



IA a.a. 2021-22. Introduzione al NLP.

Le lingue al crocevia di dati, informazioni e conoscenza

Roberto Basili

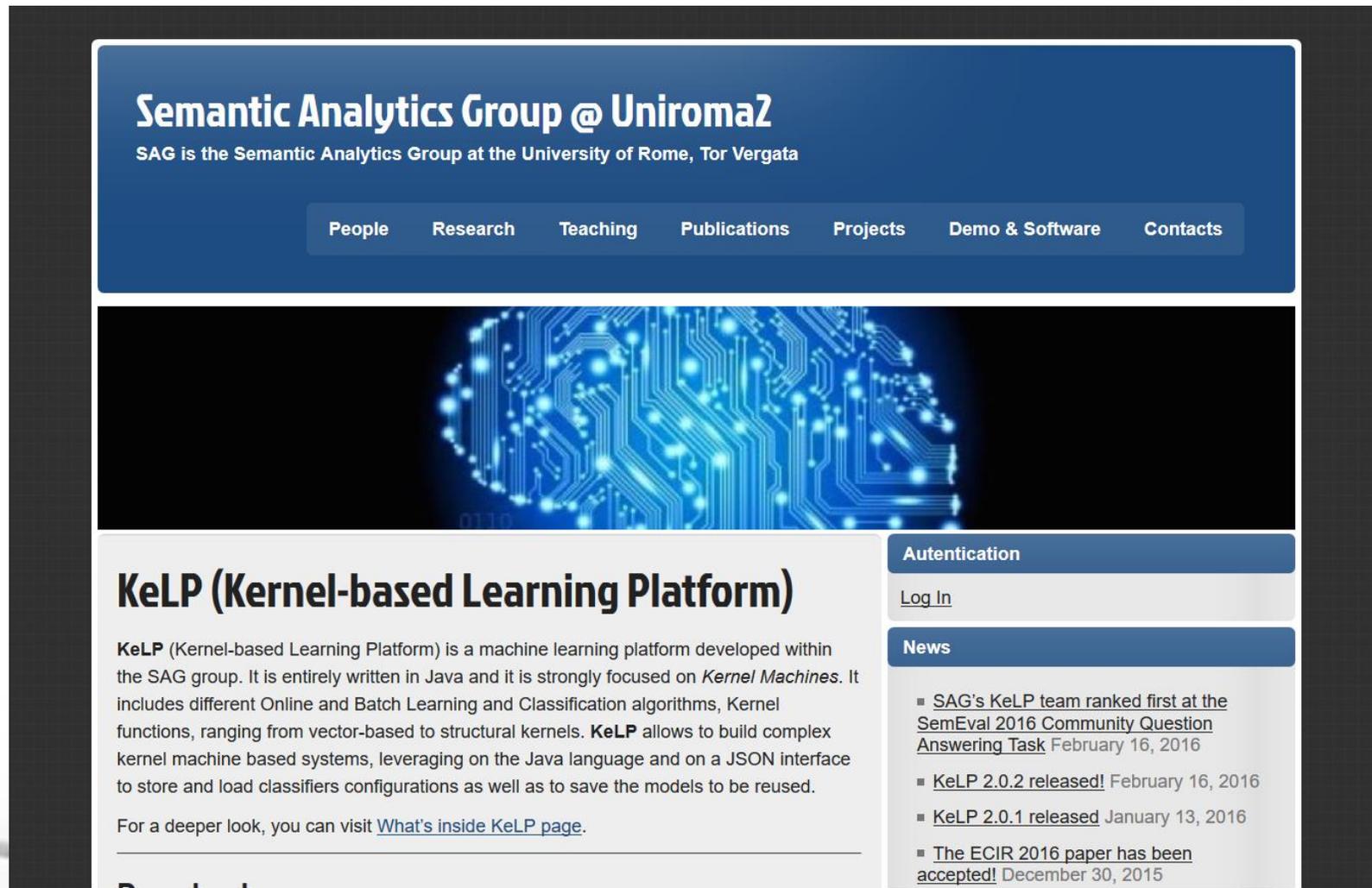
(Università di Roma, Tor Vergata, basili@info.uniroma2.it)

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>

Semantic Analytics Group

- <http://sag.art.uniroma2.it>

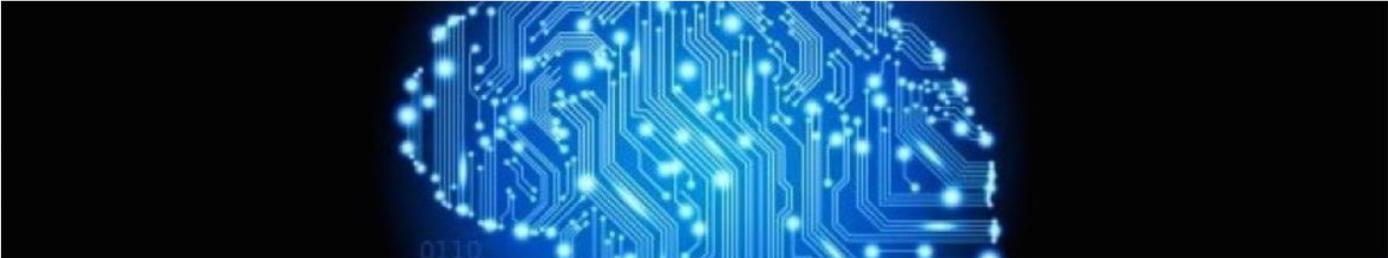


The screenshot shows the homepage of the Semantic Analytics Group website. The header is dark blue with the group name and a navigation menu. Below the header is a large image of a brain with glowing blue circuitry. The main content area features a section for KeLP (Kernel-based Learning Platform) with a description and a link to a 'What's inside KeLP' page. On the right side, there are sections for 'Authentication' with a 'Log In' link, and 'News' with a list of recent updates.

Semantic Analytics Group @ Uniroma2

SAG is the Semantic Analytics Group at the University of Rome, Tor Vergata

- People
- Research
- Teaching
- Publications
- Projects
- Demo & Software
- Contacts



KeLP (Kernel-based Learning Platform)

KeLP (Kernel-based Learning Platform) is a machine learning platform developed within the SAG group. It is entirely written in Java and it is strongly focused on *Kernel Machines*. It includes different Online and Batch Learning and Classification algorithms, Kernel functions, ranging from vector-based to structural kernels. **KeLP** allows to build complex kernel machine based systems, leveraging on the Java language and on a JSON interface to store and load classifiers configurations as well as to save the models to be reused.

For a deeper look, you can visit [What's inside KeLP page](#).

Authentication

[Log In](#)

News

- [SAG's KeLP team ranked first at the SemEval 2016 Community Question Answering Task](#) February 16, 2016
- [KeLP 2.0.2 released!](#) February 16, 2016
- [KeLP 2.0.1 released](#) January 13, 2016
- [The ECIR 2016 paper has been accepted!](#) December 30, 2015

Semantic Analytics Group @ Uniroma2

SAG is the Semantic Analytics Group at the University of Rome, Tor Vergata

People

Research

Teaching

Publications

Projects

Demo & Software

Contacts

Content Processing and Acquisition

Ontology Engineering

Machine Learning

Web & Information Retrieval

Text Processing and Natural Language Parsing

Distributional Semantics

Human-Robot Interaction

Semantic Role Labeling

Sentiment Analysis

People

[Professors](#)

[Postdocs](#)

[PhD Students](#)

[Fellows](#)



Paolo Annesi

Postdoc

annesi@info.uniroma2.it

[Details...](#)

Authentication

[Log In](#)

News

- [SAG's KeLP team ranked first at the SemEval 2016 Community Question Answering Task](#) February 16, 2016
- [KeLP 2.0.2 released!](#) February 16, 2016
- [KeLP 2.0.1 released](#) January 13, 2016
- [The ECIR 2016 paper has been accepted!](#) December 30, 2015

AI*IA: Associazione Italiana per l'IA

- Nata a Roma nel 1988 da circa 60 ricercatori Fondatori.
- Include: CNR, 45+ Università, 2 Fondazioni, Industrie e Consorzi (MPS)
- Membro autorevole del network mondiale di ricerca in AI (ECCAI, IJCAI)



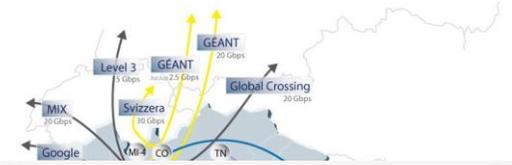
HomePage: <http://www.aixia.it/>

Twitter: https://twitter.com/ai_x_ia

- Prossimo Convegno Bari 14-17 Novembre 2017, URL: <http://aia2017.di.uniba.it/>



AILC (www.ai-ic.it)



sag.art.uniroma2.it/clic2017/it/home/

110%

CLiC It 2017 roma

EMNLP2017 Più visitati SMARRS - BIGDATA ... Didattica Usefullinks KeLP SAG Demos Books AbiliNQS RBas Bob Cloud Reader Community Detection... UtiliLinks Home WordReference.com

"IDEAS" to change the world



Home Organizzazione Call Programma Eventi Registrazione Informazioni Utili Italiano

CLIC-IT 2017

Fourth Italian Conference on Computational Linguistics

NEWS: CALL FOR PAPERS - DEADLINE EXTENSION TO July 25, 2017

AILC HOME

CALL FOR PAPERS

Outline

- **Natural Language Processing: *Task, Modelli e Metodi***
- **Un esempio: computational semantics in Prolog**
- **Trattamento delle lingue e *Machine Learning***
 - Statistical Language Processing
 - Apprendimento discriminativo per l’NLP
- **Natural Language Processing: applications**
- **Conclusions & Perspectives**

