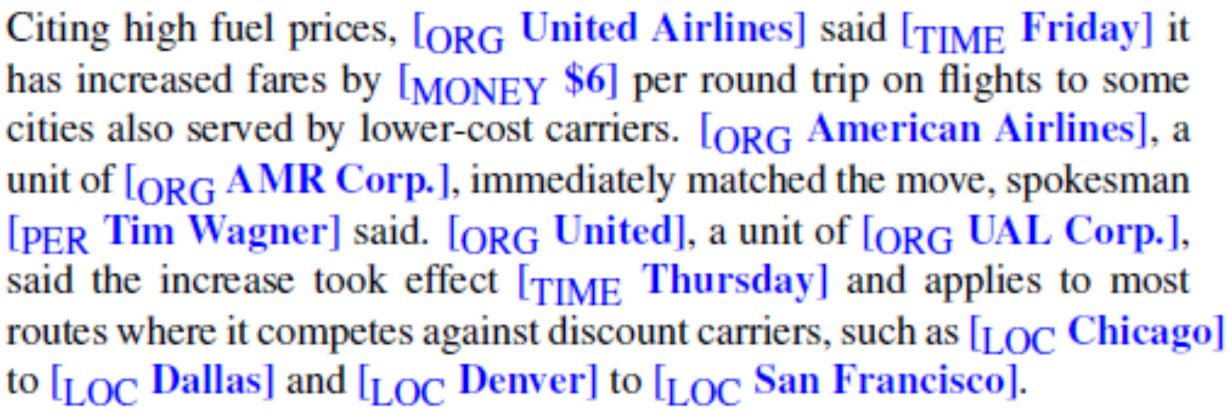
Some horror stories from past efforts :
When translating from English to Russian and then back to English:

- “The spirit is willing but the flesh is weak.” ⇒ “The liquor is good but the meat is spoiled.”
- “Out of sight, out of mind.” ⇒ “Invisible idiot.”

Named Entity Recognition

Janet, Stanford University, and Colorado are all proper nouns from a semantic perspective, these proper nouns refer to different kinds of entities:

- Janet is a person,
- Stanford University is an organization,
- Colorado is a location



Citing high fuel prices, [ORG **United Airlines**] said [TIME **Friday**] it has increased fares by [MONEY **$6**] per round trip on flights to some cities also served by lower-cost carriers. [ORG **American Airlines**], a unit of [ORG **AMR Corp.**], immediately matched the move, spokesman [PER **Tim Wagner**] said. [ORG **United**], a unit of [ORG **UAL Corp.**], said the increase took effect [TIME **Thursday**] and applies to most routes where it competes against discount carriers, such as [LOC **Chicago**] to [LOC **Dallas**] and [LOC **Denver**] to [LOC **San Francisco**].

The text contains 13 mentions of named entities including 5 organizations, 4 locations, 2 times, 1 person, and 1 mention of money

NER is very important in applications such as sentiment analysis, chatbots etc,

Language Models

Regular expressions

Simple pattern-based methods play a crucial role in natural language processing.

Regular expressions can be used to specify strings we might want to extract from a document.

Regular expressions can also help with data transformation by substitutions.

Text Normalization converts text to a more convenient, standard form. It includes tokenization, emoticon/hashtag identification, lemmatization, Stemming, Sentence segmentation

Text Normalization

Lemmatization

task of determining that two words have the same root, despite their surface differences.

- e.g., the words sang, sung, and sings are forms of the verb sing. The word sing is the common lemma of these words, and a lemmatizer maps from all of these to sing.

Lemmatization is essential for processing morphologically complex languages like Arabic.

Stemming

A simpler version of lemmatization in which we mainly just strip suffixes from the end of the word.

Sentence segmentation

Breaking up a text into individual sentences, using cues like periods or exclamation points.

