A REVIEW OF WORD SENSE DISAMBIGUATION METHOD

Mohammad Hafizi Jainal[1*], Saidah Saad[2], Rabiah Abdul Kadir[3]

[1]   Institute of IR4.0 (IIR4.0), Universiti Kebangsaan Malaysia, Malaysia
       Email: p91333@siswa.ukm.edu.my
[2]   Faculty of Information Science and Technology, Universiti Kebangsaan Malaysia, Malaysia
       Email: saidah@ukm.edu.my
[3]   Institute of IR4.0 (IIR4.0), Universiti Kebangsaan Malaysia, Malaysia
       Email: rabiahivi@ukm.edu.my
*   Corresponding Author

Article Info:

Abstract:

Background: Word Sense Disambiguation (WSD) is known to have a detrimental effect on the precision of information retrieval systems, where WSD is the ability to identify the meanings of words in context. There is a challenge in inference-correct-sensing on ambiguous words. Through many years of research, there have been various solutions to WSD that have been proposed; they have been divided into supervised and knowledge-based unsupervised. **Objective:** The first objective of this study was to explore the state-of-art of the WSD method with a hybrid method using ontology concepts. Then, with the findings, we may understand which tools are available to build WSD components. The second objective was to determine which method would be the best in giving good performance results of WSD, by analysing how the methods were used to answer specific WSD questions, their production, and how their performance was analysed. **Methods:** A review of the literature was conducted relating to the performance of WSD research, which used a comparison method of information retrieval analysis. The study compared the types of methods used in case, and examined methods for tools production, tools training, and analysis of performance. **Results:** In total 12 papers were found that satisfied all 3 inclusion criteria, and there was an anchor paper assigned to be referred. We chose the knowledge-based unsupervised approach because it has fewer word sets constraints than the supervised approaches which require training data. Concept-based ontology will help WSD in finding the semantic words concept with respect to another concept around it. **Conclusion:** Many methods was explored and compared to determine the most suitable way to build a WSD model based on semantics between words in query texts that can be related to the knowledge concept by using ontological knowledge presentation.

Introduction

In the area of Natural Language Processing (NLP) ambiguity is identified as a challenge to human language. WSD is the best solution to solve this issue (*Weaver, W*). Warren first introduced this solution in 1949. He stated that the ambiguity of the words needs to be resolved first to ensure that automatic translation between languages can be done. This step is continued and used in various linguistic fields such as Information Retrieval (*M. Sanderson, 1994*) (*Z. Zhong, H. Ng, 2012*), Machine Translation (*S. C. Yee, T. N. Hwee, C. David, 2007*) (*D. Xiong, M. Zhang, 2014*), and Information Extraction (*C.D. Bovi, L. Telesca, R. Navigli, 2015*). WSD is referred to as an AI-complete problem (*R. Navigli and P. Velardi, 2005*).

In figure 1 below in the Word Sense Structure, we have Synonym, Polyseme, Homonym, Hyponym, and Antonym. This shows many types of word units that need to be filtered. An example is polysemy, which is the state or phenomenon in which the words have more than one meaning. In NLP studies we call this semantics.
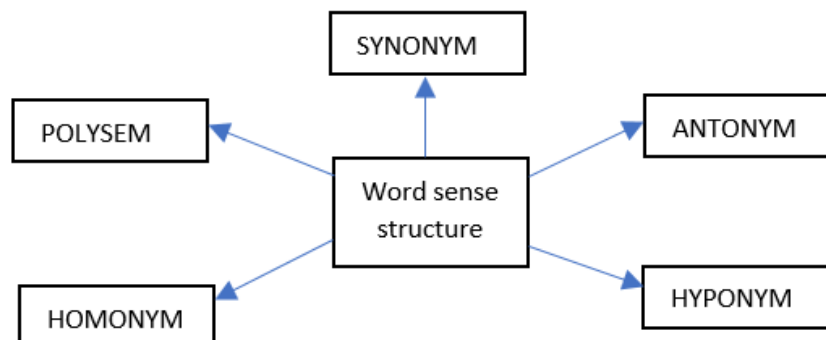
Figure 1: Figure of Word Sense Structure

The WSD step is a step to identify the meanings of words along with the context of the sentence through a computational paradigm and to distinguish a meaning between the sense of words. These words are called polysemous words, words that have many meanings and have different contexts, but with the same spelling.

