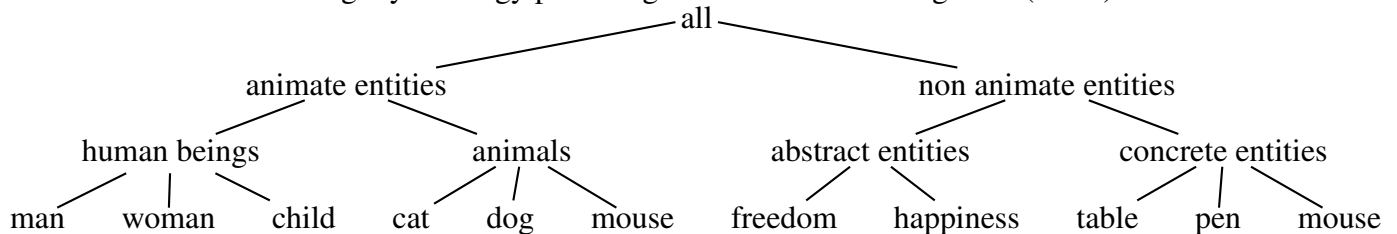


9 Lexical Semantics

Exercise IX.1

The objective of this question is to illustrate the use of a lexical semantics resource to compute lexical cohesion.

Consider the following toy ontology providing a semantic structuring for a (small) set of nouns:



- ① What is the semantic relation that has been used to build the ontology?

Cite another semantic relations that could also be useful for building lexical semantics resources.

For this semantic relation, give a short definition and a concrete example.

- ② The word "mouse" appears at two different places in the toy ontology. What does this mean? What specific problems does it raise when the ontology is used? How could such problems be solved? (just provide a sketch of explanation.)

- ③ Consider the following short text:

Cats are fighting dogs. There are plenty of pens on the table.

What pre-processing should be performed on this text to make it suitable for the use of the available ontology?

- ④ We want to use lexical cohesion to decide whether the provided text consists of one single topical segment corresponding to both sentences, or of two distinct topical segments, each corresponding to one of the sentences.

Let's define the lexical cohesion of any set of words (in canonical form) as the average lexical distance between all pairs of words present in the set³. The lexical distance between any two words is be defined as the length of a shortest path between the two words in the available ontology.

For example, "freedom" and "happiness" are at distance 2 (length, i.e. number of links, of the path: happiness → abstract entities → freedom), while "freedom" and "dog" are at distance 6 (length of the path: freedom → abstract entities → non animate entities → all → animate entities → animals → dog)

Compute the lexical distance between all the pairs of words present in the above text and in the provided ontology (there are 6 such pairs).

³Here, is actually the *lack of* cohesion that we measure: since it's a distance, the lower the more cohesion and the bigger the less cohesion.

- ⑤ Compute the lexical cohesion of each of the two sentences, and then the lexical cohesion of the whole text.

Based on the obtained values, what decision should be taken as far as the segmentation of the text into topical segments is concerned?

- ⑥ Give some examples of NLP tasks for which lexical cohesion might be useful. Explain why.