# An Innovative Taxonomy-Driven Approach for Word Sense Disambiguation Using Conceptual Relatedness

Djamel Nessah[1*], Abdelaali Bekhouche[1], Toufik Marir[2]

[1] ICOSI Laboratory, University of Khenchela, Khenchela 40002, Algeria
[2] RELA(CS)[2] Laboratory, University of Oum El Bouaghi, Oum El Bouaghi 04000, Algeria

Corresponding Author Email: nessah_djamel@univ-khenchela.dz

**ABSTRACT**

Humans share information through natural languages, where ambiguity is an inherent feature, allowing words that can be understood in various ways, often determined by the context in which they are used. Many words in languages around the world have different senses, and consequently, language resources and hierarchical organizations like WordNet highlight the notion of semantic proximity. This work presents a knowledge-based Word Sense Disambiguation approach that leverages the WordNet dictionary and principles of computing semantic similarities between concepts to facilitate this process. The model aims to contribute novel insights into the extraction of word senses, the delimitation of the context in which the ambiguous word occurs, and the proposition of an algorithm based on vector representations to compute similarities. To assess the most suitable senses of ambiguous words, an algorithm is proposed that combines the use of Wu-Palmer and Leacock-Chodorow semantic similarities on vector representations to analyze semantic tendencies and define the sense that offers the highest score. The goal is to provide a degree of certainty regarding the probable semantics of the sentence. The experimental results support the development of a semantic relationship between the occurrence context of the ambiguous word and the structure of the WordNet taxonomy. Therefore, the proposed method is expected to have a positive impact on work dealing with word semantics in various domains of artificial intelligence.

## 1. INTRODUCTION

Since the late 1940s, researchers have been aware of the key features of Word Sense Disambiguation (WSD), such as the context surrounding a target word and the use of statistical information on words, senses, and knowledge resources. In 1949, Zipf proposed the 'Law of Meaning,' stating a power-law relationship between more and less frequent words; it quickly became evident that Word Sense Disambiguation (WSD) posed a significant challenge [1]. Further works focus on their attempt to structurally encode the dependencies between a word's sense and its context [2, 3] whereas others combine some of other known models [4].

The circumstance is still getting complex, as recently the web has evolved into a vast source of diverse and dynamic information, which requires the utilization of robust tools for the online sharing of accurate and relevant content [5, 6].

In fact, there is a real lack of efficacy and robust systems to carry out these tasks. Thus, these findings have prompted researchers to seek suitable solutions, and the first ones were devoted to term's disambiguation as a central axis to achieve the desired goals.

Currently, various models are widely used for the purpose of WSD. However, there is a general hypothesis that purely syntactic representations of words possess a "psychological reality," wherein they are semantically interpreted within a specific context [7].

In this context, this paper describes a model that utilizes a knowledge corpus (WordNet ontology) and an appropriate semantic similarity measure as an attempt to formalize/mechanize two remarkable human capabilities: First, the ability to make rational decisions in an environment of ambiguous information by discerning accurate meanings of words relative to the context. Second, the capacity to undertake a diverse range of cognitive tasks that demand concentration, comprehension, adaptation, attention, and thoroughness [8].

This research manuscript discusses on how to disambiguate the term in a sentence by referring to the WordNet repository and by deploying the use of context-based similarity models [9, 10].

Following the introduction the background is presented, encompassing a general description of the WordNet taxonomy, along with various similarity measures commonly acknowledged and utilized in the literature.

The next section is largely dedicated to the diverse research works carried out in this perspective. Then we describe in section three, our approach which falls under the category of knowledge-based approaches. Last, a conclusion is provided with prospects for potential future enhancements.

## 2. STAT OF THE ART

### 2.1 WordNet taxonomy

WordNet is a comprehensive electronic lexical database for the English language; developed in 1986 at Princeton University, this database connects English nouns, verbs, adjectives, and adverbs to synonyms, further interlinked through semantic relationships that establish word definitions [11, 12].

Miller [11], an esteemed psycholinguist, drew inspiration from artificial intelligence experiments designed to explore the intricacies of human semantic memory [13].

Considering the vast amount of knowledge that speakers possess about tens of thousands of words and the concepts they represent; it appeared reasonable to assume the existence of efficient and economical storage and access mechanisms for words and concepts [14]. The Collins and Quillian model presented a hierarchical organization of concepts, allowing more specific concepts to inherit information from their more general, superordinate ones. Therefore, WordNet provides two services.

A vocabulary defines the various meanings of words.

An ontology outlines their semantic relationships.
WordNet can be regarded as an ontology for natural language [15, 16], where nouns, verbs, adjectives, and adverbs are structured into networks of synonym sets, referred to as synsets. Each synset encapsulates a fundamental lexical concept and is linked through various relationships (Figure 1).
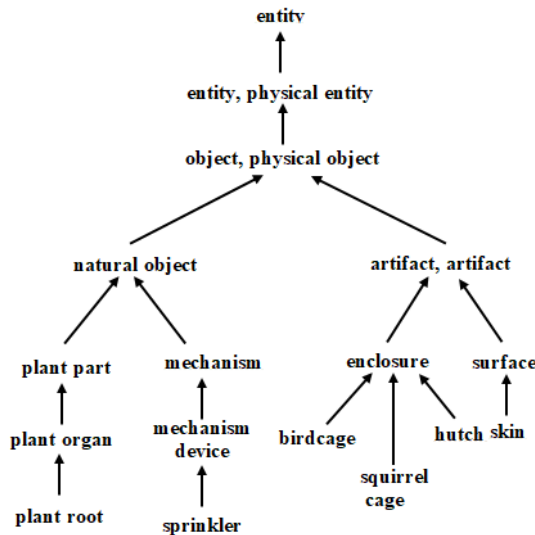


**Figure 1.** IS-A relationships in WordNet

A polysemous word will appear in one synset for each of its senses; WordNet is organized not only by the 'is-the-synonym-of' relationship. Verbs and nouns are hierarchically organized through hypernym/hyponym relation (superior / inferior), and meronym/holonym (i.e., part-of relationships).

### 2.2 Similarity measures

Similarity measures regarding their quality of services became promising in substantial development of applications dealing with semantic web issues.

In this section, is related a review of various approaches to assess similarity measurements. Foremost works initiated by researchers tend to modernize the measurements and make them more efficient and more adequate. Commonly semantic similarities are classified into two main categories:

#### 2.2.1 Edge counting approaches

This category comprises various methods for assessing similarities [17, 18]. The measurement is determined by a function of the length of the path (number of edges) separating the concepts. Additionally, in certain formulations it is determined by the positions of the concepts in the hierarchy [19, 20].

When multiple paths exist, the shortest path is selected because the closest concepts are considered the most similar [21]. Another defined similarity measure takes into account the most specific subsumer of the compared concepts [22]. The objective is to describe their cohesion and normalize their differences.