Semantics, Open Data & Natural Language

www.takungpao.com.hk

今天是 2011年11月13日 星期日 顯示器最佳分辨率1024X768

今日天氣 加入最愛 設為首頁 大公網新版

2011 中國證券金紫荊獎 Golden Bauhinia Awards

看大公報 國際短信 讀者推薦 廣告業務

首頁 國內 國際 港澳 兩岸 評論 財經 體育 教育 科技 醫學 娛樂 文化 副刊 軍事 生活 旅遊 圖片 博客

關鍵詞: 欄目: 全部 最近三個月 三個月之前 檢索

▶ 手機新聞 ▶ 手機博客 ▶ 漢語學習 ▶ 新聞點擊排行 ▶ 招聘啓事

滾動新聞:

胡總語特首:防範經濟金融風險

胡錦濤在夏威夷會見出席APEC峰會的曾蔭權。他祝賀香港區議會選舉成功,並充分肯定曾蔭權及港府工作,要求做好經濟金融風險防範

胡連會登場 共同宣示九二共識

胡錦濤第四次在APEC峰會期間會見連戰。他強調,認同「九二共識」是兩岸開展對話協商的必要前提,也是兩岸關係和平發展的重要基礎

西藏黨代會高調反「藏獨」

德國作家:外媒錯誤報道西藏

傳媒入日本福島核電站採攝

英國大裁軍 傷兵難倖免

滇礦難已30死 13人生還

礦工講述內幕 事故並不意外

范徐麗泰認民望跌最不利

選舉再獲60提名表 累積逾千人

聖保羅中學本月底截止招

選委再獲60提名表 累積逾千人

民調逆轉 藍高層 國親吵鬧地

秋門訴求多 向藍綠表不滿

世界新七奇觀 亞洲景佔四

新奇觀選舉惹爭議

中國實體書店苦苦掙扎求

加入TPP 台密集會談探路

香港人家/蔡仕榮 人生導師 活出

自我香港人家/教導子女...

債務危機紓 港ADR幾全線

歐元反彈 兌美元逼近1.38

入世十年/充分對接 華強北最

入世十年/挑戰「二次...

抽離險「雜車」 工人險生

南亞漢命案 警拘日籍妻

神八天宮对接 第七次陳江會 李克強訪港

9.1衝擊事件 中國航母試航 辛亥革命百年

2011APEC 港黑金事件 2011施政報告

區議會選舉 香港特首選舉

2011APEC 港黑金事件 2011施政報告

神八天宮对接 第七次陳江會 李克強訪港

9.1衝擊事件 中國航母試航 辛亥革命百年

I contenuti digitali sono ampiamente *opachi* dal punto di vista **semantico**

Information, Web and Natural Languages



Chinese President Hu Jintao (R) shakes hands with Honorary Chairman of the Chinese Kuomintang (KMT) Lien Chan, in Honolulu, Hawaii, the U.S., Nov. 11, 2011.

(Xinhua/Huang Jingwen)

HONOLULU, United States, Nov. 11 (Xinhua) -- Hu Jintao, general secretary of the Central

Hu meets KMT honorary chairman in Hawaii

(Xinhua)

11:10, November 12, 2011



Chinese President Hu Jintao (R) shakes hands with Honorary Chairman of the Chinese Kuomintang (KMT) Lien Chan, in Honolulu, Hawaii, the U.S., Nov. 11, 2011.
(Xinhua/Huang Jingwen)



Miao ethnic group celebrates Miao's New Year in SW China



World's first Angry Birds exclusive shop opens in Helsinki



Night life in Shanghai



China's 2011 foreign trade to grow 20 p...

Beijing house prices stumble 5.1 pct as...

Lama students start school in Tibet Col...

Police in central China crack money ca...

China-ASEAN cooperation sees notable pr...

Who is Hu Jintao?

- 1 Hu reaffirms support to Hong Kong's sta...
- 2 Hu meets KMT honorary chairman in Hawaii
- 3 China in APEC: a mutually beneficial en...
- 4 Night life in Shanghai
- 5 China's 2011 foreign trade to grow 20 p...
- 6 Beijing house prices stumble 5.1 pct as...
- 7 Lama students start school in Tibet Col...
- 8 Police in central China crack money ca...
- 9 China-ASEAN cooperation sees notable pr...



Hu Jintao



Ricerca

Circa 725.000 risultati (0,09 secondi)

Tutto

Immagini

Mappe

Video

Notizie

Shopping

Più conte

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



Lingue, Ambiguità e Conoscenza

Seguici su: [f](#) [t](#)

Sport

CERCA



HOME [CALCIO](#) [FI](#) [SCI](#) [MOTOGP](#) [TENNIS](#) [BASKET](#) [RUGBY](#) [CICLISMO](#) [RUNNING](#) [VOLLEY](#) [ALTRI SPORT](#) [LIVE](#)

Calcio / Serie A: [Calendario](#) [Dirette e Risultati](#) [Classifica](#) [Marcatori](#)



Coppa Italia, Fiorentina-Benevento 2-1: Milenkovic e Sottil portano i viola agli ottavi, applausi per la squadra di Caserta



▲ (cge)

La squadra viola non brilla, ma conquista l'accesso al turno successivo dove affronterà il Napoli. Al Franchi decidono le reti del difensore e dell'attaccante, inutile invece la firma ospite di Moncini

Content Semantics & Natural Language



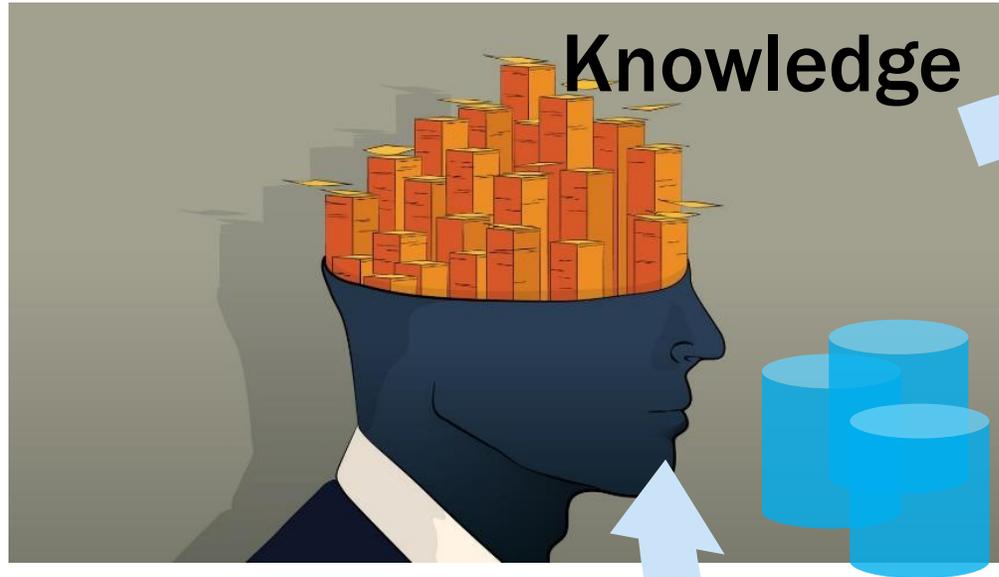
- Le lingue sono i veicoli tradizionali e consolidati della creazione, condivisione e comunicazione delle conoscenze diffuse nei contenuti del Web
- Sono le parole e le strutture sintattiche ad esprimere in modo trasparente i concetti, le attività, gli eventi, le astrazioni e le relazioni concettuali che noi scambiamo attraverso i flussi di dati
- *“Language is parasitic to knowledge representation languages but the viceversa is not true” (Wilks, 2001)*
- Apprendere la lettura (**Learning to Read**) abilita la genesi di nuova conoscenza (**Knowledge Distillation**) attraverso processi integrati di **Interpretazione semantica**

Semantica, Lingue & Learning: tasks

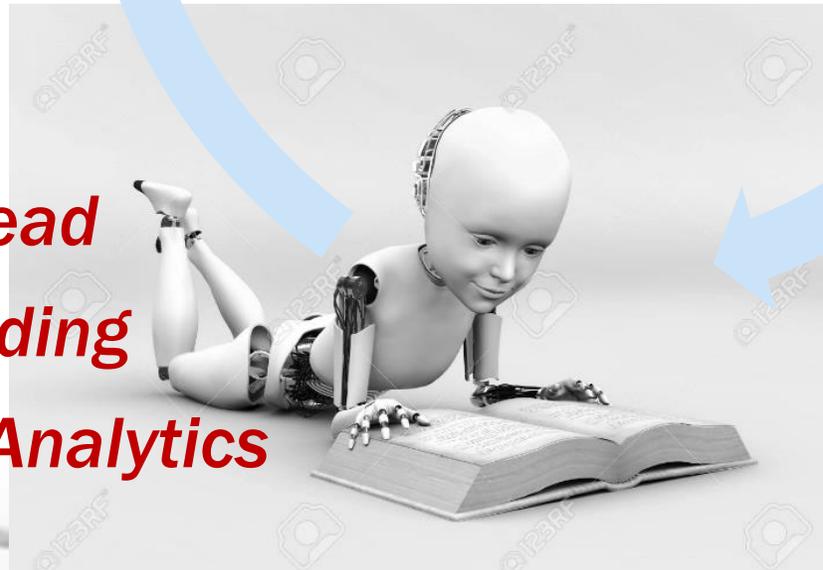


- Dal **Learning to Read** alla **Knowledge Distillation**:
 - **Information Extraction: dai testi alle KBs**
 - Entity Recognition e Classificazione
 - Estrazione di Relazioni
 - Semantic Role Labeling (Shallow Semantic Parsing)
 - **Stima della somiglianza (semantica) tra Testi diversi**
 - Funzioni di somiglianza tra parole e testi
 - Textual Entailment Recognition
 - Disambiguazione del Senso delle parole (*boa* marina vs. serpente *boa*)
 - **Ricerca dei Contenuti: Web Search, Question Classification and Selezione della Risposta (Answer Ranking)**
 - **Acquisizione Automatica della Conoscenza**, e.g. *ontology learning*
 - **Opinion Mining, Sentiment Analysis e Social Network Analysis**