Markov assumption

- The assumption that the probability of a word depends only on the previous word is called a Markov assumption.
- Markov models are the class of probabilistic models that assume we can predict the probability of some future unit without looking too far into the past.
- We can generalize the bigram (which looks one word into the past) to the trigram (which looks two words into the past) and thus to the n-gram (which looks n-1 words into the past).

$$\begin{aligned} P(w_{1:n}) &= P(w_1)P(w_2|w_1)P(w_3|w_{1:2})\ldots P(w_n|w_{1:n-1}) \\ &= \prod_{k=1}^{n} P(w_k|w_{1:k-1}) \end{aligned}$$

Parts of Speech

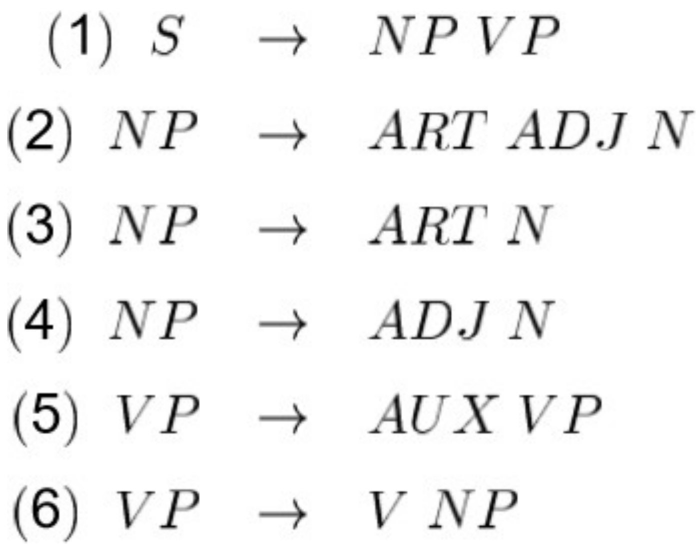
- Dionysius Thrax of Alexandria proposed set of eight parts of speech that became the basis for descriptions of European languages for the last 2000 years.
- Named entity is anything that can be referred to with a proper name: a person, a location, an organization, etc.
- Parts of speech (also known as POS) and named entities offer clues to sentence structure and meaning.
- If we know POS for a word it tells us about likely neighboring words and syntactic structure, making part-of-speech tagging a key aspect of parsing.
    - Nouns in English are preceded by determiners and adjectives, verbs by nouns
    - Verbs have dependency links to nouns

Context-free Grammars

- Context-free grammars are the backbone of many formal models of the syntax of natural language (and, of computer languages).
- CFG play a role in many computational applications, including grammar checking, semantic interpretation, dialogue understanding, and machine translation.
- They are powerful enough to express sophisticated relations among the words in a sentence, yet computationally tractable enough that efficient algorithms exist for parsing sentences with them.
- A context-free grammar consists of a set of rules or productions, which express the ways that symbols of the language can be grouped and ordered together
    - For example, NP express that an NP (or noun phrase) can be composed of either a Proper Noun or a determiner (Det) followed by a Nominal; a Nominal in turn can consist of one or more Nouns

Context Free Grammar

Example:



(1) $S \rightarrow NP\,VP$

(2) $NP \rightarrow ART\ ADJ\ N$

(3) $NP \rightarrow ART\ N$

(4) $NP \rightarrow ADJ\ N$

(5) $VP \rightarrow AUX\,VP$

(6) $VP \rightarrow V\ NP$

Representations and Understanding

Allen 1995: Natural Language Understanding - Introduction

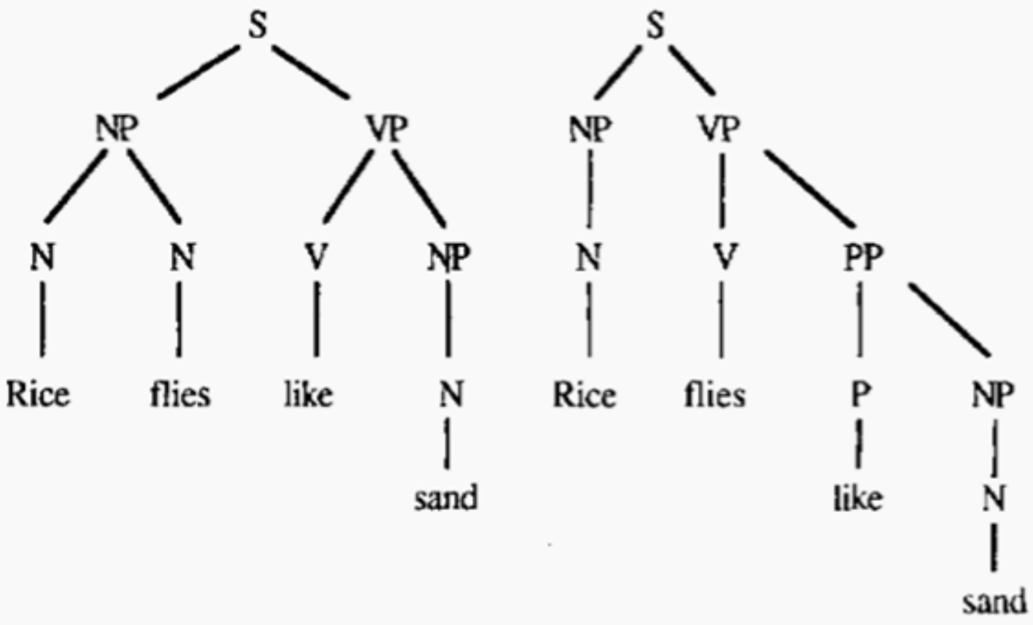


Figure 1.4 Two structural representations of *Rice flies like sand.*

Figure 1.4 Two structural representations of "*Rice flies like sand*".
Rice fly is type of a fly

CFG Rules

S -> NP VP
NP -> N N
NP -> N VP
-> V NP VP
-> V PP PP
-> P NP

Lexicon

Rice : N
Flies : N, V
Like : V, P
Sand : N

NP – Noun Phrases
VP – Verb Phrases
PP – Prepositional Phrases

innovate achieve lead

BITS Pilani

Pilani|Dubai|Goa|Hyderabad

Corpus - Information Retrieval

Text Databases and IR

Text databases (document databases)

- Large collections of documents from various sources: news articles, research papers, books, digital libraries, e-mail messages, and Web pages, library database, etc.
- Data stored is usually *semi-structured*
- Traditional information retrieval techniques become inadequate for the increasingly vast amounts of text data

Information retrieval

- A field developed in parallel with database systems
- Information is organized into (a large number of) documents
- Information retrieval problem: locating relevant documents based on user input, such as keywords or example documents

Information Retrieval

Typical IR systems

- Online library catalogs
- Online document management systems

Information retrieval vs. database systems

- Some DB problems are not present in IR, e.g., update, transaction management, complex objects
- Some IR problems are not addressed well in DBMS, e.g., unstructured documents, approximate search using keywords and relevance

