

INTELLIGENZA ARTIFICIALE

*AGENTI BASATI SU CONOSCENZA
DALLA LOGICA ALLE LINGUE NATURALI:
SEMANTICA DELLE PRODUZIONI LINGUISTICHE E TESTUALI*

Intelligenza Artificiale (a.a. 2024-2025)

Roberto Basili

J&M book, AIMA R&N



- ▶ **Natural Language Processing: *Task, Modelli e Metodi***
- ▶ **Un esempio: linguaggio naturale e computational semantics**
- ▶ **Trattamento delle lingue e *Machine Learning***
 - ▶ Statistical Language Processing
 - ▶ Apprendimento discriminativo per l'NLP
- ▶ **Natural Language Processing: applications**
- ▶ **Conclusioni & Prospettive**

OUTLINE

AI & NLP: KNOWLEDGE ACQUISITION, AND DECISION-MAKING

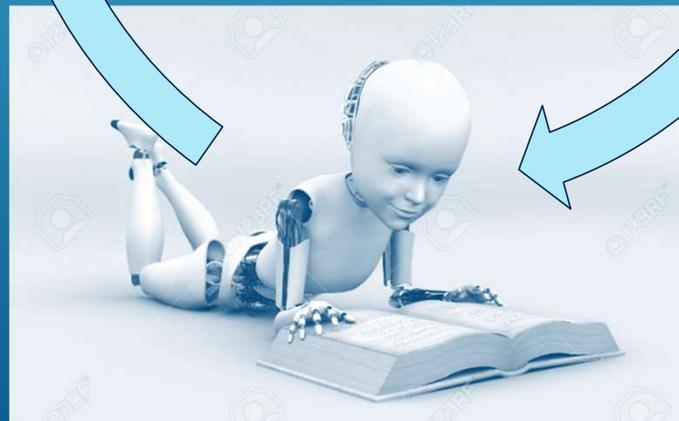
Knowledge



Inference



Learning to read
Machine Reading
Learning for Analytics



- ▶ **Natural Language Processing: *Task, Modelli e Metodi***



- ▶ **Un esempio: linguaggio naturale e computational semantics**

- ▶ **Tattamento delle lingue e *Machine Learning***

 - ▶ Statistical Language Processing

 - ▶ Apprendimento discriminativo per l'NLP

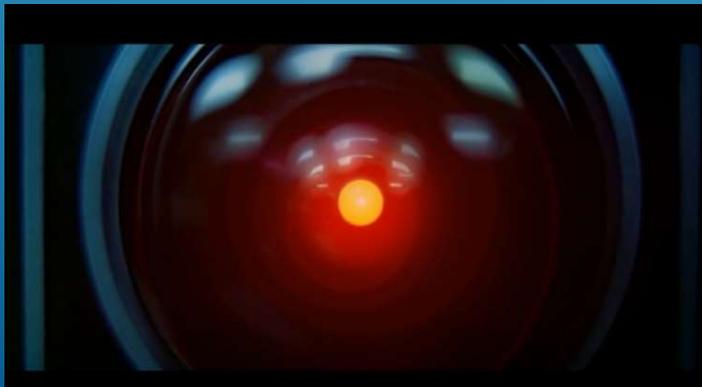
- ▶ **Natural Language Processing: applications**

- ▶ **Conclusioni & Prospettive**

OUTLINE

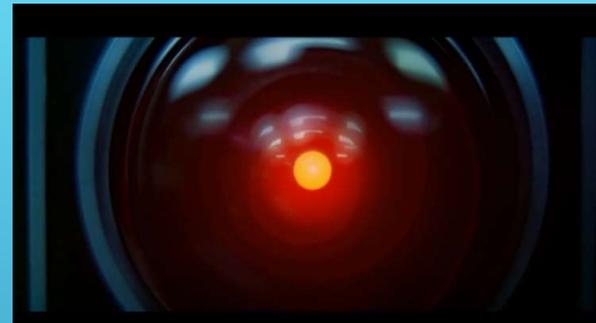
NLP: QUALI CONOSCENZE?

