# Word Network Using Lexical Database

## Ashok Kumar P[1], Vishva S[2], Gnana Murugan S[3], Dr.P. Murugeswari[4]

[1,2,3]Department of Information Technology, Karpagam College of Engineering, Coimbatore-641032,Tamilnadu, India

[4] Head of the Department, Department of Information Technology, Karpagam College of Engineering, Coimbatore-641032, Tamilnadu, India

### ABSTRACT

The Word Lexical Database is a vast collection of English words that categorizes nouns, verbs, adjectives, and adverbs into sets of cognitive synonyms. This database may seem similar to a thesaurus, as it also groups words based on their meanings, but there are key differences. Unlike a thesaurus that only links word forms, the Word Lexical Database links specific senses of words, thereby disambiguating words found in close proximity. Additionally, the Word Lexical Database annotates the semantic relationships between words, whereas a thesaurus does not follow any explicit pattern besides meaning similarity in the grouping of words.

## INTRODUCTION

Word Lexical Database is a large database of English Words. Nouns, verbs, adjectives and adverbs are grouped into sets of cognitive synonyms. Word Dictionary superficially resembles a thesaurus, in that it groups words together based on their meanings. However, there are some important distinctions. First, Word Lexical Database interlinks not just word forms—strings of letters—but specific senses of words. As a result, words that are found in close proximity to one another in the network are semantically disambiguated. Second, Word Lexical Database labels the semantic relations among words, whereas the groupings of words in a thesaurus does not follow any explicit pattern other than meaning similarity.

Word Network is a lexical database of sematic relations between word .Word Network links words into semantic relations including synonyms, hyponyms, and meronyms. The synonyms are grouped into synsets with short definitions and usage examples. Word Network can thus be seen as a combination and extension of Dictionary and thesaurus. While it is accessible to human users via a web browser, its primary use is in automatic text analysis and artificial intelligence applications. Word Network was first used for English Language. The database contains 155,327 words organized in 175,979 synsets for a total of 207,016 word-sense pairs; in compressed form, it is about 12 megabytes in size. Word Network includes the lexical categories nouns, verbs, adjectives and adverbs but ignores prepositions, determiners and other function words. Words from the same lexical category that are roughly synonymous are grouped into synsets. Synsets include simplex words as well as collocation like "eat out" and "car pool." The different senses of a polysemous word form are assigned to different synsets.

## LITERATURE SURVEY

## EXISTING SYSTEM

➤ NLTK (Natural Language Toolkit): NLTK is a popular Python library for natural language processing, and it includes a module for working with WordNet. The module provides functions for accessing WordNet's synsets, hyponyms, hypernyms, meronyms, and other semantic relationships.

➤ Open Multilingual WordNet: Open Multilingual WordNet is a project that aims to create a multilingual version of WordNet. It includes synsets and semantic relationships in multiple languages, and has been used in a variety of natural language processing tasks, such as machine translation and cross-lingual information retrieval.

Word Network Using Lexical Database can be widely used in the field of natural language processing. Over the years, many studies have been conducted on Lexical Database, examining its structure, usefulness, and limitations.

They examined the synsets and semantic relationships in Lexical Database and found that it was an effective tool for natural language processing tasks such as word sense disambiguation and information retrieval.

Several studies have also examined the effectiveness of WordNet in different domains. For example, researchers have investigated its usefulness in sentiment analysis, machine translation, and text classification. In a study by Liu and Li (2002), Lexical Database was found to be useful in improving the accuracy of text classification systems.

Other studies have focused on improving Lexical Database itself. For example, researchers have worked on expanding Lexical Database to include more words and improving the accuracy of its semantic relationships. There have also been efforts to create multilingual versions of Lexical Database, such as Lexical Database for European Language, which includes synsets and relationships in multiple languages.

While Lexical Database has been widely used and studied, it also has some limitations. For example, it is focused on English words and may not be as effective for other languages. It also has limitations in representing certain types of semantic relationships, such as those involving time and causation.

Overall, Lexical Database has been a valuable resource for natural language processing researchers and developers. Its synsets and semantic relationships have been used in a wide range of applications, and there have been efforts to improve and expand it over time.

## PROPOSED SYSTEM

Word Network using Lexical Database focused on improving the website's search functionality. I added more options to the search form, allowing users to search for synsets based on different attributes such as part of speech, definition, and examples. it Will show the Parts of Sppech , Definition , related Words and Sample Sentence for the given Word , Those result are obtained by the Existing Relationship from the Lexical Databse Data . it was able to handle a large amount of data. I also added some more visualizations of Word semantic relationships.