The basic Wu-Palmer similarity is related to the edge distance throughout the most specific subsumer. The expression is given by Eq. (1).

$$\text{Sim}_{WP}(c_1, c_2) = \frac{2*depth(c)}{depth(c_1) + depth(c_2)} \tag{1}$$

where,

C: denotes the most specific subsumer of $C_1$, $C_2$
Depth (C): length of concept C from the root
Depth (Ci): length of concept Ci through C
In Eq. (1), the probably weights of the edges in the assumption direction are not considered, nevertheless other research [23], have introduced factors $\alpha$ and $\beta$ to consider these weights which represent the strength of the relationship (specialization and/or generalization) between the linked nodes in the directed graph. The similarity function is then defined by:

$$\text{Sim}_{WP}(C_1, C_2) = \max_{J=1..k}(\alpha^{S(P_J)}\beta^{G(P_J)}) \tag{2}$$

where,

$\alpha$, $\beta$: downward and upward weights
$P_K$: paths connecting $C_1$ and $C_2$
$S(P_J)$: number of edges towards specialization
$G(P_J)$: number of edges towards generalization
Leacock and Chodorow [24] have classified their measure depending on the shortest path connecting the compared words; the formula is set by Eq. (3).

$$\text{Sim}_{lcha}(w_k, w_l) = \max_{I}\left[-\log\left[\frac{length_I(c_k, c_l)}{2D}\right]\right] \tag{3}$$

where,

$Length_I$ $(C_k, C_l)$: the length of the path 'I' connecting $C_k$, $C_l$
D: the maximum depth of the taxonomy
$W_i$: a set of concepts representing senses of the word 'i'
$C_k$, $C_l$: Concepts in the sets $W_k$, $W_l$

#### 2.2.2 Information content approaches
Information content of a concept. To quantify the information content of a concept, perceived as a crucial parameter in similarity assessment, Lin [25] introduced the use of information content. It was estimated that each concept in WordNet would possess a substantial amount of data.

Hence, concepts that share the most information are potentially the most similar. According to this insight, the amount of information they share corresponds to the

information content of their subsuming concept in the taxonomy.

The probability P of a concept C is linked to C itself along with the probability associated with all its descendant concepts. This probability is used to calculate the information content of the concept C.

$$IC_{resnik}(C) = -\log(P(C); P(C) = Freq(C))/N \qquad (4)$$

where,

N: total number of concepts

Freq(C): the number of occurrences of instances of the concept C

Information content measures. Below, in Eq. (5), we present the Resnik similarity measure associated with the information content of the least common subsumer.

$$Sim_{Resnik}(C_1, C_2) = IC(Lcs(C_1, C_2)) \qquad (5)$$

Note that in Eq. (5). The measure depends on the informative content of the least subsuming concept, therefore, another approach maximizes the informative content using Eq. (6), which uses set "$Sub_{ij}$" of all concepts subsuming $C_i$ and $C_j$.

$$Sim(C_i, C_j) = \max_{C \in Sub_{ij}}(IC(C)) \qquad (6)$$

It should be noted that there are also other hybrid measures and others based on the characteristics of the concepts, which we will not detail in this manuscript as they do not represent a significant impact in the following.

## 2.3 Related work of disambiguation approaches

WSD approaches can be classified into three main categories: knowledge-based, machine learning-based, and hybrid methods.

WSD techniques based on machine learning are applicable to tagged/untagged corpora. The training algorithm depends on the context and the structure of the corpus. Others, based on the WordNet ontology have been widely discussed.

The survey identifies four main components of a Word Sense Disambiguation (WSD) process:

### 2.3.1 Collection of word senses

Clearly, determining the lexical meanings of a word presents a core challenge in WSD. Various approaches exist, including enumerative methods relying on traditional machine-readable dictionaries, and generative methods. In this regard, ontologies such as WordNet and CoreLex [26], which assume systematic polysemy and employ semantic tagging, as well as sense induction techniques leading to fuzzier sense distinctions have been developed.

### 2.3.2 Utilization of external knowledge sources

Knowledge sources provide essential data for associating senses with words. These resources may include:

Structured resources. Thesauri provide relationships between words. Machine-readable dictionaries such as WordNet are presently the most utilized resource for WSD. Ontologies, typically comprising taxonomies and sets of semantic relationships, are also utilized [27].

Unstructured resources. Encompass raw corpora, which are collections of texts used for learning language models, and

Sense-Annotated Corpora such as SemCor [28], the largest and most widely used sense-tagged corpus, which includes 352 texts tagged with around 234,000 sense annotations.

Collocation resources. Show the tendency of words to co-occur with others.

Other resources. Such as word frequency lists, stop lists (words, like a, an, the, and so on), domain labels, etc.

### 2.3.3 Representation of the context

This process aims to convert the text into a structured format, starting with tokenization, part-of-speech tagging, lemmatization, chunking, and parsing. Each word can be characterized as a vector composed of various features. The representation of a word in context, along with additional knowledge resources, serves as the main foundation for facilitating automatic methods to identify the correct sense from a reference inventory.

There exist local attributes, which represent the immediate context of a word's usage, and topical features, that establish the main subject of a text or discourse. Syntactic features capture syntactic relations between the target word and other words, while semantic features convey semantic information.

### 2.3.4 Adoption of an automatic classification approach

In the context of WSD, two primary approaches can be broadly identified

Supervised approaches. These approaches leverage machine-learning techniques to develop a classifier based on labelled training sets [29], such as decision trees, decision lists, neural networks, and Naive Bayes. Their main limitations can be summarized as follows:

**Data scarcity:** High-quality annotated datasets are rare and expensive to create.

**Domain dependency:** Models trained on certain domains may not generalize well to others.

**Overfitting:** There is a possibility of overfitting the training data, limiting the model's applicability to unseen instances.

Unsupervised approaches. Aim to disambiguate word senses without labelled data, often using clustering techniques. These approaches face several challenges:

**Ambiguity resolution:** They may struggle with accurately resolving ambiguities in context-poor sentences.

**Evaluation complexity:** Evaluating unsupervised methods is difficult due to the lack of a standardized benchmark.

**Inconsistency:** These methods may produce inconsistent results due to their reliance on heuristics.

For knowledge-based disambiguation [30], on which relies this work, WSD seeks to utilize knowledge resources to deduce the meanings of words within a given context. Although these methods have less performance, they nevertheless have a wider coverage and this by the opportunity to use knowledge resources on a large scale [31]. The main knowledge-based techniques are:

Overlap of sense definitions. This method is often known as the Lesk algorithm, named after its creator.

Walker's approach. Based on the fact as each word is assigned to one or more categories of subjects in the thesaurus, different subjects are assigned to different senses of the word.

Selection restrictions. Utilize selection preferences to narrow down the meanings of a target word that appear in context.

Structural approaches. Supported by the presence of computational lexicons like WordNet, involve analyzing and exploiting the structure of concepts within these lexicons. Our contribution lies within this context, there are:

Similarity measurement approaches. Since the inception of WordNet, a range of measures of semantic similarity have been created to utilize the network of semantic relationships among word senses.

Graph-based approaches. These methods utilize graphical structures to identify the most suitable senses for words within their context, inspired by the concept of lexical chains [32]. They recognize contexts and enhance the consistency and continuity of meaning in discourse.

