

# Radicioni: keywords

Appunti di Federico Torrielli

## Intro, Ontologie, DOLCE

- **Ontologia:** base del sistema di rappresentazione della conoscenza, definisce un insieme di primitive con quali modellare un dominio. Cattura un range di distinzioni concettuali per un certo dominio.
- **Dolce:** ontologia fondazionale di eventi ed entità.
  - **Enduranti:** oggetti che non variano spazialmente ma possono cambiare nel tempo. Possono avere parti *non essenziali*.
  - **Perduranti:** eventi che non variano temporalmente. Gli *enduranti* partecipano ai *perduranti*.
  - **Subclass-of:** tipo di relazione di sottoclasse.
  - **Approccio moltiplicativo:** possiamo collocare nello stesso punto *spazio-tempo* due oggetti, anche se hanno caratteristiche essenziali diverse e perciò son diversi.
  - **Spazio concettuale:** spazio che contiene tutte le qualità (*qualia*) di una certa entità. Nella frase *la penna rosso scuro* ha uno spazio concettuale che è quello dei *colori* dove il rosa viene posizionato in questo spazio.
  - **Teoria della qualità:** non esistono due individui con la stessa identica qualità e non esiste la qualità senza l'oggetto a cui essa si riferisce (i colori non esistono senza oggetti a cui darli).
- **Strutture semantiche:**
  - **Grafi relazionali:** grafo con nodi ed archi. Limitato dall'espressività della struttura.
  - **Reti proposizionali:** grafo dove le proposizioni possono essere anche dei nodi, oltre che degli archi. Possiamo esprimere *De Morgan* e conoscenza logica ma non quantificatori universali
  - **Tassonomia:** struttura gerarchica. Si possono avere problemi di ereditarietà multipla che si risolvono non *shortest-path* oppure *dissonanza cognitiva*.
  - **Frame:** struttura che rappresenta le conoscenze generali riguardo situazioni, luoghi, oggetti. Un frame rappresenta la conoscenza in modo *dichiarativo* ma privo di **semantica formale**. Un frame possiede tre livelli gerarchici:
    - Livello base: oggetti ed entità
    - Livello super-ordinato: generalizzazione del livello base
    - Livello subordinato: specializzazione del livello base
  - **Prototipo:** elemento esemplare di una categoria

- **Contesto:** insieme degli elementi adiacenti ad una parola: può essere **sintattico**, **semantico** o **situazionale** (dove l'ultimo vuol dire che una parola è disambiguata dagli elementi linguistici attigui)
- **Teorie del significato:**
  - **Referenziale:** il significato fa riferimento a ciò che esiste veramente (sedia - → oggetto)
  - **Mentalista** o **Concettuale:** il significato è indiretto, mediato dall'immagine mentale che ci costruiamo dell'entità
  - **Strutturale:** il significato prende forma nel suo riferirsi ad un esempio specifico, in relazione alle altre parole della lingua (analoghi etc etc...). Qua si posiziona il **synset**.
  - **Prototipica:** il significato fa riferimento al prototipo di un oggetto od azione
  - **Distribuzionale:** il significato è determinato dall'insieme delle altre parole che co-occorrono.
- **Principio di composizionalità:** il significato degli enunciati si forma a partire dai singoli elementi sommati. Due approcci: [#composizionalità](#)
  - **Enumerazione dei sensi**
  - **Significato dinamico:** il significato della singola parola interagisce con quello delle parole adiacenti
- **Interazione Semantica:**
  - **Co-composizione:** il significato di un verbo è determinato dai suoi argomenti
  - **Forzatura del tipo semantico:** verbo + nome specifico spinge il verbo in una interpretazione semantica (*iniziare gli studi* ↔ *iniziare la cena*)
  - **Legamento selettivo:** aggettivo può selezionare una specifica porzione del significato del nome (*buon coltello*)