WSD measures are often classified into 2 types: lexical type and type of all words in one document. The first focuses on vague words versus targeted words only, while the second type involves all the words in the document. Through many years of research, there have been various solutions to the WSD that have been proposed, and these have been divided into supervised and knowledge-based unsupervised.

The supervised learning model approach is trained from the corpus, where any correct words sensed from ambiguous words are validated by human annotators (*Dirk Weissenborn, Leonhard Hennig, Feiyu Xu and Hans Uszkoreit, 2015*) (*Oren Melamud, Jacob Goldberger, and Ido Dagan, 2016*) (*Alessandro Raganato, Claudio Delli Bovi, and Roberto Navigli, 2017*).

Unsupervised approach uses lexical knowledge bases (LKBs) such as WordNet (*Satanjeev Banerjee and Ted Pedersen, 2003*) (*Devendra Singh Chaplot, Pushpak Bhattacharyya, and Ashwin Paranjape , 2015*). This method combines contextual information with semantic knowledge of LKBs. Thus, a large number of words can be obscured compared to the supervised approach.

Wordnet is a large lexical database of English (*https://wordnet.princeton.edu/*). Wordnet *(Fellbaum, C. WordNet and wordnets, 2005*), which is currently used as a standard word sense inventory for English texts, relates each word (word lemma) to a set of WordNet sense identified. Its latest version is WordNet 3.1, which contains about 155,000 words organized across 117,000 synset.

There is another corpus semantic resource other than WordNet. Some researchers use ontology learning to make the learning machine understand the contents of documents semantically. The strategy to create machine understandable content is the concept of ontologies. An ontology may be defined as a shared conceptualization of a domain (*Fensel, D. , 2001*). In another study, the ontology learning method achieved a 9% improvement over the base method (*Sin-Jae Kang , 2003*). The next study is to use the ontology learning method with WSD and semantic relations to improve the WSD performance.

This paper focuses on identifying methods in WSD that can be used in the next research study to overcome the limitation on inferencing the correct sense in ambiguous words. In the next section we explain the methodology created for this paper. Section 3 shows results, followed by discussions and references.

Method

Identification and Screening

In August 2017, when preparing the project proposal, a discussion with the supervisor was conducted. The discussion was based on the research problem which was that word ambiguity in query texts affects the performance of access to relevant information, and search results do not meet the needs of users. And for now, there is a challenge in inferencing the correct sense on ambiguous words (*Dongsuk O, Sunjae Kwon, Kyungsun Kim, Youngjoong Ko, 2018*), and following with a computerized search of the literature.

WSD with IR terms was searched using the terms. References were checked from papers and reviews, and citations were checked from included studies. The titles and abstracts were read from the retrieved search results. The retrieved paper was to identify relevant studies. Then the extended research was done by using a key journal as a main key title.

To proceed with the identification, there are 3 research questions to be solved. What is the characteristic of semantic ambiguity in query text and how existing ontological concept forms support research in semantic ambiguity of query text? The third research question is, how and what parameters are needed to identify the novel of semantic ambiguity in the query text.

Studies had to include one of the following search terms in the title, abstract, or as keywords, "WSD" or its associated term, "word sense disambiguation". There were 3 inclusion criteria

used for the study. First, the study should have WSD as its primary focus. The aim of this review was to explore research into the methods of WSD in Information Retrieval.

Second, the papers should be related to an Information Retrieval area. There is a lot of this method used in various fields like machine learning and so on. But in this research, we specifically research on Information Retrieval. We also focus on a specific domain to narrow our scope, which is Islamic knowledge. Finally, papers should have terms related to ontology. This is because the previous discussions have focused on WSD and ontology learning. We believe this state-of-art will provide an improvement over the WSD method (*Sin-Jae Kang , 2003*).

Eligibility and Inclusion

The studies were restricted to articles published in English. A total of 30 full-text articles were accessed for eligibility. Of these, 36% (11/30) were rejected because they did not meet the criteria and did not match the general design we planned to propose. For example, Boshra (*Boshra F. Zopon AL Bayaty; Shashank Joshi , 2015*) and Pranjal (*Borah P.P., Talukdar G., Baruah A. , 2019*) were rejected because we are not looking for supervised methods in our research studies, but instead are looking into unsupervised methods.