Other approaches consider the importance of the relationship between words differently. They utilize the WordNet graph, where nodes signify words connected by weighted edges ranging from [0, 1]. This approach helps generalize the WordNet graph and represents the strength of the relationship to assess the significance of a node by associating it with its intended meaning [33].

The literature presents local measures that depend on the concepts of centrality value, reflecting the relevance of a specific node within the graph and its overall integration into the structure. The frequently utilized measures include degree, closeness, betweenness, and fuzzy PageRank centrality [34].

Another type, called fuzzy global measures, aids in assessing the overall structure of the graph, with fuzzy compactness, graph entropy, and edge density being the commonly used metrics. Overall, the input text is lexically parsed, and then a WordNet graph is iteratively created where the tagged words are inserted as nodes. To create the edges, the semantic relations—hypernym, hyponym, meronym, holonym, and derivationally related forms—are taken into account. After generating the WordNet graph G, the algorithm assigns membership values or edge weights to the edges, as depicted in Figure 2, based on the significance of the corresponding semantic relations.
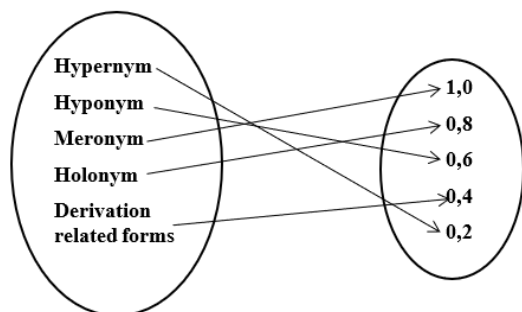


**Figure 2.** Edge weight combination

In other WSD approaches, unsupervised disambiguation algorithms were developed. They are based on two powerful constraints that words tend to exhibit:

One sense per discourse. The meaning of the target word remains highly consistent throughout the given document.

One sense per collocation. The meaning of the target word is inferred by consistent clues provided by nearby words.

The proposed algorithm is executed through several steps, beginning with the identification of all instances of the polysemous word in the corpus. Subsequently, its context is stored in an untagged training set.

Then, for each potential sense of the word, a small set of sense-learning examples is identified. This can be done either by manually tagging a subset of the training sentences or by using procedures such as identifying specific words known as seed collocations. In the third step, a supervised learning algorithm is trained using these seed sets, and the results of the classifier are applied iteratively, possibly using a decision tree.

The algorithm concludes when it converges on a stable residual set.

Various approaches explore the application of techniques to utilize one or more knowledge sources for associating the most suitable senses in a particular context. Others focus on exploiting lexical substitutes for the unsupervised induction of word senses. In this task, an annotator is presented with a target word within a context and is required to generate one or more substitutes. While useful, the most prominent method for modeling and addressing this phenomenon is word sense disambiguation. However, WSD techniques—including supervised classification, unsupervised methods, and knowledge-based approaches—face two main challenges. First, they require either large volumes of manually labelled data or extensive lexical resources with broad coverage. Second, issues such as the availability of a consistent sense inventory and questions regarding sense granularity continue to pose problems. Refer to Table 1 for more details.

**Table 1.** WSI clusters and lexical substitutions- instances of the noun "paper" [12]

| ID | Instance (Word in Context) | Cluster | - |
|---|---|---|---|
| 1 | … consistently reading *papers* with poor English … | 1 | article, manuscript |
| 2 | … while reading an item in the English *paper* today … | 2 | newspaper, periodical |
| 3 | … *papers* may use previously published materials … | 1 | article, publication |
| 4 | … the material uses fancy *paper* … | 3 | pulp, parchment |

It should be noted that context-based techniques and those using lexical substitutes are complementary. Contexts, as utilized in standard Word Sense Induction (WSI) approaches, offer a syntagmatic view of instance meaning. In contrast, lexical substitutes enable the representation of both syntagmatic and paradigmatic views of the target instance by providing contextually appropriate replacements that retain similarities to the target instance. Therefore, a method is described to address the issue of noisy contexts by utilizing lexical substitutes [35]. This method is akin to previous analyses, which have observed a significant degree of context-dependent variation below the sense level and discovered that lexical substitutes approximately mirror word sense groupings while also providing supplementary information.

Despite some methods targeting the enhancement of the query expansion process, these techniques typically overlook the context of words with multiple meanings. within the user's query. In contrast, the proposed method aims to determine the accurate sense of ambiguous terms in the query by assessing the similarity between the ambiguous term and other terms present in the query. Subsequently, weights are assigned to these similarities based on their decreasing order of distance from the ambiguous term. A similarity score is then computed for each sense of the ambiguous term based on the assigned weights, with the sense possessing the highest similarity score being considered the most appropriate one.

Although there have been advancements in WSD, current state-of-the-art machine learning approaches face several overarching limitations [36]:

**Contextual sensitivity:** Many methods struggle with capturing nuanced contextual information, leading to incorrect disambiguation.

**Scalability:** It remains a challenge, particularly for methods

relying on large annotated datasets.

**Adaptability:** The ability to adapt to new domains, languages, or evolving language use is a significant limitation.

**Interpretability:** Some advanced machine learning approaches, particularly deep learning, lack interpretability, making it difficult to understand the disambiguation process.

# 3. PRESENTATION OF THE APPROACH

The proposal outlined in this paper falls under the category of knowledge-based approaches; it can be considered a generative method, in the sense that it uses the WordNet ontology as a structured lexical reference to explore:

The senses inventory for the ambiguous word.

The tendency of words to co-occur with others within a given context.

Identification of the context of occurrence: The semantic relations established by the connections of WordNet synsets, as well as the use of semantic similarities between these nodes, allow for outlining the context to increase the probability of obtaining the uppermost score corresponding to the most appropriate sense.

Further elaboration on this disambiguation process will be provided in the subsequent sections.

## 3.1 Selection of similarity measures

An evaluation of various existing similarity approaches based on WordNet led to the combination of both the Wu-Palmer and the Leacock-Chodorow measures as given by Eq. (1) and Eq. (3), respectively.

The first measure is path-based and thus relies on the length of the shortest path between the compared concepts in WordNet, considering the "IS-A" relationship. This measure is inversely proportional to the depth of the hierarchy, which represents the longest path from the leaf node to the root of the hierarchy. This measure offers several advantages:

The Wu-Palmer measure is intriguing because it utilizes the depth of the compared concepts as well as that of their least common subsume (LCS) to perform the measurement. This formulation ensures adherence to certain aspects as described in Section 2.2.1.

Similarly, the Leacock-Chodorow measure relies on the path concept and thus hinges on the shortest path length between compared concepts in WordNet, according to the "IS-A" relationship. This measure is inversely proportional to the hierarchy depth, representing the longest path from the leaf node to the hierarchy's root. It offers the following advantages:

-Simplicity.

-Utilizes a logarithmic convex function, featuring a smooth and regular curve that is more sensitive to small variations than linear functions, hence providing more precise results.

-Several studies have demonstrated its correlation with human judgment, reaching 82%.

-According to its formulation, the closer two concepts are in proximity, the more similar they tend to be in nature, aligning well with the WordNet structure.