## WordNet

- **WordNet:** reference system lessicale on-line. Il suo compito è quello di associare dei possibili sensi ad un termine (granularità sul singolo term). WN è organizzato per relazioni semantiche, ovvero puntatori tra synset in un grafo. Esso contiene:
  - **nomi (nouns):** organizzati in **gerarchie**, partendo da **25 campi semantici distinti**, possiedono queste features (guardiamo la frase *il canarino è un uccello piccolo, colorato, che sa cantare e vola, con un piccolo becco ed ali*):
    - **attributi** (*piccolo, colorato*): modificatori del nome, espressi come aggettivi
    - **parti** (*becco, ali*): usano la **meronimia** o **olonimia**
    - **funzioni** (*canta, vola*): cosa serve, cosa fa.
  - **verbi:** organizzati in una pluralità di relazioni, partendo da **15 alberi inizializzatori**. Hanno un grado di polisemia maggiore, ed il suo significato viene completato *quasi sempre* dai suoi argomenti. Per i verbi come

relazione abbiamo anche la **troponimia**, ovvero se nel fare l'attività X facciamo anche l'attività Y (*mormorare, parlare*)

- **aggettivi e avverbi** organizzati in **iperspazi N-dimensionali**
- **Synset**: concetto specifico. Al singolo lemma corrisponde un certo *synset set*. Inoltre, un synset contiene una breve definizione ("gloss") e, nella maggior parte dei casi, una o più brevi frasi che illustrano l'uso dei membri del synset.
- **Form-Meaning pair**: coppia univoca.
  - **Word Meaning**: forma lemmatizzata. Se una parola ha più word meanings allora è una parola **polisemica**.
  - **Word Form**: come la parola può comparire in tutte le sue forme (*utterance* ed altro)
- **Relazioni**:
  - **Sinonimia**: due parole che sostituite in una frase non cambiano il suo valore di verità. Dato che sinonimi puri sono rari rilassiamo la relazione vincolandoli al contesto della frase.
  - **Antonimia**: opposto → ricco, povero.
  - **Iponomia/Iperonimia**: relazione gerarchica
  - **Meronimia**: relazione di tipo *part-whole*, transitiva e simmetrica.
- **WordNet, in più al dizionario ha**:
  - Specificazioni futili
  - Informazione sugli elementi coordinati (puntatori ai fratelli)
  - Puntatori alle sottoclassi (iponimia)

## FrameNet

- **FrameNet**: database lessicale basato sulla teoria della **semantica per frame** → il significato della parola può essere dato dal proprio *frame semantico* di riferimento. Il significato di cucinare è derivato dal Frame *Apply\_heat* (l'atto di applicare calore a qualcosa), dove il cucinare è semplicemente un suo **frame element**.
  - **Frame**: situazioni, oggetti ed eventi stereotipati. *Frame = Definition + FEs + LUs*. Un frame come Revenge potrebbe avere il seguente vocabolario:
    - Nomi: revenge, vengeance, reprisal...
    - Verbi: avenge, revenge, retaliate...
    - Aggettivi: vengeful, vindictive...
    - Verb-Noun phrases: take revenge, weak vengeance...
  - **Frame elements**: parti che compongono il frame e lo rendono completo nel contesto (per *Apply\_heat* i suoi FE sono: container (la pentola), cook (chi sta cucinando), food (cosa sta cucinando), heating instrument (dove lo sta cucinando)). Essi sono usati per **annotare** le parole delle frasi.
  - **Lexical Unit**: ciò che richiama il frame (per *Apply\_heat* degli esempi potrebbero essere *fry, boil, sautee*... ovvero tutti i singoli modi e richiami al cucinare) → coppia termine-senso che risolve la **polisemia** in FN.

# ConceptNet

- **ConceptNet**: missione è quella di contenere la conoscenza del senso comune delle cose per le KB delle macchine → tutta quella conoscenza che è **omessa** dalle comunicazioni sociali, nonostante ne sia la base. Nasce dal progetto di Open Mind Common Sense. Ogni nodo della rete non è perfettamente univoco e specificato, come in WN.
- **ConceptNet KB**: è **rivedibile**, ovvero, esattamente come la conoscenza e le cose umane, può essere cambiata in un secondo momento.
- **Utilizzo CN**: ragionamento di tipo contestuale →
  - **Contextual Neighborhoods**: misurare la relazione tramite path e densità dei nodi nel percorso
  - **Analogia**: due nodi sono analoghi se hanno due archi entranti sovrapposti
  - **Projection**: *Los Angeles è in California che è in USA che è nel mondo* (ognuno ha un arco esattamente uguale da uno all'altro)
  - **Disambiguazione**
  - **Analisi di affetti**: individuazione categorie emotive
- **Costruzione CN**:
  - **Estrazione**: 50-ina regole di estrazione applicate a OMCS con regex
  - **Rilassamento**:
    - Rimozione dei **duplicati**
    - **Lifting di conoscenza**: lifting delle informazioni a partire dalla relazione *isA* → *Se la banana è un frutto e il frutto è dolce allora la banana è dolce.*
    - **Generalizzazione tematiche e lessicali**: *buy food / purchase food* → *buy*
    - **Parsing delle proprietà**: si aggiungono ai noun le relazioni *PropertyOf* come *PropertyOf* → *('apple', 'red')*
    - **Riconciliazione discrepanze morfologiche**: *bici* viene collegata a *bicicletta*