There are four papers under the ontology field which we have included in our criteria because ontology learning will be implemented in our knowledge base or as an answers scheme. So, the criteria included should at least have a minimal number of research studies. There were only 13% (4/30) from other research studies, which included Brank (*Janez Brank , Marko Grobelnik , Dunja Mladenić , 2005*), Kulkani (*Manasi Kulkarni; Suneeta Sane , 2011*), Hakkoum (*Aimad Hakkoum & Said Raghay, 2016*) and Sadi (*A. B. M. Shamsuzzaman Sadi;Towfique Anam;Mohamed Abdirazak;Abdillahi Hasan Adnan;Sazid Zaman Khan;Mohamed Mahmudur Rahman;Ghassan Samara , 2016*).

For the main topic criteria, which is unsupervised methods, we got 20% (6/30) matching results, including IR Dongsuk (*Dongsuk O, Sunjae Kwon, Kyungsun Kim, Youngjoong Ko, 2018*), Pachenko (*Alexander Panchenko, Fide Marten, Eugen Ruppert, Stefano Faralli, Dmitry Ustalov, Simone Paolo Ponzetto, Chris Biemann, 2017*), Rufai (*R. A. Kadir, R. A. Yauri, A. Azman, 2018*), Meysam (*Meysam Arab; Mansoor Zolghadri Jahromi; Seyed Mostafa Fakhrahmad , 2016*), Agirre (*Eneko Agirre, Aitor Soroa, Mark Stevenson , 2010*) and Pederson (*Pedersen, Ted et al. , 2005*); and these were based on our proposal design in a future paper.

Result

In total 12 papers were found that satisfied all 3 inclusion criteria, and there was an anchor paper for reference. We chose knowledge-based unsupervised approach because it has fewer word set constraints than the supervised approaches, which require training data. The concept-based ontology will help WSD in finding the semantic words concept with respect to other concepts around it.

Knowledge Based WSD

We know that WSD has a detrimental effect on the precision of information retrieval systems. To overcome this challenge there are some methods that have been published by some researchers in NLP studies. Dongsuk (*Dongsuk O, Sunjae Kwon, Kyungsun Kim, Youngjoong Ko, 2018*) is a new method used to generate the context of an ambiguous word by using

similarity between an ambiguous word and words in the input documents. Dongsuk (*Dongsuk O, Sunjae Kwon, Kyungsun Kim, Youngjoong Ko, 2018*) created a new similarity calculation method based on the semantic network structure of BabelNet (*https://babelnet.org/*). The outcome showed that WSD significantly improved when compared with the existing method. This shows that their method had perfectly improvised the precision of IR.

On the other hand, the other researches used the knowledge-based method as interpreted from its meaning; Pachenko (*Alexander Panchenko, Fide Marten, Eugen Ruppert, Stefano Faralli, Dmitry Ustalov, Simone Paolo Ponzetto, Chris Biemann, 2017*) came out with the knowledge-free method as interpretable. As an improvisation from previous research, it made the present WSD tools feature as a web interface for all-word disambiguation of texts. This tool makes the sense predictions human readable by providing interpretable word sense inventories, sense representations and disambiguation results. A public API is provided, enabling seamless integrations.

There is another way to increase the precision of IR by modifying the original query. Rufai (*R. A. Kadir, R. A. Yauri, A. Azman, 2018*) is a method technique that uses ontology-based query expansion. The unsupervised method is implemented to resolve ambiguous natural language. The disambiguation process implements usage of an external dictionary such as WordNet and uses a statistical linguistic technique for solving the problem of ambiguity automatically.

To achieve semantic query disambiguation, Rufai (*R. A. Kadir, R. A. Yauri, A. Azman , 2018*) attempts to represent natural language query into semantic web. The data on semantic webs are represented in RDF format, which is Subject, Predicate, Object. To achieve the objectives, Rufai performs concept notation using n-grams, which involves the annotation of several ontology concepts and addition of relationships between the concepts. These ontology concepts are stored and form a repository called knowledge base. The method shows that the overall result increases by 6% and 10% for precision and recall.

In the Meysam (*Meysam Arab; Mansoor Zolghadri Jahromi; Seyed Mostafa Fakhrahmad , 2016*) research area on unsupervised methods, they are using a graph-based algorithm to select the final sense of the target word. Meysam extracts all available senses of the context words as nodes. It then draw the edges of the graph by calculating semantic similarity relations between sense of the words using a combination of semantic similarity methods. Finally, by exploiting the constructed graph, the final sense of the target words may be selected. On benchmark datasets senseval-2 and senseval-3, the result shows that the proposed model outperforms all other graph-based methods.

