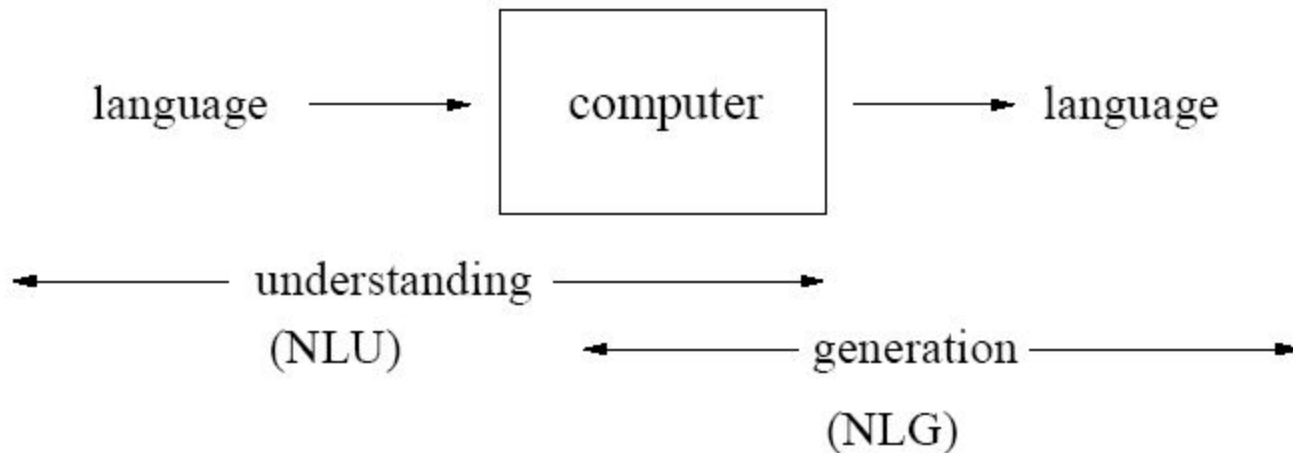# Introduction to Voice Computing(NLP)

By: Fahim Uz Zaman

What is NLP?

- Natural Language Processing (NLP) is a field in Artificial Intelligence (AI) devoted to creating computers that use natural language as input and/or output.
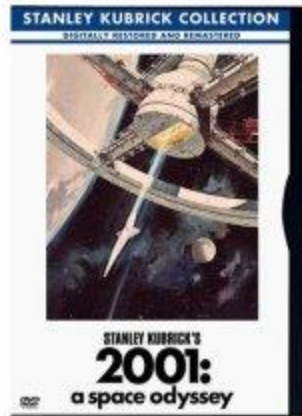
Why NLP?

- To interact with computing devices using human (natural) languages. For example,
    - Building intelligent robots (AI).
    - Enabling voice-controlled operation.
- To access (large amount of) information and knowledge stored in the form of human languages quickly.
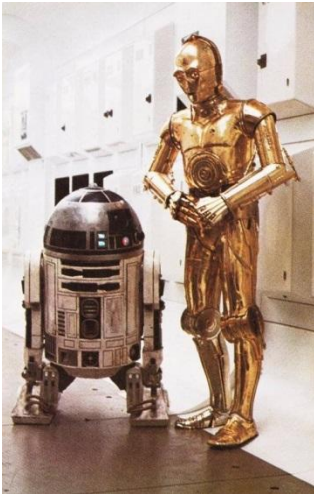
Early days of NLP:
Machines that Can Speak

- HAL 9000 in "2001: A Space Odyssey"

Machines that Can Speak (cont.)

- C3PO
  in Star Wars

- KITT
  in Knight Rider

But Still a Sci-Fi…

- "However, as 2001 approached it became clear that *2001*'s predictions in computer technology were far fetched. Natural language, lip reading, planning and plain common sense in computers were still the stuff of science fiction."
HAL 9000 - Wikipedia

Test for Intelligence – Turing Test

- Alan Turing (1950) proposed a test of a machine's capability to perform human-like conversation.

  - A human judge engages in a natural language conversation with two other parties, one a human and the other a machine; if the judge cannot reliably tell which is which, then the machine is said to pass the test.

Early Conversational Programs

- ELIZA (by Joseph Weizenbaum), 1966
  - A psychotherapist
  - No real understanding; simple pattern-matching to respond to user input -> canned responses

(my ?x depresses me) (why does your ?x depress you) 5

(life ?x) (why do you say it ?x) 3

(I could ?x) (you could ?x) 2

(because ?x) (that is a good reason) 3

(?x) (tell me more) 0

>> Hello.

How do you do. Please state your problem.

>> Men are all alike.

In what way ?

>> They're alwas bugging us about something or other.

I'm not sure I understand you fully.

>> They're always bugging us about something or other.

Can you think of a specific example ?

>> Well, my boyfriend made me come here.

Your boyfriend made you come here ?

>> He says I'm depressed much of the time.