- ▶ 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.*



- ▶ Riconoscimento e Sintesi del linguaggio parlato

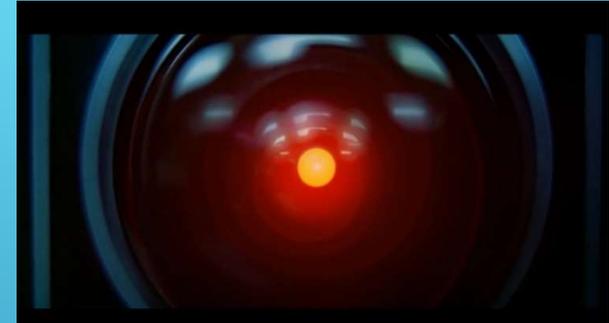
- ▶ Dizionari (spelling)
- ▶ Regole fonetiche (come i suoni vengono riconosciuti e prodotti)



- ▶ Comprensione

- ▶ Conoscenze Lessicali
 - ▶ Qual'è 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'È 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*?

QUAL'E' LA CONOSCENZA DI HAL?

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)
 - ▶ **Obbiettivi Operativi** (e.g. e-mail spam)
 - ▶ **Entità coinvolte**, ad esempio *persone* or *luoghi*
 - ▶ **Eventi potenziali** (e.g. fatti raccontati dal testo)
 - ▶ **Obbiettivi 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 nei motori di ricerca, la pianificazione, la 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 le conoscenze linguistiche ed i Modelli del Mondo nei diversi tipi di inferenza?

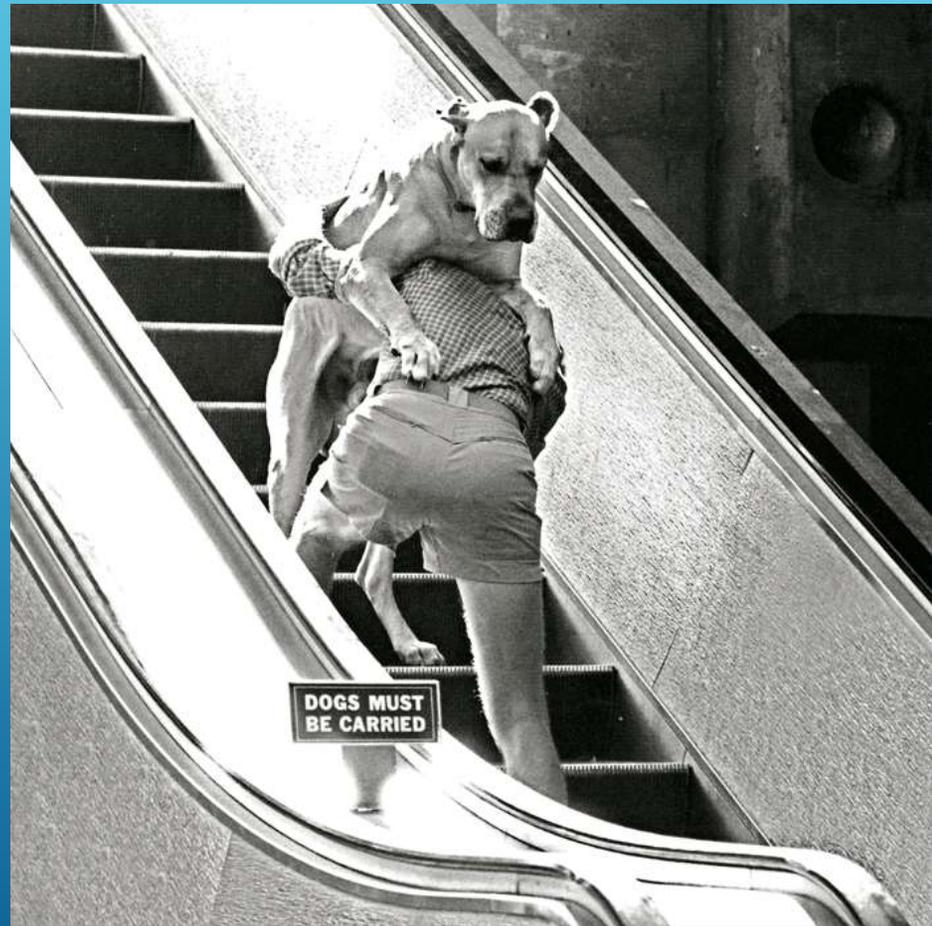
▶ FLESSIBILITÀ

- ▶ 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



La frase inglese

"Dogs must be carried on this escalator"