AI & NLP: knowledge acquisition, and decision-making



Learning to read
Machine Reading
Learning for Analytics

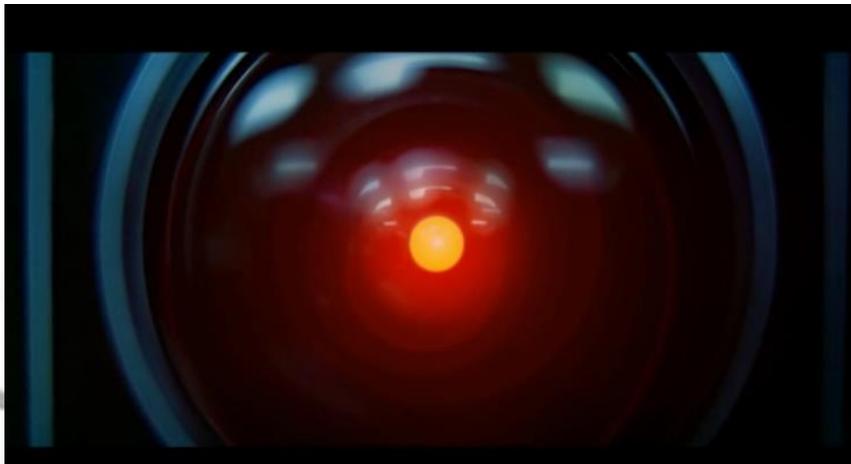


Outline

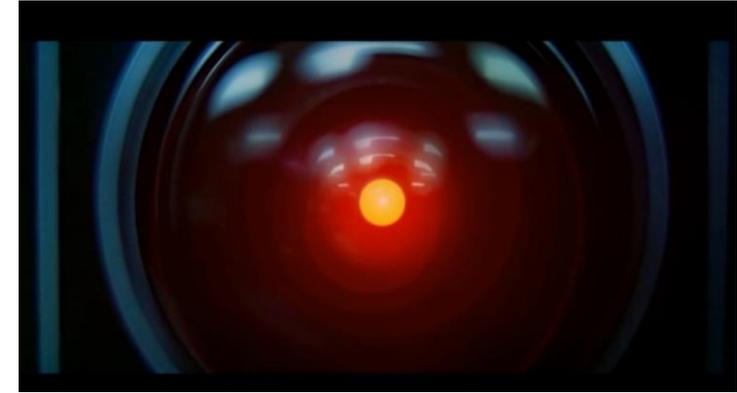
- **Artificial Intelligence & Natural Language Processing**
 - Comunicazione linguistica & Conoscenza
 - Il ruolo dei dati
-  **Natural Language Processing: *Task*, Modelli e Metodi**
- **Trattamento delle lingue e *Machine Learning***
 - Statistical Language Processing
 - Apprendimento discriminativo per l'NLP
- ***Natural Language Processing*: applicazioni**
- **Conclusioni & Prospettive**

NLP: quali conoscenze?

- 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.*

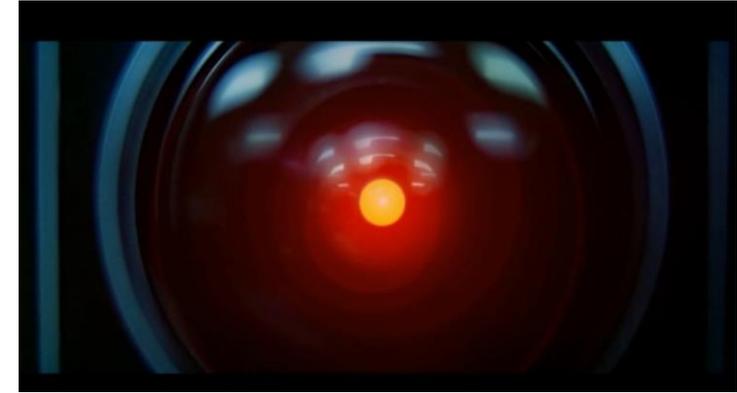


Qual'e' la conoscenza di HAL?



- **Riconoscimento e Sintesi del linguaggio parlato**
 - Dizionari (*spelling*)
 - Regole fonetiche (come i suoni vengono riconosciuti e prodotti)
- **Comprensione**
 - **Conoscenze Lessicali**
 - Qual'e' il significato delle parole?
 - Come tali significati si combinano (*`pod bay door'*)
 - **Competenza riguardo alla struttura sintagmatica delle frasi**
 - *I'm I do, Sorry that afraid Dave I'm can't*

Qual'e' la conoscenza di HAL?



- **Dialogo & pragmatica**

- “*open the door*” è una richiesta (e non una dichiarazione o una interrogazione)
- La replica implica una azione ed è necessario usare modi gentili (anche a fronte dell'intenzione di uccidere ...)
- E' utile comportarsi in modo cooperativo (*I'm afraid, I can't...*)
- Infine: cosa significa *that* in *I can't do that*?

Trattamento delle lingue come processo di interpretazione (semantica)

- Elaborare un testo corrisponde a comprendere diversi aspetti relativi al suo significato:
 - Dominio tematico (e.g. scienze/economia/sport)
 - Obiettivi Operativi (e.g. e-mail spam)
 - Entità coinvolte, ad esempio *persone* or *luoghi*
 - Eventi potenziali (e.g. fatti raccontati dal testo)
 - Obiettivi Comunicativi (e.g. dialogo, ordini/dichiarazioni/pianificazione)
- RISULTATO: una *rappresentazione esplicita del significato del testo* con lo scopo di *sostenere tipi diversi di decisioni* (inferenze) (e.g. ranking nell'IR, pianificazione, acquisizione di nuova conoscenza, ...)

Sfide Principali

- **Accuratezza Linguistica**
 - Qual è il grado di approssimazione della performance dei parlanti nativi?
- **Robustezza (errori/rumore/incompletezza)**
 - Quanti e quali errori compiamo
- **Scala**
 - Qual è la copertura dei fenomeni (Lessici/Grammatiche)?
- **Expressività**
 - Di quale informazione semantica di base abbiamo bisogno nei Dizionari, Lessici e Thesauri?
 - Come interagiscono la conoscenze linguistiche ed i Modelli del Mondo nei diversi tipi di inferenza?
- **Flessibilità**
 - Come variano le prestazioni rispetto alla variabilità linguistica (e.g. *producer vs. consumer*)?
- **Naturalezza**
 - La qualità della produzione o del riconoscimento linguistico è indistinguibile da quella dei parlanti della lingua

Lingue & Ambiguità



Ambiguità

La frase inglese

"Dogs must be carried on this escalator"

attiva le seguenti differenti possibili interpretazioni:

- *All dogs should have a chance to go on this wonderful escalator ride*
- *This escalator is for dog-holders only*
- *You can't carry your pet on the other escalators*
- *When riding with a pet, carry it*