I am sorry to hear that you are depressed.

Modern NLP:
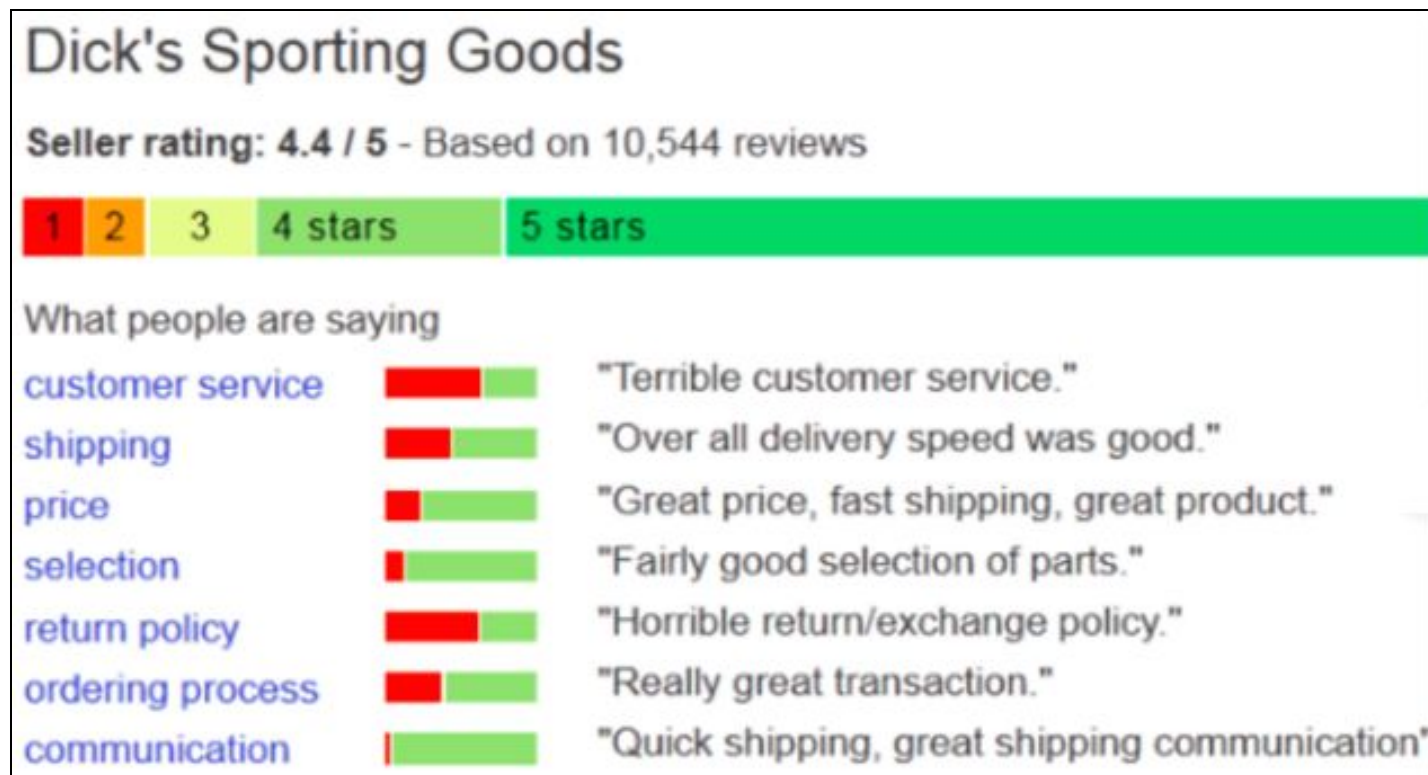# NLP in the Days of Big Data

Three trends:

1. An **enormous amount of information** is now available in machine readable form as natural language text (newspapers, web pages, medical records, financial filings, product reviews, discussion forums, etc.)

2. Conversational agents are becoming an important form of human-computer **communication**

3. Much of human-human interaction is now mediated by computers via **social media**

NLP Applications

- Three prominent application areas:
  - Text analytics/mining (from *"**unstructured data**"*)
    - Sentiment analysis
    - Topic identification
    - Digital Humanities (*"new ways of doing scholarship that involve collaborative, transdisciplinary, and computationally engaged research, teaching, and publishing."*)
  - Conversational agents
    - Siri, Cortana, Amazon Alexa, Google Assistant
    - Chatbots
  - Machine translation

Text Analytics

- Data-mining of weblogs, microblogs, discussion forums, user reviews, and other forms of user-generated media.

Dick's Sporting Goods

Seller rating: 4.4 / 5 - Based on 10,544 reviews

| 1 | 2 | 3 | 4 stars | 5 stars |

What people are saying

customer service		"Terrible customer service."
shipping		"Over all delivery speed was good."
price		"Great price, fast shipping, great product."
selection		"Fairly good selection of parts."
return policy		"Horrible return/exchange policy."
ordering process		"Really great transaction."
communication		"Quick shipping, great shipping communication"

Text Analytics (cont.)