## ADVANTAGES OF PROPOSED SYSTEM

➤ Better Visualization of word results,

➤ Better Gloss / Definition,

➤ Grammatic Suggestions,

➤ Also include Sample Sentence, Parts of Speech ,etc.

## STRUCTURE AND DEFINITIONS

The Word Network, which uses a lexical database, is structured around synsets, which are collections of words that share a common meaning. Each synset is connected to other synsets through a network of semantic relations. For instance, the synset for "dog" is linked to the synset for "animal" through the hyponymy relation, since a dog is a specific type of animal. Likewise, the synset for "leg" is connected to the synset for "body" through the meronymy relation, because a leg is a component of a body.
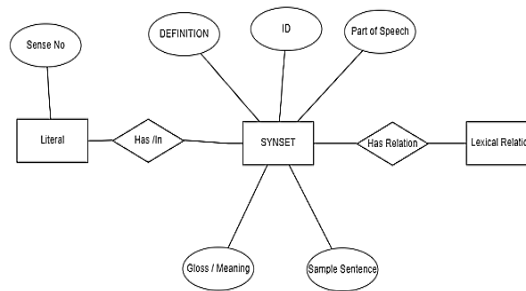


**Fig 1.Synset Block Diagram**

The image below depicts the basic structure of the Word Network. The fundamental concept underlying the relationship between words in the network is that of synonyms, such as sad and unhappy, or benefit and profit. These words share the same concept and can be used interchangeably in similar contexts.

These types of words are organized into synsets, which are unordered sets. Synsets are linked together if they share even small conceptual relationships. Each synset in the network has its own brief definition and many are illustrated with examples of how they can be used in a sentence. This emphasis on definitions and examples sets the Word Network apart from other networks.
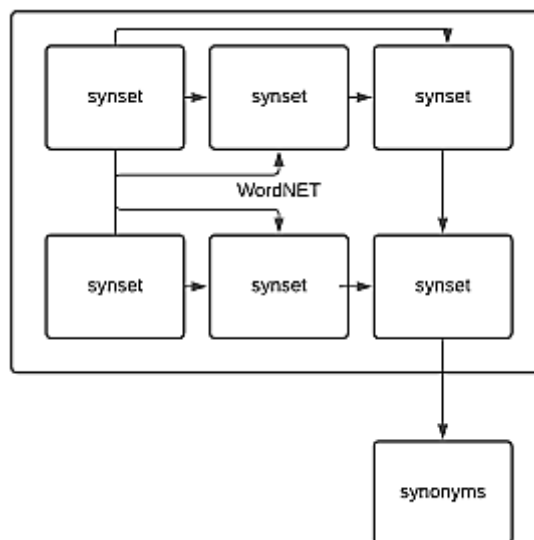


**Fig 2.Structure of Word Network**

## RELATIONS IN THE WORD NETWORK

**Hyponym**: Linguistically, a word with a broad meaning is considered a hypernym, which encompasses more specific meanings or categories of words, known as hyponyms. For instance, color is a hypernym

of red. Hyponymy refers to the relationship between a hypernym and a specific instance of a hyponym. A hyponym is a word or phrase whose semantic field is narrower and more specific than its hypernym. The semantic field of a hypernym is also known as a superordinate.
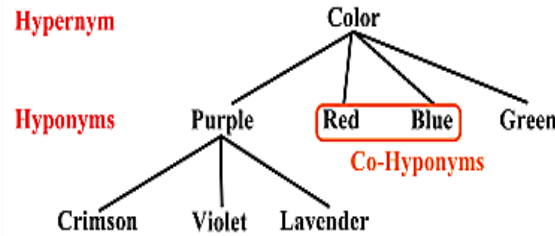


**Fig 3.Relation between Hypernym and Hyponyms**

**Meronymy**: The Word Network adheres to the meronymy relation, which establishes the relationship between a whole and its parts. For instance, a bike has parts such as two wheels, handle, and petrol tank, which are inherited by its subordinate categories. If a bike has two wheels, then a sports bike also has wheels. In linguistics, this type of relationship is utilized for adverbs that describe the characteristics of the noun. The parts are inherited downwards, as all bikes and their types have two wheels, whereas not all types of automobiles have two wheels.

**Toponymy**: Linguistically, toponymy refers to the existence of a "manner" relationship between two lexical items. Within the Word Network, verbs that describe events that necessarily and unidirectionally imply one another are connected, such as {buy}-{pay}, {succeed}-{try}, and {show}-{see}. The hierarchy of verbs towards the bottom showcases the manners that characterize the events, such as communication-talk-whisper.

**Antonymy**: Under the Word Network, adjective words are organized into antonymy pairs such as wet and dry, smile and cry. Each antonym pair is associated with sets of semantically related words. For example, cry is linked to words like weep, shed tears, sob, and wail, making them all indirect antonyms of a smile.

