

## Ontologie lightweight e fondazionali

Casi di studio:  
SUMO e WordNet

---

---

---

---

---

---

---

## WordNet

- WordNet è una risorsa linguistica che consiste in un lessico organizzato secondo relazioni di significato
  - Progetto dell'Università di Princeton (<http://wordnet.princeton.edu>)
- Gli elementi del lessico (sostantivi, verbi, aggettivi, ecc.) sono raggruppati in insiemi denominati synset
  - Ogni synset corrisponde a un significato
  - L'idea è che i termini che fanno parte di uno stesso synset siano sinonimi

---

---

---

---

---

---

---

## Esempio

Il sostantivo inglese "road" appartiene a due synset diversi, uno letterale, come 'strada' e l'altro metaforico, come 'mezzo per'

- S: (n) road, route (an open way (generally public) for travel or transportation)
- S: (n) road (a way or means to achieve something) "the road to fame"

---

---

---

---

---

---

---

### Synset

- Nello stesso synset si trovano più termini.
  - Ad esempio, il primo dei due synset in cui si trova il sostantivo 'road' contiene anche 'route'
- Lo stesso sostantivo può essere collocato in posizioni del tutto diverse della rete.

```

    graph TD
      A["{implementation, effettuation}"] --> B["{means, agency, way}"]
      B --> C["{instrument, tool}"]
      B --> D["{road}"]
      B --> E["{...}"]
      F["{artifact, artefact}"] --> G["way"]
      G --> H["{road, route}"]
      G --> I["{path}"]
      G --> J["{access, approach}"]
      G --> K["{...}"]
    
```

---

---

---

---

---

---

---

---

### Altre relazioni

- Oltre alle relazioni tassonomiche (iperonimia e iponimia) in wordnet sono presenti altre relazioni semantiche:
  - Mereonimia (per i sostantivi, *window* is a part of *building*)
  - Implicazione (per i verbi, *snore* implica *sleep*)
  - Troponimia (per i verbi, *lisp* is a troponym of *talk*)

---

---

---

---

---

---

---

---

### WordNet come light weigth ontology

- Le relazioni verticali (iper- e iponimia) e orizzontali (mereonimia, implicazione, ecc.) tra i lessemi fanno di WordNet una 'rete di parole'.
- Word Net è definita anche un'ontologia "leggera", perchè ha la struttura di un'ontologia senza avere l'apparato formale proprio delle vere ontologie.

---

---

---

---

---

---

---

---

## Disambiguazione

- E' disponibile uno strumento per disambiguare il significato di un termine in base al contesto.
- Nella frase
  - “la finestra del programma” per esempio il synset della parola “finestra” è ben diverso da quello che essa ha nella frase “Giulia si affacciò dalla finestra”,
  - mentre “Giulia aprì la finestra” è ambigua tra i due significati.

---

---

---

---

---

---

---

## Sigma Browser

Il tool per navigare l'ontologia SUMO

---

---

---

---

---

---

---

## A cosa serve

- Sigma Browser è uno strumento per cercare concetti nell'ontologia SUMO usando come chiavi di ricerca i termini della lingua inglese
  - WordNet è un lessico elettronico in cui le parole sono raggruppate in insiemi (synset) che esprimono un certo concetto
  - Concetti (o synset) diversi sono collegati tra di loro da relazioni di specificità/generalità, quindi WordNet è più propriamente una tassonomia
- Il browser funziona perché ad ogni concetto di SUMO corrispondono uno o più termini di WordNet
  - Cercando in WordNet, si trovano tutti i significati della parola cercata e, per ogni significato, il concetto di SUMO corrispondente

---

---

---

---

---

---

---

## Interfaccia di ricerca

Sigma knowledge engineering environment [ Home | Graph | ]  
Browsing Interface

KR:  Language:

KB Term:  Show

English Word:  Noun Show

**Knowledge base statistics:**  
Total Terms Total Axioms Total Rules  
10537 54595 3869

Relations: 1121

Parola da cercare

---

---

---

---

---

---

---

---

## Risultati della ricerca

La parola cercata (*shout*) è presente in quattro accezioni diverse (o *synset*) in Wordnet. Ogni *synset* è identificato da un numero ed è accompagnato da una descrizione. Ad ogni *synset* corrisponde un diverso concetto di SUMO. Per esempio, *shout* inteso come *gridare sgarbatamente all'indirizzo di qualcuno* (20084529) corrisponde al concetto *Communication* in SUMO, mentre *gridare nel senso di parlare a voce molto alta* (200912473) corrisponde al concetto *Vocalizing* in SUMO.

Synset diversi possono corrispondere allo stesso concetto!

---

---

---

---

---

---

---

---

## Attenzione!

- Sinset diversi possono corrispondere allo stesso concetto.
  - Nell'esempio, le due accezioni di urlare come "parlare a voce molto alta" (200912473) e "fare un urlo" corrispondono entrambe al concetto *Vocalizing* in SUMO, cioè il loro significato corrisponde al concetto di 'vocalizzare'
- Ci sono due modi in cui un *synset* di WordNet può corrispondere a un concetto SUMO
  - Il concetto SUMO corrisponde esattamente al significato del *synset* (equivalent mapping)
  - Il concetto SUMO include al suo interno ('sussume') anche il significato del *synset* (subsumption mapping)

---

---

---

---

---

---

---

---

## Informazioni su un concetto

Cliccando su un concetto SUMO si arriva a una pagina in cui viene descritto il concetto. L'immagine può essere assente. A destra, compare l'elenco dei termini Wordnet associati al concetto.


La parentesi (Documentation ...) contiene la descrizione informale del concetto così come è data nella documentazione associata a SUMO.

Segue l'elenco dei concetti di rispetto a cui il concetto è qualche modo più specifico (subclass) o che sono più specifici rispetto a esso. Per esempio, vocalizing è 'subclass' di body motion e Radiating Sound.

KB Term:  Show

English Word:  Show

**Vocalizing** (vocalizing)



Tutti i concetti di SUMO e le relazioni che li collegano sono linkati alla loro definizione

appearance as argument number 1

93

Merge.kif 8787-8795

9322-9327 externalImage vocalizing

8782-8792 vocalizing is a subclass of body motion

8781-8781 vocalizing is a subclass of radiating sound

---

---

---

---

---

---

---

---

---

---

## Informazioni su un concetto (2)

(=>

(instance ?VOCAL Vocalizing)

(exists (HUMAN

(and

(instance ?HUMAN Human)

(instance ?VC VocalCords)

(part ?VC ?HUMAN)

(agent ?VOCAL ?HUMAN)

(instantiated ?VOCAL ?VC)))

Merge.kif 8787-8795

- if a process is an **instance of** **vocalizing**
- then there exists an agent so that agent is an **instance of** **human** and an object is an **instance of** **VocalCords** and agent is a **part of** agent and object is an **instrument** for process

Per capire meglio i concetti che rappresentano azioni sono molto utili le informazioni contenute nelle regole di ragionamento (assiomi) associate al concetto.

Per esempio, la regola qui sopra dice che se un processo appartiene al concetto (= *is an instance of*) di vocalizzare, allora ci dev'essere un agente umano che lo esegue e una parte dell'agente, le corde vocali, che serve come strumento per la vocalizzazione

---

---

---

---

---

---

---

---

---

---