Elaborazione del Linguaggio Naturale: Interpretazione, Ragionamento automatico e Apprendimento delle macchine

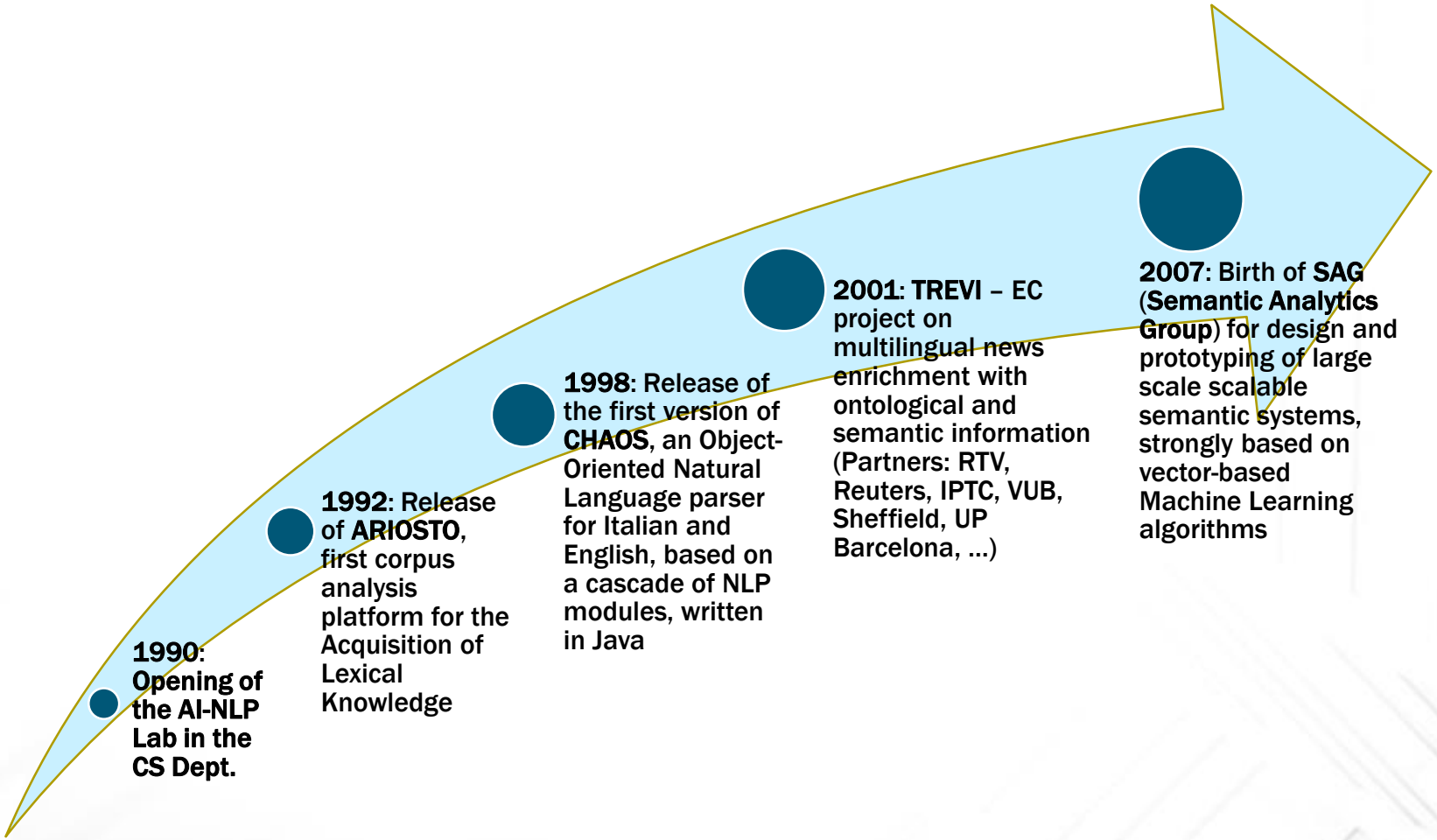
Roberto Basili
(Università di Roma, Tor Vergata)
dblp: http://dblp.uni-trier.de/pers/hd/b/Basili:Roberto.html
Google scholar: https://scholar.google.com/citations?user=U1A22fYAAAAJ&hl=it&oi=sra

Università di Bologna, 16 Maggio 2016

Overview

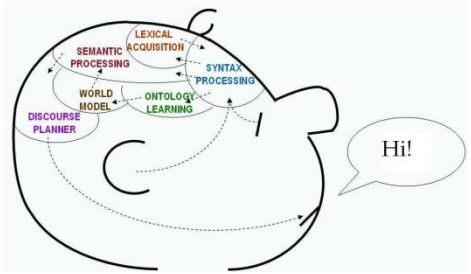
- Intelligenza Artificiale e Lingue parlate e scritte
  - Informazioni, Rappresentazioni coinvolte, Sfide (ri)correnti, success (and unsuccessful) stories
- Elaborazione Automatica delle Lingue: Modelli, Metodi e *Risultati*
- break
- Ruolo delle Tecnologie dell'Apprendimento ed Applicazioni:
  - Sviluppo Automatico di Dizionari, Lessici Semantici ed Ontologie
  - Trattamento Semantico della Documentazione Investigativa
  - Sistemi Web-based di Opinion Mining, Market Watch & Brand Reputation Management

SAG: a not-so-short history



REVEAL

2012:
Birth of
Reveal srl



Research@ART
Artificial intelligence at RTV

Robust Statistical Language Processing

Lexical Semantics

- Lex
- Cor
- Sem
- Verb Argument Str
- Frame Semantics

Machine Learning for Knowledge Acquisition

- Terminological Dictionaries
- IE Template Filling Rules
- Relation Extr
- Automatic Te
- Distributional
- Semantic Wo

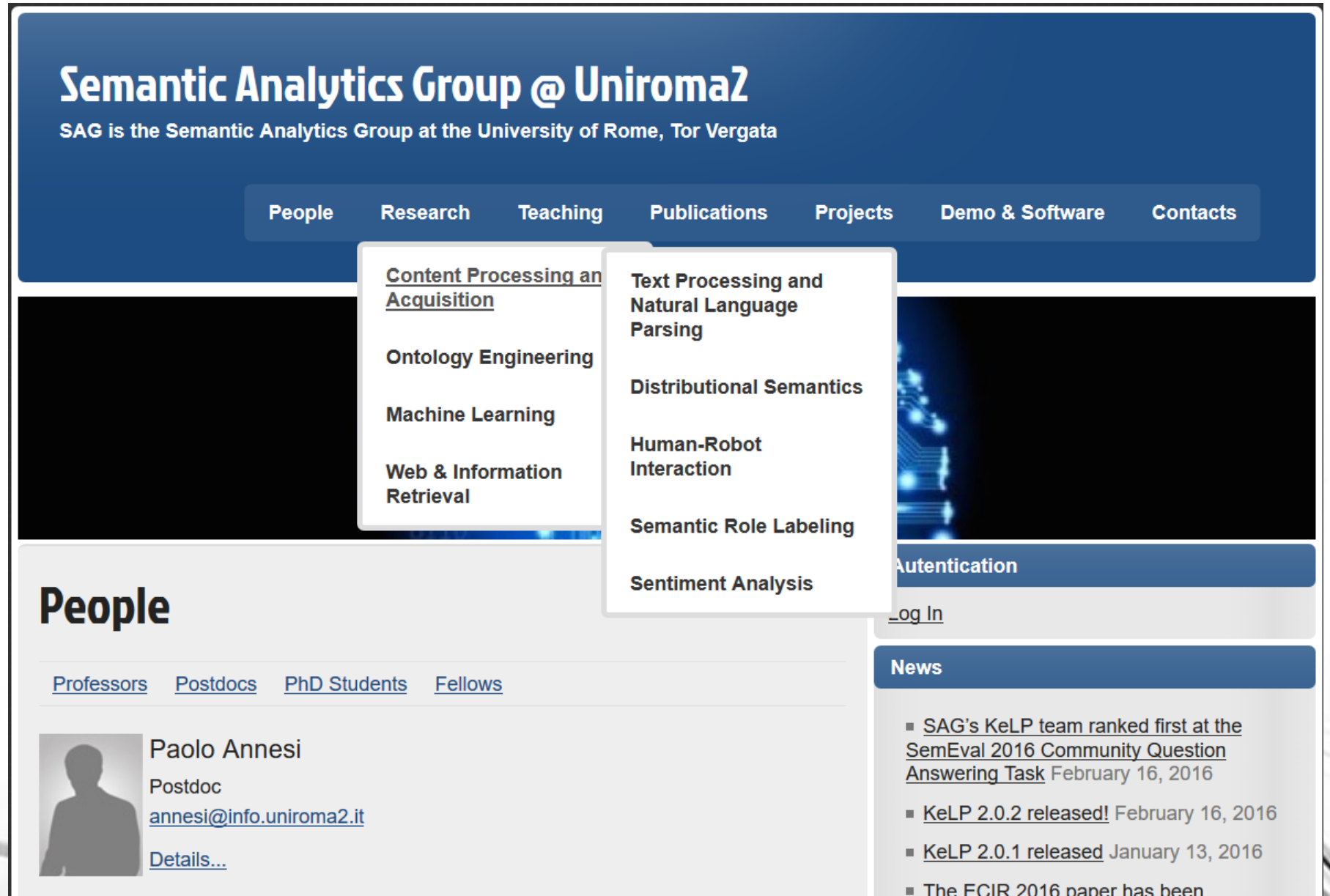
Ontology Engineering

- Acquisition
- Ontology Le
- NERC and O
- Ontology Ma
- Ontology Li
- Multilingual Ontology Representation

Web IR and IE

- *Information Extraction*
- *Information Retrieval*
- *Learning to Rank*
- *Automatic Document Categorization*
- *Question Answering*
- *Multimedia Indexing and Retrieval*
- *Dialogue & Interactive Question Answering*
- *Social Web Analytics (Opinion Mining)*