Livelli di Ambiguità

- Semantica



- Grammaticale



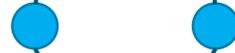
- Morfologica



- Fonologica



dei/dèi



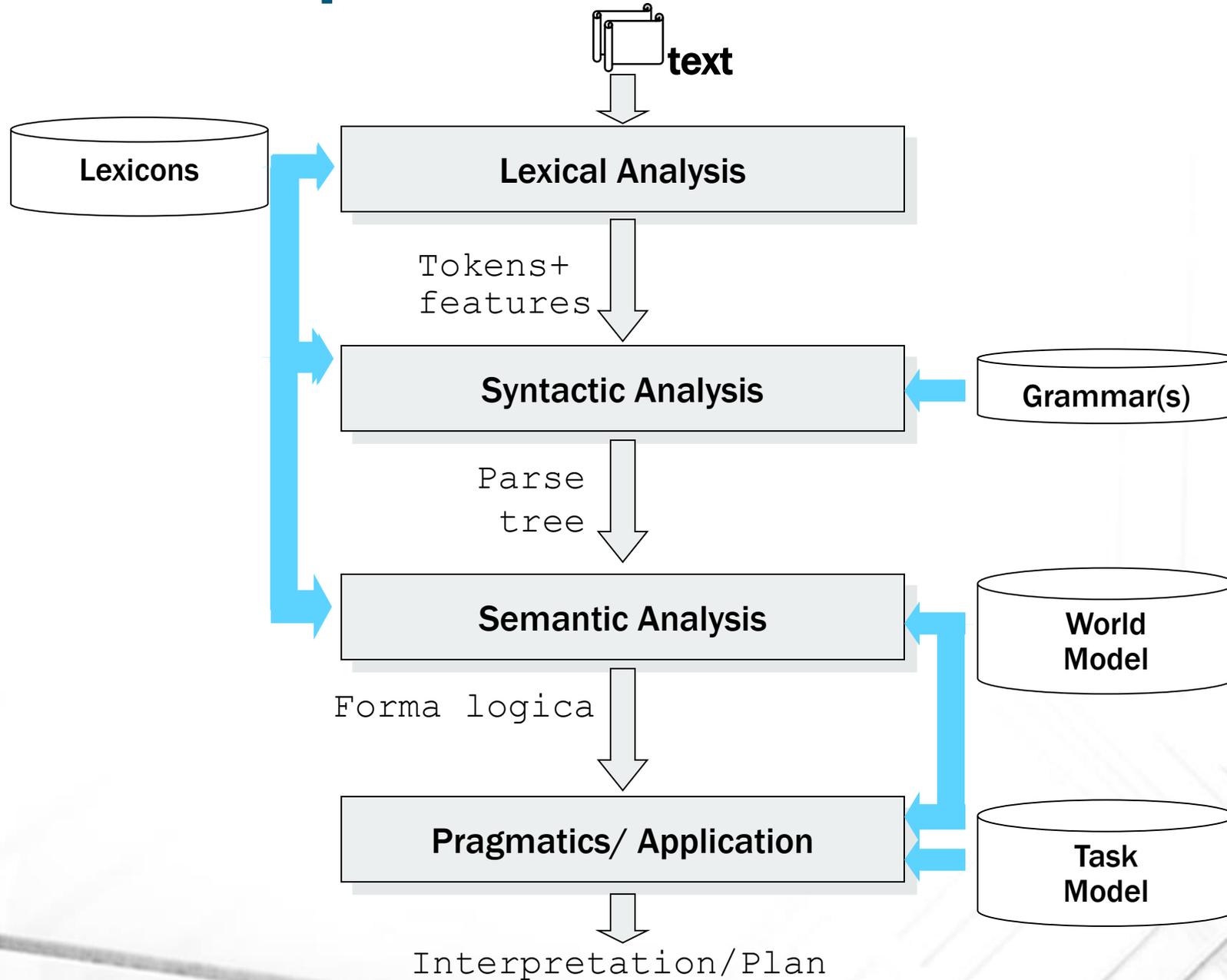
*compro la borsa
in pelle*



il timore dei manager



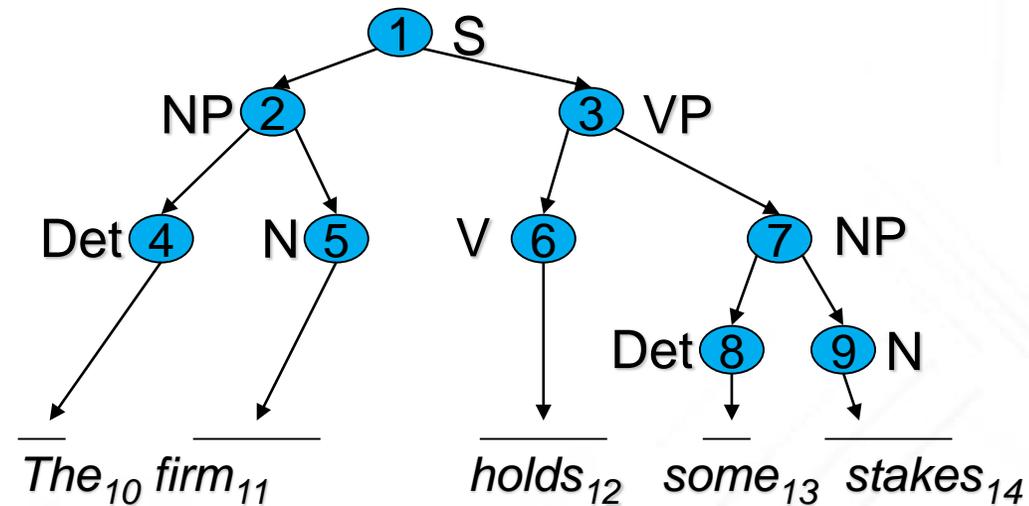
NLP: il processo



Sintassi: Grammatiche a struttura sintagmatica

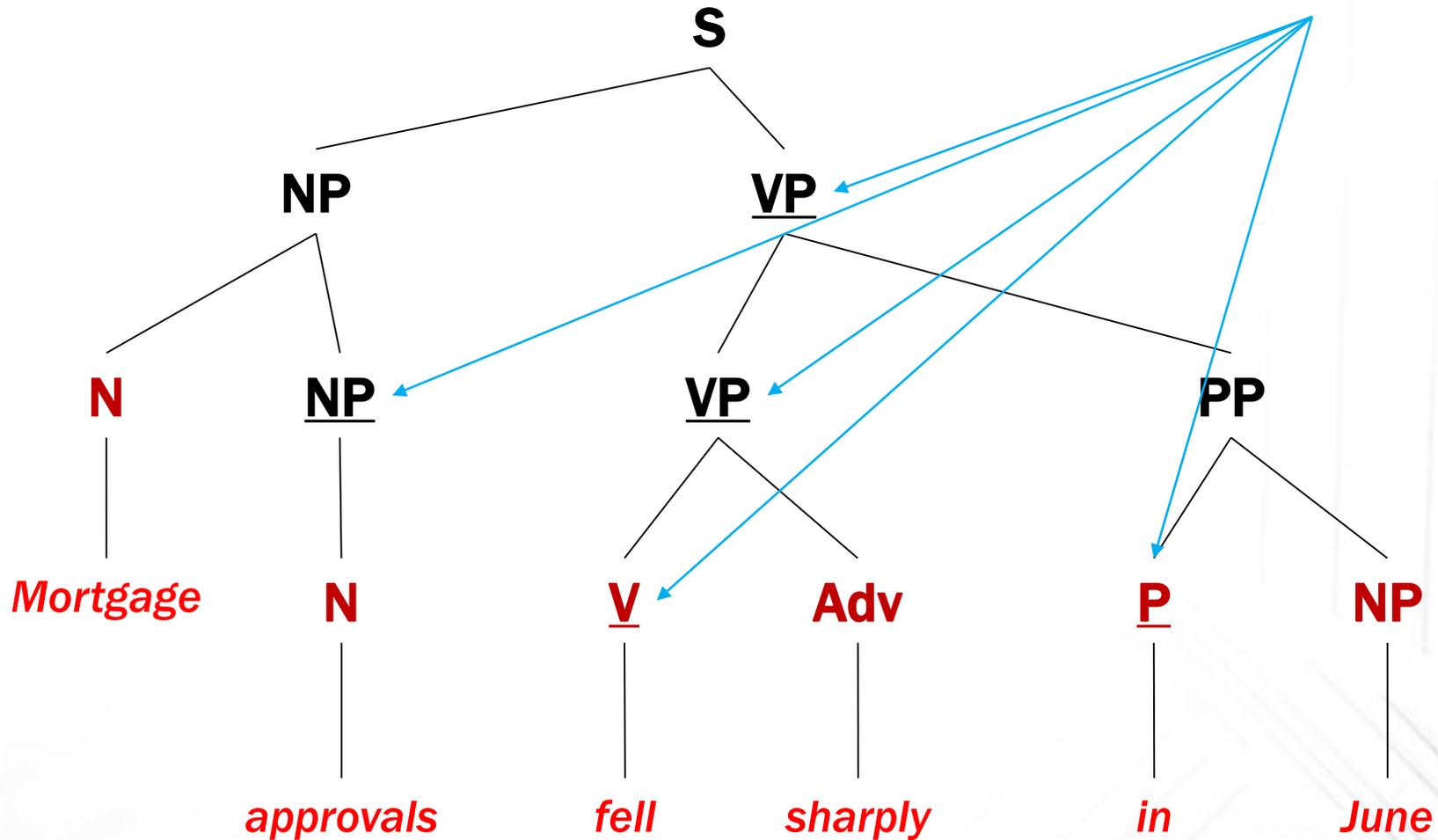
“The firm holds some stakes”

- $V_n = \{S, NP, VP, Det, N\}$, Axiom: S
- Productions: $\{S \rightarrow NP VP, VP \rightarrow V NP, NP \rightarrow Det N\}$
- Derivation:
 - $S > NP VP > Det N VP > The N VP > The firm VP > The firm V NP > The firm holds NP > The firm holds Det N > The firm holds some N > The firm holds some stakes$



Parsing a Costituenti

Teste Grammaticali marcate



Altre strutture: Dependency Parsing

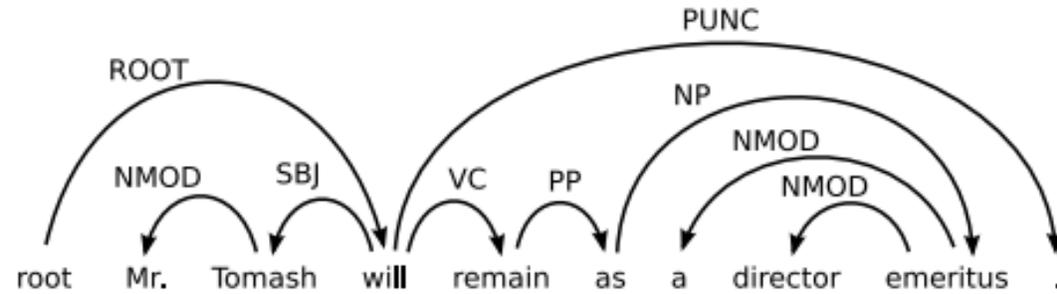


Figure 1: A projective dependency graph.