**Cross – PoS Relations**: The majority of relations within the Word Network pertain to the same part of speech. Based on these part-of-speech relations, the network can be divided into four subnets, one for each of the following: nouns, verbs, adjectives, and adverbs. Additionally, the network contains some cross-part-of-speech pointers, including morphosemantic links that connect words with the same meaning and shared stems. For instance, several pairs like (reader read) have been specified, in which the noun of the pair has a semantic layer with respect to the verb.

## WORKING AND USES

Word Network can be Implemented and Accessed using a Web Service. The website is Build using Java, JavaScript, Html and Rest.

Word Network using Lexical Database provide a User Interface Using HTML and CSS. It is allowing the User to enter the Keyword to Search and Keyword is prompt as input and the keyword is processed in backend the Processed
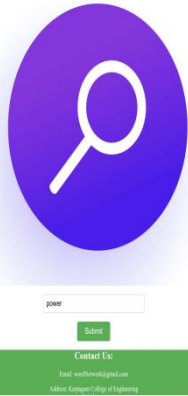
Search result is Displayed to the User using HTML and CSS in a meaningful way, thus is make the User to access or View the Search Result Efficiently by the User.

Word Network using Lexical Database works by providing a rich representation of the meaning of words, which can be used in a variety of natural language processing applications. For example, in information retrieval, Word Network using Lexical Database can be used to expand queries and improve search results by including related words and concepts. In machine translation, Word Network using Lexical Database can be used to disambiguate words with multiple meanings and to generate more accurate translations.

Overall, Word Network using Lexical Database is a powerful tool for English language processing that can be used in a wide range of applications. Its network of semantic relations and synsets provides a rich representation of the meaning of words, which can be used to improve the accuracy and effectiveness of natural language processing algorithms. One of the most common uses of Word Network using Lexical Database is in information retrieval, where it is used to expand queries and improve search results.

## RESULT



## CONCLUSION

The use of Word Network in information systems has been diverse, serving various purposes such as word-sense disambiguation, information retrieval, automatic text classification, automatic text summarization, machine translation, and even automatic crossword puzzle generation. Unlike traditional dictionaries that only provide the definition of a word and thesauruses that only provide synonyms and antonyms, the Word Network project offers a combination of both dictionary and thesaurus results in its

search output. Along with synonyms and antonyms, Word Network displays similar words to the input word and a sentence using the search word to help understand its meaning.

## REFERENCES

1. Abd-Elwasaa, Ahmed. 2016. "WORDNET: A Database of Lexical Relations." SlideShare, January 2. Accessed 2020-08-02.

2. Camacho-Collados, Jose, and Mohammad Taher Pilehvar. 2018. "From Word to Sense Embeddings: A Survey on Vector Representations of Meaning." arXiv, v3, October 26. Accessed 2020-08-03.

3. Collins, Allan M. and M. Ross Quillian. 1969. "Retrieval Time from Semantic Memory." J. Verbal Learning and Verbal Behavior, vol. 8, no. 2, pp. 240-247, April. Accessed 2020-08-03.

4. Educative. 2020. "How to use WordNet in Python." Educative, Inc. Accessed 2020-08-03.

5. Fellbaum, C. 2006. "WordNet(s)." In: Keith Brown (ed), Encyclopedia of Language & Linguistics, Second Edition, vol. 13, pp. 665-670. Oxford: Elsevier. Accessed 2020-08-02.

6. Fellbaum, C. 2012. "WordNet." First Interdisciplinary Summer School on Ontological Analysis, Trento, Italy, July 16-20. Accessed 2020-08-03.

7. Fellbaum, C. and G. A. Miller. 2006. "Whither WordNet?" Presentation, LREC. Accessed 2020-08-03.

8. Global WordNet Association. 2020. "English WordNet: Releases." GitHub, April 17. Accessed 2020-08-03.

9. Jurafsky, Daniel and James H. Martin. 2019. "WordNet: Word Relations, Senses, and Disambiguation." Chapter C in: Speech and Language Processing, Third Edition draft, October 16. Accessed 2020-08-03.

10. Khoi, Nguyen. 2012. "WordNet Introduction." SlideShare, October 31. Accessed 2020-08-02.

11. Miller, George A. 1995. "WordNet: A LexicalDatabase for English." Comm. of the ACM, vol. 38, no. 11, pp. 39-41, November. Accessed 2020-08-02.

12. Miller, George A., Richard Beckwith, Christiane Fellbaum, Derek Gross, and Katherine Miller. 1993. "Introduction to WordNet." August. Accessed 2020-08-02.