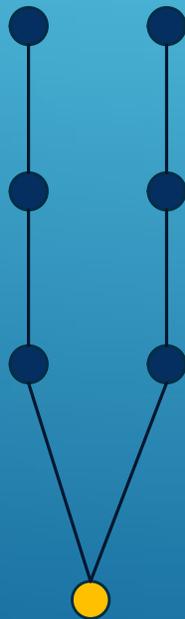
attiva le seguenti differenti possibili interpretazioni:

- ▶ *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*

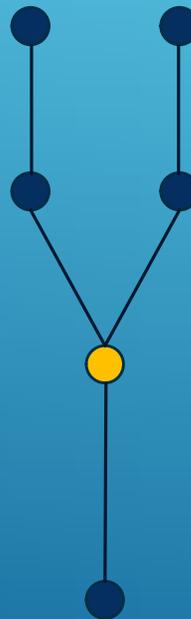
AMBIGUITÀ

LIVELLI DI AMBIGUITÀ

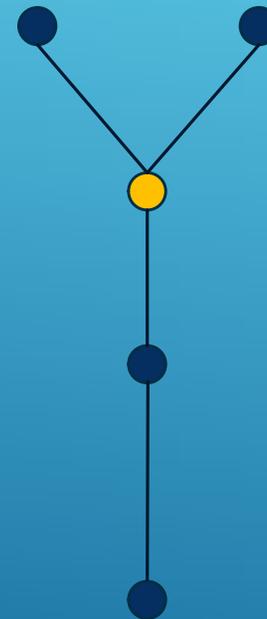
- ▶ SEMANTICA
- ▶ GRAMMATICALE
- ▶ MORFOLOGICA
- ▶ FONOLOGICA