Information Retrieval Techniques

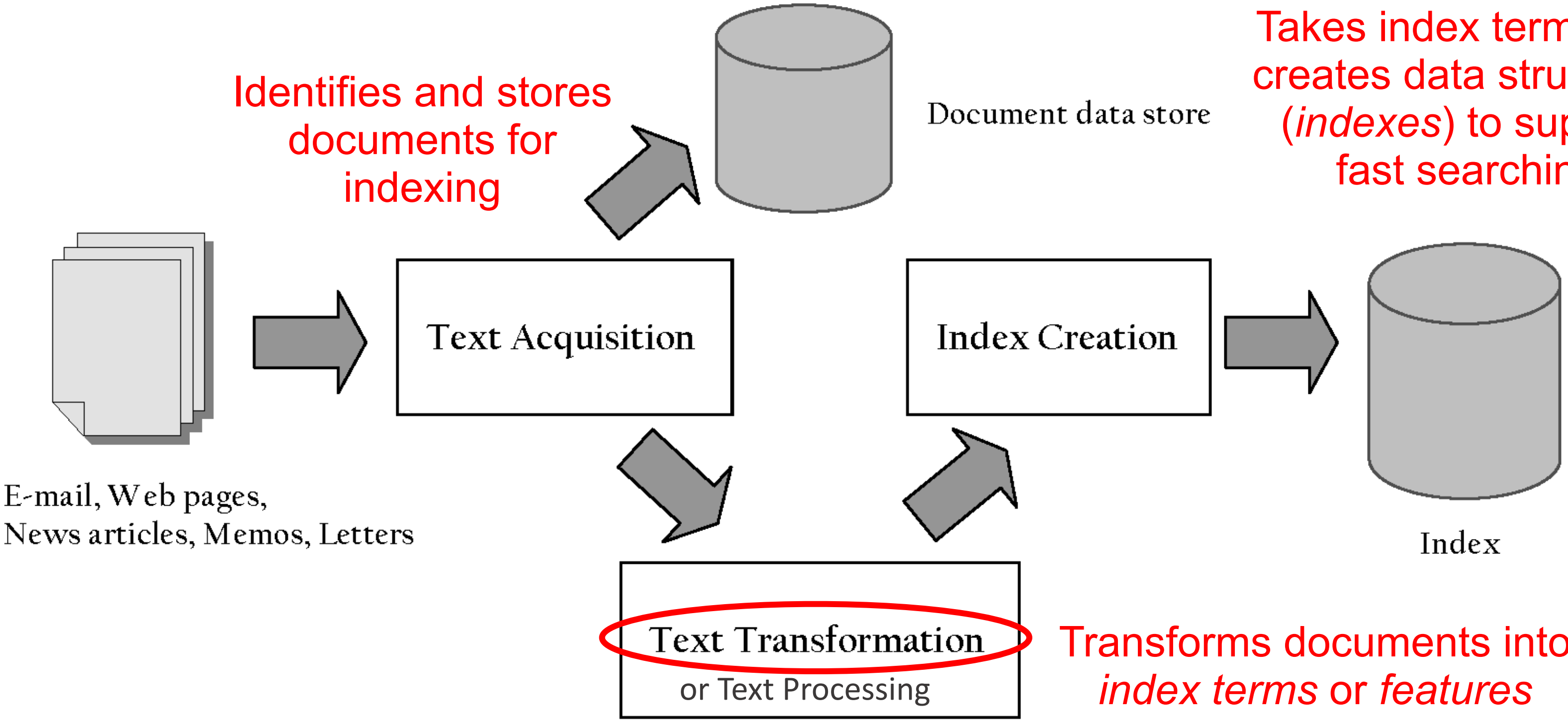
Basic Concepts

- A document can be described by a set of representative keywords called index terms.
- Different index terms have varying relevance when used to describe document contents.
- This effect is captured through the assignment of numerical weights to each index term of a document. (e.g.: frequency, tf-idf)

DBMS Analogy

- Index Terms → Attributes
- Weights → Attribute Values

Indexing Process



Text + Meta data
(Doc type, structure,
features, size, etc.)

Identifies and stores
documents for
indexing

Takes index terms and
creates data structures
(*indexes*) to support
fast searching

Transforms documents into
index terms or *features*

Zipf's Law

Distribution of word frequencies is very *skewed*

- Few words occur very often, many hardly ever occur
- e.g., "the" and "of", two common words, make up about 10% of all word occurrences in text documents