Indeed, when L (length of the shortest path) $\rightarrow$ 2D then the similarity tends to 0, the concepts are quite distant.

When L $\rightarrow$ 2 (the shortest possible path between two concepts) we get

sim $(C_k, C_l)$=-log(2/2D)=Log(D), this represents the potential maximum similarity. Therefore, we can normalize the obtained measurements in [0, 1].

## 3.2 Methodology

As mentioned earlier, structural approaches utilize the concept of lexical chains to facilitate disambiguation by offering context for resolving ambiguous terms and maintaining the coherence of discourse.

In the initial phase, a method will be suggested to delimit the context of the ambiguous term. To achieve this, the City Block distance will be employed to assess the similarity of word feature vectors. Subsequently, a disambiguation approach cantered on computing similarity measures will be elaborated upon; the objective is to establish the association between words in the given context.

### 3.2.1 Delimitation of the context

The purpose is to construct vectors representing sentences or fragments by incorporating data that reflects their essential characteristics. Then, a similarity measure is employed to evaluate their proximity or distance in the text.

Initially, to identify the terms within the context, we focus on nouns and verbs in their nominal forms. To achieve this, specialized scripts and libraries dedicated to natural language processing (NLP) are utilized to convert verbs into their most suitable nominal forms.

Indeed, accounting for verbs contributes to delineating the context as comprehensively as feasible, thereby enhancing the precision of measurements between terms.

Let's consider two adjacent sentences, $S_1$ and $S_2$. Suppose the set of terms belonging to these sentences is $\{T_1, T_2, ..., T_n\}$ where $T_i \in S_1$ or/and $S_2$, for i = 1, …, n, with n terms.

Expressed as tuples, $S_1$ can be represented as $(T_1, T_2, ... T_n)$, and $S_2$ as $(T_1, T_2, ... T_n)$.

TF-IDF (Term Frequency-Inverse Document Frequency) is indeed a widely used statistical method in information retrieval and natural language processing. It quantifies the importance of a term within a specific document or context.

The Term Frequency (TF) is defined as the relative frequency of a term Ti within a sentence S. See Eq. (7).

$$Tf_S(Ti) = \frac{f(Ti)}{\sum f(Tj)} j = 1, ..., n \qquad (7)$$

The Inverse Document Frequency (IDF) measures the quantity of information a term carries across a corpus of sentences or fragments. It is calculated by taking the logarithm of the total number of sentences or fragments (N) divided by the number of sentences or fragments that contain the term.

$$Idf(Ti) = \frac{N}{|S_j|} \ Ti \in S_j, j = 1,2,… \qquad (8)$$

The formula Tf×Idf is applied to all terms within the context. As a result, each sentence is represented by a vector with a dimension equal to the total number of terms present in the context. In this representation each component "i" of the vector corresponds to the Tf×Idf value associated with term "i". This vector representation captures the importance of each term within the context of the sentences.

To delimit the context, the proposed approach aims to assess the proximity between vectors $V_1$ and $V_2$. A resultant vector, denoted Vr, is constructed as follows:

$$\vec{V_r} = \vec{V_1} + \vec{V_2}$$
$$Vr(x) = V_1(x) + V_2(x), x = 1...n; n: terms$$

To determine whether the meaning of the second sentence extends the meaning of the first, indicating the need for context extension, we calculate the distance between the two vectors representing the two sentences and the vector Vr (refer to Figure 3).
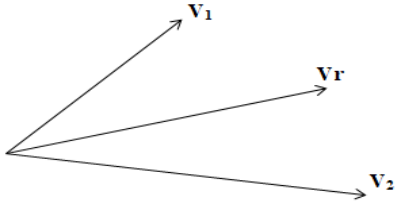


**Figure 3.** Vectors $V_1$, $V_2$ and the resultant vector $V_r$

For this purpose, the city block distance is utilized, which is a distance measure calculated as the average difference between dimensions. In most cases, this distance yields results similar to those obtained by the Euclidean distance.

$$D_{CityB}(V_r, V_{ji}) = \sum_{i=1}^{P} |V_{ri} - V_{ji}| \quad j = 1,2, \ldots \quad (9)$$

$$Sim_1(V_r, V_{1i}) = \frac{1}{D_{CityB}(V_r, V_{1i})} \quad (10)$$

$$Sim_2(V_r, V_{2i}) = \frac{1}{D_{CityB}(V_r, V_{2i})} \quad (11)$$

**Algorithm 1**
**Begin**
While (not end of text)
- Get all nouns in sentences $S_1$ and $S_2$
- Get all appropriate nouns related to verbs in $S_1$ and $S_2$
- Let $T_1$, $T_2$, …, $T_n$ be the obtained terms.
- Put the terms in any order (e.g. alphabetical)
      For I=1 to n
    Calculate X=TF-IDF for each term
    If ($T_i \in S_1$) then
      $V_1(I)$=X
    Else // $T_i \in S_2$
      $V_2(I)$=X
    End
   End
     For I=1 to N
         $V_r(I)$= $V_1(I)$+ $V_2(I)$
    End
// calculate distance measures
- Apply Eq. (9) above to calculate distance measure according to City Block formula for both Vr, $V_1$ and Vr, $V_2$

- Apply Eq. (10) and (11) to measure similarities for both $Sim_1$=(Vr, $V_1$) and $Sim_2$=(Vr,$V_2$)
   If ($Sim_1 \geq Sim_2$) then
     Merge $S_1$ and $S_2$ // extended meaning
     $S_1$= $S_1 \cup S_2$
     $S_2$=next_sentence
   End
End while
$S_1$={$W_1$,$W_2$, ..., $W_n$}
Return (Context=$S_1$)
**End.**

If $Sim_1 (V_r, V_{1i}) \geq Sim_2(V_r, V_{2i})$: The sense of the first sentence should be extended to the second. In this case, the context will be formed by merging the two sentences. Else, the similarity is lower, there is a break in the sense and the second sentence is supposed to carry another sense.

The algorithms used for identifying the context and choosing the most appropriate sense of the ambiguous word are not time-consuming, and their scalability depends on the context's size. In these circumstances, the semantics supported by the context in terms of the number of words are unlikely to be large enough to challenge the complexity and efficiency of the algorithms.