http://sag.art.uniroma2.it/



Overview

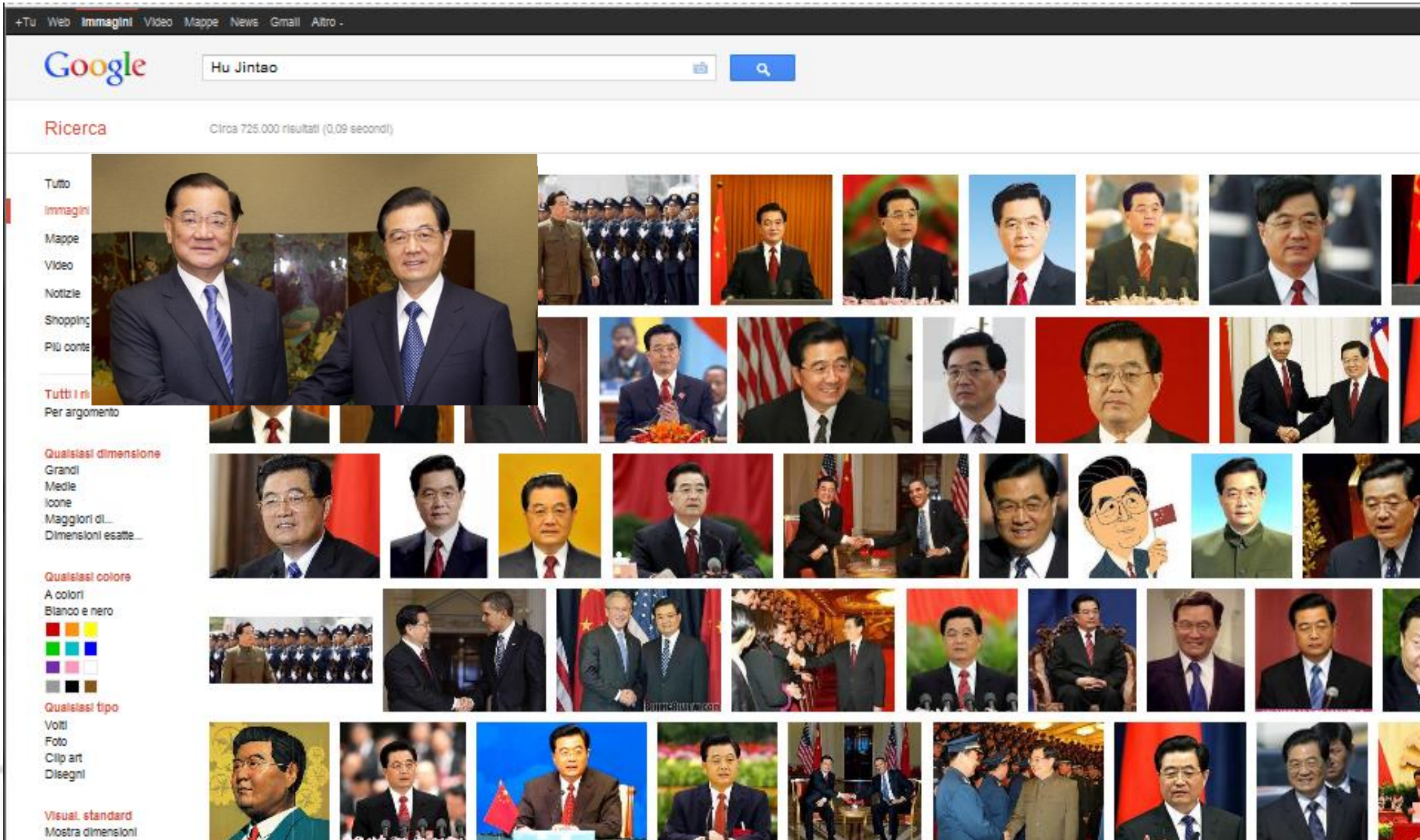
- Intelligenza Artificiale e Lingue parlate e scritte
  - Informazioni e Rappresentazioni coinvolte, Sfide (ri)correnti, …
- Elaborazione Automatica delle Lingue: Modelli, Metodi e *Risultati*
- break
- Ruolo delle Tecnologie dell'Apprendimento ed Applicazioni:
  - Sviluppo Automatico di Dizionari, Lessici Semantici ed Ontologie
  - Trattamento Semantico della Documentazione Investigativa
  - Sistemi Web-based di Opinion Mining, Market Watch & Brand Reputation Management

Semantics, Open Data and Natural Language

Web contents, characterized by rich multimedia information, are mostly opaque from a semantic standpoint

Information, Web and Natural Languages

Who is Hu Jintao?


+Tu Web Immagini Video Mappe News Gmail Altro -
Google
Hu Jintao
Ricerca
Circa 725.000 risultati (0,09 secondi)
Tutto
Immagini
Mappe
Video
Notizie
Shopping
Più corte
Tutti i ri
Per argomento
Qualsiasi dimensione
Grandi
Medie
Icone
Maggiori di...
Dimensioni esatte...
Qualsiasi colore
A colori
Bianco e nero
Qualsiasi tipo
Volti
Foto
Clip art
Disegni
Visual. standard
Mostra dimensioni

Content Semantics and Natural Language

- Human languages are the main carrier of the information involved in processes such as *retrieval*, *publication* and *exchange* of knowledge as it is associated to the open Web contents
- Words and NL syntactic structures express concepts, activities, events, abstractions and conceptual relations we usually share through data
- "*Language is parasitic to knowledge representation languages but the viceversa is not true*" (Wilks, 2001)
- From **Learning to Read** to **Knowledge Distillation** as a (integrated pool of) Semantic interpretation Task(s)

Semantics, Natural Language & Learning: tasks

- From **Learning to Read** to **Knowledge Distillation** as a (integrated pool of)

  Semantic interpretation Task(s)

  - **Information Extraction**
    - Entity Recognition and Classification
    - Relation Extraction
    - Semantic Role Labeling (Shallow Semantic Parsing)
  - **Estimation of Text Similarity**
    - Structured Text Similarity/Textual Entailment Recognition
    - Sense disambiguation
  - **Semantic Search, Question Classification and Answer Ranking**
  - **Knowledge Acquisition**, e.g. ontology learning
  - **Social Network Analysis, Opinion Mining**

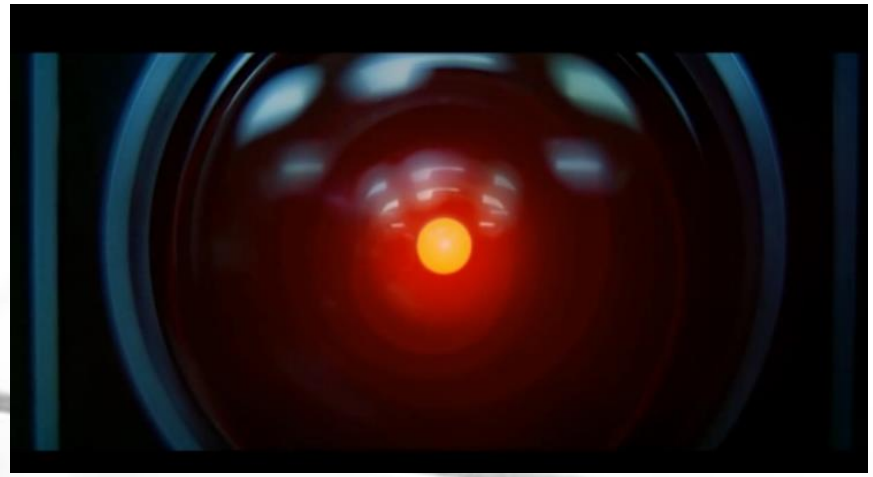
Overview

- Intelligenza Artificiale e Lingue parlate e scritte
  - Informazioni e Rappresentazioni coinvolte
  - Sfide (ri)correnti, battaglie (già) vinte e rischi inerenti …
- Elaborazione Automatica delle Lingue: Modelli, Metodi e *Risultati*
- break
- Ruolo delle Tecnologie dell'Apprendimento ed Applicazioni:
  - Sviluppo Automatico di Dizionari, Lessici Semantici ed Ontologie
  - Trattamento Semantico della Documentazione Investigativa
  - Sistemi Web-based di Opinion Mining, Market Watch & Brand Reputation Management

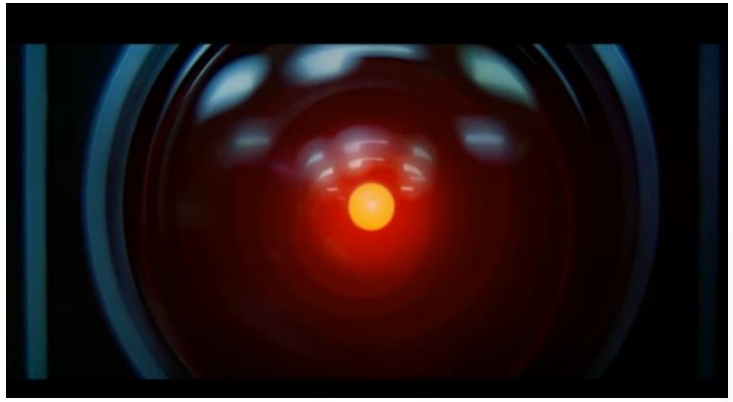
Which Knowledge?

- HAL 9000, da "*2001: A Space Odyssey*"

- Dave: *Open the pod bay doors, Hal*.
- HAL: *I'm sorry Dave, I'm afraid I can't do that*.



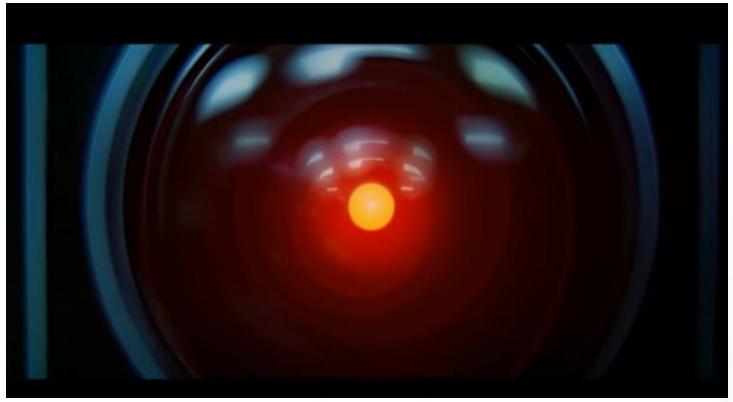
What's HAL knowledge?



- **Recognition & Synthesis of spoken language**
  - Dictionaries (spelling)
  - Phonetics (how to produce/recognize sound)

- **Understanding**
  - **Lexical Knowledge**
    - What do the words mean?
    - How they combine (`*pod bay door*')
  - **Knowledge about the syntagmatic structure of sentences**
    - *I'm I do, Sorry that afraid Dave I'm can't*

What's HAL knowledge?



- **Dialogue & pragmatics**
  - "*open the door*" is a request (and not a declaration or a search query)
  - Replying is a type of action that imply kindness (even if a planning to kill is in progress …)
  - It is useful to behave cooperatively (*I'm afraid, I can't…*)
  - What about `*that*' in `*I can't do that*'?

Language Processing as a (*semantic*) *interpretation process*

- Processing a text corresponds to the understanding of a number of aspects related to its *meaning*
  - Thematic Domain (e.g. science/economics/sport)
  - Operational Objectives (e.g. e-mail *spam*)
  - Involved Entitites, such as people or locations
  - Potential events described (e.g. facts told by news)
  - Communicative Objectives (e.g. dialogue, orders/declarations/planning)
- Outcome: an explicit *representation of the text meaning* ...
- able to trigger different inferences
  (e.g. IR *relevance, planning*, knowledge updates, ...)