## Cover

- **Cover**: risorsa linguistica sviluppata ad UniTO, dati estratti da ConceptNet, NASARI e BabelNet. In Cover abbiamo un **vettore** per concetto dove ogni entry del vettore ha un synset inerente in relazione alla entry
- **Vettore**: un esempio del vettore "fork" (*forchetta*), le relazioni derivano da *ConceptNet*

| Relations | Entries                |
|-----------|------------------------|
| isA       | tool, cutlery, utensil |

| Relations  | Entries                              |
|------------|--------------------------------------|
| atLocation | table, desk plate                    |
| usedFor    | eating                               |
| madeOf     | metal                                |
| relatedTo  | tool, food, utensil, cutlery, eating |

- **Costruzione della risorsa:**

1. **Input del sistema:** per ogni concept abbiamo il BabelSynset ed il suo NASARI vector
2. **Estrazione semantica:** per ogni elemento del concetto, con i suoi elementi del vettore NASARI, recuperiamo tutti i nodi di ConceptNet e le relazioni per ogni lessicalizzazione del synset in input.
3. **Rilevazione della rilevanza:** usiamo NASARI per capire quanto sia importante il singolo termine in relazione al concetto scelto
  1. se è nel vettore di NASARI
  2. oppure almeno  $\beta$  elementi intorno a  $t$  nel nodo CN sono dentro il vettore NASARI
4. **Identificazione del BabelSynsetID:**
  1. Se il termine l'abbiamo trovato grazie al suo vettore NASARI allora abbiamo già il suo BabelSynsetID
  2. Se invece il termine è stato trovato nell'intorno allora prendiamo tutti i possibili candidati NASARI e facciamo la similarità tra quelli ed il vettore NASARI in input, se supera una certa soglia, allora abbiamo trovato il senso corretto
5. **Vector Injection:** ogni concetto estratto (sotto forma di *BabelSynsetID*) viene iniettato nella corretta dimensione utilizzando la relazione di ConceptNet

## BabelNet

- **BabelNet:** allargamento di *WordNet* multilinguistico includendo risorse estratte con processi automatici da *Wikipedia*.
- **Struttura:** BN è un grafo diretto costruito da nodi (BNSynset, che contiene tutte le sue lessicalizzazioni, anche in altre lingue) ed archi (Relazioni semantiche del tipo *is-a*, *part-of*...)
- **Costruzione BN:**
  - **WordNet:** concetti come nodi e puntatori semantici tra synset come relazioni
  - **Wikipedia:** wikipages come nodi e hyperlinks come relazioni
  - Merge: si fa un merge di ciò che è uguale tra le due risorse

# NASARI

- **NASARI**: risorsa vettoriale che contiene le similarità tra termini.
- **Vettori**: i vettori vengono distinti in:
  - **Vettori lessicali**: le features sono *termini*
  - **Vettori unified**: le features sono costituite da *BabelSenseID*
- **Semantic similarity**: per costruire la similarità è stata utilizzata la **weighted overlap**
- **Weighted overlap**:

$$WO(v_1, v_2) = \frac{\sum_{q \in O} (rank(q, v_1) + rank(q, v_2))^{-1}}{\sum_{i=1}^{|O|} (2i)^{-1}}$$