- Typically this involves the extraction of **limited** kinds of semantic and pragmatic information from texts
  - Entity mentions
  - Concept identification
  - Sentiment

Conversational Agents

- Combine
  - Speech recognition/synthesis
  - Question answering
    - From the web and from structured information sources (freebase, dbpedia, yago, etc.)
  - Simple agent-like abilities
    - Create/edit calendar entries
    - Reminders
    - Directions
    - Invoking/interacting with other apps

Hi Noah, have you taken your medication today?

I will. Thanks for the reminder.

Do you still have a runny nose and fever?

I'm feeling better, but I think I should see Dr. Thompson.

I'll schedule an appointment. Let me know a convenient time slot.

AMAZON ALEXA
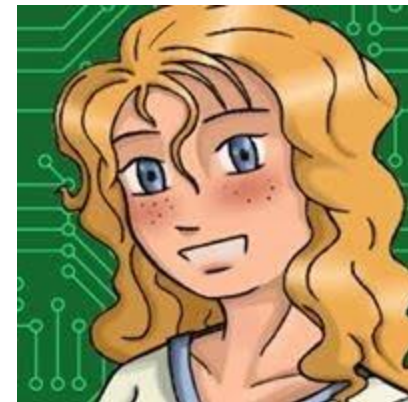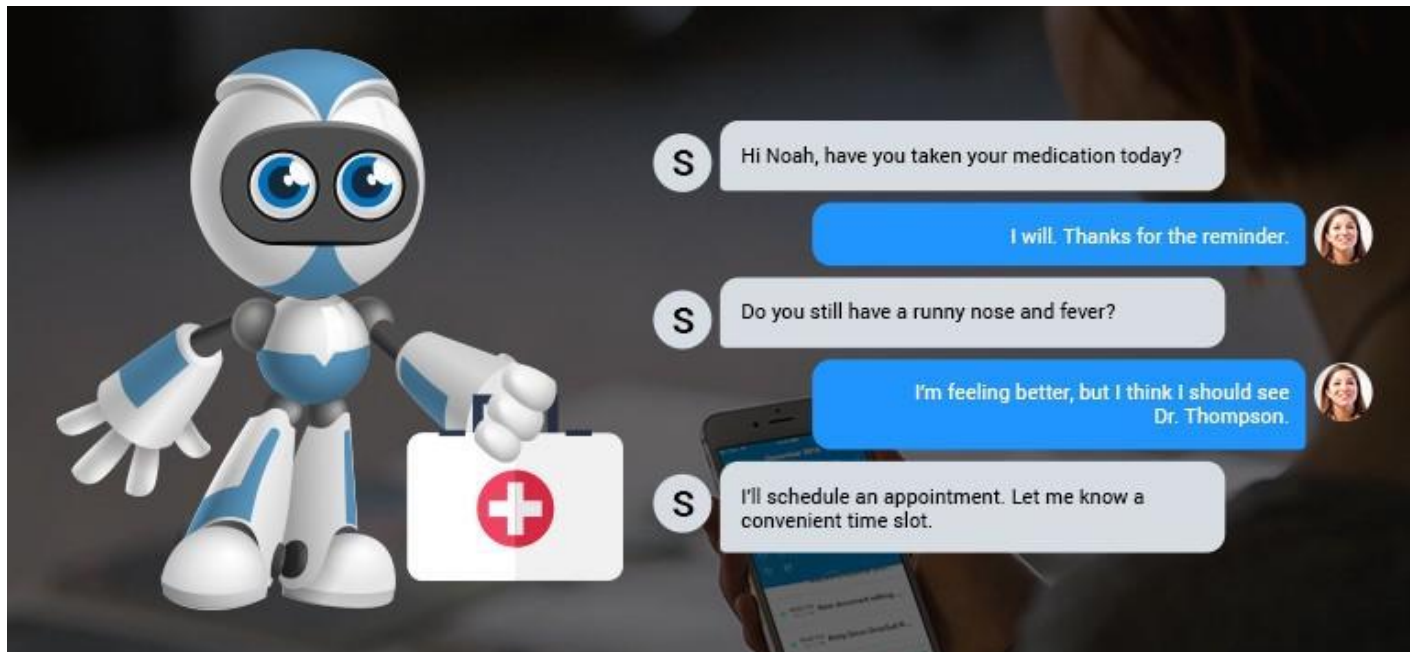
Mitsuku

Question Answering

- Traditional *information retrieval* provides documents/resources that provide users with what they need to satisfy their information needs.

- *Question answering* on the other hand directly provides an answer to information needs posed as questions.

IBM Watson

https://www.youtube.com/watch?v=WFR3lOm_xhE

Machine Translation

- The automatic translation of texts between languages is one of the oldest non-numerical applications in Computer Science.