3.2.2 Disambiguation methodology
After setting the context where the ambiguous word occurs, the disambiguation process involves the following steps:
- Utilizing the WordNet database to retrieve a list of various senses of the ambiguous word 'Wa' along with their respective depths.
Let {($Wa_1$, $Da_1$), ($Wa_2$, $Da_2$), …, ($Wa_N$, $Da_p$)} the set of those pairs.
- For each sense of the word 'Wa', identify its least common subsumer (LCS) with all words in the context ($W_1$, $W_2$, ... $W_n$), along with the depth of their LCS and the length of the path connecting 'Wa' to all the words in the context towards their LCS.
- For each sense of the word "Wa", calculate both two similarities with all the words in the context ($W_1$, $W_2$,..., $W_N$)
First, the Wu-Palmer formula is applied:

$$Sim_{WP}(W_a, W_I) = \frac{2*D(W)}{2*D(W)+Spath(W_a,W_I)} \quad (12)$$
$$I=1,2,\ldots,n$$

where,
W: the least common subsumer of $W_a$, $W_I$
$W_I$: word of the context (I=1, …, n)
$W_a$: the ambiguous word Wa.
Spath: length of shortest path relating $W_a$, $W_I$
D(W): depth of word "W"
Second, it is applied the Leacock-Chodorow formula:

$$Sim(W_a, W_I) = -LOG\left(\frac{Spath(W_a, W_I)}{2*D}\right) \quad (13)$$
$$I = 1,2,\ldots,n$$

D=19 for WordNet 3.1
- For each sense of the word "Wa", two similarity vectors are constructed:
$Vs_1$: composed of similarities calculated using the Wu-Palmer formula.
$Vs_2$: consisting of similarities evaluated according to the Leacock-Chodorow formula.
- Calculate the angle between $Vs_1$ and $Vs_2$, by the cosine formula, in order to determine the degree of proximity or distance of the two vectors. Figure 4 shows the overall process.

$$Cos(\propto) = \frac{\sum_{i=1}^{n} Vs_1(i) * Vs_2(i)}{\sqrt{\sum_{i=1}^{n} Vs_1(i)^2} * \sqrt{\sum_{i=1}^{n} Vs_2^2}} \quad (14)$$

The calculated cosines should be evaluated for each sense of "Wa," resulting in a set equal in size to the number of senses of "Wa." The maximum cosine value among these will be

selected, reflecting that the vectors Vs1 and Vs2 with the smallest angle between them represent the most appropriate sense in the context (Figure 4).
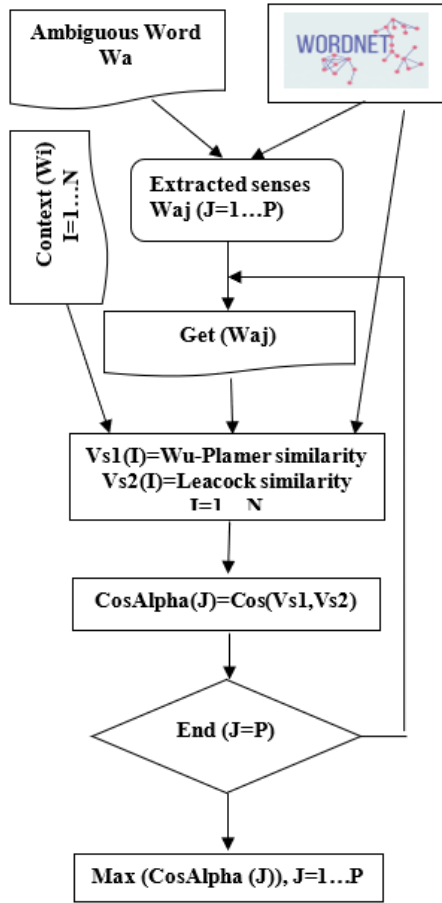


**Figure 4.** Overall process execution

**Algorithm 2**
**Begin**
- Use libraries to get From WordNet:
- Senses of ambiguous Word "Wa" and place them in a vector Wa[1..P]
- Place their Depths in a vector VdWa[1…P], P senses.
- Place in a vector Wc[1...N] the words of context.
- Place in a vector VdWc[1...N] their depths, N: words
- Use a vector VdCs[1...N] to store depths of Lcs
- Use a vector Vlen [1...N] to store length of shortest paths
- Use a vector CosAlpha to store cosines
- Use Vs1 [1...N] a vector similarities according to Wu-P
- Use Vs2 [1...N] a vector similarities according to

Leacock
For J=1 to P
    prod=1; sum1=0; sum2=0;
    For K=1 to N
        Seek(Cs)=LCS(Wa(J),Wc(K));
        VdCs(K)=depth(Cs);
        Vlen(K)=Length(Path(Wa(J),Wc(K)));

$$Vs1(K) = \frac{2*depth(Cs)+1}{2*depth(Cs)+Vlen(K)}$$

$$Vs2(K) = -LOG\left(\frac{Vlen(K)}{2*D}\right)$$

prod=prod*Vs1(K)*Vs2(K);
sum1=sum1+ (Vs1(K))²;
sum2=sum2+ (Vs2(K))²;
End For
$sum1 = \sqrt{sum1}$
$sum2 = \sqrt{sum2}$

$$CosAlpha(J) = \frac{prod}{sum1 * sum2}$$

End For
MaxCos=Max (CosAlpha(J); J=1...P)
**Return** (Wa(J)) // the selected sense
**End.**

## 4. EXPERIMENTATION AND DISCUSSION

The case studies presented have been meticulously selected to demonstrate the effectiveness of this approach by accurately assigning the correct sense to ambiguous words within specific texts or sentences. The metrics employed were sourced from WordNet and include:
-The depths of the synsets.
-The lengths of the paths.
-The senses of an ambiguous word.
Table 2 provides sentences that were used to assign the appropriate sense. Table 3 lists various senses of the ambiguous word "Mouse" as extracted from WordNet 3.1. Sentence 1 is highlighted as an example to explain the method; it achieved the highest score due to the closeness of its similarity vectors. It is associated with synset 4, where the angle between the two vectors is smallest.
Table 4 includes all words from the context related to the ambiguous word "Mouse". Tables 5-8 present various metrics such as the depths of words in the context, the least common subsumer (LCS), the shortest paths, similarities, and the application of the cosine measure to determine the appropriate sense.

**Table 2.** Submitted sentences and assigned senses

| Sentence | Ambiguous Word "Wa" | Related Synsets in WordNet | Max (Cosine) | Selected Synset | Assigned Sense | Precision |
|---|---|---|---|---|---|---|
| The child was doing his homework in his room, suddenly the mouse of his computer felt on the floor. He had to use the keyboard of the computer to complete his homework". | Mouse | 1-mouse.n.01 depth=12 2-shiner.n.01 depth=10 3-mouse.n.03 Depth=4 4-mouse.n.04 depth=9 | 0,948 | Synset 4 | Synset Name: mouse.n.04 Definition: a hand-operated electronic device that controls the coordinates of a cursor on your computer screen Depth: 8 | 98% |
| At the river bank, the environmentalists set up their equipment to monitor water | Bank | number=10 | 0,973 | Synset 1 | Synset Name: bank.n.01 Definition: sloping land | 99% |

| | | quality. | | | | (especially the slope beside a body of water) Depth: 5 | |
|---|---|---|---|---|---|---|---|

| The housekeeper was wiping the dust from the paintings in the exhibition room, when she saw a tear on a famous painting | Tear | 1-tear.n.01 depth=6 2-rip.n.02 depth=6 3-bust.n.04 depth=10 4-tear.n.04 depth=9 | 0,960 | Synset 2 | Synset Name: rip.n.02 Definition: an opening made forcibly as by pulling apart Depth: 6 | 97% |
|---|---|---|---|---|---|---|

Precision is assessed regarding the closest measure of the other senses.