- **Similarity**:

$$sim(w_1, w_2) = \max_{v_1 \in C_{w_1}, v_2 \in C_{w_2}} \sqrt{WO(v_1, v_2)}$$

## VerbNet

- **VerbNet**: rete di verbi che collega pattern sintattici ed aspetti di significato. Gli elementi sono rappresentati *gerarchicamente* ed in maniera *indipendente dal dominio*. Lessici sono mappati in altre risorse lessicali come WN, PropBank o FN. Fornisce un'associazione chiara tra livello sintattico e semantico
- **Classi Verbal**: insieme di verbi che mostrano un comportamento semantico simile e pattern morfo-sintattici simili. Si ispirano direttamente alle cosiddette **classi di Levin**. Ciascuna classe è descritta da:
  - **Ruolo semantico**: serve per precisare il comportamento e gli elementi partecipanti coinvolti nei vari tipi di azioni. Esempi sono: *Actor, Agent, Attribute, Beneficiary, Cause, Location...*
  - **Preferenze Selezionali**: ciascun verbo ha degli elementi che devono/possono/non devono caratterizzarlo (*mangiare* richiede *esser vivo, senziente, avere del cibo*). Funzionano come una tassonomia che permette di specificare preferenze e vincoli.
  - **Frame sintattici (funziona sinattica)**: consistono in una descrizione sintattico e una rappresentazione semantica con una struttura sub-eventuale (ci serve per differenziare *john left the room* (uscire) da *john left the book on the table* (lasciare)). Ogni frame include:
    - **Nome**: es. *Transitive*
    - **Sintassi**: es. *Agent V Patient: Paula HIT the ball*
    - **Predicati Semantici**: hanno una funzione temporale che può essere espressa tramite una variabile *E* che ci permette di specificare la sua struttura eventuale *start(E), during(E), end(E)*. es:  
*cause(Agent, E) ∧ manner(during(E), directedmotion, Agent) ∧ ...*

## Membri della classe verbale CUT:

| MEMBERS                                     |                           |                                      |
|---|---------------------------|--------------------------------------|
| CHIP (FN 1, 2; WN 1, 2, 5; G 1)             | HEW (WN 2)                | SCRAPE (FN 1, 2, 3; WN 1, 3, 6; G 1) |
| CHOP (FN 1; WN 1; G 1)                      | REAM (WN 2, 3)            | SCRATCH (FN 1; WN 2; G 1)            |
| CLIP (WN 1, 4; G 1)                         | RIP (FN 1, 2; WN 1; G 1)  | SLASH (WN 1, 3; G 1)                 |
| CUT (FN 1, 2, 3; WN 1, 24, 25, 30, 32; G 1) | SAW (WN 1; G 1)           | SLICE (FN 1, 2; WN 1; G 1)           |
| HACK (WN 1, 3; G 1)                         | SCARIFY (WN 1, 2, 3; G 1) | SNIP (WN 1, 2; G 1)                  |

## Classe verbale HIT, Overview completa:

This is not the parent class. For full class view click below

[Full Class View](#)

hit-18.1  
hit-18.1-1

Member Verb Lemmas:

|        |      |       |      |      |        |       |        |       |
|--------|------|-------|------|------|--------|-------|--------|-------|
| BATTER | BEAT | BUMP  | BUTT | DRUM | HAMMER | HIT   | JAB    | KICK  |
| KNOCK  | LASH | POUND | RAP  | SLAP | SMACK  | SMASH | STRIKE | STOMP |
| WHANG  | TAP  |       |      |      |        |       |        |       |

ROLES:

Instrument [ +body\_part | +refl ]

NP V PP

NP V PP PP

EXAMPLE:

Paul hit at the window.

[SHOW DEPENDENCY PARSE TREE](#)

SYNTAX:

Agent VERB (at) Patient

SEMANTICS:

→ CONTACT( e1 , Agent , Patient )

DO( e2 , Agent )

MANNER( e2 , Directedmotion , Agent )

FORCE DYNAMICS:

Volitional Attend FD representation

- **Informazioni della singola classe verbale:**

- **Sommario**
- **Membri della classe**
- **Proprietà** (lista delle possibili alternanze)

- **Alternanza:** possiamo alternare gli argomenti dei verbi nelle frasi per ottenere una semantica diversa (oppure no):

- **Alternanza locativa:** *Sharon sprayed water on the plants* ↔ *Sharon sprayed the plants with water*
- **Alternanza causale:** *Tony broke the window* ↔ *The window broke*
- **Alternanza middle:** *Tony broke the crystal vase* ↔ *Crystal vase break easily*

- **Levin Classes:** i verbi possono essere raggruppati in classi, ogni classe caratterizzata da un set di pattern sintattici. Ipotesi: **la sintassi riflette componenti semantici impliciti**. Le classi di Levin:

- **Non** sono **semanticamente omogenee**
- Non sono *completamente* sintatticamente omogenee
- Verbi possono appartenere a più classi