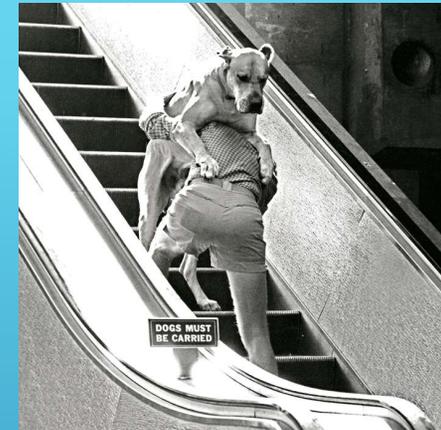
dei/dèi



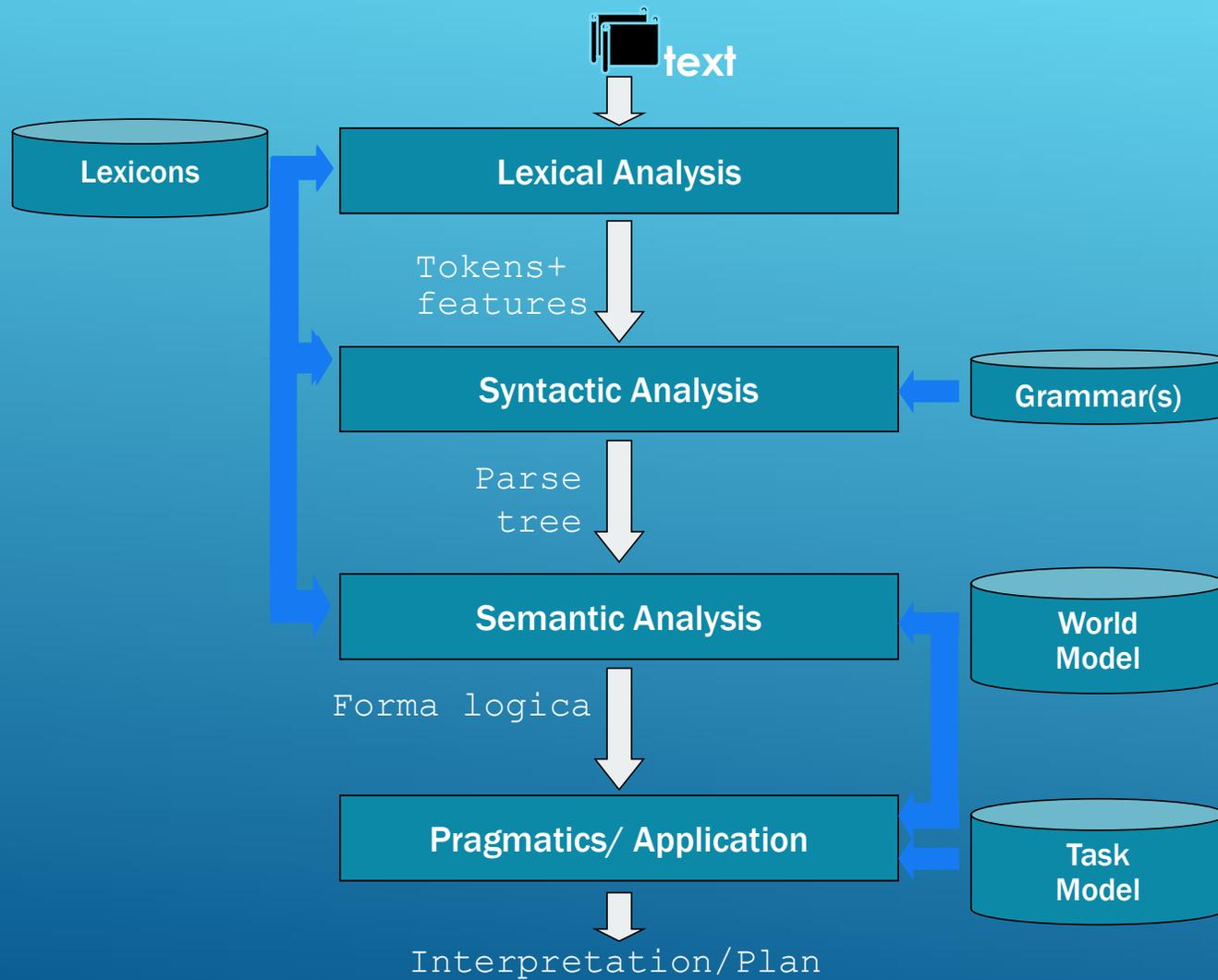
*compro la borsa
in pelle/in negozio*



il timore dei manager



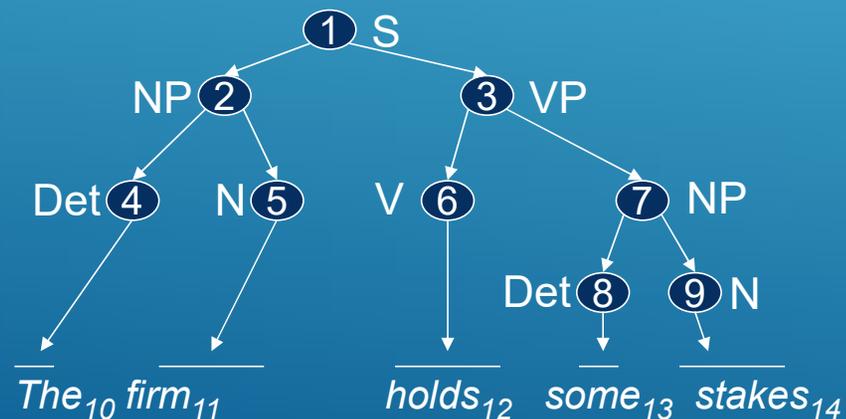
NLP: IL PROCESSO