- In the past 15 years or so, MT has gone from a niche academic curiosity to a robust commercial industry.
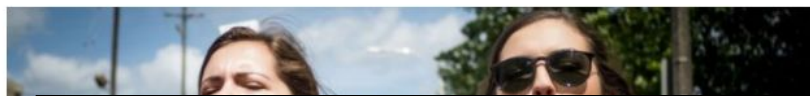
Text Mining Applications – Unsupervised

• Text clustering
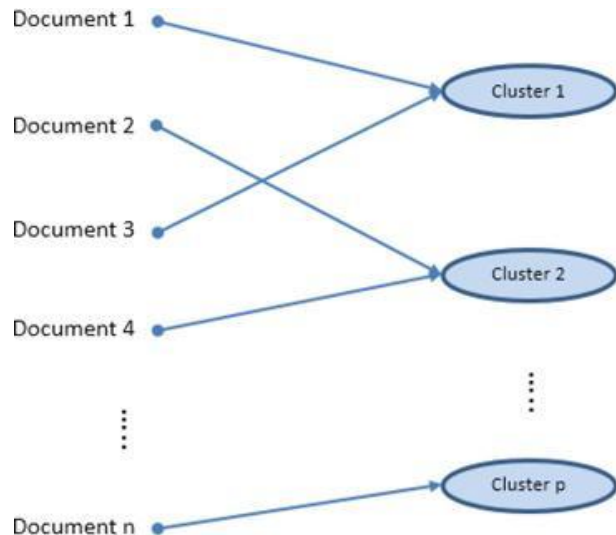
Cluster No.	Comment	Key Words
1	1, 3, 4	doctor, staff, friendly, helpful
2	5, 6, 8	treatment, results, time, schedule
3	2, 7	service, clinic, fast

• Trend analysis
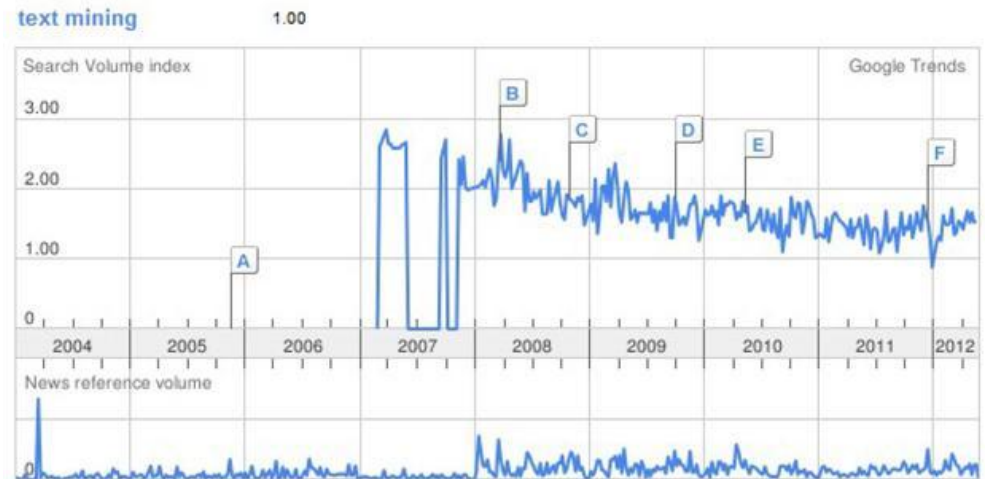
Trend for the Term "text mining" from Google Trends

Text Mining Applications – Supervised

- Many typical **predictive modeling** or classification applications can be enhanced by incorporating textual data in addition to traditional input variables.
  - churning propensity models that include customer center notes, website forms, e-mails, and Twitter messages
  - hospital admission prediction models incorporating medical records notes as a new source of information
  - insurance fraud modeling using adjustor notes
  - sentiment categorization (next page)
  - stylometry or forensic applications that identify the author of a particular writing sample
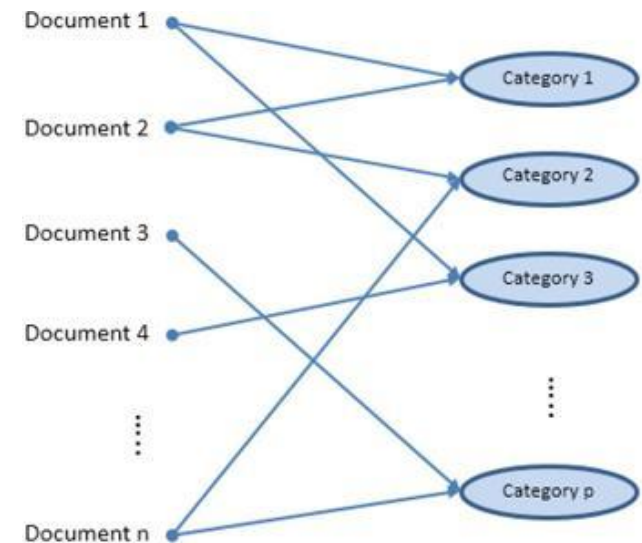
Sentiment Analysis

- The field of sentiment analysis deals with categorization (or classification of opinions expressed in textual documents

> The TV is **wonderful**. **Great size**, **great picture**, **easy interface**. It makes a **cute** little song when you **boot** it up and when you shut it off. I just want to point out that the **43"** **does not** in fact play videos from the USB .This is really **annoying** because that was one of the major perks I wanted from a new **TV**. Looking at the product description now, I realize that the feature list applies to the **X758** series as a whole, and that each model's capabilities are listed below. Kind of a **dumb** oversight on my part, but it's equally **stupid** to put a description that does not apply on the listing for a very specific model.