- **Struttura eventuale:** i verbi fanno implicitamente riferimento ad una struttura temporale che può essere divisa in tre:
  - **Preparatory Process:** verbi di tipo *run*
  - **Culmination:** verbi di tipo *hit*
  - **Consequent State:** verbi di tipo *break*

## PropBank

- **PropBank:** corpus annotato con proposizioni verbali ed argomenti.
- **RoleSet:** insieme di ruoli che corrispondono ad un utilizzo distinto di un verbo e che può essere associato ad un insieme di frame sintattici. L'idea è di far corrispondere la sua **firma sintattica** (RoleSet) ad un significato quando viene attivato.
- **FrameSet:** *RoleSet* a cui viene associato un frame (un verbo polisemico possiede diversi frameset). Due *FrameSet* son distinti se:
  - Criteri sintattico-semantici
  - Alternanze non preservano i significati di un verbo
  - Sono combinazioni verbo-particella (*set-in, set-off, go-on*)
- **Differenze con FrameNet:**
  - Costruzione di alternation simmetriche-asimmetriche intercambiabili, meno specifici che in FrameNet
  - Ci si concentra di più su i verbi piuttosto che su i noun e le loro connessioni
- **Esempio:** *John[arg.0] rings the bell[arg.1]* (ring.01) | *Tall aspen trees[arg.0] ring the lake[arg.1]* (ring.02) →

| ring.01 | Make sound of bell | ring.02 | To surround        |
|---------|--------------------|---------|--------------------|
| Arg0    | causer of ringing  | Arg0    | Sorrounding entity |
| Arg1    | thing rung         | Arg1    | Sorrounded entity  |
| Arg2    | ring for           | x       | x                  |

## FrameNet vs. VerbNet vs. PropBank

Dal più specifico al più generico:

- **FrameNet:** **Cynthia**(*ingestor*) **ate**(*predicate*) **the peach**(*ingestible*) with a **fork**(*instrument*).
- **VerbNet:** **Cynthia**(*agent*) **ate**(*predicate*) **the peach**(*patient*) with a **fork**(*instrument*).
- **PropBank:** **Cynthia**(*arg0*) **ate**(*predicate*) **the peach**(*arg1*) **with a fork**(*argm-manner*).



# Risorse distribuzionali

- Risorse basate sul concetto di **co-occorrenza** e **frequenza** di una parola in un certo contesto → risorse vettoriali → word embeddings

| Reti semantiche                           | Rappresentazione distribuzionale       |
|---|--|
| Alta specificità                          | Bassa specificità                      |
| Multi-lingua                              | Una sola lingua                        |
| Difficile comparare sensi                 | Facile comparare sensi                 |
| Difficile integrazione in app. downstream | Facile integrazione in app. downstream |

## Lesslex

- **Lesslex**: risorsa in cui le unità lessicali sono i synset ma la rappresentazione è *vettoriale*
- **Costruzione risorsa**: si accede a BabelNet, si ottengono i significati di un termine (*synset*), si estraggono i termini interessanti dai significati, si cerca il vettore per ognuno dei termini (con *ConceptNet NumberBatch*) e si ottiene una serie di vettori per ogni *senso* → Viene fatta poi la **media** dei vettori per cogliere il vettore del senso

## Task di similarity

- **Semantic similarity**: con vettori possiamo fare la cosine similarity →

$$\text{Similarity}(s_i, s_j) = \cos(\theta) = \frac{s_i \cdot s_j}{\|s_i\| \cdot \|s_j\|}$$

- **Ranked Semantic Similarity**:

$$\text{rank}_{sim}(t_1, t_2) = \max_{\vec{c}_i \in s(t_1), \vec{c}_j \in s(t_2)} [((1 - \alpha) \cdot (\text{rank}(\vec{c}_i) + \text{rank}(\vec{c}_j))^{-1}) + (\alpha \cdot \cos_{sim}(\vec{c}_i, \vec{c}_j))]$$

- **Sense Identification**: due termini sono comparati ed al sistema viene chiesto quale senso è quello connesso quando si seleziona la coppia di parole. (es. *<greeting, wave>* e *<weather, wave>*)
- **Max Similarity**: massimo della cosine similarity tra due termini
- **Neighbourhood similarity**: una volta trovato il vettore di un senso si fa clustering con vettori vicini e si considera il vettore centroide per calcolare la similarity