The unsupervised method also gives benefits for a specific domain, especially in important fields in the health care domain, like in the biomedical field. Agirre (*Eneko Agirre, Aitor Soroa, Mark Stevenson , 2010*) creates an unsupervised method in a biomedical domain using a graph-based method. This method does not use any labelled training data and instead uses knowledge from a Unified Medical Language System (UMLS). The research studies use Personalized Page Ranked for WSD implementations. The results show that the method outperforms other methods using UMLS thesauruses as knowledge sources.

Lastly, in other literature focusing on the unsupervised method, Pederson (*Pedersen, Ted et al. , 2005*) introduced a method of WSD that selects the sense of the target word that has the

maximum relatedness. The maximum relatedness with the content words was found in a large window of surrounding context. Pederson showed that an algorithm can be used with any measure that computes a relatedness score between two concepts and found that in general the performance of an algorithm improved as the window of context increased. The result is that the extended gloss overlap measure (lesk-e) was overall the most effective, and the gloss vector measure fared particularly well with adjectives, which are essentially impossible for path based and information content measures.

Table 1: List of Unsupervised Methods

Year	Author	Approach	Evaluation/Result	Field
2005	Ted Pedersen et al	A method of word sense disambiguation that assigns a targeted word the sense that is most related to the senses of its neighbouring words. Measures similarity and relatedness that are based on finding paths in a concept network, information content derived from a large corpus, and word sense glosses.	They find that the two gloss-based measures that have been developed are particularly flexible and effective measures for word sense disambiguation.	Unsupervised WSD
2010	Eneko Agirre	A graph-based approach to WSD in the biomedical domain. Personalized PageRank algorithm to select the best sense for each ambiguous word. Combination of all vocabularies in the MRREL table of the Metathesaurus.	Evaluated on the NLM-WSD dataset, the algorithm outperforms other methods that rely on the UMLS Metathesaurus alone.	Unsupervised WSD
2016	Meysam Arab et al	New method of combining similarity metrics that uses higher order relations between words to assign appropriate weights to each edge in the graph. And a new approach for selecting the most	Senseval-2 and Senseval-3 shows that the proposed model outperforms all other graph-based methods.	Unsupervised WSD

			appropriate sense of the target word that makes use of the indegree centrality algorithm and senses of the neighbour's words.	
2017	Alexander Panchenko et al	Unsupervised, knowledge-free, and interpretable at the same time. The user interface of the system provides efficient access to the produced WSD models via a RESTful API or via an interactive Web-based graphical user interface.	All evaluated models outperform both random and most frequent sense baselines.	Unsupervised WSD
2018	Dongsuk O et al	New WSD method to generate the context of an ambiguous word by using similarities between an ambiguous words and words in the input document. New word similarity calculation method based on the semantic network structure of BabelNet.	It has higher performance than the state-of-the-art unsupervised knowledge based WSD system in the average performance of both datasets. Use SemEval-2013 and SemEval-2015.	Unsupervised WSD
2018	Rufai et al	Based on the Query Expansion ontology. A statistical linguistic technique for solving the problem of ambiguity automatically.	Increases the result of precision and recall by 6% and 10%, respectively.	Unsupervised WSD
2021	Sm Zobaed et al	Propose SensPick, a type of stacked bidirectional Long Short-Term Memory (LSTM) network to perform the WSD task. Neural network-based WSD approaches rely on a sense-annotated corpus	SensPick outperforms traditional and state-of-the-art models on most of the benchmark datasets with a relative improvement of 3:5% in the F-1 score.	Unsupervised WSD

since they do not utilize
lexical resources.

Concept Based Ontology

Concept based WSD means that the method uses machine learning for WSD. What we focus on this sub section is making use of ontology concepts to extract the conceptual knowledge information. The trend now focusses on modern information systems that use "concept processing" rather than "data processing". This explains that the basic unit of processing is less of an atomic piece of data and is becoming more of a semantic concept which carries an interpretation and exists in a context with other concepts. Ontology is commonly used as a structure to capture knowledge about certain areas by providing relevant concepts and relations between them.