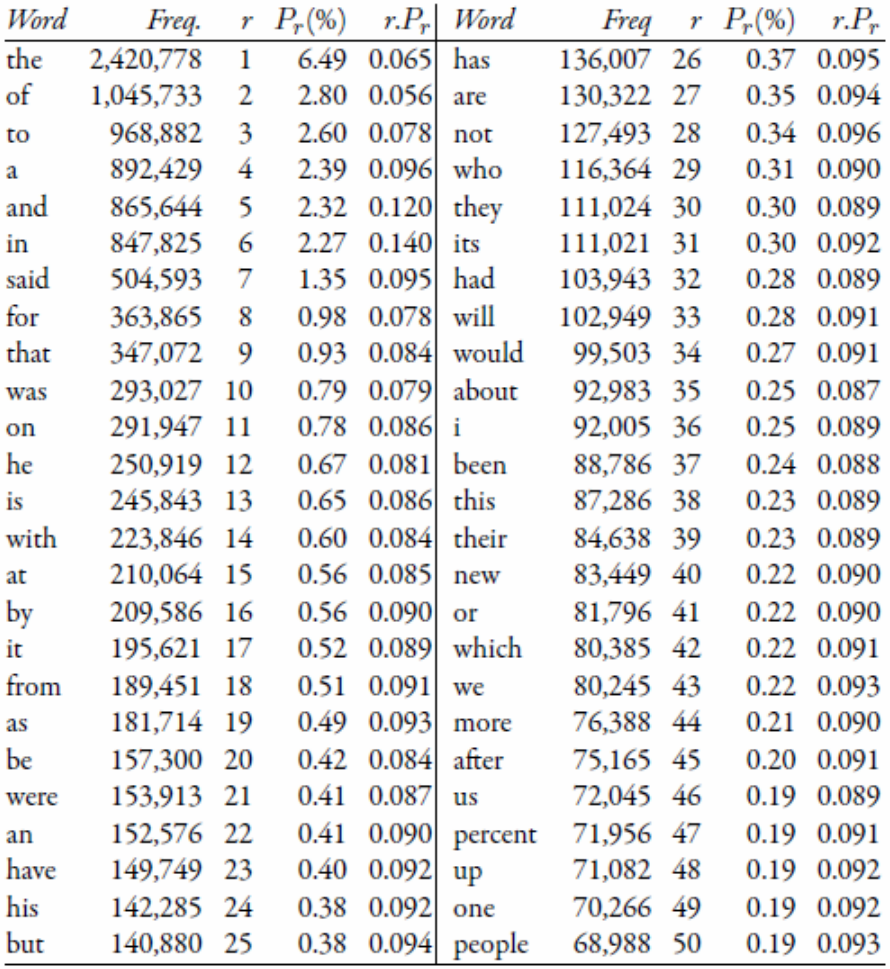
Zipf's law:

- The frequency *f* of a word in a corpus is inversely proportional to its rank *r* (assuming words are ranked in order of *decreasing* frequency)