## Principio di composizionalità

- **Principio di composizionalità:** Il **principio di Frege** è un principio della semantica che stabilisce che il significato di un'espressione linguistica è determinato solo dal significato delle sue espressioni costituenti e dalle modalità con cui sono combinate. Per esempio, il significato della frase "Socrate è un uomo" è completamente dato dal significato delle parole "Socrate" e "uomo" e della connessione logica "è un". Il principio quindi asserisce che vale per le espressioni linguistiche quello che vale, per esempio, per le espressioni aritmetiche, in cui il valore di un'espressione dipende dal valore degli operandi e dal significato dell'operatore con cui sono combinati. La prima formulazione esplicita di questo principio è normalmente attribuita a *Gottlob Frege*. La vogliamo perché possiamo teoricamente elaborare infinite stringhe ma la mente ha capacità finite di memoria.
  - **Productivity argument:** Dato che esistono finite parole nel dizionario e finiti modi di combinarle la composizionalità ci rende funzionante questo principio. Vogliamo essere *produttivi* e la composizionalità è ottimale. Ha solo il problema che non riesce a distinguere cose come la parola *flat* in *flat note*, *flat tyre*, *flat beer*.
  - **Systematicity argument:** la composizionalità aiuta con la sistematicità del linguaggio, che aiuta nella comprensione del linguaggio stesso. Il problema è che il linguaggio spesso *non* è sistematico!
  - **Modularity argument:** individuiamo due tipi di modularità
    - **Modularità di Chomsky e Fodor:** il potere generativo di una grammatica può essere catturato da un solo modulo (conoscenze lessicali + regole sintattiche)
    - **Modularità rilassata:** basta un modulo per l'analisi sintattica e un modulo per significati e regole di combinazione
- **Composizionalità incrementale:** il significato di un'espressione complessa ad uno stage di processing è computata basandosi sulle espressioni costituenti a quello stage e sulla struttura sintattica costruita fino a quello stage. La composizionalità deve essere passata in maniera incrementale.
- **Composizionalità semplice:**
  - **I significati** di un'espressione elementare sono i soli vincoli nella *computazione* del senso di un'espressione complessa
  - **La sintassi** delle espressioni elementari sono i soli vincoli della *struttura* della computazione del senso di un'espressione complessa
  - **Conseguenze:** il significato dell'espressione non dovrebbe dipendere dal contesto in cui occorre + **principio inside-out** (l'analisi semantica procede *bottom-up*)
  - **Problema delle conseguenze:** esistono frasi **progressive** dove invece non è possibile che il significato non dipenda dal contesto (*contraddizione*) → La composizionalità semplice non funziona!
- **Principio del contesto:** espressioni elementari non hanno significato isolate, ma solo nel contesto di espressioni complesse. Questo principio è **opposto** a quello

della composizionalità → memorizzare le informazioni dato che il significato non è costruibile a priori dalla frase.

- **Perché la composizionalità non funziona?:**

- **Illusioni semantiche:** arca di Mosè
- **Misinterpretazioni:** frasi cortocircuitanti come "*when Anna dressed the baby played in the crib*"

- **Event-Related brain potentials (ERP):** risposte cerebrali a stimoli del linguaggio particolari:

- **N400:** impulso nel cervello riguardante il processing semantico di una frase (sia positivo che negativo)
- **P600:** impulso riguardante violazioni di vincoli sintattici

## Tabella Riassuntiva Risorse

| Nome       | Struttura  | A cosa serve   | Come viene costruita                        | Composizione Risorsa  |
|------------|--|--|---|---|
| WordNet    | Synsets: Nomi (organizzati in gerarchie) + verbi (in relazioni) + aggettivi e avverbi (iperspazi n-dimensionali) | word-sense disambiguation, information retrieval, automatic text classification, automatic text summarization, machine translation | A mano                                      | Form-meaning pairs, synsets collegati in un grafo attraverso relazioni di sinonimia, (ip-iper)onimia... |
| FrameNet   | Frames (Nomi, verbi, aggettivi e verb-noun phrases) + frame elements + lexical units                             | question answering, paraphrasing, recognizing, textual entailment, information extraction  | Esempi annotati + regole                    | Frame (situazioni): frame elements (azioni) + lexical units (richiami)                                  |
| ConceptNet | Grafo (Concetti come nodi, asserzioni di senso comune come archi)  | contextual-neighborhoods, analogia, projection, disambiguazione, analisi di affetti  | Estrazione da OMCS con regole, rilassamento | frasi nominali, frasi verbali, frasi aggettivi o clausole   |