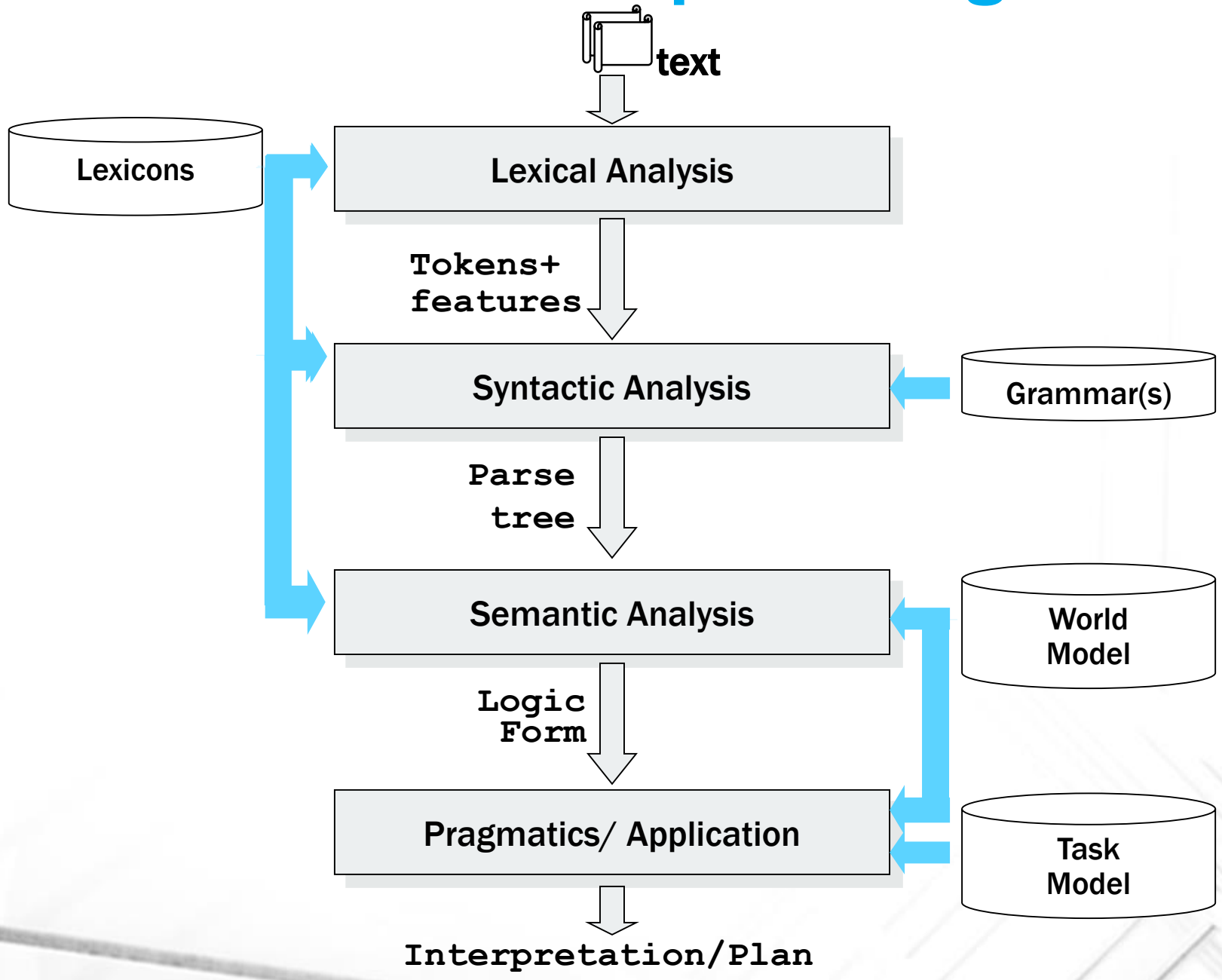
Some Reflections

- Understading *linguistic information* requires specific knowledge about:
  - The natural language itself (e.g. *grammar*)
  - The world (e.g. *bay door*, *Dave* or *opening*)
  - How language make **reference** to the world
- NLP applications deals with texts by exploiting the specific context:
  - Application purposes, e.g. document search
  - The domain and the operational context of an application
  - The distinction between language producer (speaker/writer) and consumer (hearer/reader)

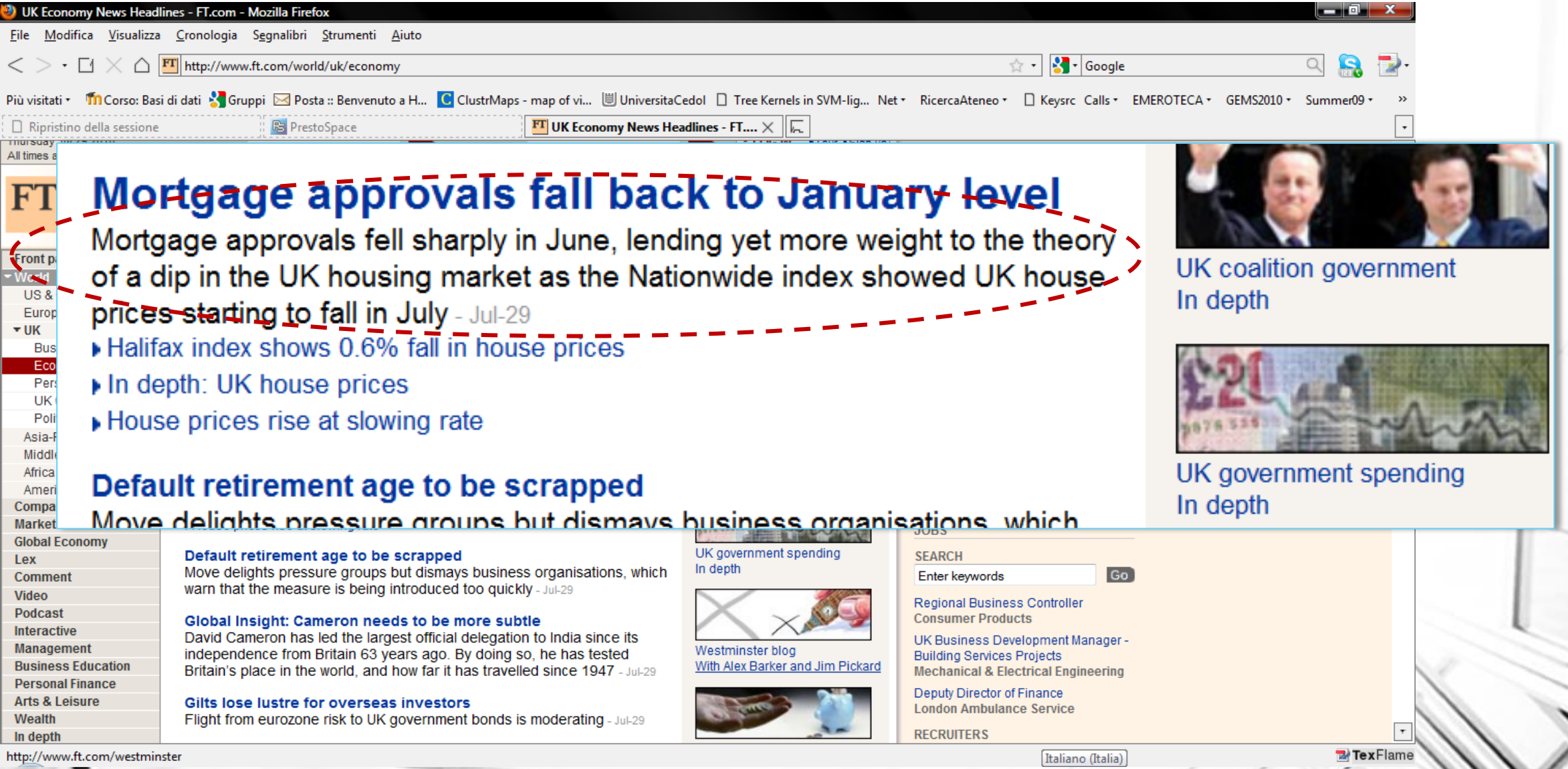
Major Challenges

- *Linguistic Accuracy* in approximating the human-level of performance
- *Robustness* (errors/noise/incompleteness)
- *Scale*
  - Coverage of the phenomena (Lexicons/Grammars)
- *Expressivity*
  - Dictionaries, Lexicons and *Thesauri*
  - World Models and types of inference
- *Flexibility*
  - Adequate performance across linguistic variability (e.g. producer vs. consumer)
- *Naturalness*

NLP: the standard processing chain



Grammatical Analysis



Syntax and Semantics in textual data

- Compositionality
- The meaning of a complex expression is solely determined by the meanings of its constituent expressions and the rules used to combine them.
- "*I will consider a language to be a set (finite or infinite) of sentences, each finite in length and constructed out of a finite set of elements. All natural languages are languages in this sense. Similarly, the set of "sentences" of some formalized system of mathematics can be considered a language*"
  Chomsky 1957

Syntax

- In linguistics, **syntax** is the study of the rules that govern the structure of sentences, and which determine their relative grammaticality.
- Such rules govern a number of language phenomena as systems for phonology, morphology, syntax as well as discourse

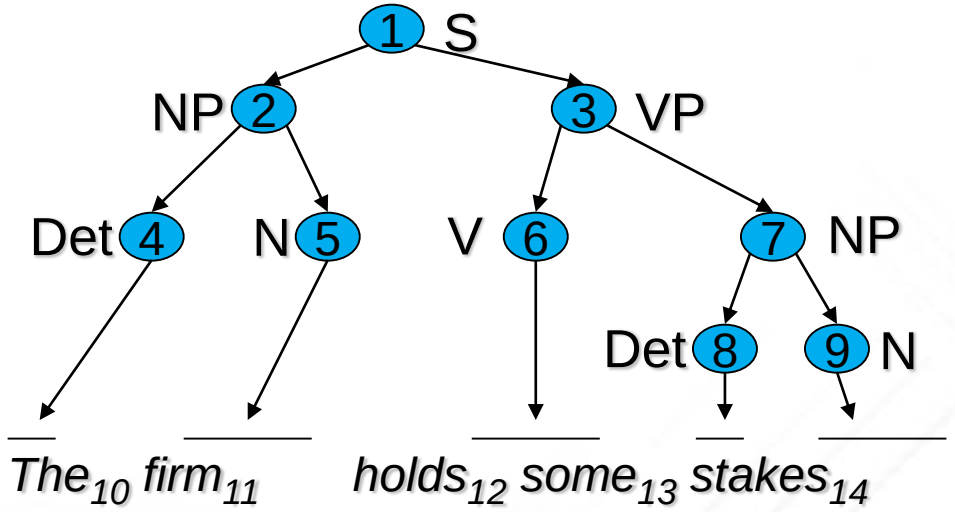
Parse Trees

- The representation of the parsing result is a structure that expresses:
  - The order of constituent elements in the sentence
  - The grammatical type of constituents
  - The hierarchical organization of constitunts
- The structures able to express these properties are the derivation trees also called parse trees