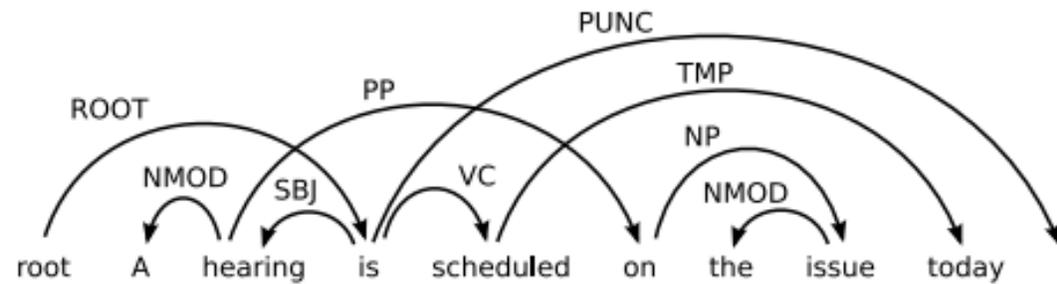
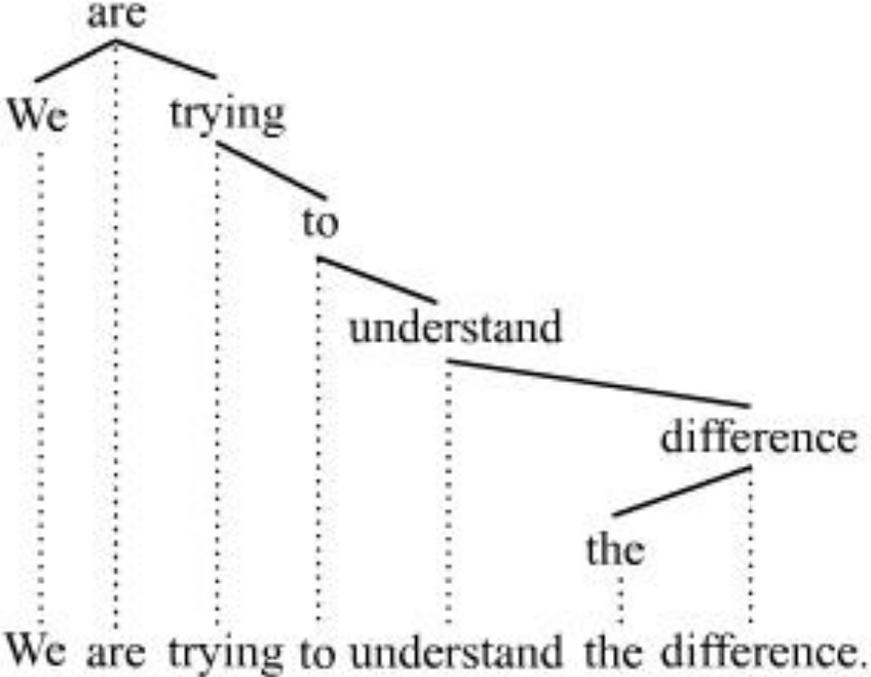
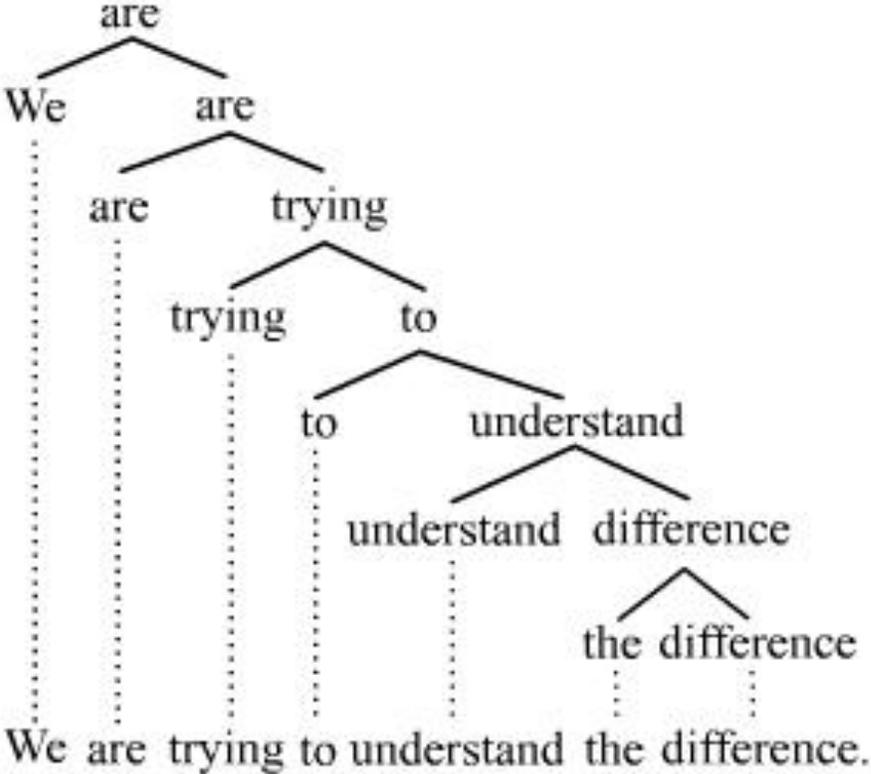


Figure 2: Non-projective dependency graph.

Constituency vs. Dependency



Dependency



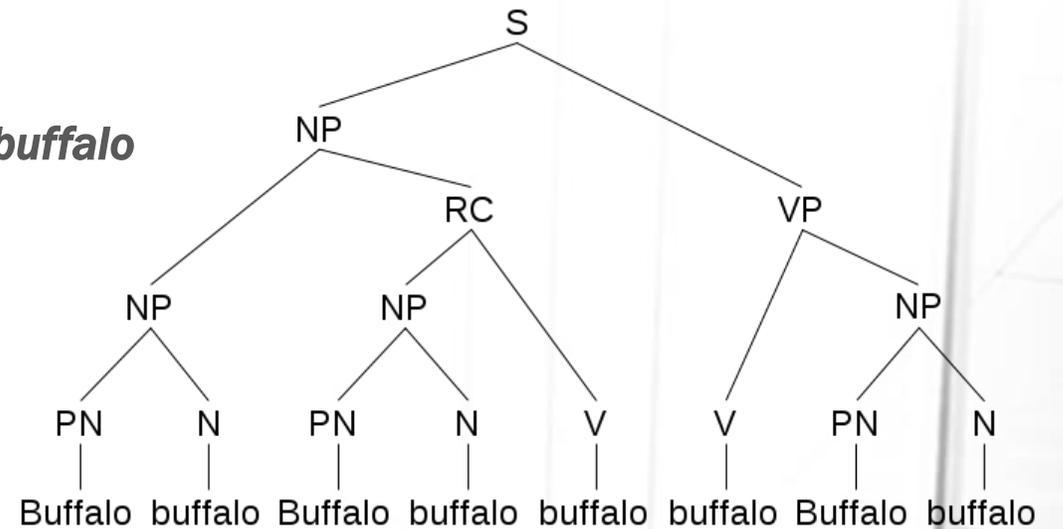
Constituency (BPS)

Parsing e Ambiguità

- Lo spazio di ricerca di un parser è enorme a causa delle molteplici ambiguità che interagiscono in modo combinatorio
 - E.g. *La vecchia porta la sbarra,*