$$f = \frac{k}{r} \Rightarrow f \times r = k$$

where *k* is a constant for the corpus

Top 50 Words from AP89



Word	*Freq.*	*r*	$P_r(\%)$	$r.P_r$	*Word*	*Freq*	*r*	$P_r(\%)$	$r.P_r$
the	2,420,778	1	6.49	0.065	has	136,007	26	0.37	0.095
of	1,045,733	2	2.80	0.056	are	130,322	27	0.35	0.094
to	968,882	3	2.60	0.078	not	127,493	28	0.34	0.096
a	892,429	4	2.39	0.096	who	116,364	29	0.31	0.090
and	865,644	5	2.32	0.120	they	111,024	30	0.30	0.089
in	847,825	6	2.27	0.140	its	111,021	31	0.30	0.092
said	504,593	7	1.35	0.095	had	103,943	32	0.28	0.089
for	363,865	8	0.98	0.078	will	102,949	33	0.28	0.091
that	347,072	9	0.93	0.084	would	99,503	34	0.27	0.091
was	293,027	10	0.79	0.079	about	92,983	35	0.25	0.087
on	291,947	11	0.78	0.086	i	92,005	36	0.25	0.089
he	250,919	12	0.67	0.081	been	88,786	37	0.24	0.088
is	245,843	13	0.65	0.086	this	87,286	38	0.23	0.089
with	223,846	14	0.60	0.084	their	84,638	39	0.23	0.089
at	210,064	15	0.56	0.085	new	83,449	40	0.22	0.090
by	209,586	16	0.56	0.090	or	81,796	41	0.22	0.090
it	195,621	17	0.52	0.089	which	80,385	42	0.22	0.091
from	189,451	18	0.51	0.091	we	80,245	43	0.22	0.093
as	181,714	19	0.49	0.093	more	76,388	44	0.21	0.090
be	157,300	20	0.42	0.084	after	75,165	45	0.20	0.091
were	153,913	21	0.41	0.087	us	72,045	46	0.19	0.089
an	152,576	22	0.41	0.090	percent	71,956	47	0.19	0.091
have	149,749	23	0.40	0.092	up	71,082	48	0.19	0.092
his	142,285	24	0.38	0.092	one	70,266	49	0.19	0.092
but	140,880	25	0.38	0.094	people	68,988	50	0.19	0.093

Associated Press collection of news stories from 1989 (called AP89)

Vocabulary Growth

Heaps' Law, another prediction of *word occurrence*

As *corpus* grows, so does *vocabulary size.* However, fewer new words when corpus is already large

Observed relationship (Heaps' Law):

$$v = k \times n^{\beta}$$

where

v is the *vocabulary size* (number of *unique words*)

n is the *total number* of *words* in corpus

k, β are parameters that vary for each corpus

(typical values given are $10 \leq k \leq 100$ and $\beta \approx 0.5$)

- Predicting that the number of new words increases very rapidly when the corpus is small

Heaps' Law Predictions

Number of new words *increases* very rapidly when the corpus is small, and continue to increase indefinitely

Predictions for TREC collections are accurate for large numbers of words, e.g.,

- First 10,879,522 *words* of the AP89 collection scanned
- Prediction is 100,151 *unique words*
- Actual number is 100,024

Predictions for *small* numbers of words (i.e., < 1000) are much worse

Heaps' Law on the Web

Heaps' Law works with very *large* corpora

- New words occurring even after seeing 30 million!
- Parameter values different than typical TREC values

New words come from a variety of sources

- *Spelling errors*, *invented words* (e.g., product, company names), *code*, *other languages*, *email addresses*, etc.

Search engines must deal with these *large* and *growing vocabularies*

Vector Space Model

Represent a doc by a term vector

- Term: basic concept, e.g., word or phrase
- Each term defines one dimension
- N terms define a N-dimensional space
- Element of vector corresponds to term weight
- E.g., d = $(x_1,…,x_N)$, x_i is “importance” of term i

New document is assigned to the most likely category based on vector similarity.

How to Assign Weights

Two-fold heuristics based on frequency

- TF (Term frequency)
    - More frequent ***within*** a document → more relevant to semantics
- IDF (Inverse document frequency)
    - Less frequent ***among*** documents → more discriminative

TF-IDF Weighting

TF-IDF weighting : **weight(t, d) = TF(t, d) * IDF(t)**

- Freqent within doc → high tf → high weight
- Selective among docs → high idf → high weight

Recall VS model

- Each selected term represents one dimension
- Each doc is represented by a feature vector
- Its *t*-term coordinate of document *d* is the TF-IDF weight
- This is more reasonable