Green color represents positive tone, red color represents negative tone, and product features and model names are highlighted in blue and brown, respectively.

Structured + Text Data in Predictive Models

- Use of both types of data in building predictive models.
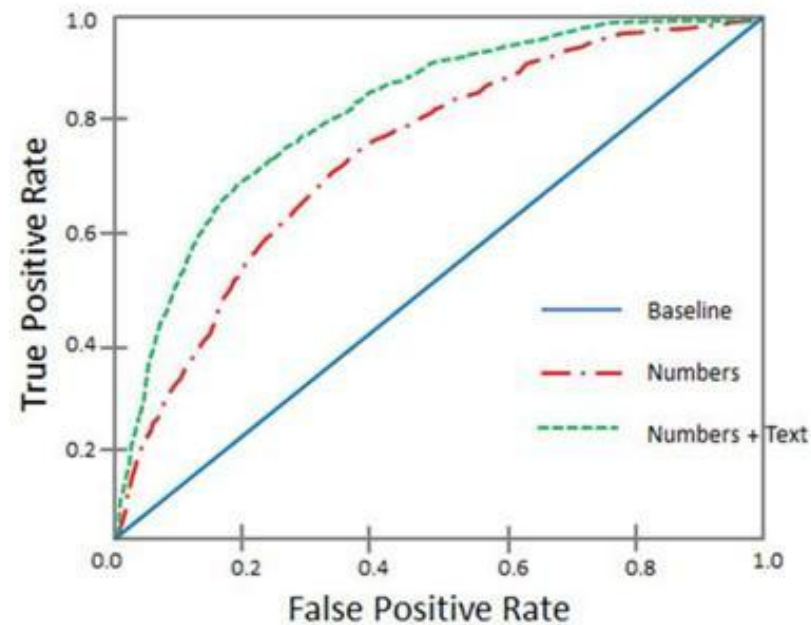
ROC Chart of Models With and Without Textual Comments

NLP Tasks

- NLP applications require several NLP analyses:
  - Word tokenization
  - Sentence boundary detection
  - Part-of-speech (POS) tagging
    - to identify the part-of-speech (e.g. noun, verb) of each word
  - Named Entity (NE) recognition
    - to identify proper nouns (e.g. names of person, location, organization; domain terminologies)
  - Parsing
    - to identify the syntactic structure of a sentence
  - Semantic analysis
    - to derive the meaning of a sentence

1. Part-Of-Speech (POS) Tagging

- POS tagging is a process of assigning a POS or lexical class marker to each word in a sentence (and all sentences in a corpus).

    Input: **the lead paint is unsafe**

    Output: **the/Det lead/N paint/N is/V unsafe/Adj**

2. Named Entity Recognition (NER)

- NER is to process a text and identify named entities in a sentence
  - e.g. "U.N. official Ekeus heads for Baghdad."

[ORG U.N.] official [PER Ekeus] heads for [LOC Baghdad] .

3. Shallow Parsing

- Shallow (or Partial) parsing identifies the (base) syntactic phases in a sent
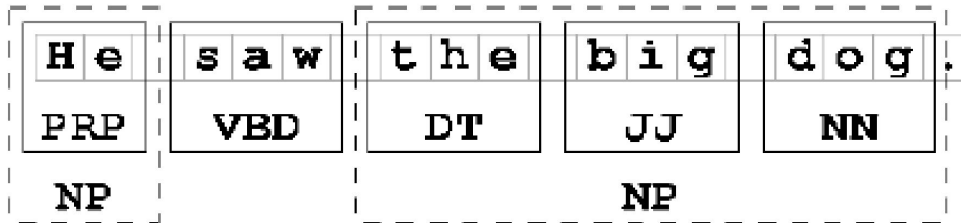
```
[NP He] [v saw] [NP the big
dog]
```

- After NEs are identified, **dependency parsing** is often applied to extract the syntactic/dependency relations between the NEs.