**Table 3.** Different senses of ambiguous word "mouse" extracted from WordNet 3.1

| Senses | Definition | Depth |
|---|---|---|
| Synset 1 | any of numerous small rodents | 12 |
| Synset 2 | a swollen bruise caused by a blow to the eye | 10 |
| Synset 3 | person who is quiet or timid | 4 |
| Synset 4 | a hand-operated electronic device | 8 |

**Table 4.** Words of the context related to "Mouse"

| Noon | Depth | Verb/Noon Derivation |
|---|---|---|
| Child | 7 | |
| Make | 9 | doing/make |
| Homework | 10 | |
| Room | 7 | |
| Computer | 8 | |
| Fall | 6 | felt/fall |
| Floor | 7 | |
| Use | 9 | use/use |
| Keyboard | 7 | |
| Completion | 6 | Complete/completion |

**Table 5.** Computed similarities related to the word "mouse" in synset 1

| Noon | Depth | Depth LCS (Mouse, Noon) | Shortest Path Length (Mouse, Noon) | Sim1 Wu-P $Vs_1$ [1...10] | Sim2 Leacock-C $Vs_2$ [1...10] | Vs1*Vs2 | $\|Vs_1\|$ | $\|Vs_2\|$ |
|---|---|---|---|---|---|---|---|---|
| Child | 7 | 5 | 7 | 0.647 | 0.735 | 0.475 | | |
| Make | 9 | 0 | 21 | 0.048 | 0.258 | 0.012 | | |
| Homework | 10 | 0 | 22 | 0.045 | 0.237 | 0.011 | | |
| Room | 7 | 3 | 13 | 0.368 | 0.466 | 0.172 | | |
| Computer | 8 | 3 | 14 | 0.350 | 0.434 | 0.152 | 0.980 | 1.334 |
| Fall | 6 | 0 | 18 | 0.056 | 0.325 | 0.018 | | |
| Floor | 7 | 3 | 13 | 0.368 | 0.466 | 0.172 | | |
| Use | 6 | 0 | 18 | 0.056 | 0.325 | 0.018 | | |
| Keyboard | 7 | 3 | 13 | 0.368 | 0.466 | 0.172 | | |
| Completion | 9 | 0 | 21 | 0.048 | 0.258 | 0.012 | | |

$Cos(\alpha)=Cos(Vs_1,Vs_2)=0,928, \alpha=21,86$

**Table 6.** Computed similarities related to the word "mouse" in synset 2

| Noon | Depth | Depth LCS (Mouse, Noon) | Shortest Path Length (Mouse, Noon) | Sim1 Wu-P $Vs_1$ [1…10] | Sim2 Leacock-C $Vs_2$ [1…10] | Vs1*Vs2 | $\|Vs_1\|$ | $\|Vs_2\|$ |
|---|---|---|---|---|---|---|---|---|
| Child | 7 | 0 | 15 | 0.067 | 0.404 | 0.027 | | |
| Make | 9 | 1 | 17 | 0.158 | 0.349 | 0.055 | | |
| Homework | 10 | 1 | 18 | 0.150 | 0.325 | 0.049 | | |
| Room | 7 | 0 | 17 | 0.059 | 0.349 | 0.021 | | |
| Computer | 8 | 0 | 18 | 0.056 | 0.325 | 0.018 | 0.401 | 1.166 |
| Fall | 6 | 1 | 14 | 0.188 | 0.434 | 0.081 | | |
| Floor | 7 | 0 | 17 | 0.059 | 0.349 | 0.021 | | |
| Use | 6 | 1 | 14 | 0.188 | 0.434 | 0.081 | | |
| Keyboard | 7 | 0 | 17 | 0.059 | 0.349 | 0.021 | | |
| Completion | 9 | 1 | 17 | 0.158 | 0.349 | 0.055 | | |

$Cos(\alpha)=Cos(Vs_1,Vs_2)=0,916, \alpha=23,61$

**Table 7.** Computed similarities related to the word "mouse" in synset 3

| Noon | Depth | Depth LCS (Mouse, Noon) | Shortest Path Length (Mouse, Noon) | Sim1 Wu-P Vs₁ [1...10] | Sim2 Leacock-C Vs₂ [1...10] | Vs1*Vs2 | \|\|Vs₁\|\| | \|\|Vs₂\|\| |
|------|-------|-------------------------|-------------------------------------|-------------------------|------------------------------|---------|---------|---------|
| Child | 7 | 3 | 3 | 0.778 | 1.103 | 0.858 | | |
| Make | 9 | 0 | 13 | 0.077 | 0.466 | 0.036 | | |
| Homework | 10 | 0 | 14 | 0.071 | 0.434 | 0.031 | | |
| Room | 7 | 3 | 5 | 0.636 | 0.881 | 0.561 | | |
| Computer | 8 | 3 | 6 | 0.583 | 0.802 | 0.468 | 1.339 | 2.169 |
| Fall | 6 | 0 | 10 | 0.100 | 0.580 | 0.058 | | |
| Floor | 7 | 0 | 17 | 0.059 | 0.349 | 0.021 | | |
| Use | 6 | 0 | 10 | 0.100 | 0.580 | 0.058 | | |
| Keyboard | 7 | 3 | 5 | 0.636 | 0.881 | 0.561 | | |
| Completion | 9 | 0 | 13 | 0.077 | 0.466 | 0.036 | | |

$$Cos(\alpha)=Cos(Vs_1,Vs_2)=0,925, \alpha=22,35$$

**Table 8.** Computed similarities related to the word "mouse" in synset 4

| Noon | Depth | Depth LCS (Mouse, Noon) | Shortest Path Length (Mouse, Noon) | Sim1 Wu-P Vs₁ [1…10] | Sim2 Leacock-C Vs₂ [1…10] | Vs1*Vs2 | \|\|Vs₁\|\| | \|\|Vs₂\|\| |
|------|-------|-------------------------|-------------------------------------|-------------------------|------------------------------|---------|---------|---------|
| Child | 7 | 3 | 7 | 0.538 | 0.735 | 0.396 | | |
| Make | 9 | 0 | 17 | 0.059 | 0.349 | 0.021 | | |
| Homework | 10 | 0 | 18 | 0.056 | 0.325 | 0.018 | | |
| Room | 7 | 4 | 7 | 0.600 | 0.735 | 0.441 | | |
| Computer | 8 | 6 | 4 | 0.813 | 0.978 | 0.794 | 1.563 | 2.125 |
| Fall | 6 | 0 | 14 | 0.071 | 0.434 | 0.031 | | |
| Floor | 7 | 4 | 7 | 0.600 | 0.735 | 0.441 | | |
| Use | 6 | 0 | 14 | 0.071 | 0.434 | 0.031 | | |
| Keyboard | 7 | 6 | 3 | 0.867 | 1.103 | 0.956 | | |
| Completion | 9 | 0 | 17 | 0.059 | 0.349 | 0.021 | | |

$$Cos(\alpha)=Cos(Vs_1,Vs_2)=0,948, \alpha=18,54$$
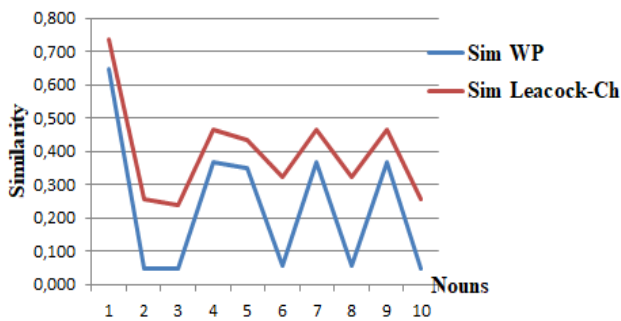