| <b>Nome</b> | <b>Struttura</b>   | <b>A cosa serve</b>            | <b>Come viene costruita</b>   | <b>Composizione Risorsa</b>                   |
|-------------|--|--------------------------------|---|---|
| BabelNet    | BabelSynsets   | come WN ma multilingua         | WordNet (synsets) + Wikipedia (infoboxes) + altre risorse semantiche miste                    | Babelsynsets con glosse                       |
| Cover       | Vettore per ogni concetto di BabelNet collegato a BabelSynsets simili usando NASARI  | conceptual similarity ed altri | Concetto di BabelNet unito a simili con NASARI e iniettato attraverso relazioni di ConceptNet | Vettore (relazioni + entries)                 |
| Nasari      | Vettori lessicali (termini) + unified (bablesenseID)   | similarity                     | BabelSynsetID confrontati attraverso la weighted overlap similarity con altri synsets         | Vettore (nome;similarity points)              |
| VerbNet     | Classi verbali (nome + ruolo semantico + preferenze selettionali + frame sintattico). Frame sintattico (nome, sintassi, predicati semantici) | ...                            | Verbi di PropBank mappati sulle loro classi di Levin corrispondenti                           | Classi verbali contenenti frame sintattici    |
| PropBank    | Roleset + frameset   | ...                            | Machine Learning su verbi   | Proposizioni verbali organizzati in un corpus |

| Nome               | Struttura           | A cosa serve | Come viene costruita  | Composizione Risorsa |
|--------------------|---------------------|--------------|---|----------------------|
| LessLex            | Vettori per synsets | ...          | BabelNet synsets, si estraggono i vettori dai significati con ConceptNet<br>Cumberbatch si fa la media dei vettori per cogliere il vettore di senso | Vettore (synsets)    |
| Unified Verb Index | ...                 | ...          | PropBank + VerbNet + FrameNet   | ...                  |

## Domande Radicioni

- Inizia sempre da un'esercitazione. O a scelta sua o a scelta dello studente
- Spiegazione dell'esercitazione e tool usato (Wordnet, Framenet ecc.)
- Nel testo dell'esercitazione vi ho chiesto di produrre un output particolare, discutiamolo un momento, cosa c'è dentro a questo file?
- Wordnet: cos'è un synset? (e poi annessa discussione su Wordnet)
- Wordnet è uno strumento per una annotazione funzionale? Nel senso, per i task svolti nelle esercitazioni, non sarebbe meglio qualcos'altro?
- Come si usa Framenet in un contesto applicativo?
- Se io le dico "spazi concettuali" gli viene in mente qualcosa? Cosa sono? Dove li abbiamo visti? (lo studente non se lo ricordava ed è passato oltre) -- DOLCE
- Mi definisca in modo sintetico cos'è Verbnet.
- Partenza dall'esercitazione sulla summarization: Come valutare un riassunto?
- Cos'è Nasari?
- Nasari è una perfetta copia vettoriale di Babelnet? *No, Babelnet è più ricco, (iponimi, iperonimi...) è quasi un'ontologia, mentre Nasari è una rappresentazione degli oggetti.*
- Cos'ha BabelNet in più di Nasari? *I verbi.*
- Cos'è un enduring? Un perdurante?
- Com'è strutturato FrameNet? A cosa serve?

- Com'è strutturato VerbNet? A cosa serve?
  - Quali sono le differenze fra FrameNet, VerbNet e ConceptNet?
  - Che differenza c'è in Babelnet tra Named Entities e Concetti? (trabocchetto, in un nodo Babelnet possono essere presenti entrambi)
  - Mi parli della teoria dei prototipi.
  - Cosa le viene in mente se le dico "ruolo semantico" e "funzione sinattica"?  
*Verbnet*. Perfetto, mi parli di *Verbnet* allora.
  - *Verbnet*: cos'è l'alternation?
  - eventuali problemi dell'alta granularità di FrameNet. Differenze con quella di WordNet
  - Come fare disambiguazione con FrameNet
  - Cos'è l'approccio moltiplicativo -- DOLCE
  - Cos'è la semantica procedurale
  - Cosa sono gli spazi concettuali
  - Differenza tra teoria dei prototipi e teoria degli esemplari
-