SINTASSI: GRAMMATICHE A STRUTTURA SINTAGMATICA

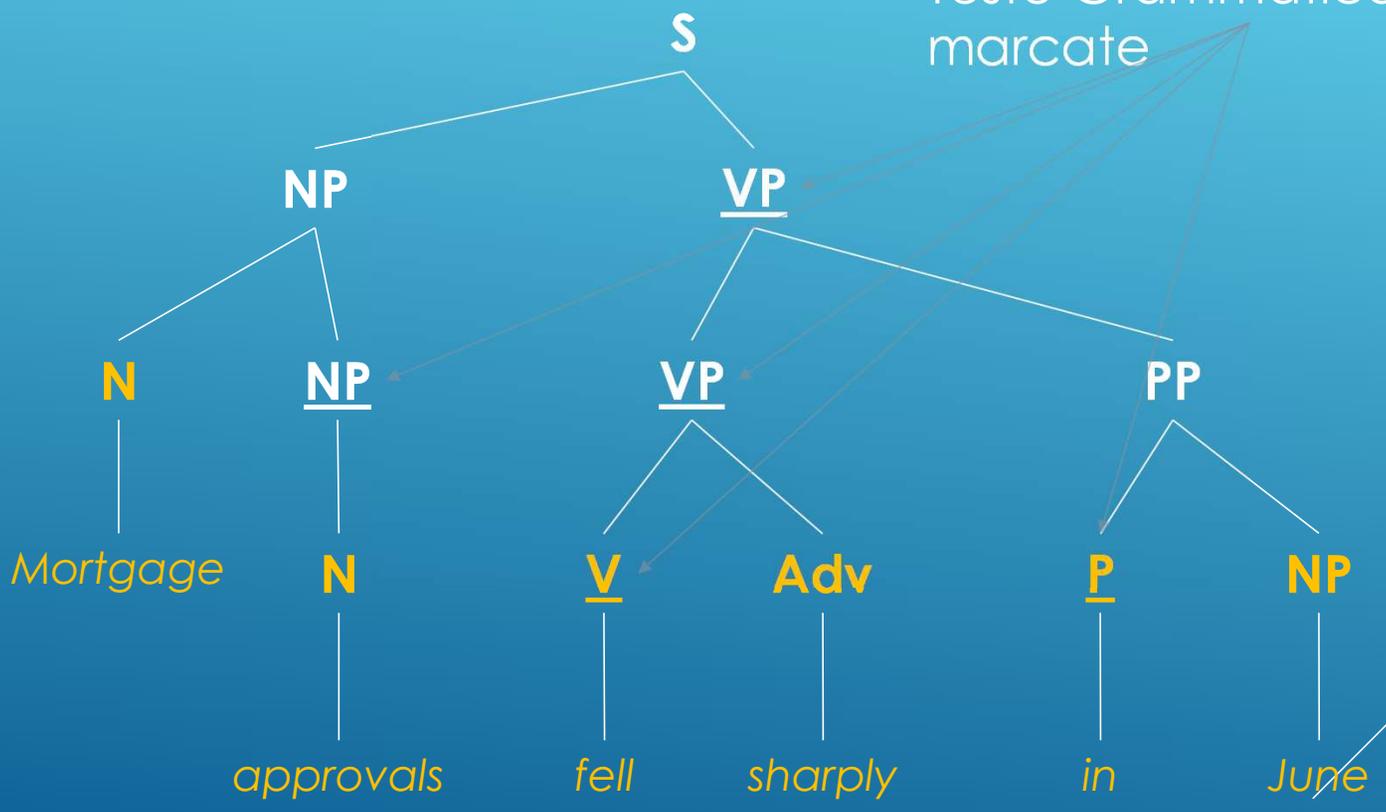
"The firm holds some stakes"

- ▶ $V_n = \{S, NP, VP, Det, N\}$, 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