Buffalo buffalo Buffalo buffalo buffalo buffalo Buffalo buffalo

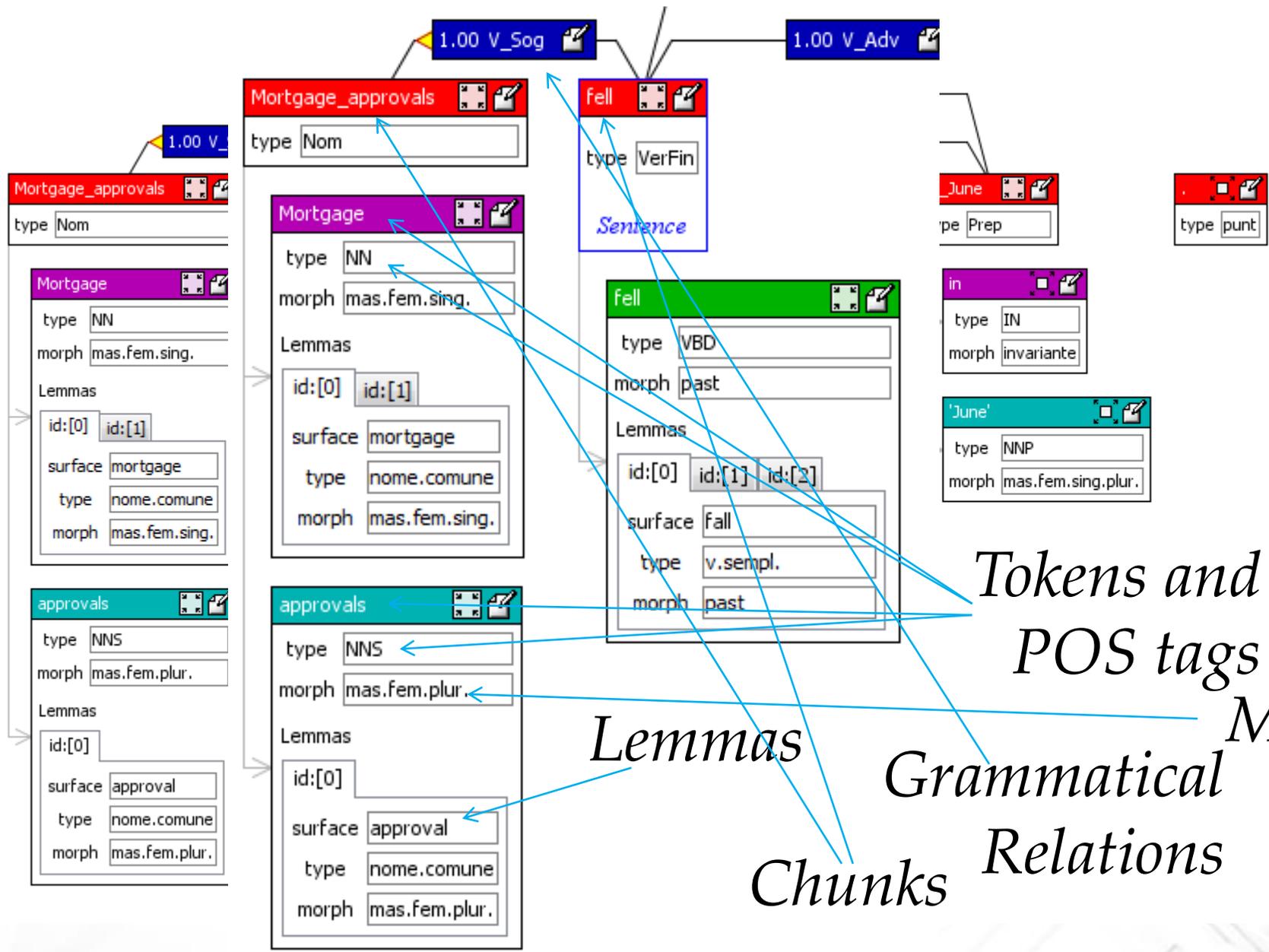
- C'è una dipendenza stretta con la semantica
 - La maggior parte delle sorgenti di ambiguità non possono essere risolte a livello grammaticale
 - L'informazione lessicale (i sensi) sono cruciali
 - *Operare in un mercato* ≠ *Operare un paziente*



Bison from Buffalo, New York, who are intimidated by other bison in their community also happen to intimidate other bison in their community



**(A(SHIP SHIPPING)SHIP)
SHIPPING(SHIPPING SHIPS))**



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

Un esempio per la lingua italiana

The image displays two windows from the RevNLT system. The left window, titled "RevNLT System GUI - Reveal srl", shows a parse tree for the sentence "Matteo corre verso la La Spezia. Poi passa con il rosso al semaforo." The tree structure is as follows:

- Root node: **V_Sog 1.00** (type: Nom) and **V_PP 1.00** (type: VerFin).
- Under **V_Sog 1.00**:
 - Node: **Matteo** (type: Nom, morph: invaria...)
 - Node: **Matteo** (type: SP, morph: invaria...)
- Under **V_PP 1.00**:
 - Node: **corre** (type: VerFin, morph: 3.sing.ind.p...)
 - Node: **verso_la_La_Spe...** (type: Prep, morph: invaria...)
 - Node: **verso** (type: E, morph: invaria...)
 - Node: **la** (type: RD, morph: fem.s...)
 - Node: **La_Spe...** (type: SP, morph: invariante)
 - Node: **punt** (type: punt)

The right window, titled "RevNLT Server/Client Manager", shows the server configuration. It includes fields for "hostname" (10.6.0.178) and "port number" (9.991), with a "Connect" button. A status indicator shows "Active server found - used language is: [it]". Below this is a "Server Launcher" section with a "Language" dropdown set to "it". There are also checkboxes for "parse text interactively" and "parse many files (batch processing)". The "Input text" field contains the sentence: "Matteo corre verso la La Spezia. Poi passa con il rosso al semaforo." At the bottom, there is a "Select output file" button, a checked checkbox for "Parse to anonymous temporary file and open in the viewer", and "Clear text" and "Start parsing" buttons.

Semantica

- Qual è il significato di una frase come
John saw Kim?

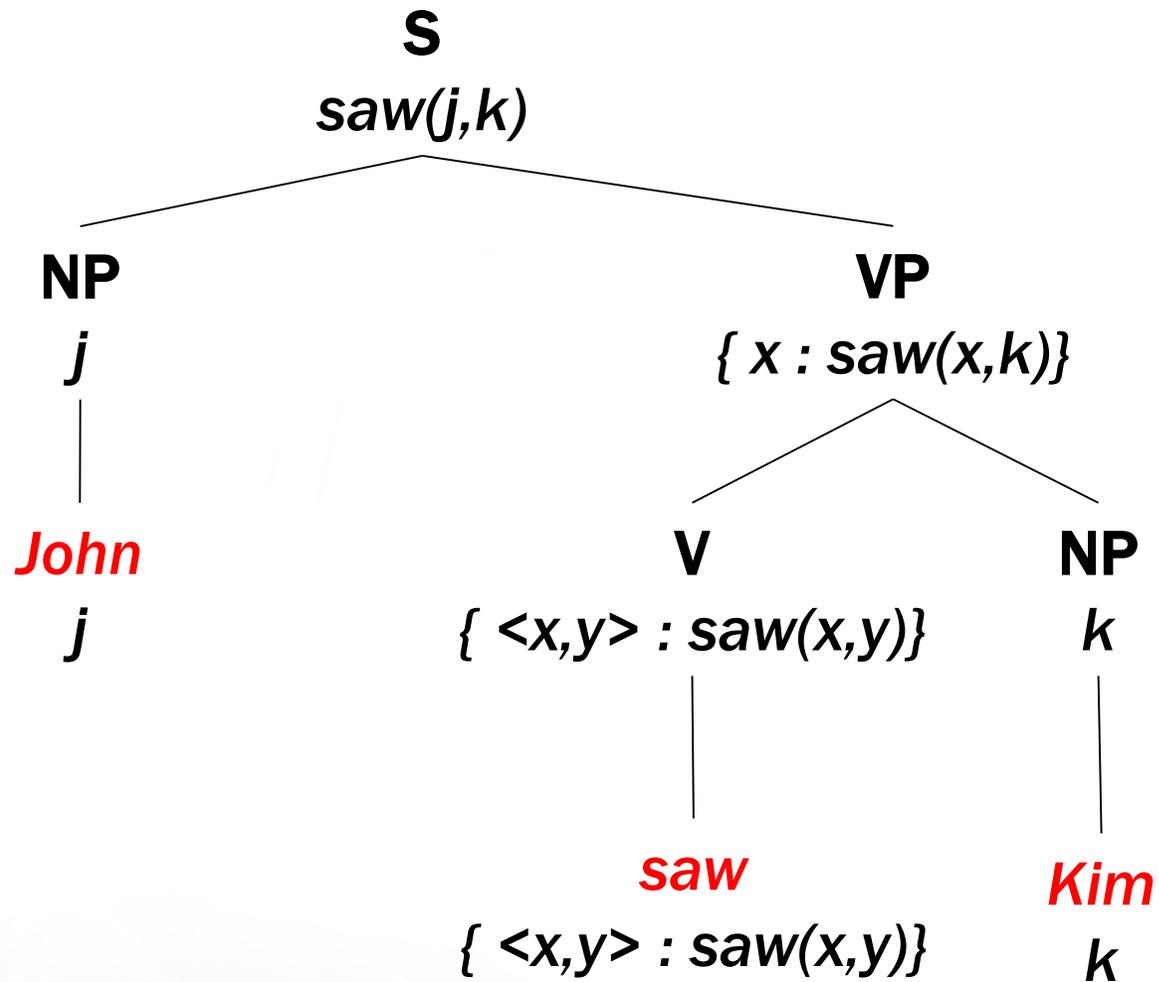


- **Proprietà desiderabili di una rappresentazione del significato:**
 - Deve essere **composizionale**, i.e. il significato deve essere funzione dei costituenti come **Kim, John** e il verbo **see**
 - **Indipendente dai fenomeni della sintassi**, e.g. **Kim was seen by John** è una frase sinonima (parafrasi)
 - Deve essere usata per **derivare delle conseguenze**:
 - **RISPONDERE A DOMANDE (QA):** **Who** was seen by John? **Kim!**
 - **COMPRENDERE UNA STORIA:** John saw Kim. **He** started running to **her**.