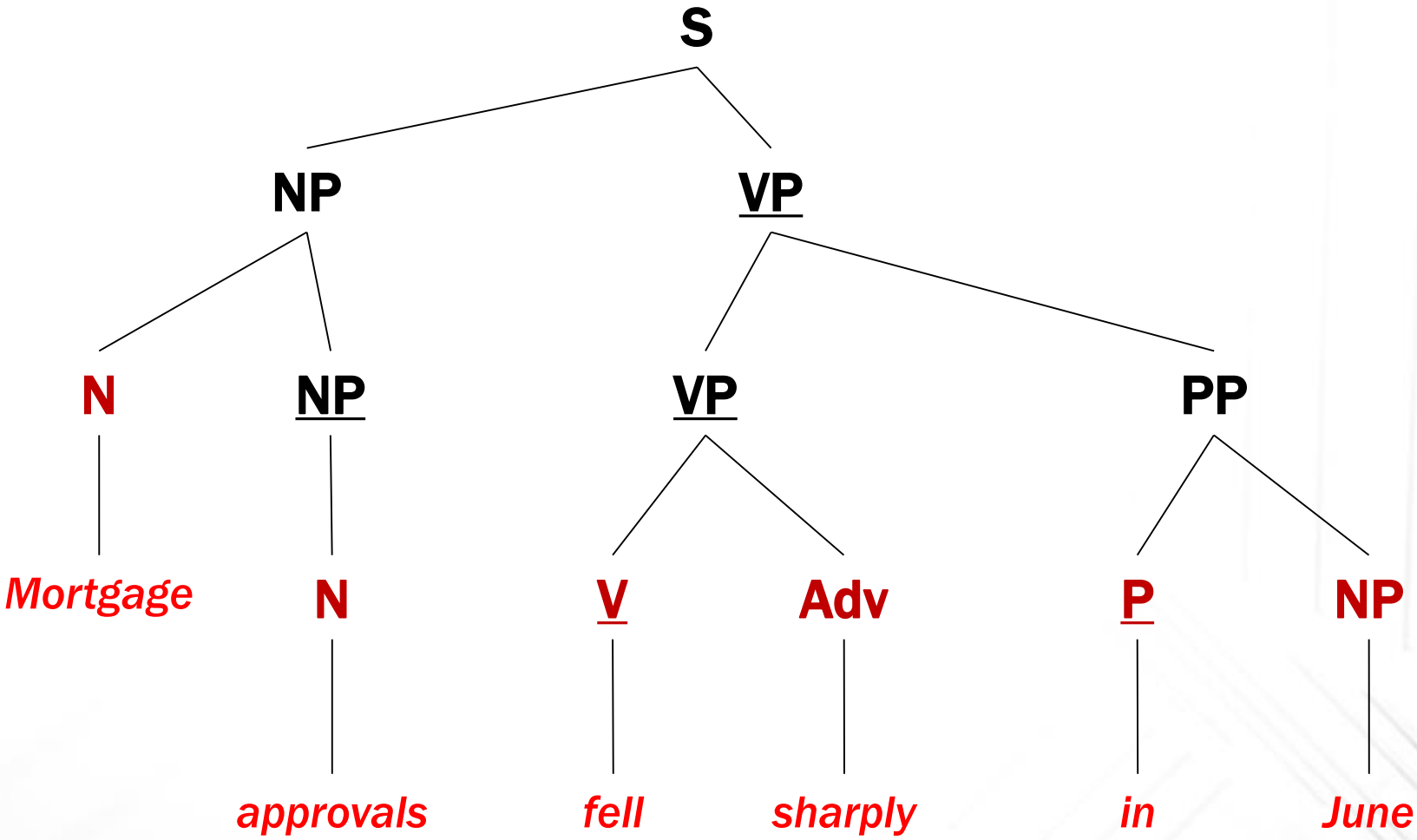
Grammars and Trees

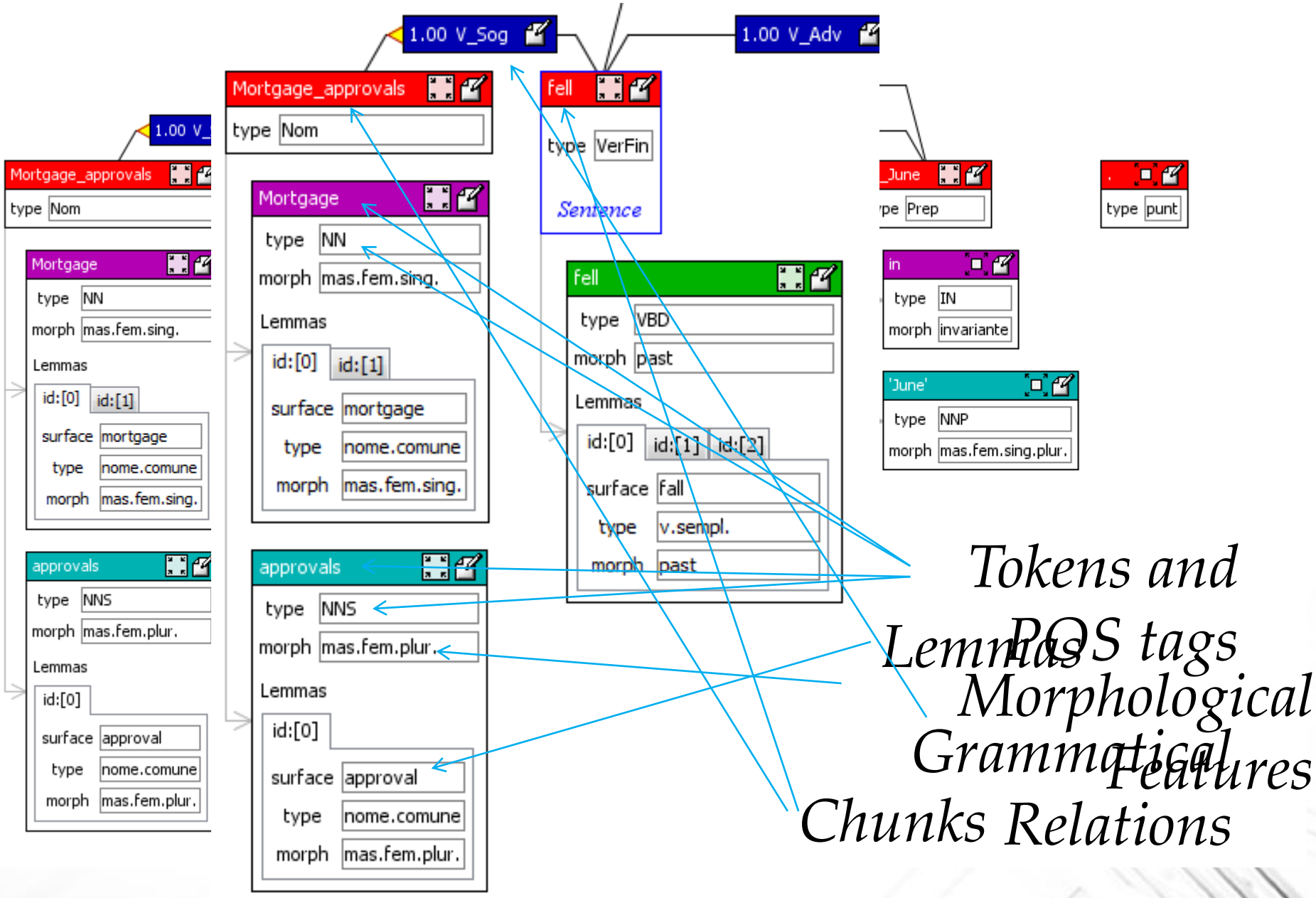
"The firm holds some stakes"

- Vn={S,NP,VP,Det,N}, Axiom: S
- Productions: {S→NP VP, VP→V NP, NP→Det N}
- Derivation:
  - S > NP VP > Det N VP > *The* N VP > *The firm* VP > *The firm* V NP > *The firm holds* NP > *The firm holds* Det N > *The firm holds some* N > *The firm holds some stakes*



Constituency-based Parsing





FT (July, 29): *Mortgage approvals fell sharply in June.*

Challenges for Parsing

- Huge complexity as for the ambiguity in the morphosyntactic descriptions of words
  - E.g. La vecchia porta la sbarra
- Interdependency with semantic information
  - Most ambiguity cannot be solved only at the grammatical level
  - Lexical Semanic information is crucial as grammatical structures are constrained by word senses
    - Operating in a market vs. Operating a patient

Semantics

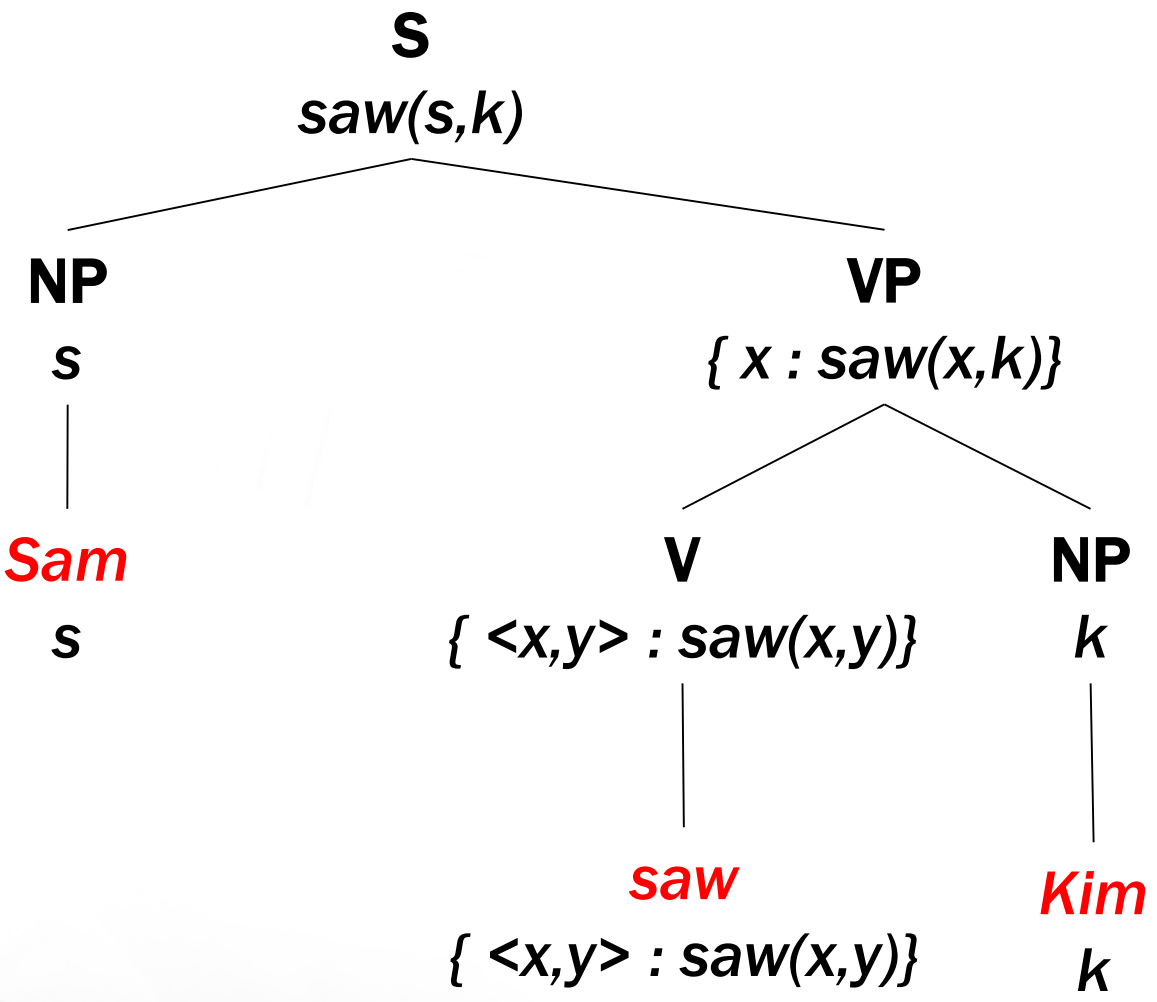
- What is the meaning of the sentence

  John saw Kim?

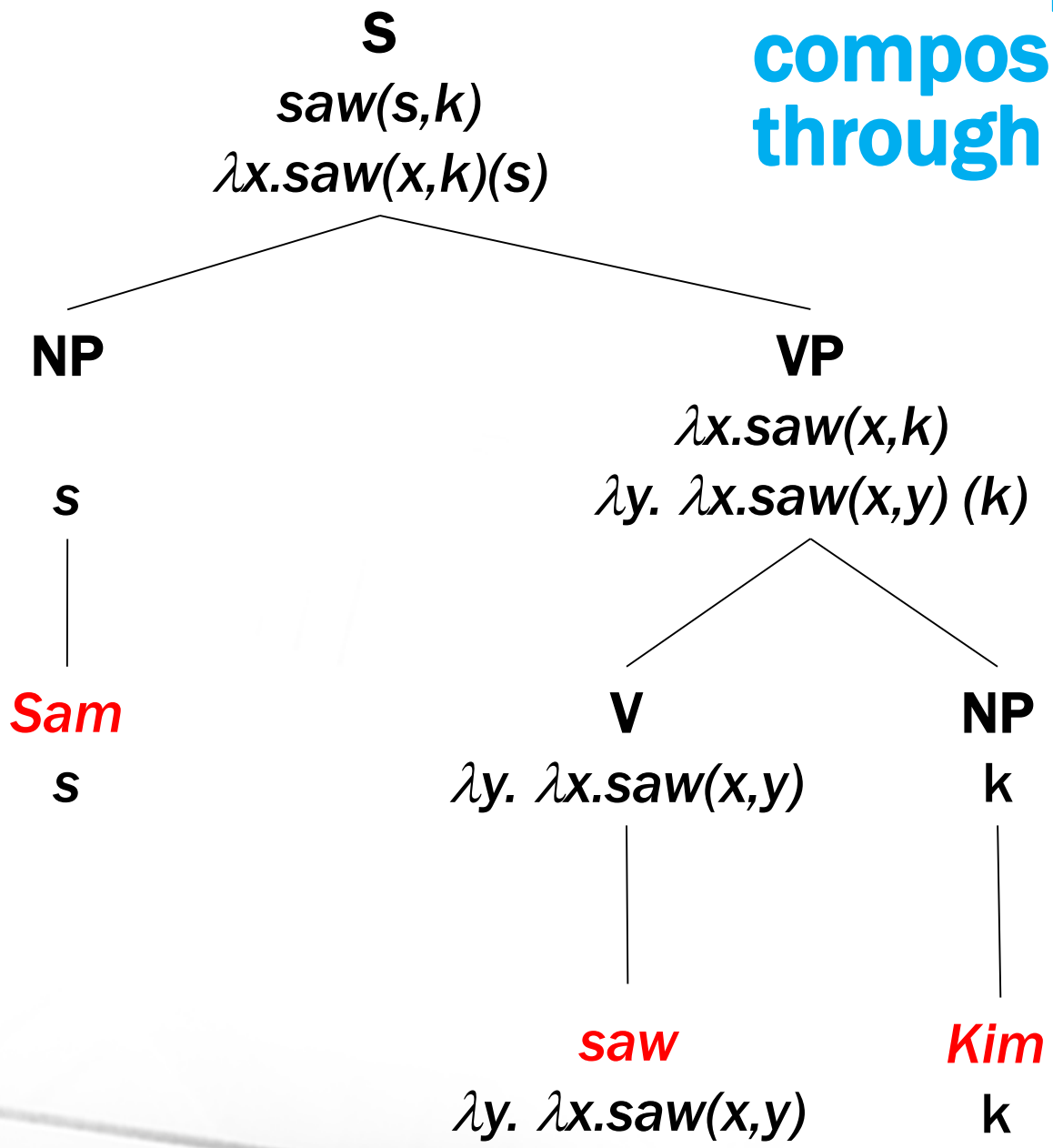
- Desirable Properties:
  - **It should be derivable as a function of the indivdual constituents,** i.e. the meanings of costituents such as ***Kim***, ***John*** and ***see***
  - **Independent from syntactic phenomena**, e.g. ***Kim was seen by John*** is a paraphrasis
  - It must be directy used to trigger some **inferences:**
    - ***Who*** *was seen by John?* ***Kim!***
    - *John saw Kim.* ***He*** *started running to* ***her****.*