Table 2: List of Concepts Based with Ontology

Year	Author	Approach	Evaluation/Result	Field
2009	Janez Brank et al	A survey of the state of the art in ontology evaluation	There is no single best or preferred approach to ontology evaluation.	Ontology, Evaluation
2011	Kulkarni et al	Method for word sense disambiguation based on Lesk algorithm which uses lexical database WordNet as a knowledge base. Ontology as a clarification tool accepts English statement as input and gives the best possible meaning of the given word.	Experiment with Senseval-2 test data for lexical sample task. The results show betterment over the original Lesk algorithm.	Ontology
2016	Hakkoum et al	A semantic-based search engine for the Qur'an; it is based on creating an ontology that represents the knowledge of the Qur'an in Web Ontology Language format, and a natural language interface that transforms user queries expressed in Arabic into SPARQL queries and then retrieves answers from the ontology.	To evaluate the performance and effectiveness of an approach, they asked a researcher in Islamic studies to provide them with a set of questions. Two performance metrics were used: precision and recall. Precision 95%, Recall 73%.	Ontology

2016	Sadi et al	Modelled nature related concepts of the holy Quran using OWL/RDF.	SPARQL queries can retrieve verses and concepts of interest.	Ontology

The future research in the proposal design will mostly be using the conceptual knowledge to retrieve a better result. This means that WSD in this research paper was using ontology knowledge as the answer scheme.

Kulkarni (*Manasi Kulkarni; Suneeta Sane , 2011*) makes the ontology as a clarification tool that retrieves the synonym set for target as well as neighbouring words from the term. It uses various components to achieve this goal. The WSD is based on Lesk algorithm, which uses the lexical database WordNet as a knowledge base, and the results show betterment over the original Lesk algorithm.

To be focused on research studies on Islamic domain the ontology needs to be defined as Islamic knowledge. As such, the accurate information will be retrieved from Islamic conceptual knowledge. There are some researchers that have developed an Islamic conceptual ontology. Sadi (*A. B. M. Shamsuzzaman Sadi;Towfique Anam;Mohamed Abdirazak;Abdillahi Hasan Adnan;Sazid Zaman Khan;Mohamed Mahmudur Rahman;Ghassan Samara , 2016*) makes use of Quranic information to develop an ontological modelling on Quranic Nature Domain, and Hakkoum (*Aimad Hakkoum & Said Raghay, 2016*) uses Islamic ontology as Quranic Knowledge in a Q&A System.

Sadi (*A. B. M. Shamsuzzaman Sadi;Towfique Anam;Mohamed Abdirazak;Abdillahi Hasan Adnan;Sazid Zaman Khan;Mohamed Mahmudur Rahman;Ghassan Samara, 2016*) has modelled nature related concepts of the holy Quran using OWL (Web Ontology Language) / RDF (Resource Description Framework). The method involves identifying relations among those concepts. These concepts and relations are represented as classes/instances and properties of an OWL ontology. The outcomes show that the retrieval queries give a relevant answer, thereby indicating that the model is working.

Hakkoum (*Aimad Hakkoum & Said Raghay , 2016*) uses ontology in Islamic domain, representing the Quranic knowledge in Web Ontology Language format, and a natural language interface that transforms user queries expressed in Arabic into SPAQL queries. The answers for the user questions come from the ontology. Hakkoum uses experts to evaluate the effectiveness and use two performance metrics: precision and recall. This means the evaluations need to use both methods rather than only one. Therefore, to hide the complexity of the ontology to the end user, a Natural Language Interface (NLI) is needed (*Damljanovic, D. , 2008*).

Discussion and Recommendation
Many methods were explored and compared to find out the most suitable way to build a WSD model based on semantics between words in query text that can be related on knowledge concepts by using ontological knowledge presentation.