Semantica delle Lingue: la forma logica



John saw Kim



Computational Semantics

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

Analisi Semantica in Prolog

```
SemAnalisiDemo
[0] Giuseppe corre
[1] Giuseppe usa il prolog
[2] Giuseppe usa il linguaggio C
[3] Giuseppe diede il libro a Mario
[4] Giuseppe corre in Gennaio
[5] Giuseppe diede il libro a Mario
[6] Giuseppe diede il libro a Mario in campagna
[7] Mario diede il libro a Giuseppe in campagna
[8] Giuseppe diede il libro a Mario in Gennaio
[9] lo zio di Mario corre
[10] lo zio di Mario diede a Giuseppe il libro in Gennaio
[11] lo zio di Mario corre con Giuseppe in Gennaio
[12] il libro di Mario tratta il prolog
[13] il libro tratta di storia a Gennaio
[14] il libro di Mario tratta di storia a Gennaio
[15] il libro di storia tratta di Mario con Giuseppe
[16] lo zio di Mario preferisce il libro di storia a Gennaio
[18] lo zio di Mario preferisce il libro di matematica in Gennaio
The logic form of the example 0 is:

    corre('Giuseppe')

FL = corre('Giuseppe') .
3 ?- _
```

```
SemAnalisiDemo
[0] Giuseppe corre
[1] Giuseppe usa il prolog
[2] Giuseppe usa il linguaggio C
[3] Giuseppe diede il libro a Mario
[4] Giuseppe corre in Gennaio
[5] Giuseppe diede il libro a Mario
[6] Giuseppe diede il libro a Mario in campagna
[7] Mario diede il libro a Giuseppe in campagna
[8] Giuseppe diede il libro a Mario in Gennaio
[9] lo zio di Mario corre
[10] lo zio di Mario diede a Giuseppe il libro in Gennaio
[11] lo zio di Mario corre con Giuseppe in Gennaio
[12] il libro di Mario tratta il prolog
[13] il libro tratta di storia a Gennaio
[14] il libro di Mario tratta di storia a Gennaio
[15] il libro di storia tratta di Mario con Giuseppe
[16] lo zio di Mario preferisce il libro di storia a Gennaio
[18] lo zio di Mario preferisce il libro di matematica in Gennaio
The logic form of the example 4 is:

    corre('Giuseppe')&tempo('Gennaio')

FL = corre('Giuseppe')&tempo('Gennaio') .
4 ?-
```

```
SemAnalisiDemo
[0] Giuseppe corre
[1] Giuseppe usa il prolog
[2] Giuseppe usa il linguaggio C
[3] Giuseppe diede il libro a Mario
[4] Giuseppe corre in Gennaio
[5] Giuseppe diede il libro a Mario
[6] Giuseppe diede il libro a Mario in campagna
[7] Mario diede il libro a Giuseppe in campagna
[8] Giuseppe diede il libro a Mario in Gennaio
[9] lo zio di Mario corre
[10] lo zio di Mario diede a Giuseppe il libro in Gennaio
[11] lo zio di Mario corre con Giuseppe in Gennaio
[12] il libro di Mario tratta il prolog
[13] il libro tratta di storia a Gennaio
[14] il libro di Mario tratta di storia a Gennaio
[15] il libro di storia tratta di Mario con Giuseppe
[16] lo zio di Mario preferisce il libro di storia a Gennaio
[18] lo zio di Mario preferisce il libro di matematica in Gennaio
The logic form of the example 5 is:

    dare('Giuseppe',libro,'Mario')

FL = dare('Giuseppe', libro, 'Mario') .
5 ?- _
```

Semantica da un punto di vista linguistico

- **Semantica Lessicale**
 - Studio del significato delle parole individuali
- **Semantica Formale** (o Compositional Semantics, Sentential Semantics)
 - Come i significati individuali si compongono al fine di determinare il significato degli enunciati individuali
- **Discorso o Pragmatica**
 - Come i significati frasali si combinano tra loro e con altri fatti riguardo a diversi tipi di contesti in modo da comporre significati di un testo o di un discorso
 - Il Dialogo o la Conversazione spesso sono integrati nella interpretazione di un Discorso



Semantica Lessicale: Sinonimia

- Parole che hanno lo stesso significato in alcuni o tutti i contesti
 - *couch / sofa*
 - *big / large*
 - *automobile / car*
 - *vomit / throw up*
 - *Water / H₂O*
- Due lessemi sono sinonimi se possono essere sostituiti l'uno all'altro in tutte le situazioni
 - Hanno in questo modo lo stesso **significato proposizionale**

Semantica Lessicale: Sinonimia (2)

- Gli esempi di sinonimia perfetta sono pochi (o inesistenti)
 - Perché è così?
 - Anche se gli aspetti del significato si preservano potrebbero differire nella accettabilità legata al genere, allo slang, al tono o allo stile
- Example:
 - Water vs. H₂O
 - **Non diremmo mai:**
I like fresh H₂O after the tennis

Semantica Lessicale: Terminologia

- Lemmi e parole
 - Un **lessema** è la relazione stabile (astratta) tra significato e forma
 - Un **lemma** è la forma grammaticale usata per rappresentare un lessema.
 - *Carpet* è il lemma di *carpets*, *Dormir* è il lemma di *duermes*.
 - Le forme superficiali quali *carpets*, *duermes* sono le **parole flesse (wordforms)**
- Il lemma *bank* ha due **sen**si:
 - **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.**
- Un **senso** è la rappresentazione discreta di un aspetto del significato di una parola

Sinonimia come relazione tra i sensi piuttosto che tra *parole*

- Ad esempio: *big* and *large*
- Sono sinonimi?
 - *How big is that plane?*
 - *Would I be flying on a large or small plane?*
- E qui?
 - *Miss Nelson, for instance, became a kind of big sister to Benjamin.*
 - *?Miss Nelson, for instance, became a kind of large sister to Benjamin.*
- Infatti:
 - *big* ha un senso che significa *being older*, or *grown up*
 - *large* manca del tutto di un tale senso

WordNet

- Home page:
 - <http://wordnetweb.princeton.edu/perl/webwn>

WordNet Search - 3.1

- [WordNet home page](#) - [Glossary](#) - [Help](#)

Word to search for:

Display Options:

Key: "S:" = Show Synset (semantic) relations, "W:" = Show Word (lexical) relations

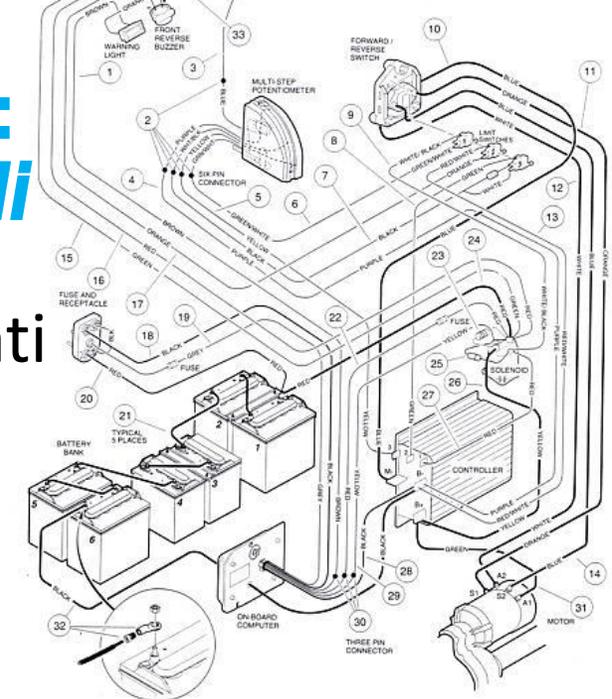
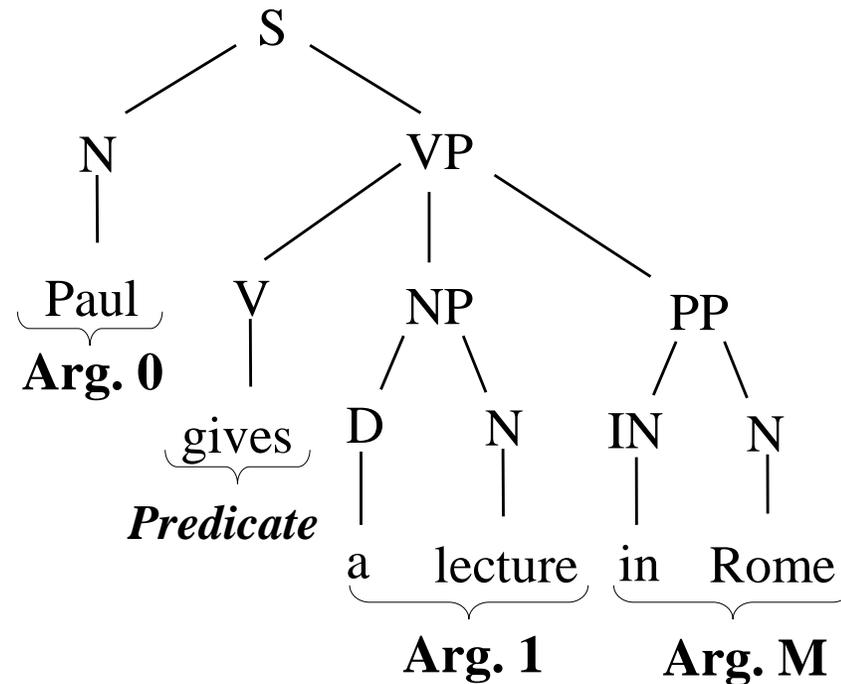
Display options for sense: (gloss) "an example sentence"

Noun

- **S: (n) meaning, [significance](#), [signification](#), [import](#)** (the message that is intended or expressed or signified) *"what is the meaning of this sentence"; "the significance of a red traffic light"; "the signification of Chinese characters"; "the import of his announcement was ambiguous"*
 - [direct hyponym](#) / [full hyponym](#)
 - [direct hypernym](#) / [inherited hypernym](#) / [sister term](#)
 - [derivationally related form](#)
- **S: (n) meaning, [substance](#)** (the idea that is intended) *"What is the meaning of this proverb?"*

Dalla semantic lessicale alla logica formale: comporre i significati dei costituenti lessicali