[PER Bill Gates] founded [ORG Microsoft].

found

nsubj   *dobj*

Bill Gates   Microsoft
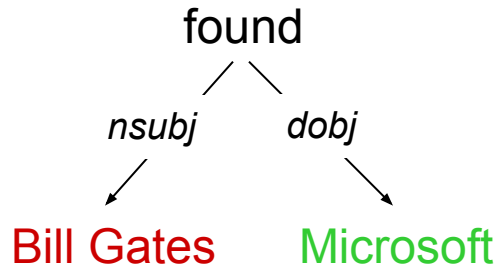
Dependency Relations
nsubj(Bill Gates, found)
dobj(found, Microsoft)

4. Information Extraction (IE)

- Identify specific pieces of information (data) in an unstructured or semi-structured text

- Transform unstructured information in a corpus of texts or web pages into a structured database (or templates)

- Applied to various types of text, e.g.
  - Newspaper articles
  - Scientific articles
  - Web pages
  - etc.

Bridgestone Sports Co. said Friday it had set up a joint venture in Taiwan with a local concern and a Japanese trading house to produce golf clubs to be supplied to Japan.

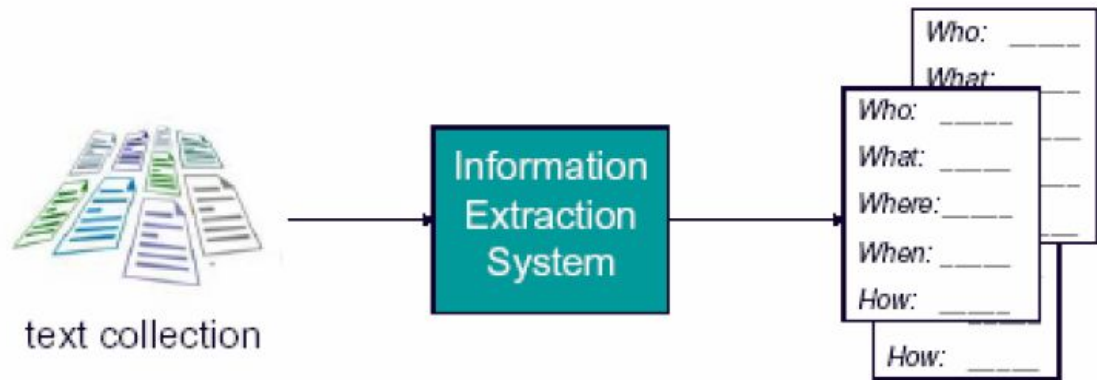The joint venture, Bridgestone Sports Taiwan Co., capitalized at 20 million new Taiwan dollars, will start production in January 1990 with production of 20,000 iron and "metal wood" clubs a month.

template filling

TIE-UP-1
Relationship: TIE-UP
Entities: "Bridgestone Sport Co."
          "a local concern"
          "a Japanese trading house"
Joint Venture Company:
          "Bridgestone Sports Taiwan Co."
Activity: **ACTIVITY-1**
Amount: NT$200000000

ACTIVITY-1
Activity: PRODUCTION
Company:
          "Bridgestone Sports Taiwan Co."
Product:
          "iron and 'metal wood' clubs"
Start Date:
          DURING: January 1990

But NLP very is hard..

- Understanding natural languages is hard … because of inherent *ambiguity*
- Engineering NLP systems is also hard … because of:
  - Huge amount of data resources needed (e.g. grammar, dictionary, documents to extract statistics from)
  - Computational complexity (intractable) of analyzing a sentence

Ambiguity (1)

"Get the cat with the gloves."

Ambiguity (2)

Find at least 5 meanings of this sentence:
"I made her duck"

1. I cooked waterfowl for her benefit (to eat)
2. I cooked waterfowl belonging to her
3. I created the (plaster?) duck she owns
4. I caused her to quickly lower her head or body
5. I waved my magic wand and turned her into undifferentiated waterfowl

Ambiguity (3)

Some ambiguous headlines

- Juvenile Court to Try Shooting Defendant
- Teacher Strikes Idle Kids
- Kids Make Nutritious Snacks
- Bush Wins on Budget, but More Lies Ahead
- Hospitals are Sued by 7 Foot Doctors

Ambiguity is Pervasive

- **Phonetics**
  - I mate or duck
  - I'm eight or duck
  - Eye maid; her duck
  - Aye mate, her duck
  - I maid her duck
  - I'm aid her duck
  - I mate her duck
  - I'm ate her duck
  - I'm ate or duck
  - I mate or duck

Sound like
"I made her duck"

- **Lexical category** (part-of-speech)
  - "duck" as a noun or a verb
- **Lexical Semantics** (word meaning)
  - "duck" as an animal or a plaster duck statue
- Compound nouns
  - e.g. "dog food", "Intelligent design scores …"
- **Syntactic ambiguity**

"I saw   a man   on the hill   with a telescope"

- [But semantics can sometimes help disambiguate]

"I saw   a man   on the hill   with a hat"

The Bottom Line

- Complete NL Understanding (thus general intelligence) is impossible.

- But we can make incremental progress.

- Also we have made successes in **limited domains**.

The Big Picture Approach

All of these applications operate by **exploiting underlying regularities** in human languages.