In our paper we make Dongsuk (*Dongsuk O, Sunjae Kwon, Kyungsun Kim, Youngjoong Ko, 2018*) as our main reference study, because the paper meets our criteria to close a gap on our problem statement. But there is a need of improvements in some areas. Figure 2 is the overall picture on the Dongsuk WSD. This figure shows the concept of WSD to improve the performance of WSD and how they are integrated into the WSD system. The important component to focus on here is the word similarity calculation method and the iterative subgraph construction method.
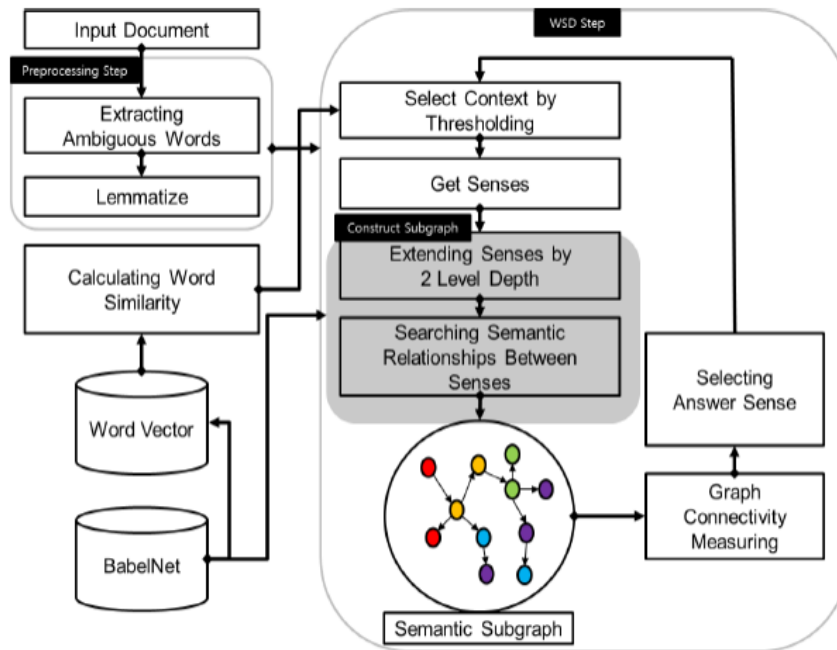

Figure 2: Figure of The Concept of WSD by Dongsuk

If we investigate the figure 2 diagram, we can see that there is an initial step to filter out the input document. This processing step is to extract an ambiguous word, lemmatize and do other processing steps. We also investigate the ontology concept diagram given in figure 3 below.
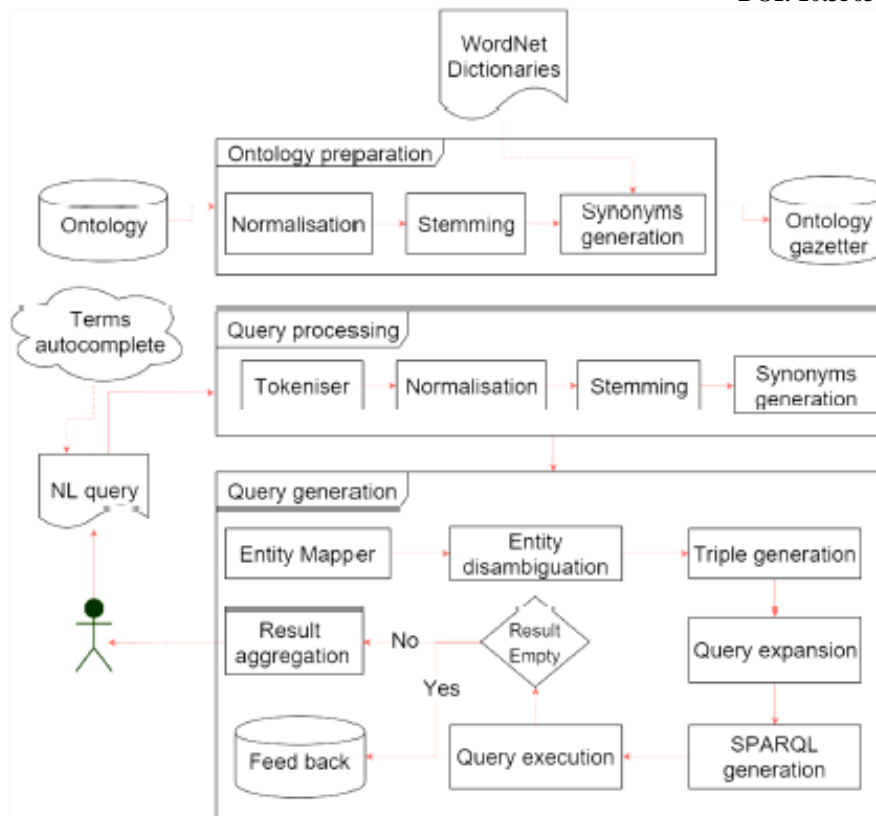
Figure 3: Figure of WSD Ontology Concept by Hakkoum

In figure 3, we have gotten inspired from the existing concept by Hakkoum and will make some enhancements to adapt it to our criteria. There are three main subprocesses: (1) Knowledge base preparation, which is an offline task, (2) Query processing and (3) Formal query generation.