A Truth conditional semantics

John saw Kim



NL Interpretation as compositional processing through *lambda espressions*



Target Semantic Phenomena

- **Entities.** Entità descritte nei testi (persone, luoghi, organizzazioni, date, espressioni numeriche o monetarie)

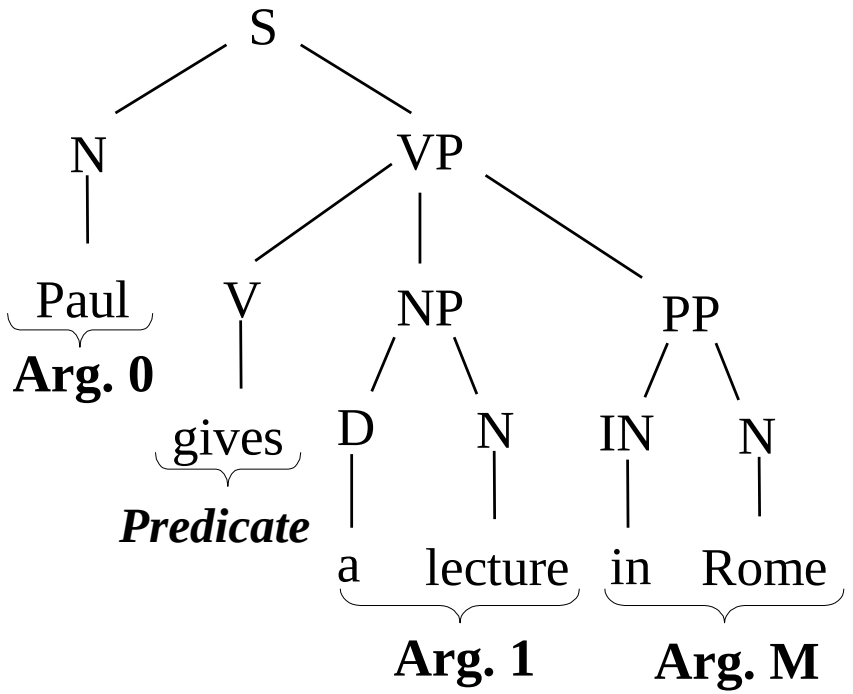
- **Relations.** Relazioni / Associazioni tra entità

- **Facts.** Fatti ed Eventi

- **Topics.** Temi / Contesto / Dominio

Predicazione ed Argomenti

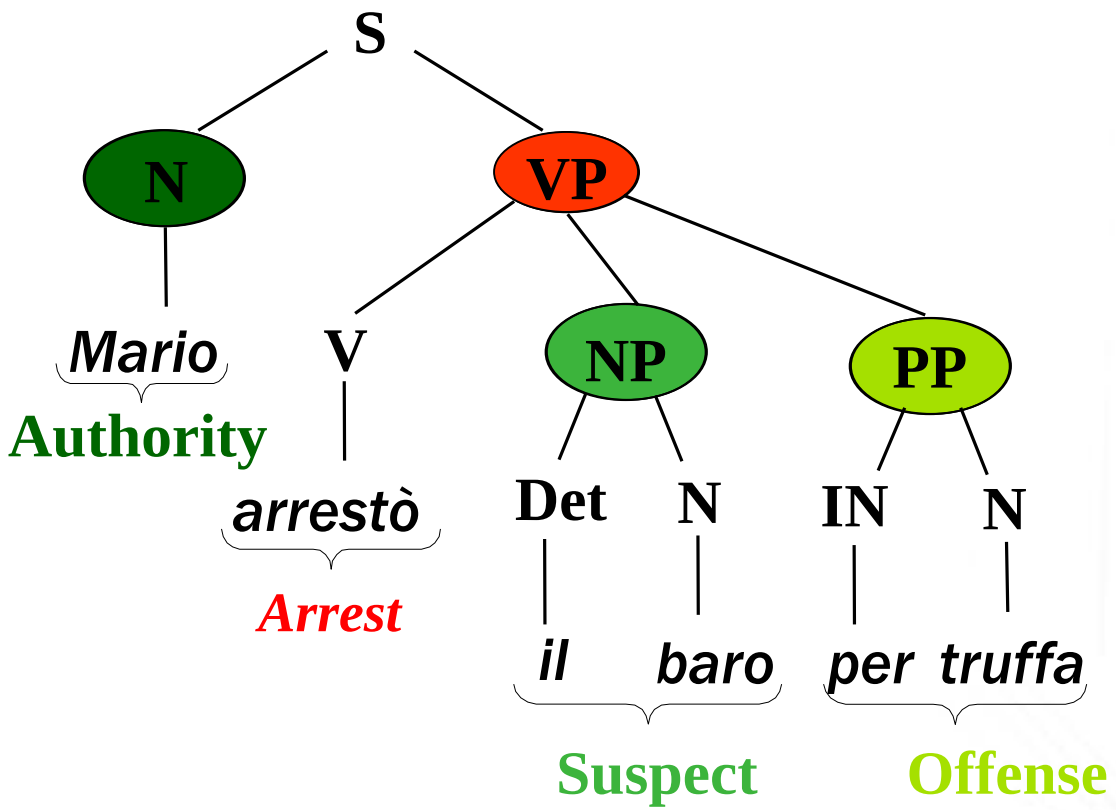
- Il *mapping* sintassi-semantica



Annotazioni Semantiche diverse: PropBank vs. FrameNet

Linking syntax to semantics: see later slides on Semantic Role Labeling

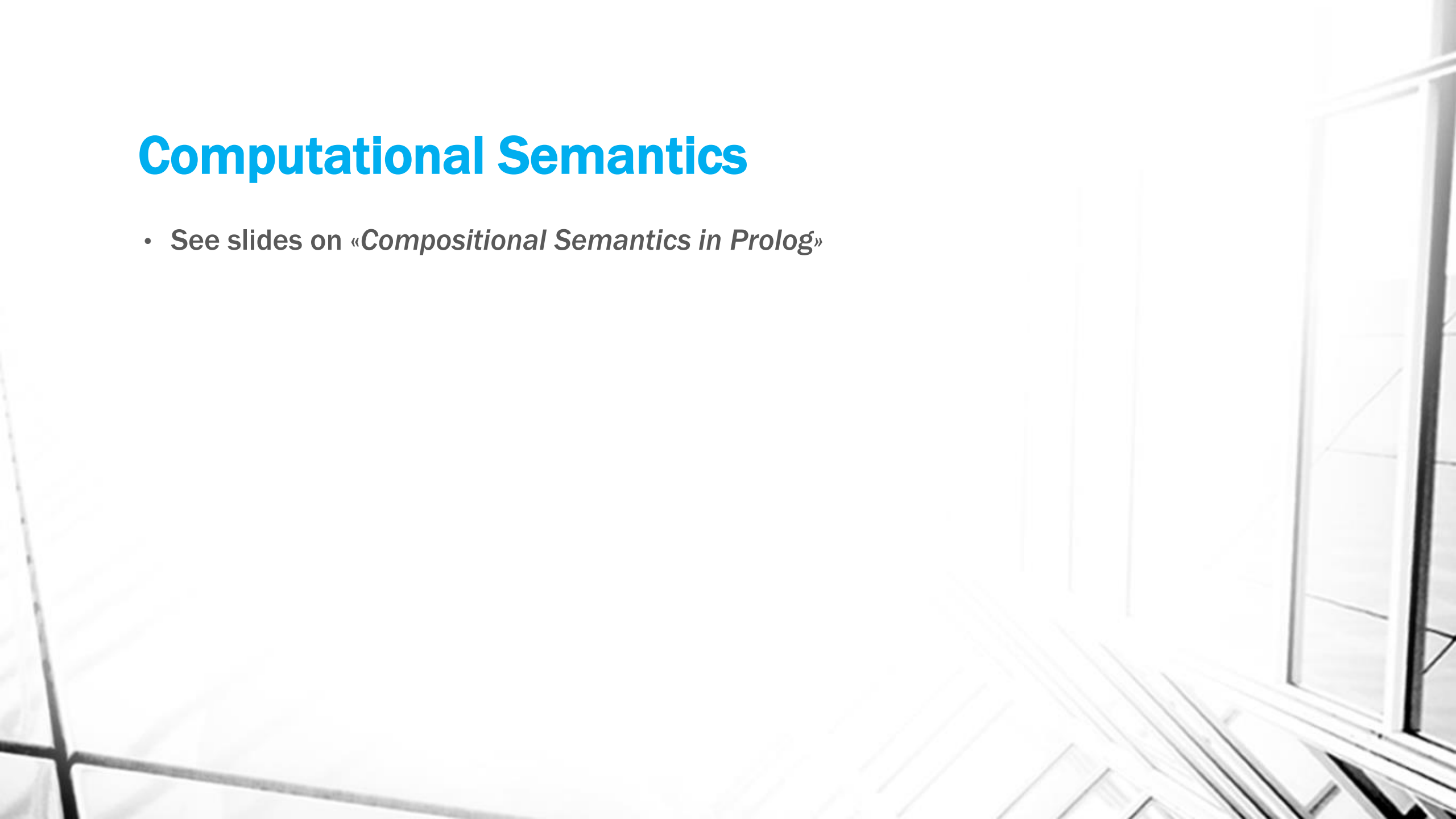
Mario arrestò il baro per truffa



[Il baro]$_{\text{Suspect}}$ *[fu arrestato]*$_{\text{Arrest}}$ *[da Mario]*$_{\text{Authority}}$ *[per truffa]*$_{\text{Offense}}$

Computational Semantics

- See slides on «*Compositional Semantics in Prolog*»



Three Linguistic Perspectives on Meaning

- Lexical Semantics
  - The meanings of individual words
- Formal Semantics (or Compositional Semantics or Sentential Semantics)
  - How those meanings combine to make meanings for individual sentences or utterances
- Discourse or Pragmatics
  - How those meanings combine with each other and with other facts about various kinds of context to make meanings for a text or discourse
  - Dialog or Conversation is often lumped together with Discourse

Lexical Semantic:
Relationships between word meanings

- Homonymy
- Polysemy
- Synonymy
- Antonymy
- Hypernomy
- Hyponomy
- Meronomy

Homonymy

- **Homonymy:**
  - Lexemes that share a form
    - Phonological, orthographic or both
  - But have unrelated, distinct meanings
  - Clear example:
    - Bat (wooden stick-like thing) vs
    - Bat (flying scary mammal thing)
    - Or bank (financial institution) versus bank (riverside)
  - Can be also homophones, homographs, or both:
    - Homophones:
      - *Write* and *right*
      - *Piece* and *peace*

Polysemy

- The **bank** is constructed from red brick
  I withdrew the money from the **bank**
- Are those the same sense?
- Or consider the following WSJ example
  - While some banks furnish sperm only to married women, others are less restrictive
  - Which sense of bank is this?
    - Is it distinct from (homonymous with) the river bank sense?
    - How about the savings bank sense?