Just for illustration …

- Many complex and more effective weighting variants exist in practice

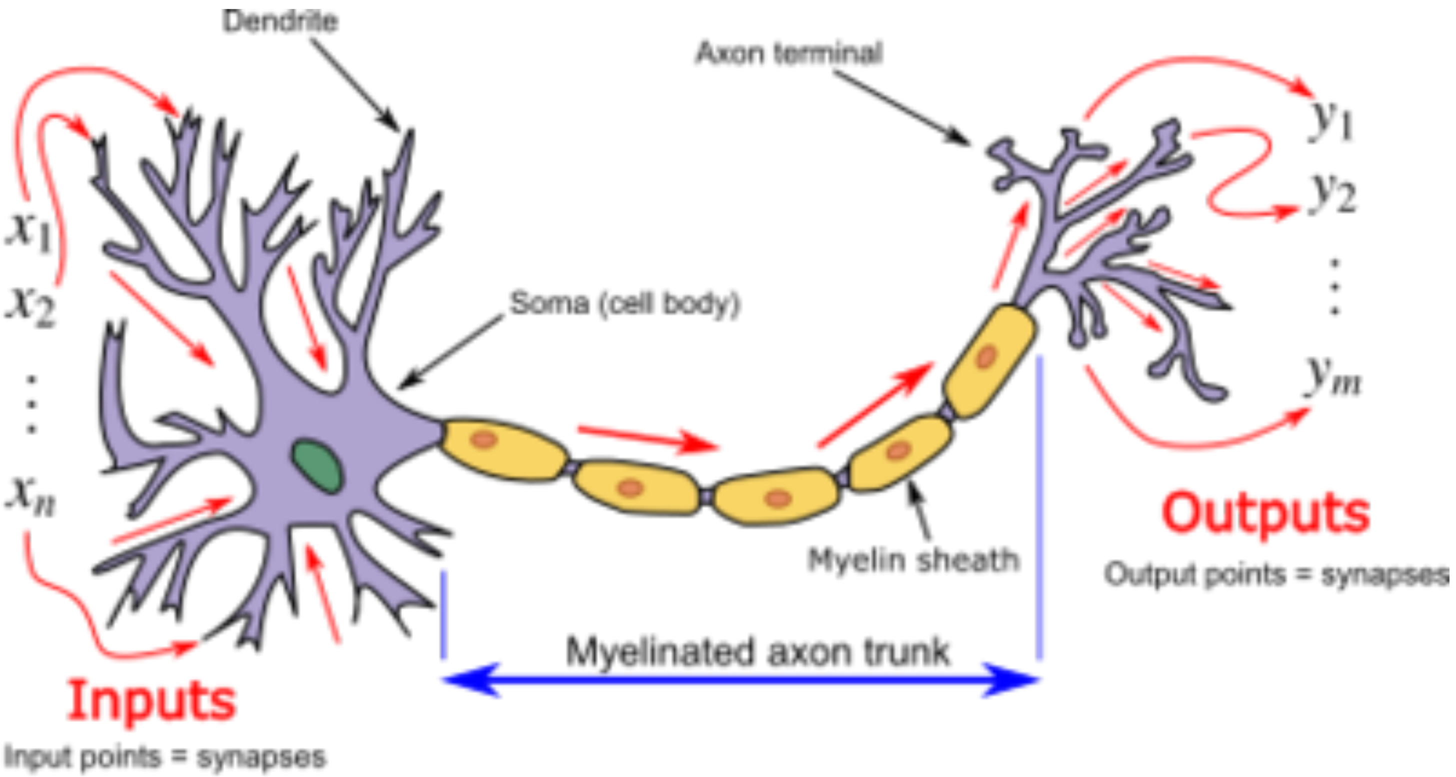
innovate achieve lead

BITS Pilani

Pilani|Dubai|Goa|Hyderabad

Neural Networks

Biological Neuron Model



https://en.wikipedia.org/wiki/Biological_neuron_model

Artificial Neural Network

- It is logical to look at the brain's architecture for inspiration on how to build an intelligent machine.
- Artificial neural networks (ANNs) are machine learning models inspired by the networks of biological neurons found in our brains
  - Very similar to how birds inspired aero plane design. But aero planes evolved without consideration of birds.
- Some researchers argue that we should drop the biological analogy altogether (e.g., call "units" rather than "neurons")

The Basic Ideas in Neural Networks, Communications of ACM 1994 by David E. Rumelhart et al.