- Il *mapping* sintassi-semantica: assemblare significati

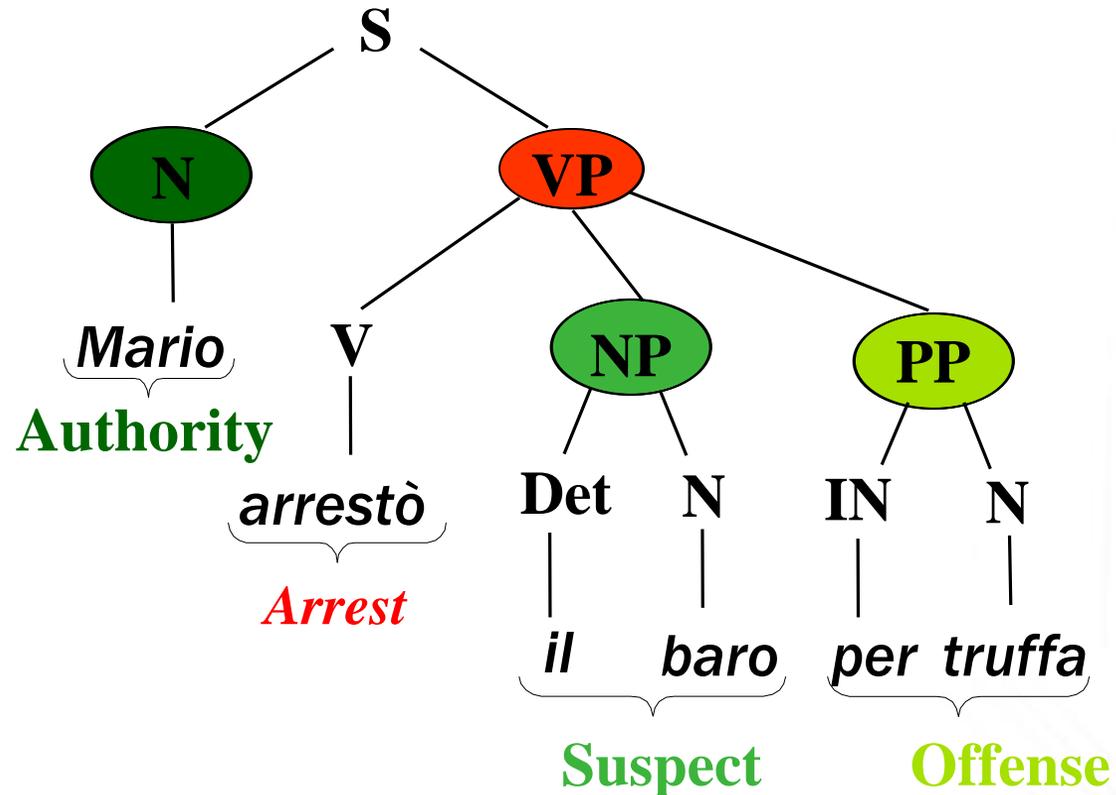


E' richiesta una teoria della predicazione e dei ruoli semantici di riferimento (Teorie *lessicali*: PropBank, Teorie *Cognitive*: **FrameNet**)

L'interfaccia sintassi-semantica: Frame Semantics

(Fillmore, 1975)

Mario arrestò il baro per truffa



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

Frame Semantics

- Per alcuni studiosi (i cosiddetti empiristi in semantica) le **parole rappresentano “classi di esperienza”** (cioè *situazioni*)
- Un **frame** è una struttura cognitiva (i.e. una sorta di prototipo) indicizzato dalle *parole* e usato a supporto dei processi di comprensione (Fillmore, 1975)
 - Le Lexical Units (LU) **evocano** un Frame in una frase
- La struttura del Frame comprende gli elementi che esprimono i partecipanti alla situazione (detti Frame Elements)
- Durante la comunicazione le LU evocano i frame necessari

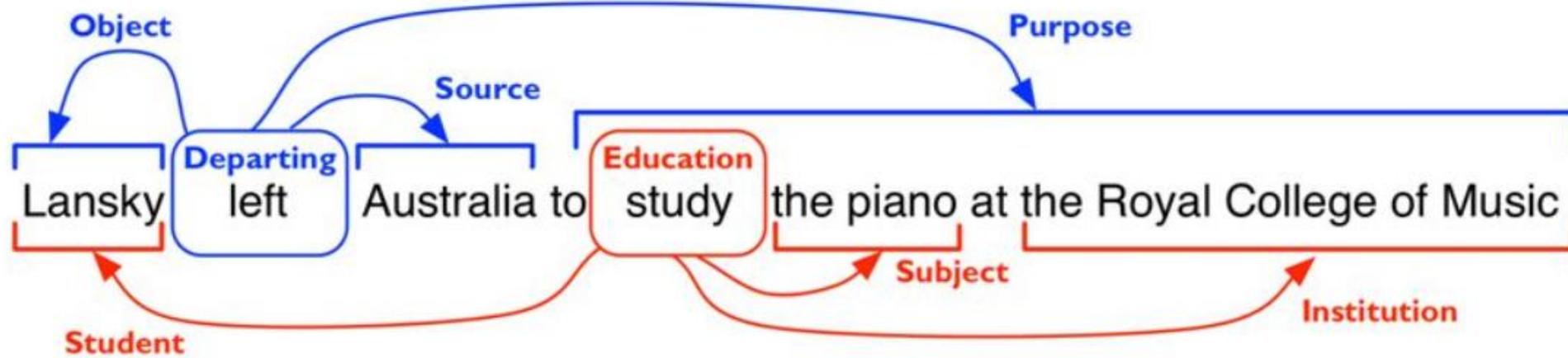
Frame Semantics: un esempio

Frame: KILLING	
A KILLER or CAUSE causes the death of the VICTIM.	
Frame Elements	KILLER John <u>drowned</u> Martha.
	VICTIM John <u>drowned</u> Martha .
	MEANS The flood <u>exterminated</u> the rats by cutting off access to food .
	CAUSE The rockslide <u>killed</u> nearly half of the climbers.
	INSTRUMENT It's difficult to <u>suicide</u> 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

- Una descrizione lessicale dovrebbe definire il frame indicizzato con i suoi frame element, insieme ai vincoli della sua *realizzazione* a livello sintattico:
 - *John bought a computer from Janice for 1000 \$*
- Ne scaturisce un sistema di mappature verso le posizioni gramaticali, cioè gli argomenti sintattici
 - Ad es. Il buyer occupa (in genere) la posizione del sogetto della frase
- Argomenti Obbligatori ed Argomenti opzionali
- Selectional preferences (Restrizioni selettive)
 - Il seller ed il buyer in genere appartengono alla categoria di “UMANI” o “GRUPPI SOCIALI”

Frames: Nuove strutture



Outline

- **Artificial Intelligence & Natural Language Processing**
 - Comunicazione linguistica & Conoscenza
 - Il ruolo dei dati
- **Natural Language Processing: *Task*, Modelli e Metodi**
- **Esempio: Computational Semantics in Prolog**
- **Trattamento delle lingue e *Machine Learning***
 - Statistical Language Processing
 - Apprendimento discriminativo per l’NLP
- **Natural Language Processing: applications**
- **Conclusions & Perspectives**



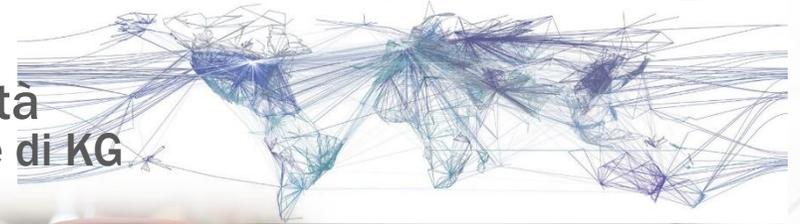
Applicazioni: Fenomeni Semantici di interesse



- **Entità.** Individui, luoghi, organizzazioni citate nei testi
 - *Indicizzazione semantica nei motori di ricerca, knowledge graphs (KG) di Google*



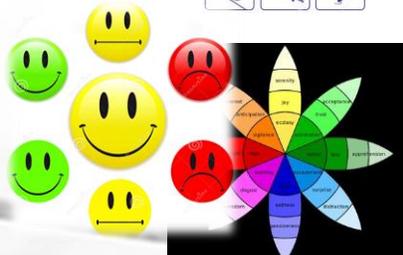
- **Relazioni.** Associazioni tra entità
 - Machine Reading: acquisizione di KG



- **Fatti.** Fenomeni o Eventi
 - Acquisizione di ontologie
 - Fact checking



- **Topic.** Argomenti di Discussione / Community / Niche Domains
 - Classificazione automatica di testi, notizie e mail



- **Tratti Emotivi e Psicologici.** Social Science, Profilazione
 - Sentiment Analysis, Recommending prodotti, CRM
 - Marketing in real-time, Brand Reputation Amanagement

Outline

- **Artificial Intelligence & Natural Language Processing**
 - Comunicazione linguistica & Conoscenza
 - Il ruolo dei dati
- **Natural Language Processing: *Task*, Modelli e Metodi**
-  **Esempio: Computational Semantics in Prolog**
- **Trattamento delle lingue e *Machine Learning***
 - Statistical Language Processing
 - Apprendimento discriminativo per l'NLP
- **Natural Language Processing: applications**
- **Conclusions & Perspectives**

Semantica & Calcolo

- Vedi slide su «[Computational Semantics in Prolog](#)»

References

- 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 & Schütze, MIT Press 2001.
- Sitografia:
 - SAG, Univ. Roma Tor Vergata: <http://sag.art.uniroma2.it/>
 - Reveal s.r.l.: <http://www.revealsrl.it/>

