

Evalutation

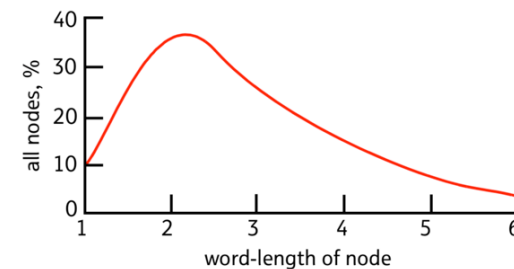
Valutare le performance di ConceptNet non è facile poiché, come abbiamo già accennato, definire il senso comune non è banale.

Un possibile approccio potrebbe essere quello di definire alcune condizioni minimali mediante le quali definire un “senso comune”, ma anche questo task non risulta immediato.

Quanto è evocativo un nodo?

Una delle domande che potremmo porci è quanto i concetti siano espressi in maniera semplice piuttosto che oscura, ossia, quanto un nodo è evocativo.

Possiamo usare la lunghezza delle locuzioni contenute nei nodi per quantificare questa feature (Figura)



Tendenzialmente, quanto più la locuzione di un nodo è corta tanto più il nodo sarà semplice. Approssimativamente il 70% delle locuzioni dei nodi ha una lunghezza minore o uguale alle tre parole

Osservando il grafico possiamo notare che la maggior parte delle locuzioni dei nodi sono corte, tendenzialmente più semplici di una qualsiasi frase composta (*verb-noun phrase-prepositional phrase*. e.g., *take dog for walk*)

Rapporto
fra
concetti
asseriti
ed
inferiti

Giacché la maggior parte delle asserzioni sono semplici potremmo anche chiederci quanto queste siano ripetute.

Per rispondere a questa domanda
calcoliamo la frequenza con cui le
asserzioni contenute in ConceptNet
sono presenti nel corpus OCMS:

Circa il 32% delle asserzioni di ConceptNet non so
mai state asserite in OCMS (i.e., sono stati inferiti)

Il 58% delle asserzioni è stato asserito solo una volta

Il restante 10% (ossia 160.000 statements) è stato asserito due o più volte.

Nota:

Avere il 32% di statement inferiti indica che ConceptNet dispone di un buon
sistema di estrazione e generazione delle informazioni (*come abbiamo visto
nella seconda sezione di questo capitolo "Costruzione di ConceptNet"*)