SEMANTICA

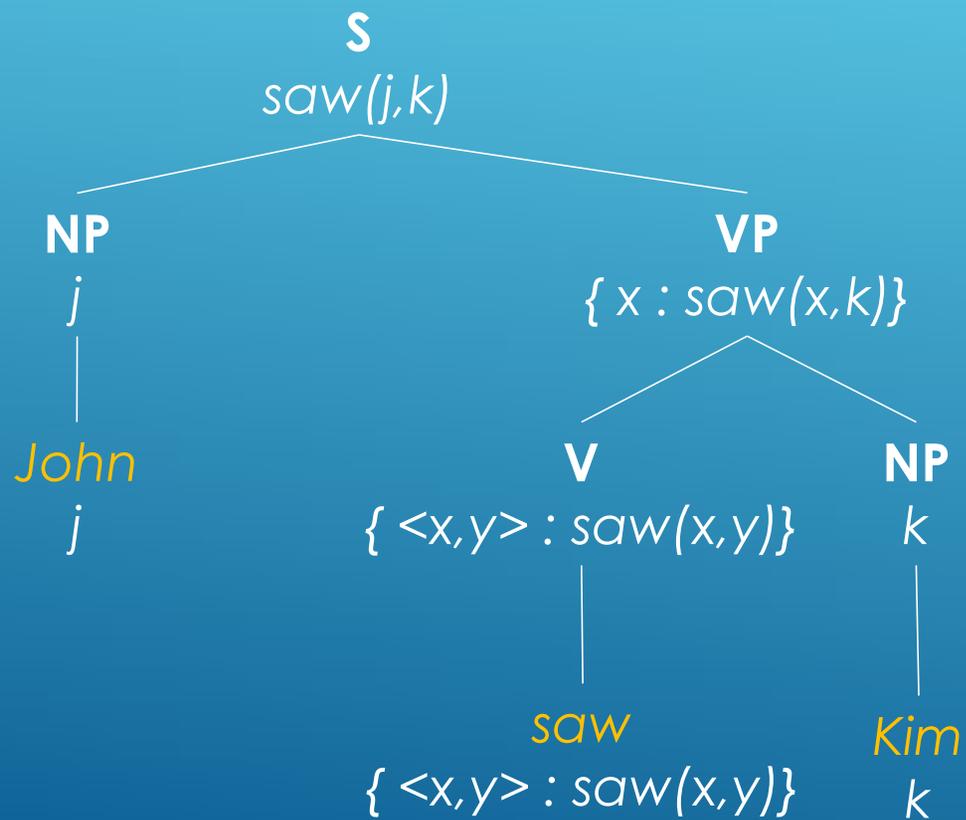
- ▶ Qual è il significato di una frase come

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 (parafraresi)
 - ▶ 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

```
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')

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

```
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')

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

```
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')

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

APPLICAZIONI: FENOMENI SEMANTICI DI INTERESSE



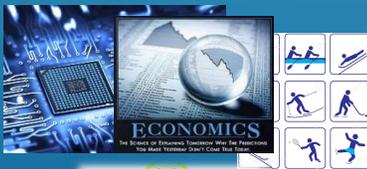
- ▶ **Popolamento Entità.** Individui, luoghi, organizzazioni citate nei testi
 - ▶ *Indicizzazione semantica nei motori di ricerca, knowledge graphs (KG) di Google*



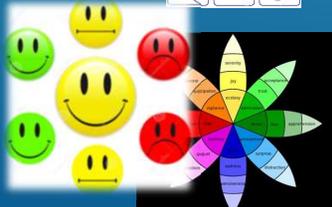
- ▶ **Relazioni.** Associazioni tra entità
 - ▶ Machine Reading: acquisizione di KG



- ▶ **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



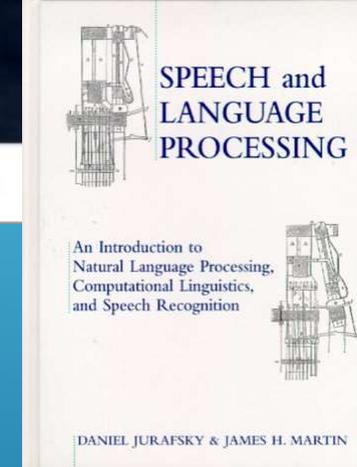
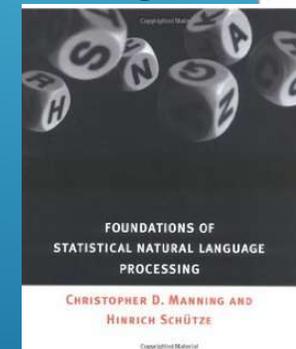
▶ AI & Robotics. «Robot Futures», Ilah Reza Nourbakhsh, MIT Press, 2013

▶ 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/>



REFERENCES

- ▶ **Natural Language Processing: *Task, Modelli e Metodi***
- ▶ **Un esempio: linguaggio naturale e computational semantics**



Tattamento delle lingue e *Machine Learning*

- ▶ Statistical Language Processing
- ▶ Apprendimento discriminativo per l'NLP
- ▶ **Natural Language Processing: applications**
- ▶ **Conclusioni & Prospettive**

OUTLINE