Sometimes in complex ways, sometimes in pretty trivial ways.

Language structure → Formal models → Practical applications
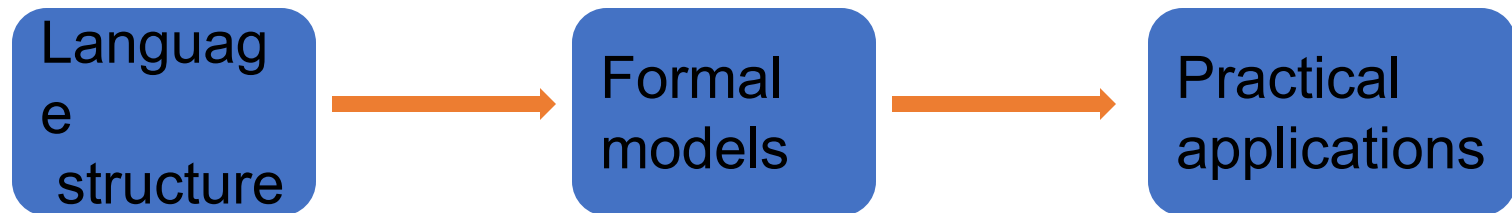
Topics: Linguistics

- Word-level processing
- Syntactic processing
- Lexical and compositional semantics
- Discourse structure

Different Levels of Linguistic Analysis

- Phonology
  - Speech audio signal to phonemes
- Morphology
  - Inflection (e.g. "I", "my", "me"; "eat", "eats", "ate", "eaten")
  - Derivation (e.g. "teach", "teacher", "nominate", "nominee")
- Syntax
  - Part-of-speech (noun, verb, adjective, preposition, etc.)
  - Phrase structure (e.g. noun phrase, verb phrase)
- Semantics
  - Meaning of a word (e.g. "book" as a bound volume or an accounting ledger) or a sentence
- Discourse
  - Meaning and inter-relation between sentences

Topics: Techniques

- Finite-state methods
- Context-free methods
- Probabilistic models
- Neural network models
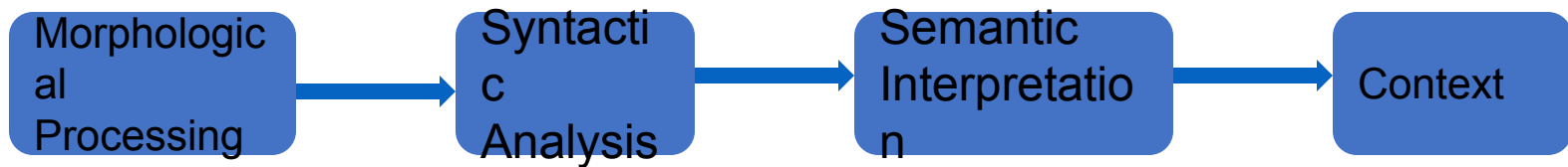
Supervised machine learning methods
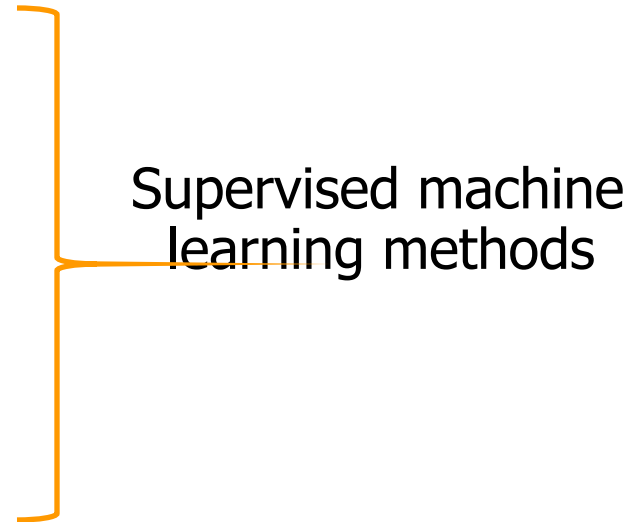
Process Pipeline

- Phonology
- Morphology
- Syntax
- Semantics
- Pragmatics
- Discourse

Each kind of knowledge has associated with it an encapsulated set of processes that make use of it.

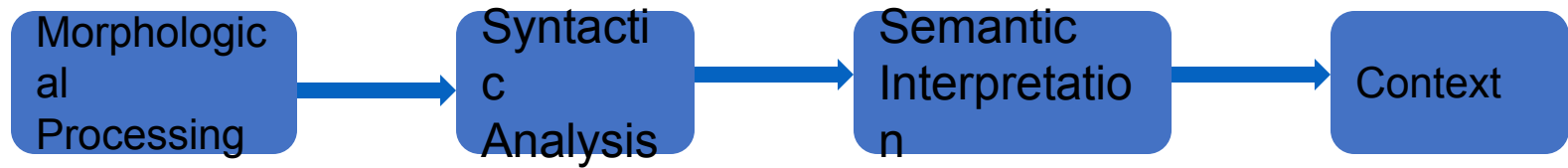Interfaces are defined that allow the various levels to communicate.

