**Esercizio di laboratorio n.1**

Trovare 4 concetti

Generico concreto: PERSON

Specifico concreto: BRICK

Generico astratto: EMOTION

Specifico astratto: REVENGE

Diamo delle definizioni dei termini

Person: A person is a human being

Brick: A brick is an object used in construction of buildings

Revenge: Revenge is a reaction in return of another one

Emotion: Emotion is a psychic state

A questo punto scarichiamo il csv con tutte le definizioni della classe

Dobbiamo fare un calcolo di similarità fatto come sovrapposizione lessicale tra le varie definizioni.

Prendiamo tutte le definizioni del concetto e proviamo a calcolare la sovrapposizione lessicale tra tutte le definizioni, è possibile che per un concetto questa intersezione sia vuota, un modo per farlo puo essere provare a estrarre le parole più frequenti e vediamo quelle parole quante volte ci sono in tutte le definizioni, e calcoliamo uno score.

Aggregare poi sulle due dimensioni concretezza e specificità

In genere c’è piu sovrapposizione su concetti concreti rispetto a concetti astratti

Il risultato finale sarà una tabella come quella delle slide con degli indicatori di sovrapposizione calcolata come vogliamo noi

I valori vengono bassi, perché la sovrapposizione lessicale sarà bassa, anche per concetti comuni,

c’è una sovrapposizione solo per pochi termini

human compare 29

being 16

ss 30

usiamo questo valore normalizzato facendo la media

poi aggreghiamo questa metrica per le dimensioni, concetti concreti, concetti astratti e poi specifici e poi…

Es 2

Abbiamo delle definizioni

Vogliamo fare una ricerca onomasiologica attraverso uno script

Per ogni concetto prendiamo tutte le definizioni e le usiamo come input del programma che deve cercare di trovare il termine di partenza usando wordnet

Es emotion: range of confort.., ecce cc

Le prendiamo come input e il programma deve andare su wordnet e trovare il sinset giusto

Tip: usare il genus, capire quali sono i genus, in genere sono iperonimi, lavoriamo sulle def, prendiamo il termine piu frequenti, tra quelli spesso c’è un genus o piu genus, tra quelli interroghiamo wordnet e ci facciamo dare solo i sinset associabili a quei genus e relativi iponimi