

# Dynamic Word-Sense Disambiguation with Semantic Similarity

Sam Zhang  
Computer Systems Lab 2009-2010

## Abstract

How do we assign meaning to words? This project investigates semantics from a lexical perspective, using the WordNet and OpenCyc ontologies to create a semiotic map of our consensus reality. Given a list of words, how can we find the word least like the others? Through a heuristical search across the hypernym ontology, computational semantics can discover the contextual meaning of words, even when the only context given is the other words from which it must differentiate itself. This method, which has not been given a name previously, will hitherto be known as dynamic word sense disambiguation.

## Background and Introduction

To create a similarity metric for semantics, this author uses lexical ontologies available publicly such as WordNet. The relationships most relevant in this scenario are hypernyms and hyponyms, which are when a word “is of the class of” another. See Fig 1.

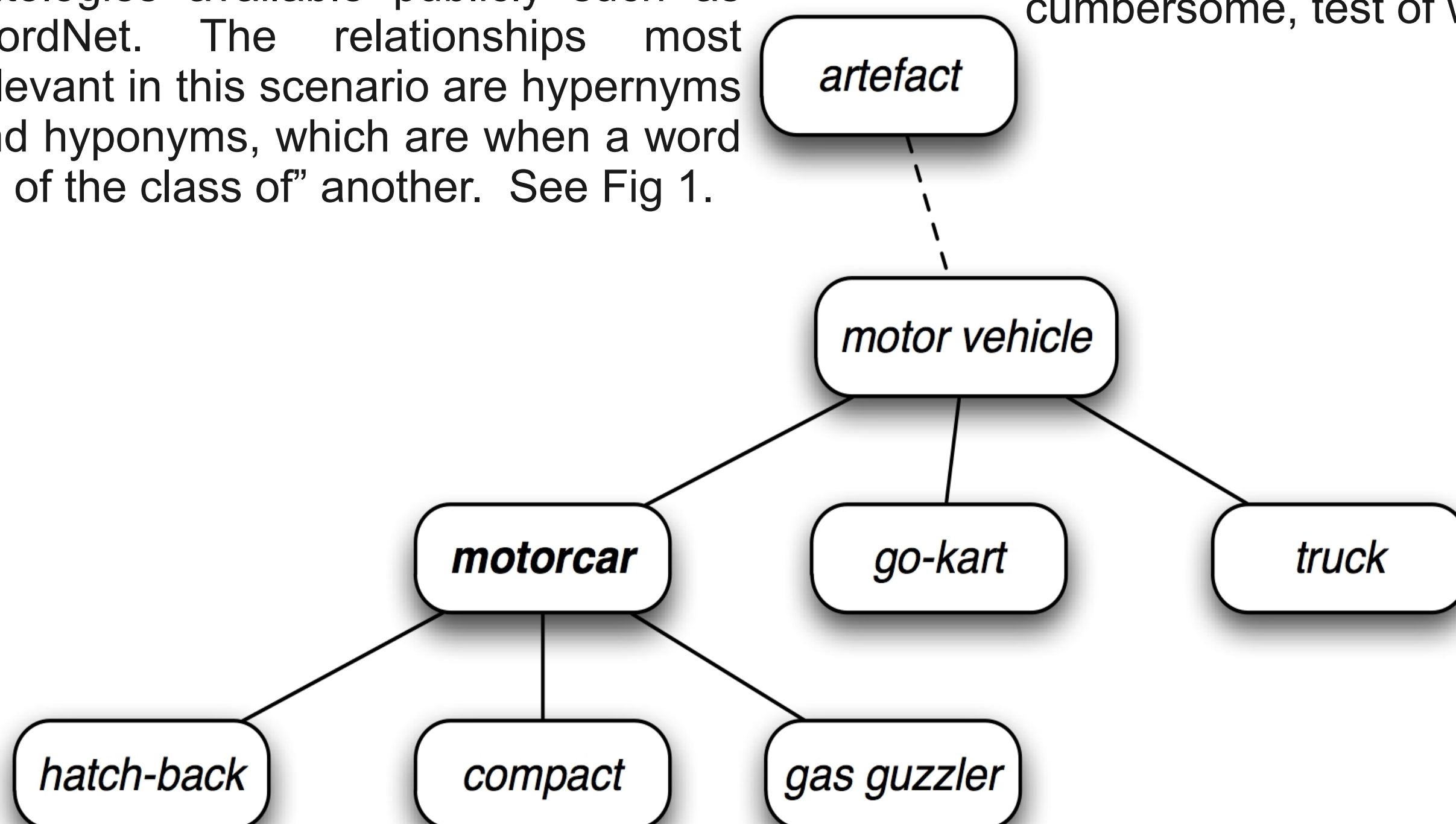


Fig 1: Holonyms and hypernyms  
Source: <http://nltk.googlecode.com>

## Discussion

To approach the issue, this project researched the feasibility of the semantic web to perform algorithmic searches across the lexical ontology, using the hypernym depth, least common hypernym, and distance to the root. The current working implementation is in Python, and will soon be transitioned to the Web Ontology Language to facilitate the transition of the Internet into the so-called Semantic Web, or Web 3.0.

## Results and Conclusions

The project successfully disambiguates word-sense around the 90% benchmark, failing most often due to human inconsistency as to the similarity of different words. Thus semantic similarity has been proved to be an accurate, if not cumbersome, test of word sense.