Artificial Neural Network

- ANNs are at the very core of deep learning.
- They are versatile, powerful, and scalable to tackle large and highly complex machine learning tasks such as
  - Classifying billions of images (e.g., Google Images),
  - Powering speech recognition services (e.g., Apple’s Siri),
  - Recommending the best videos to watch to hundreds of millions of users every day (e.g., YouTube), or
  - Learning to beat the world champion at the game of Go (DeepMind’s AlphaGo)

NLP Validation – Turing Test

Turing Test

The Turing Test is a method of inquiry in artificial intelligence for determining whether or not a computer is capable of thinking like a human being

The original Turing Test requires three terminals, each of which is physically separated from the other two. One terminal is operated by a computer, while the other two are operated by humans.

During the test, one of the humans functions as the questioner, the second human and the computer function as respondents.

> The questioner interrogates the respondents within a specific subject area, using a specified format and context.
>
> After a preset length of time or number of questions, the questioner is then asked to decide which respondent was human and which was a computer.

ELIZA

User: I am unhappy.

ELIZA: DO YOU THINK COMING HERE WILL HELP YOU NOT TO BE UNHAPPY

User: I need some help, that much seems certain.

ELIZA: WHAT WOULD IT MEAN TO YOU IF YOU GOT SOME HELP

User: Perhaps I could learn to get along with my mother.

ELIZA: TELL ME MORE ABOUT YOUR FAMILY

User: My mother takes care of me.

ELIZA: WHO ELSE IN YOU FAMILY TAKES CARE OF YOU

User: My father.

ELIZA: YOUR FATHER

User: You are like my father in some ways.

The above dialogue sequence is from ELIZA, an early(1960) natural language processing system that could carry on a limited conversation with a user by imitating the responses of a psychotherapist

ELIZA

ELIZA is a surprisingly simple program that uses pattern matching to recognize phrases like “I need X” and translate them into suitable outputs like “What would it mean to you if you got X?”.

Eliza’s mimicry of human conversation was very successful: many people who interacted with ELIZA came to believe that it really understood them and their problems

ELIZA could pass the Turing Test by manipulating symbols it does not understand fully. Some experts argue that this does not determine intelligence comparable to humans.

More on Turing Test

There are supporters, detractors for Turing Test

The Loebner Prize is an annual competition in artificial intelligence that awards prizes to the computer programs considered by the judges to be the most human-like.

The format of the competition was that of a standard Turing test.

In each round, a human judge simultaneously holds textual conversations with a computer program and a human being via computer. Based upon the responses, the judge must decide which is which.

(The prize is reported as defunct since 2020)

More on Turing Test

There have been a number of variations/alternatives to the Turing Test to make it more relevant.

Reverse Turing Test -- where a human tries to convince a computer that it is not a computer. An example of this is a CAPTCHA.

Total Turing Test -- where the questioner can also test perceptual abilities as well as the ability to manipulate objects.

Minimum Intelligent Signal Test -- where only true/false and yes/no questions are given.

Alternatives to Turing Tests were later developed because many see the Turing test to be flawed. These alternatives include tests such as:

The Marcus Test -- in which a program that can 'watch' a television show is tested by being asked meaningful questions about the show's content.

The Lovelace Test 2.0 -- which is a test made to detect AI through examining its ability to create art (e.g. fabrication of fictional stories).

Winograd Schema Challenge -- which is a test that asks multiple-choice questions in a specific format.

Thank You

References

	Author(s), Title, Edition, Publishing House
T1	Speech and Language processing: An introduction to Natural Language Processing, Computational Linguistics and speech Recognition by Daniel Jurafsky and James H. Martin[3rd edition]
T2	Foundations of statistical Natural language processing by Christopher D.Manning and Hinrich schutze