Metaphor and Metonymy

- Specific types of polysemy
- Metaphor:
  - Germany will pull Slovenia out of its economic slump.
  - *I spent 2 hours on that homework*.
- Metonymy
  - The White House announced yesterday.
  - This chapter talks about part-of-speech tagging
  - Bank (building) and bank (financial institution)

Synonyms

- Word that have the same meaning in some or all contexts.
  - *filbert / hazelnut*
  - *couch / sofa*
  - *big / large*
  - *automobile / car*
  - *vomit / throw up*
  - *Water / H_20*
- Two lexemes are synonyms if they can be successfully substituted for each other in all situations
  - If so they have the same **propositional meaning**

Synonyms

- But there are few (or no) examples of perfect synonymy.
  - Why should that be?
  - Even if many aspects of meaning are identical still may not preserve the acceptability based on notions of politeness, slang, register, genre, etc.
- Example:
  - Water and H_2O
  - I would not say

    I like fresh H_2O after the tennis

Some terminology

- Lemmas and wordforms
  - A **lexeme** is an abstract pairing of meaning and form
  - A **lemma** or **citation form** is the grammatical form that is used to represent a **lexeme**.
    - ***Carpet*** is the lemma for ***carpets, Dormir*** is the lemma for ***duermes***.
  - Specific surface forms *carpets, sung, duermes* are called **wordforms**

- The lemma *bank* has two **senses:**
  - Instead, a **bank** can hold the investments in a custodial account in the client's name
  - But as agriculture burgeons on the east **bank**, the river will shrink even more.

- A **sense** is a discrete representation of one aspect of the meaning of a word

Antonyms

- Senses that are opposites with respect to one feature of their meaning
- Otherwise, they are very similar!
  - dark / light
  - short / long
  - hot / cold
  - up / down
  - in / out
- More formally: antonyms can
  - define a binary opposition or opposite ends of a scale (*long/short, fast/slow*)
  - Be **reversives**: *rise/fall, up/down*

Hyponymy

- One sense is a **hyponym** of another if the first sense is more specific, denoting a subclass of the other
  - *car* is a hyponym of *vehicle*
  - *dog* is a hyponym of *animal*
  - *mango* is a hyponym of *fruit*
- Conversely
  - *vehicle* is a hypernym/superordinate of *car*
  - *animal* is a hypernym of *dog*
  - *fruit* is a hypernym of *mango*

superordinate	vehicle	fruit	furniture	mammal
hyponym	car	mango	chair	dog

Hypernymy more formally

- Extensional:
  - The class denoted by the superordinate extensionally includes the class denoted by the hyponym
- Entailment:
  - A sense A is a hyponym of sense B if being an A entails being a B
- Hyponymy is usually transitive
  - (A hypo B and B hypo C entails A hypo C)

II. WordNet

- A hierarchically organized lexical database
- On-line thesaurus + aspects of a dictionary
  - Versions for other languages are under development

Category	Unique Forms
Noun	117,097
Verb	11,488
Adjective	22,141
Adverb	4,601

WordNet

- Home page:
  - http://www.cogsci.princeton.edu/cgi-bin/webwn

How is "*sense*" defined in WordNet?

- The set of near-synonyms for a WordNet sense is called a **synset (synonym set)**; it's their version of a sense or a concept
- Example: **chump** as a noun to mean
  - 'a person who is gullible and easy to take advantage of'

{chump[1], fool[2], gull[1], mark[9], patsy[1], fall guy[1], sucker[1], soft touch[1], mug[2]}

- Each of these senses share this same gloss
- Thus for WordNet, the meaning of this sense of **chump** *is* this list.

Format of Wordnet Entries

The noun "bass" has 8 senses in WordNet.

1. bass[1] - (the lowest part of the musical range)
2. bass[2], bass part[1] - (the lowest part in polyphonic music)
3. bass[3], basso[1] - (an adult male singer with the lowest voice)
4. sea bass[1], bass[4] - (the lean flesh of a saltwater fish of the family Serranidae)
5. freshwater bass[1], bass[5] - (any of various North American freshwater fish with lean flesh (especially of the genus Micropterus))
6. bass[6], bass voice[1], basso[2] - (the lowest adult male singing voice)
7. bass[7] - (the member with the lowest range of a family of musical instruments)
8. bass[8] - (nontechnical name for any of numerous edible marine and freshwater spiny-finned fishes)

The adjective "bass" has 1 sense in WordNet.

1. bass[1], deep[6] - (having or denoting a low vocal or instrumental range)
   "a deep voice"; "a bass voice is lower than a baritone voice"; "a bass clarinet"

WordNet Noun Relations

Relation	Also called	Definition	Example
Hypernym	Superordinate	From concepts to superordinates	$breakfast^1 \rightarrow meal^1$
Hyponym	Subordinate	From concepts to subtypes	$meal^1 \rightarrow lunch^1$
Member Meronym	Has-Member	From groups to their members	$faculty^2 \rightarrow professor^1$
Has-Instance		From concepts to instances of the concept	$composer^1 \rightarrow Bach^1$
Instance		From instances to their concepts	$Austen^1 \rightarrow author^1$
Member Holonym	Member-Of	From members to their groups	$copilot^1 \rightarrow crew^1$
Part Meronym	Has-Part	From wholes to parts	$table^2 \rightarrow leg^3$
Part Holonym	Part-Of	From parts to wholes	$course^7 \rightarrow meal^1$
Antonym		Opposites	$leader^1 \rightarrow follower^1$

WordNet Verb Relations

Relation	Definition	Example
Hypernym	From events to superordinate events	$fly^9 \rightarrow travel^5$
Troponym	From a verb (event) to a specific manner elaboration of that verb	$walk^1 \rightarrow stroll^1$
Entails	From verbs (events) to the verbs (events) they entail	$snore^1 \rightarrow sleep^1$
Antonym	Opposites	$increase^1 \Longleftrightarrow decrease^1$

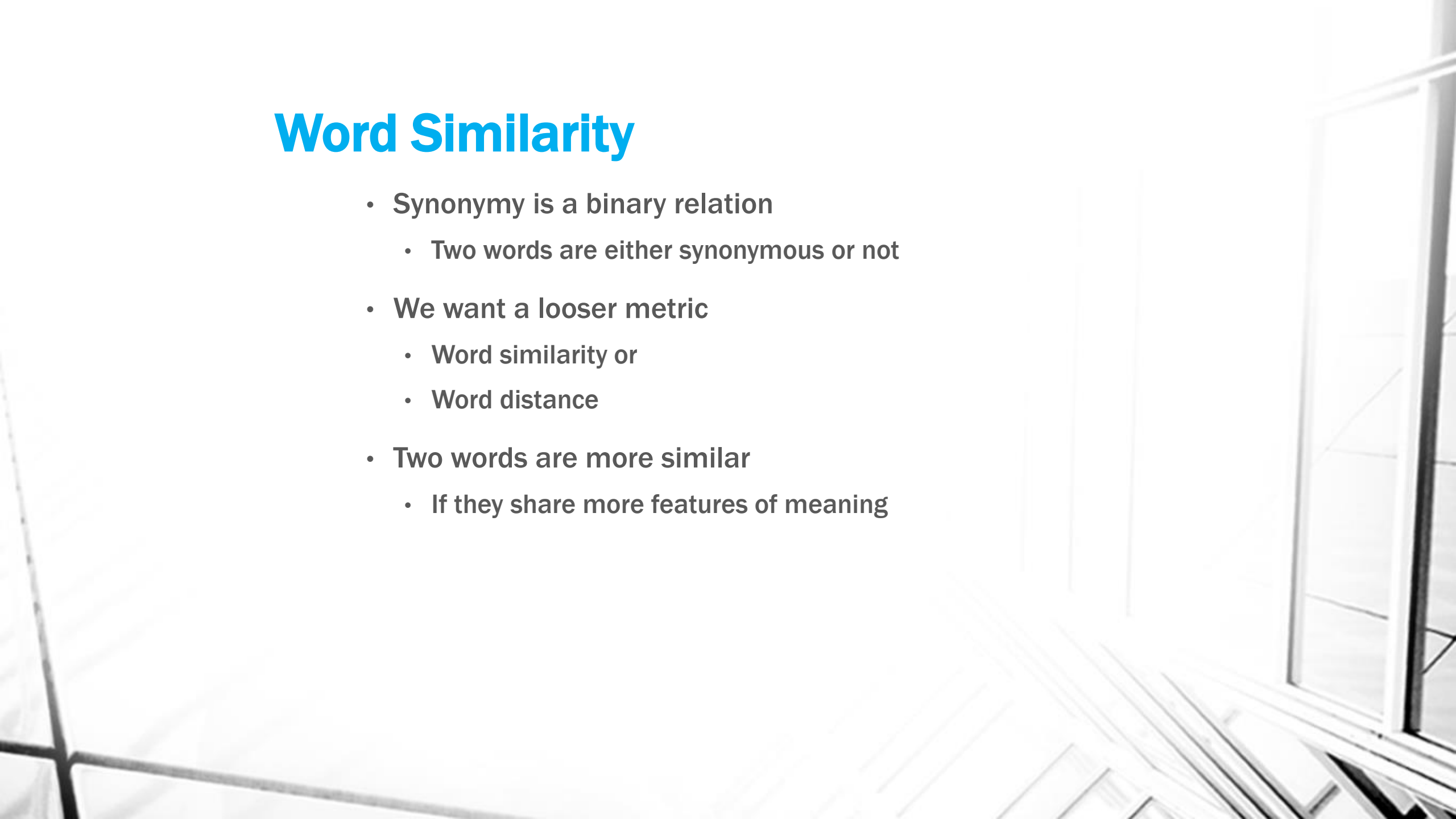
WordNet Hierarchies

```
Sense 3
bass, basso --
(an adult male singer with the lowest voice)
=> singer, vocalist, vocalizer, vocaliser
   => musician, instrumentalist, player
      => performer, performing artist
         => entertainer
            => person, individual, someone...
               => organism, being
                  => living thing, animate thing,
                     => whole, unit
                        => object, physical object
                           => physical entity
                              => entity
               => causal agent, cause, causal agency
                  => physical entity
                     => entity


Sense 7
bass --
(the member with the lowest range of a family of
musical instruments)
=> musical instrument, instrument
   => device
      => instrumentality, instrumentation
         => artifact, artefact
            => whole, unit
               => object, physical object
                  => physical entity
                     => entity
```

Word Similarity

- Synonymy is a binary relation
  - Two words are either synonymous or not
- We want a looser metric
  - Word similarity or
  - Word distance
- Two words are more similar
  - If they share more features of meaning



Word Similarity

- Actually these are really relations between **senses**:
  - Instead of saying "*bank is like fund*"
  - We say
    - Bank1 *is similar to* fund3
    - Bank2 *is similar to* slope5
- Similarity are computed over both words and senses