**Figure 5.** Similarities for "sense 1"
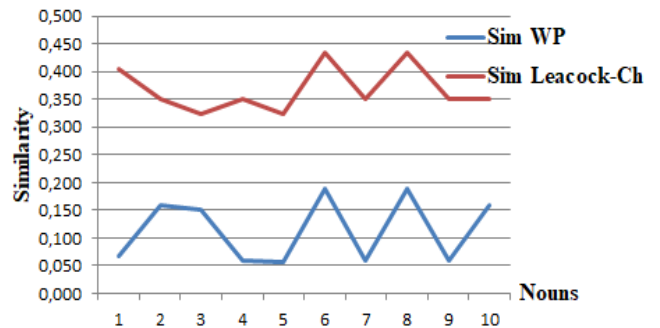


**Figure 6.** Similarities for "sense 2"
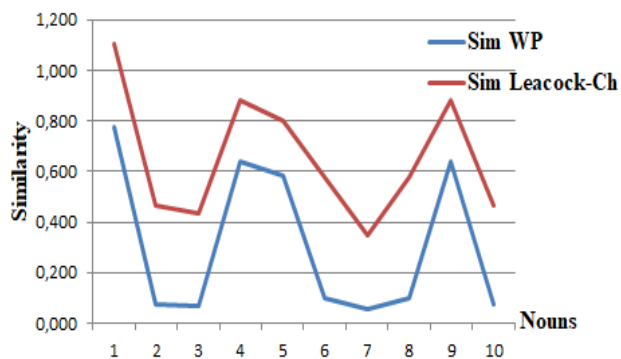


**Figure 7.** Similarities for "sense 3"
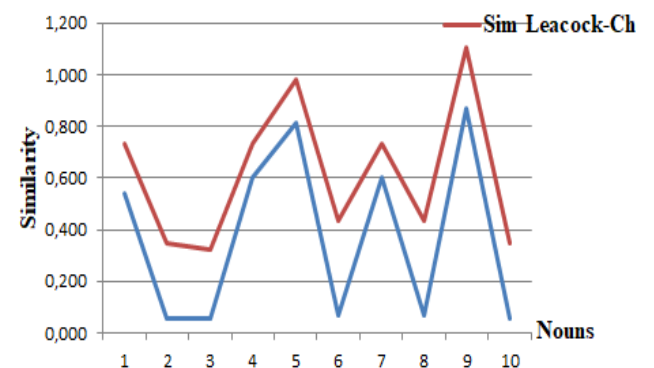


**Figure 8.** Similarities for "sense 4"

Therefore, the sense associated with synset 4, "a hand-operated electronic device", should be considered the most appropriate. Figures 5-8, provide an overview of the variations in the similarity vectors Vs1 and Vs2, depending on the depths and lengths of the paths.

This approach heavily relies on the structure of the WordNet hierarchy. The depths of the concepts, their relationships, and the lengths of the paths connecting them through the most specific subsumer form the foundation of this method.

It is important to note that defining the context of the ambiguous word remains a challenging task, as it significantly influences the processing of measurements. These elements often define and facilitate the management of meaningful connections within the network. Thus, the more accurately the context is identified, the more reliable the results will be.

The precision of the selected sense is measured relative to the nearest sense. Therefore, smaller cosine similarities between the similarity vectors indicate closer proximity of the vectors. In this context, precision reflects the degree of similarity among the closest senses.

Looking ahead, we consider using adjectives as an alternative to enhance the identification of discourse context. Furthermore, improving the process of transforming verbs into their corresponding nouns poses a significant challenge, as this conversion can produce various nouns whose application might lead to markedly different outcomes.

## 5. CONCLUSION AND FUTURE WORKS

Word sense disambiguation is a complex task integral to ongoing research in the field of natural language processing. The goal is to accurately predict the intended meaning of words using lexical resources, such as knowledge bases, or datasets that utilize various machine learning techniques. This innovative disambiguation approach is embedded in knowledge resources and semantic similarity measures, with detailed discussions on the algorithms, model components, and experimental aspects.

The model is founded on two key concepts: firstly, it leverages the WordNet taxonomy as a hierarchical base of knowledge; secondly, it employs semantic similarity measures between the nodes of the tree to evaluate their proximity and distance.

The disambiguation has focused on nouns and nominal forms of verbs, driven by the understanding that nouns' lexemes are semantically rich and thus highly beneficial for the disambiguation process. Additionally, the choice of similarity methods (such as Wu Palmer, Leacock and Chodorow, and cosine) along with other parameters like assigned weights, proportions, and path lengths, play crucial roles in influencing the accuracy of the results.

Experiments were carried out using recent, widely recognized software packages known for their efficacy in artificial intelligence applications. The results were highly encouraging, showing acceptable and promising effectiveness. Therefore, further enhancements, especially in context identification, are expected to yield even better outcomes.

Looking to the future, we plan to broaden the scope from nominal to adjectival phrases. This expansion is anticipated to enhance accuracy and facilitate the handling of more complex aspects, such as the figurative meanings of words, through the integration of functions, verbs, and adjectives.

Additionally, in terms of context identification, a concept known as "selectional restriction" involves verifying the exclusion of terms that do not appropriately fit the sense of the ambiguous word. Often, these terms share a least common subsumer with the root or a node very close to the root, despite their depth in the taxonomy.

## REFERENCES

[1] Navigli, R. (2009). Word sense disambiguation: A survey. ACM Computing Surveys (CSUR), 41(2): 10. https://doi.org/10.1145/1459352.1459355

[2] Vij, S., Jain, A., Tayal, D., Castillo, O. (2018). Fuzzy logic for inculcating significance of semantic relations in word sense disambiguation using a WordNet graph. International Journal of Fuzzy Systems, 20: 444-459. https://doi.org/10.1007/s40815-017-0433-8

[3] Komninos, A., Manandhar, S. (2016). Structured generative models of continuous features for word sense induction. In Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers, Osaka, Japan, pp. 3577-3587. https://aclanthology.org/C16-1337

[4] Başkaya, O., Jurgens, D. (2016). Semi-supervised learning with induced word senses for state of the art word sense disambiguation. Journal of Artificial Intelligence Research, 55: 1025-1058. https://doi.org/10.1613/jair.4917

[5] Nessah, D., Kazar, O. (2013). An improved semantic information searching scheme based multi-agent system and an innovative similarity measure. International Journal of Metadata, Semantics and Ontologies, 8(4): 282-297. https://doi.org/10.1504/IJMSO.2013.058411

[6] Roller, S., Erk, K. (2016). PIC a different word: A simple model for lexical substitution in context. In Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, San Diego, California, pp. 1121-1126. https://doi.org/10.18653/v1/N16-1131