This often leads to a **pipeline architecture**.

Morphological Processing	→	Syntactic Analysis	→	Semantic Interpretation	→	Context

But Problem..

- Remember our pipeline...

| Morphological Processing | → | Syntactic Analysis | → | Semantic Interpretation | → | Context |

It really looks like this
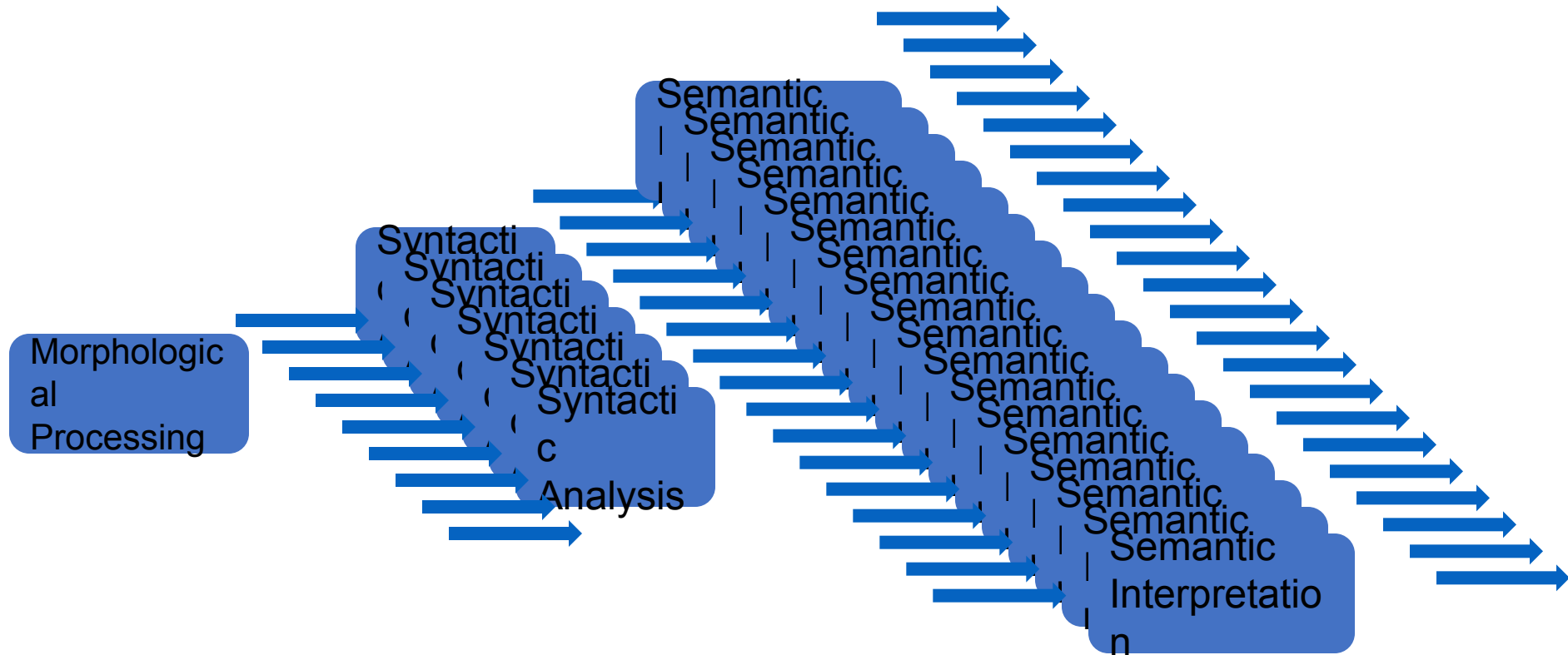
Dealing with Ambiguity

Four possible approaches:

1. **Formal approaches** -- Tightly coupled interaction among processing levels; knowledge from other levels can help decide among choices at ambiguous levels.

2. Pipeline processing that ignores ambiguity as it occurs and hopes that other levels can eliminate incorrect structures.

3. **Probabilistic approaches** based on making the most likely choices

4. Don't do anything, maybe it won't matter

Models and Algorithms

- By models we mean the formalisms that are used to capture the various kinds of linguistic knowledge we need.

- Algorithms are then used to manipulate the knowledge representations needed to tackle the task at hand.

Various Models

- Finite state machines
- Rule-based and logic-based approaches
- Probabilistic models
- Neural network models

Various Algorithms

- In particular..
  - State-space search
    - To manage the problem of making choices during processing when we lack the information needed to make the right choice
  - Dynamic programming
    - To avoid having to redo work during the course of a state-space search
      - CKY, Earley, Minimum Edit Distance, Viterbi, Baum-Welch
  - Classifiers
    - Machine learning based classifiers that are trained to make decisions based on features extracted from the local context