Why word similarity

- Spell Checking
- Information retrieval
- Question answering
- Machine translation
- Natural language generation
- Language modeling
- Automatic essay grading

Lexical Semantics: toward predicates
Syntactic Argument Structures

- (Verbal) Relations require a fixed number of participants, called arguments
- The syntactic structure predicts the number and type of arguments through subcategorization frames
  - (Bob (gave (Mary) (the book) (on Monday)))
  - (Bob (gave (the book) (to Mary) (on Monday)))

Lexical Semantics: Predicates & Thematic roles

- Arguments play specific roles, called thematic roles, depending on the predicate but invariant across different syntactic structures giving rise to predicate argument structures
  - `give(Agent:` *Bob*, `Theme:` *the_book*, `Recipient:` *Mary*`)`

- Thematic roles of individual arguments are indexed by their predicates

- *General* and *lexicalized* roles have been introduced

THEMATIC ROLES

AGENT: Deliberately performs the action described by the verb

THEME (PATIENT): Undergoes the action of the verb or is in the state described by the verb

EXPERIENCER: Experiences the emotional or mental state or change described by the verb

INSTRUMENT: Entity used to carry out the action described by the verb

LOCATION: Place where action or state occurs

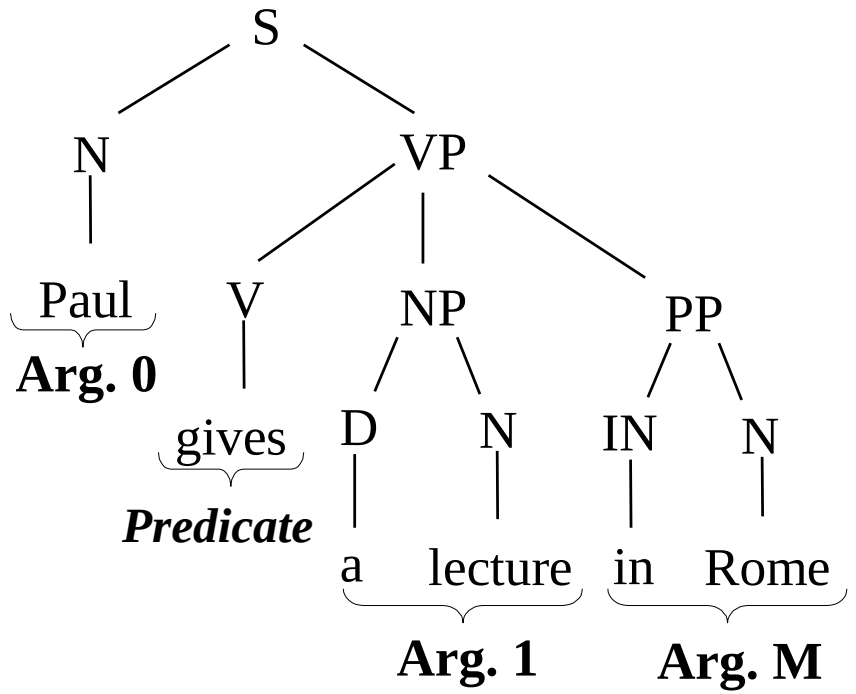
GOAL: Place toward which action is directed

SOURCE: Place from which action originates

ASSOCIATIVE: Performs action with Agent.

Lexical Semantics: Predicates, Arguments & Roles

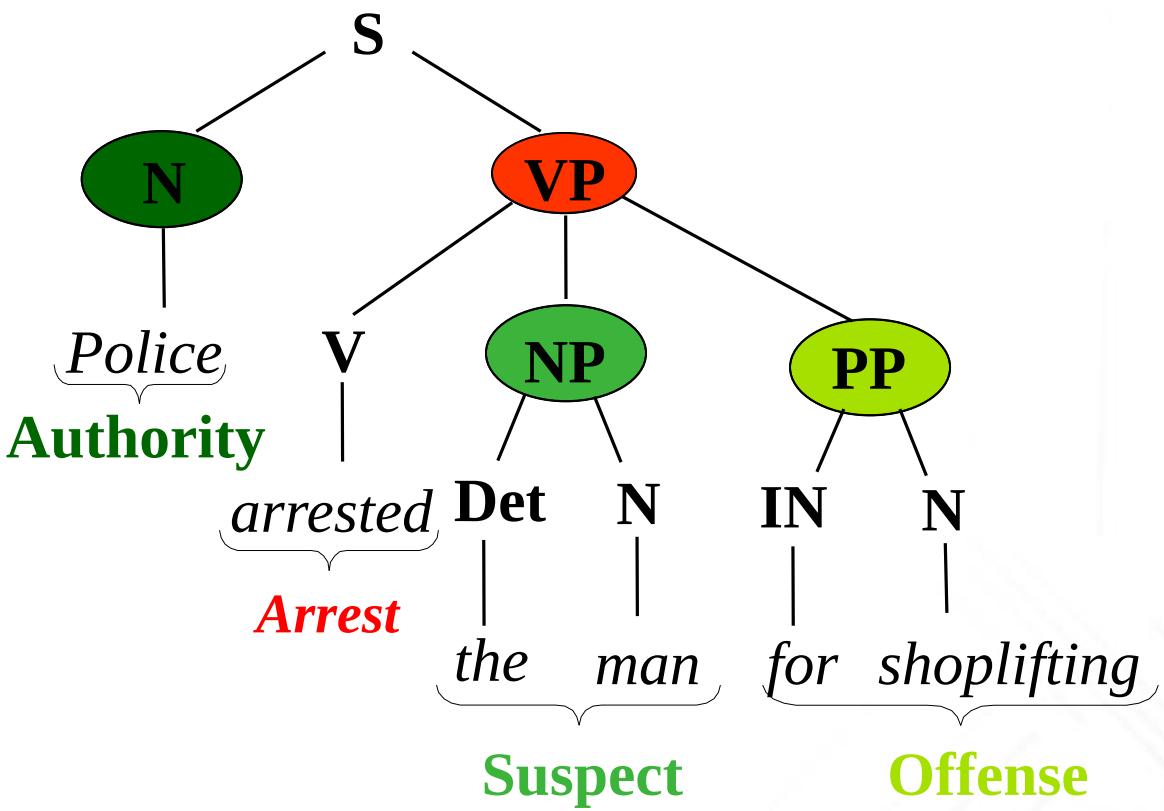
- The syntax-semantic mapping



- Different semantic theories
  (e.g. PropBank vs. FrameNet)

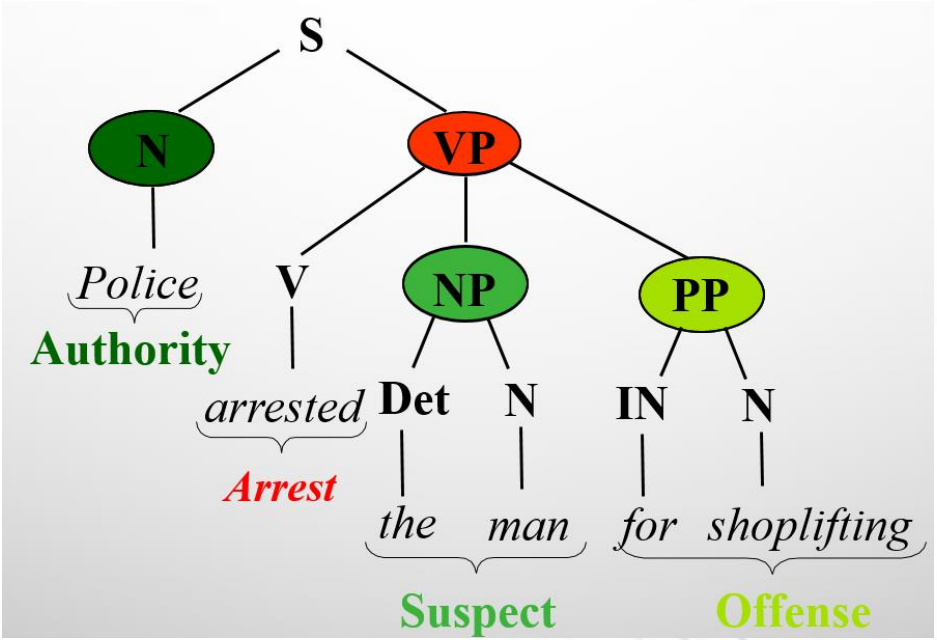
Linking syntax to semantics (Framenet)

- *Police arrested the man for shoplifting*



A tabular vision

Word	Predicate	Semantic Role
Police	-	Authority
arrested	Target	Arrest
the	-	SUSPECT
man	-	SUSPECT
for	-	OFFENSE
Shoplifting	-	OFFENSE



Semantics in NLP: Resources

- Lexicalized Predicate Models
  - Propbank
  - NomBank
- Framenet
  - Inspired by frame semantics
  - Frames are lexicalized prototoypes for real-world situations
  - Participants are called frame elements (roles)

Frame Semantics

- Research in Empirical Semantics suggests that **words represents categories of experience** (*situations*)
- A ***frame*** is a cognitive structuring device (i.e. a kind of prototype) indexed by *words* and used to support understanding (Fillmore, 1975)
  - *Lexical Units* **evoke** a Frame in a sentence
- Frames are made of *elements* that express participants to the situation (*Frame Elements*)
- During communication LUs evoke the frames

Fram

Frame: KILLING		
A KILLER or CAUSE causes the death of the VICTIM.		
Frame Elements	KILLER	**John** drowned Martha.
	VICTIM	John drowned **Martha**.
	MEANS	The flood exterminated the rats **by cutting off access to food**.
	CAUSE	**The rockslide** killed nearly half of the climbers.
	INSTRUMENT	It's difficult to suicide **with only a pocketknife**.
Predicates	annihilate.v, annihilation.n, asphyxiate.v,assassin.n, assassinate.v, assassination.n, behead.v, beheading.n, blood-bath.n, butcher.v, butchery.n, carnage.n, crucifixion.n, crucify.v, deadly.a, decapitate.v, decapitation.n, destroy.v, dispatch.v, drown.v, eliminate.v, euthanasia.n, euthanize.v, ...	

Frame Semantics

- Lexical descriptions are expected to define the indexed frame and the frame elements with their realization at the syntactic level:
  - *John bought a computer from Janice for 1000 $*
- Mapping into syntactic arguments
  - the buyer is (usually) in the subject position
- Obligatory vs. optional arguments
- Selectional preferences
  - The seller and the buyer are usually “humans” or “social groups”

The FrameNet project

- The aims
  - Create a lexical resource by describing a significant portion of English in terms of precise and rich frame semantics
- The output
  - Frame Database: a structured system of Frames and Fes
  - Lexical database: syntactic and semantic descriptions of frame-evoking words (N,V,A)
  - Annotated Corpus: wide coverage examples


FrameReport - Mozilla Firefox
Frame Report (recent data)
| Top of Frame Index | Top of Lexical Unit Index |
Committing_crime
Definition:
A Perpetrator (generally intentionally) commits a Crime, i.e. does something not permitted by the laws of society.
They PERPETRATED a felony by substituting a lie for negotiations.
The suspect had allegedly COMMITTED the crime to gain the attention of a female celebrity.
FEs:
Core:
Crime [Cr] An act, generally intentional, that has been formally forbidden by law.
How can he COMMIT treason against the King of England in a foreign country , if he is not English?
He PERPETRATED a crime against mother nature.
Perpetrator [Perp] The individual that commits a Crime.
How can he COMMIT treason against the King of England in a foreign country , if he is not English?
He PERPETRATED a crime against mother nature.
Non-Core:
Frequency [Freq] The frequency with which a Crime is committed.
The average serial killer COMMITS a crime every five years.
Instrument [Inst] The Instrument used in committing the crime.
Most crimes are COMMITTED with a firearm.