[7] Luan, Y., Hauer, B., Mou, L., Kondrak, G. (2020,). Improving word sense disambiguation with translations. In Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP), pp. 4055-4065. https://doi.org/10.18653/v1/2020.emnlp-main.332

[8] McRoy, S. (2005). Word sense disambiguation: The case of combination of knowledge sources. Computational Linguistics, 30(2): 247–249 https://doi.org/10.1162/coli.2004.30.2.247

[9] Mittal, K., Jain, A. (2015). Word sense disambiguation method using semantic similarity measures and OWA operator. ICTACT Journal on Soft Computing, 5(2): 896-904. https://doi.org/10.21917/IJSC.2015.0126

[10] Budanitsky, A., Hirst, G. (2006). Evaluating wordnet-based measures of lexical semantic relatedness. Computational Linguistics, 32(1): 13-47. https://doi.org/10.1162/coli.2006.32.1.13

[11] Miller, G.A. (1992). WordNet: A lexical data base for English. Association for Computing Machinery, 38(11): 483. https://doi.org/10.3115/1075527.1075662

[12] Alagić, D., Šnajder, J., Padó, S. (2018). Leveraging lexical substitutes for unsupervised word sense induction. Proceedings of the AAAI Conference on

Artificial Intelligence, 32(1): 5004-5011. https://doi.org/10.1609/aaai.v32i1.12017

[13] Pustejovsky, J. (1998). The Generative Lexicon. MIT Press. https://mitpress.mit.edu/9780262661409/the-generative-lexicon/.

[14] McCrae, J.P., Rademaker, A., Bond, F., Rudnicka, E., Fellbaum, C. (2019). English WordNet 2019-an open-source WordNet for English. In Proceedings of the 10th Global Wordnet Conference. Wroclaw, Poland, pp. 245-252. https://aclanthology.org/2019.gwc-1.31.

[15] Meng, L., Huang, R., Gu, J. (2013). A review of semantic similarity measures in wordnet. International Journal of Hybrid Information Technology, 6(1): 1-12.

[16] Madani, Y., Erritali, M., Bengourram, J. (2019). Sentiment analysis using semantic similarity and Hadoop MapReduce. Knowledge and Information Systems, 59: 413-436. https://doi.org/10.1007/s10115-018-1212-z

[17] Ho Lee, J., Ho Kim, M., Joon Lee, Y. (1993). Information retrieval based on conceptual distance in IS-A hierarchies. Journal of Documentation, 49(2): 188-207. https://doi.org/10.1108/eb026913

[18] Rada, R., Bicknell, E. (1989). Ranking documents with a thesaurus. Journal of the American Society for Information Science, 40(5): 304-310. https://doi.org/10.1002/(SICI)1097-4571(198909)40:5%3C304::AID-ASI2%3E3.0.CO;2-6

[19] Fellbaum, C., Miller, G. (1998). Combining local context and WordNet similarity for word sense identification. In WordNet: An Electronic Lexical Database, pp. 265-283. https://doi.org/10.7551/mitpress/7287.003.0018

[20] Li. Y., Bandar. Z.A., McLean. D. (2003). An approach for measuring similarity between words using multiple information sources. IEEE Transaction on Knowledge and Data Engineering, 15(4): 871-882. https://doi.org/10.1109/TKDE.2003.1209005

[21] Resnik, P. (1995). Using information content to evaluate semantic similarity in a taxonomy. arXiv preprint cmp-lg/9511007. https://doi.org/10.48550/arXiv.cmp-lg/9511007

[22] Wu. Z., Palmer. M. (1994). Verb semantics and lexical selection. In 32nd Annual Meeting of the Association for Computational Linguistics, United States, pp. 133-138. https://doi.org/10.3115/981732.981751

[23] Seddiqui, M.H., Aono, M. (2010). Metric of intrinsic information content for measuring semantic similarity in an ontology. In Proceedings of the Seventh Asia-Pacific Conference on Conceptual Modelling, Brisbane, Australia, pp. 89-96.

[24] Leacock, C., Chodorow, M. (1998). Combining Local Context and WordNet Similarity for Word Sense Identification. https://doi.org/10.7551/mitpress/7287.003.0018

[25] Lin, D. (1998). Automatic retrieval and clustering of similar words. In Proceedings of the 36th Annual Meeting of the Association for Computational Linguistics and 17th International Conference on Computational Linguistics, pp. 768-774.

[26] Buitelaar, P.P. (1998). CoreLex: Systematic polysemy and under specification. Brandeis University. https://www.proquest.com/openview/14de0ccab1950c48112d8d7a3c267ea8/1?pq-origsite=gscholar&cbl=18750&diss=y.

[27] Borah, P.P., Talukdar, G., Baruah, A. (2014). Approaches for word sense disambiguation–A survey. International Journal of Recent Technology and Engineering, 3(1): 35-38.

[28] Miller, G.A., Leacock, C., Tengi, R., Bunker, R. (1993). A semantic concordance. Human Language Technology, 1993: 303-308. https://doi.org/10.3115/1075671.1075742

[29] Bevilacqua, M., Maru, M., Navigli, R. (2020). Generationary or "how we went beyond word sense inventories and learned to gloss". In Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP), pp. 7207-7221. https://doi.org/10.18653/v1/2020.emnlp-main.585

[30] Bevilacqua, M., Pasini, T., Raganato, A., Navigli, R. (2021). Recent trends in word sense disambiguation: A survey. In International Joint Conference on Artificial Intelligence, Montreal, Canada, pp. 4330-4338. https://doi.org/10.24963/ijcai.2021/593

[31] Pal., A.R., Saha, D. (2015). Word sense disambiguation: A survey. International Journal of Control Theory and Computer Modeling, 5(3): 1-16. https://doi.org/10.5121/ijctcm.2015.5301

[32] Barba, E., Pasini, T., Navigli, R. (2021). ESC: Redesigning WSD with extractive sense comprehension. In Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, pp. 4661-4672. https://doi.org/10.18653/v1/2021.naacl-main.371

[33] Bouhsissin, S., Sael, N., Benabbou, F. (2023). Evaluating data sources and datasets in intelligent transport systems through a Weighted Scoring Model. International Journal of Transport Development and Integration, 7(4): 353-365. https://doi.org/10.18280/ijtdi.070409

[34] Yarowsky, D. (1995). Unsupervised word sense disambiguation rivaling supervised methods. In 33rd Annual Meeting on Association for Computational Linguistics, Cambridge, Massachusetts, pp. 189-196. https://doi.org/10.3115/981658.981684

[35] Kremer, G., Erk, K., Padó, S., Thater, S. (2014). What substitutes tell us-analysis of an "all-words" lexical substitution corpus. In Proceedings of the 14th Conference of the European Chapter of the Association for Computational Linguistics, Gothenburg, Sweden, pp. 540-549. https://doi.org/10.3115/v1/E14-1057

[36] Sadatcharam, Y., Muruganadam, D. (2023). A review of deep learning algorithms for anomaly detection in videos. International Journal of Safety and Security Engineering, 13(2): 205-211. https://doi.org/10.18280/ijsse.130203