

WORD NET: A LEXICAL PERSPECTIVE OF ARRANGEMENT OF WORDS

Khatri Imran Idrishbhai
Research Scholer,
Department of Linguistics,
Gujarat University-Ahmedabad, India

Abstract:

Word Net is an online lexical reference system. It is an attempt to understand the human cognition with the help of computational help. Linguists developed this concept and created an online lexical reference system, which was first called psycho lexicography and later knows as Word Net. This paper is based on Word Net. This paper will give some ground level information of a concept.

The 20th century has seen the emergence of psycholinguistics, an interdisciplinary field of research concerned with the cognitive bases of linguistics competence. Bothe linguists and psycholinguists have explored in considerable depth the factor determining the contemporary structure of linguistic knowledge in particular. Miller and Johnson-Laird (1976) have proposed that research concerned with the lexical component of language should call psycho lexicography. They claim that words are connected with each other through lexical relations. – (Miller and Johnson-Laird, 1976).

To support the claim, linguists developed a concept and created an online lexical reference system, which was first called psycho lexicography and later known as Word Net. English Word Net uses different lexical relations like synonymy, meronymy, hyponymy, antonymy, polysemy and other part of speech.

What is Word Net?

“Word Net is an online lexical reference system, whose, design is inspired by current psycholinguistic theories of human lexical memory. English nouns, verbs, and adjectives are organized into synonym sets, each representing one underlying lexical concept. Different relations link the synonym sets.” (Miller, 1985)

In 1985 a group of psychologists and linguists at Princeton University undertook to develop a lexical database along lines suggested by these investigations – (Miller, 1985). The initial idea was to provide an aid to use in searching dictionaries conceptually, rather than merely alphabetically. As the work proceeded, however, it demanded a more ambitious formulation of its own principles and goals. Word Net is the result.

Inasmuch as it instantiates hypotheses based on result of psycholinguistic research, Word Net

Express, an International Journal of Multi Disciplinary Research

ISSN: 2348 – 2052 , Vol. 1, Issue 1, January 2014

Available at: www.express-journal.com

can be said to be a dictionary based on psycholinguistic principles. The most ambitious feature of

Word Net, however, is its attempt to organize lexical information in terms of word meanings, rather than word form.

Relational Semantics:

Relational semantics disambiguates the word by using another word, which is related through lexical relations. That means, if we put synonyms together they will help to disambiguate each other. In the same manner, other lexical relations like hypernymy, meronymy, antonymy etc. are also helpful to disambiguate the word.

Synonymy is the basic relation of Word Net. To create a good synonym set or synset, there are some basic principles, which should be kept in mind. The basic principle includes the following:

- (i) Principle of Minimality.
- (ii) Principle of Coverage.
- (iii) Principle of Replaceability.

(cf. http://www.cse.iitb.ac.in/~cs460_it632/lecture_notes/29jan07.pdf)

Let's have a brief glimpse of these principles.

(i) Principle of Minimality:

Principle of minimality suggests that to disambiguate the word by using minimum amount of its synonym word. That means if the word disambiguates itself by using only one synonym of a given word then we don't necessarily need another synonym.

Example: House 1: {house, building}

House 2: {house, family}

In above example the lexical item "HOUSE" holds to different senses. The synonym of House-1 defines lexical item "HOUSE" as a structure of residence. And the synonym of House-2 defines it as „a group of people who are related by kinship relations.“ In this way principle of minimality is very important in creating Word Net.

(ii) Principle of Coverage:

Principle of coverage suggests that, the synset should cover all the possible synonyms of a word. Here, it is also important to keep in mind that, a synonym word should convey same sense. An example makes it clearer:

Example: House 1: {house, family, home...}

The word "House" in the example given above is disambiguated by its synonym "family" and explicitly creates its meaning as „a group of people who are related by kinship relations.“ But the synonym word „home" also implies the same meaning, so it would also be included in synset, to satisfy the principle of coverage.

(iii) Principle of Replaceability:

According to the principle each and every synonym must be able to replace itself with another one, without changing the meaning. The sense conveyed by the synsets has to be same.

Example: House: {house, family, home ... building}

In the given example the word “House” means „a group of people who are related by kinship relations.“ And from the synset first three synonyms are able to replace each other. But the last one “building” does not able to replace the word “House”. Because of this problem the last synonym “building” should be removed from the synset of the “House” which means „a group of people who are related by kinship relations.“

Word Net versus Thesaurus and Standard Dictionary

The problem with an alphabetical thesaurus is redundant entries: if word W1 and word W2 are synonyms, the pair should be entered twice, once alphabetically under W1 and again alphabetized under W2. The problem with typical thesaurus is that two look-ups are required, first on alphabetical list and again in the thesaurus proper, thus doubling a user’s search time.

The most obvious difference between Word Net and standard dictionary is that Word Net divides the lexicon into five categories: Nouns, Verbs, Adjectives, Adverbs, and Function words. While the standard dictionary contains all the lexical items alphabetically. Actually, Word Net contains only nouns, verbs, adjectives, and adverbs. The relatively small set of English function words is omitted on assumption that they are syntactic component of language.

As it is discussed above, that Word Net contains nouns, verbs, adjectives, adverbs, it deals in different way with different categories of the words. Word Net uses lexical relations like Hyponymy, Meronymy, and Antonymy to distinguish the nouns, verbs in Word Net, distinguished by the help of the arguments between nouns and verbs. Lexical relations of Polysemy and Synonymy are helpful in distinguishing verbs. Adjectives distinguished by Antonymy and Synonymy relations. (Miller, 1993)

The gloss plays a very important role in the Word-Net since it is through this that the synsets are linked across Word Nets. The linguists emphasized that glosses in the Word Net explicate the synset senses, but cannot really be encyclopedic, scientific or legal definitions. In explicating the senses, they are assisted by the members of the synset and also the accompanying example sentence. Since the gloss is used for linking and creating the synsets it was decided that;

- The glosses will be short and simple,
- They will be expressed both in specific language and in English as well as the in IPA,
- In example sentences idiomatic and poetic expressions will be avoided.

References:

- Battacharya, P. (2007) Word Net.
Available at: http://www.cse.iitb.ac.in/~cs460_it632/lecture_notes/29jan07.pdf
- Miller, G.A., Beckwith, R., Fellaubum, C., Gross, D., Miller, K. (1993) „Introduction to WordNet: An On-line Lexical Database“.
Available at: <http://courses.media.mit.edu/2002fall/mas962.MAS962.miller.pdf>

Express, an International Journal of Multi Disciplinary Research

ISSN: 2348 – 2052 , Vol. 1, Issue 1, January 2014

Available at: www.express-journal.com

- Khatri, Imran (2007) „*Gujarati Law Net: Word Net of Law Terms*“ Unpublished Dissertation. The Maharaja Sayajirao University of Baroda, Vadodara.