Killing

FEs:

Non-Core:

Beneficiary [ben]	This extra-thematic FE applies to participants that derive a benefit from the occurrence of the event specified by the target predicate.
Circumstances []	Circumstances describe the state of the world (at a particular time and place) which is specifically independent of the event itself and any of its participants.

Semantic Type: Physical_entity
Excludes: Cause

It's difficult to SUICIDE with only a pocketknife.

Killer [Kill]	The person or sentient entity that causes the death of the Victim.
Means []	The method or action that the Killer or Cause performs resulting in the death of the Victim.
Victim []	The living entity that dies as a result of the killing.

Excludes: Cause

Semantic Type: State_of_affairs
Excludes: Cause

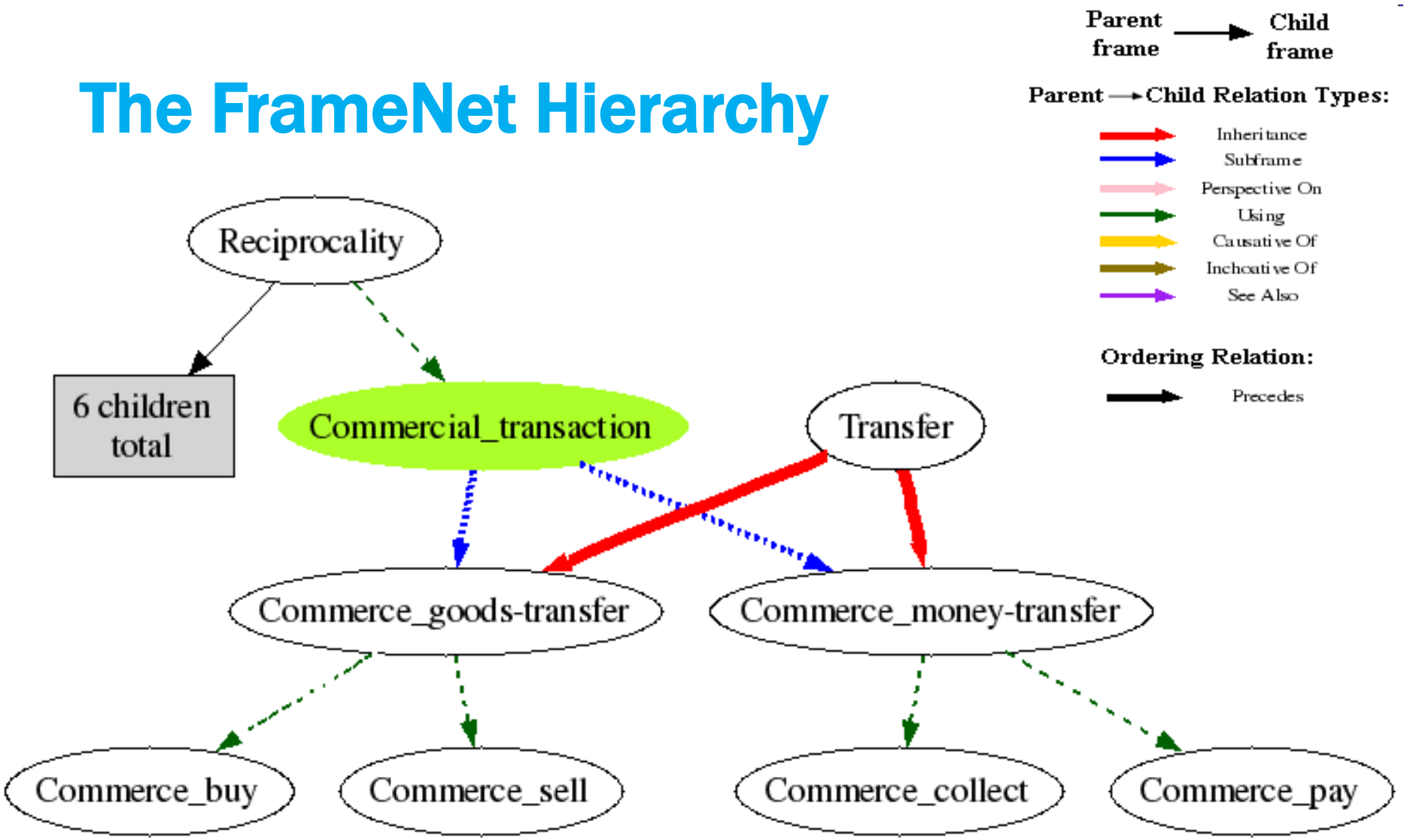
The flood EXTERMINATED the rats by cutting off access to food.

Semantic Type: Sentient

Non-Core:

Beneficiary [ben] This extra-thematic FE applies to participants that derive a benefit from the occurrence of the event specified by the target predicate.

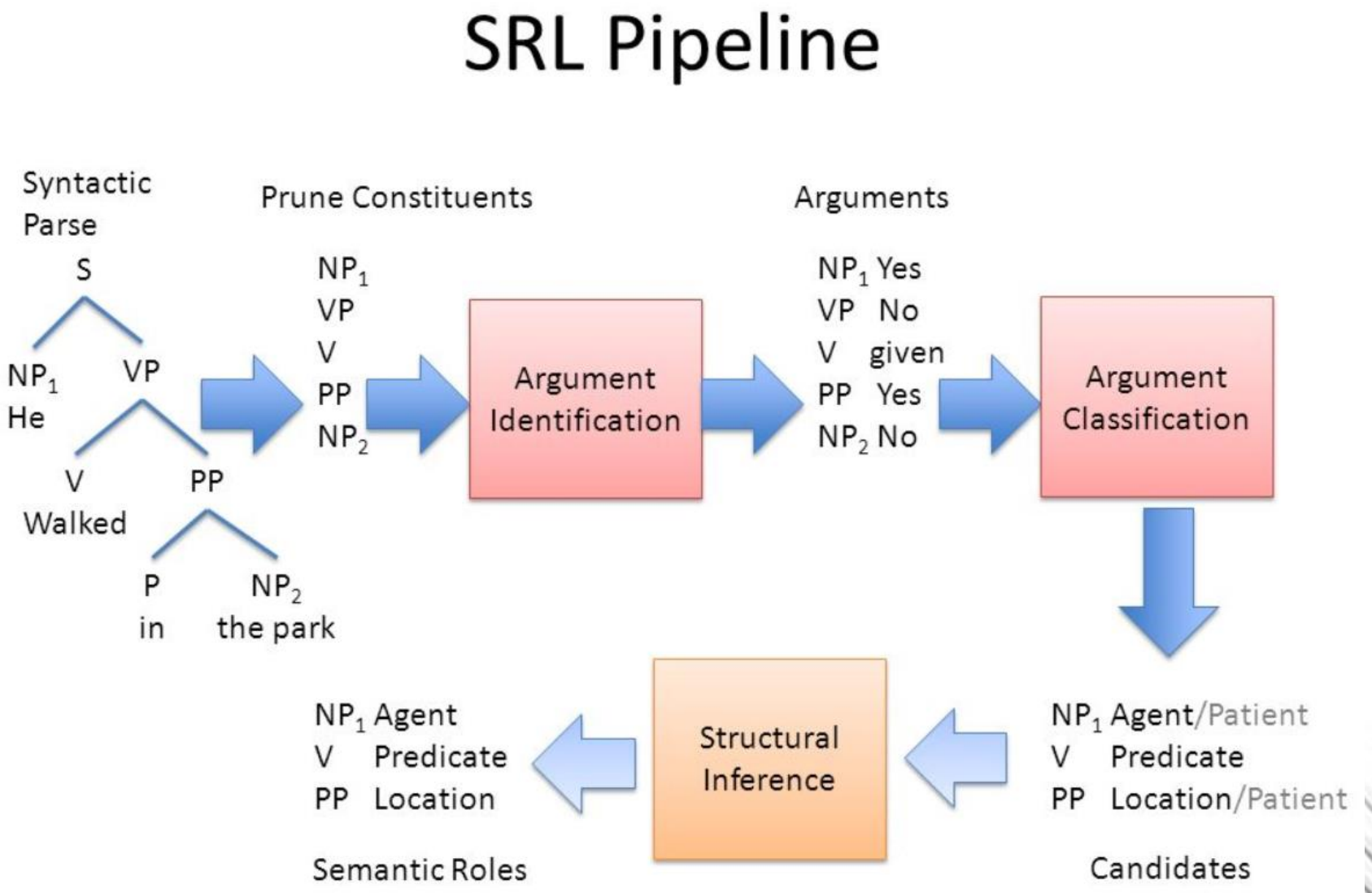
The FrameNet Hierarchy



Framenet - Data

- Methodology of constructing FrameNet
  - Define/discover/describe frames
  - Decide the participants (frame elements)
  - List lexical units that evoke the frame
  - Find example sentences in the BNC and annotate them
- Corpora
  - FrameNet I -British National Corpus only
  - FrameNet II -LDC North American Newswire corpora
- Size
  - >10,000 lexical units, >825 frames, >135,000 sentences
- **http://framenet.icsi.berkeley.edu**

Using Framenet/PropBank



Overview

- Intelligenza Artificiale e Lingue parlate e scritte
  - Informazioni e Rappresentazioni coinvolte
  - Sfide (ri)correnti, battaglie (già) vinte e rischi inerenti …
- Elaborazione Automatica delle Lingue: Modelli, Metodi e *Risultati*
- break
- Ruolo delle Tecnologie dell'Apprendimento ed Applicazioni:
  - Sviluppo Automatico di Dizionari, Lessici Semantici ed Ontologie
  - Trattamento Semantico della Documentazione Investigativa
  - Sistemi Web-based di Opinion Mining, Market Watch & Brand Reputation Management

Overview

- Intelligenza Artificiale e Lingue parlate e scritte
  - Informazioni e Rappresentazioni coinvolte
  - Sfide (ri)correnti, battaglie (già) vinte e rischi inerenti …
- Elaborazione Automatica delle Lingue: Modelli, Metodi e *Risultati*
- break
- Ruolo delle Tecnologie dell'Apprendimento ed Applicazioni:
  - Sviluppo Automatico di Dizionari, Lessici Semantici ed Ontologie
  - Trattamento Semantico della Documentazione Investigativa
  - Sistemi Web-based di Opinion Mining, Market Watch & Brand Reputation Management

Conclusioni

- I dati della odierna società della conoscenza **sono opachi dal punto di vista epistemologico** e l'intermediazione dei sistemi di calcolo deve sostenere **processi complessi di interpretazione**
- Le tecnologie del linguaggio possono svolgere un ruolo fondamentale nel **sostenere in modo accurato i processi agenti sui Big Data** e nel renderli **economicamente sostenibili**
- In NLP le metodologie di **Rappresentazione della Conoscenza e Reasoning** conoscono una specifica **sinergia** con le metodologie di **Machine Learning**
  - Strutture Dati particolarmente **complesse** (alberi e grafi etichettati)
  - **Enormi volumi** di conoscenza coinvolti
  - **Vaghezza** ed **Incompletezza** caratteristici delle diverse inferenze necessarie
- Questi processi di AI (NLP&ML) costituiscono una branca attiva dell'Informatica che determina in modo rilevante il successo di processi innovativi della automazione in diversi ambiti industriali
  - Gestione Documentale
  - Semantic Search
  - Opinion Analysis & Brand Reputation

Riferimenti

- AI & Robotics. «Robot Futures», Ilah Reza Nourbakhsh, MIT Press, 2013
- NLP & ML:
  - «Statistical Methods for Speech Recognition», F. Jelinek, MIT Press, 1998
  - «Speech and Language Processing", D. Jurafsky and J. H .Martin, Prentice-Hall, 2009.
  - "Foundations of Statistical Natural Language Processing, Manning & Schtze, MIT Press 2001.
- Sitografia:
  - SAG, Univ. Roma Tor Vergata: http://sag.art.uniroma2.it/
  - Reveal s.r.l.: http://www.revealsrl.it/