To conclude, the model will be discussed as follows. Figure 4 shows the relation between Knowledge Based and Concept based methods.
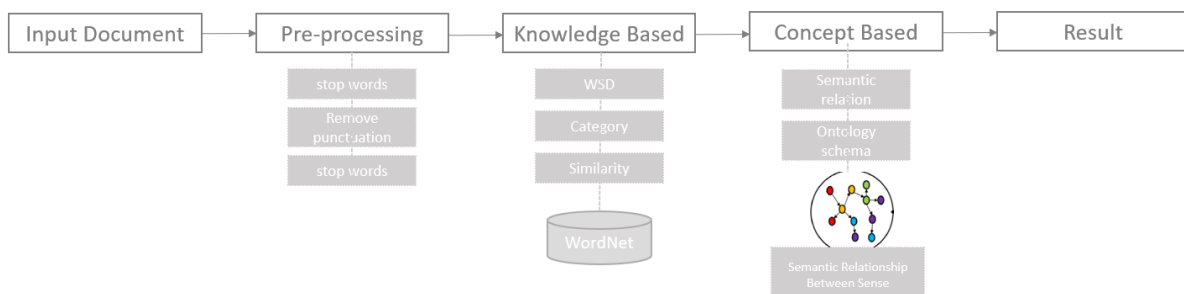

Figure 4: Relation between Knowledge Based, and Concept Based.

The next step of our research is to create details on a new WSD method to improvise the main method referred before by using two types of knowledge: contextual knowledge and conceptual knowledge of the selected domain-based ontology.

References

Aimad Hakkoum & Said Raghay (2016). Semantic Q&A System on the Quran , *Arabian Journal for Science and Engineering, 5205–5214(2016).*

A. B. M. Shamsuzzaman Sadi;Towfique Anam;Mohamed Abdirazak;Abdillahi Hasan Adnan;Sazid Zaman Khan;Mohamed Mahmudur Rahman;Ghassan Samara (2016) Applying Ontological Modeling on Quranic Nature Domain , 2016 *7th International Conference on Information and Communication Systems (ICICS), IEEE, DOI: 10.1109/IACS.2016.7476102.*

Alessandro Raganato, Claudio Delli Bovi, and Roberto Navigli (2017) Neural sequence learning models for word sense disambiguation. *In proceedings of the 2017 Conference on Empirical Methods in Natural Language Pro-cessing, pp. 1156-1167.*

Alexander Panchenko, Fide Marten, Eugen Ruppert, Stefano Faralli, Dmitry Ustalov, Simone Paolo Ponzetto, Chris Biemann (2017). Unsupervised, Knowledge-Free, and InterpretableWord Sense Disambiguation, *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing: System Demonstrations, Association for Computational Linguistics, 2017 , 91–96.*

Borah P.P., Talukdar G., Baruah A. (2019). WSD for Assamese Language. In: Kalita J., Balas V., Borah S., Pradhan R. (eds) Recent Developments in Machine Learning and Data Analytics. *Advances in Intelligent Systems and Computing, vol 740. Springer, Singapore. https://doi.org/10.1007/978-981-13-1280-9_11.*

A. Raganato, C.D. Bovi, R. Navigli (2016) Automatic construction and evaluation of a large semantically enriched wikipedia, in: *Proceedings of the 25th International Joint Conference on Artificial Intelligence, IJCAI '16, 2016, pp. 2894–2900.*

Boshra F. Zopon AL Bayaty; Shashank Joshi (2015). Empirical Comparative Study to Supervised Approaches for WSD Problem: Survey 2015 *IEEE Canada International Humanitarian Technology Conference (IHTC2015), IEEE, DOI: 10.1109/IHTC.2015.7238056.*

C.D. Bovi, L. Telesca, R. Navigli (2015) Large-scale information extraction from textual definitions through deep syntactic and semantic analysis, *Trans. Assoc. Comput. Linguist. 3 (2015) 529–543.*

C. Hung, S.J. Chen (2016) Word sense disambiguation based sentiment lexicons for sentiment classification, *Knowl.-Based Syst. 110 (2016) 224–232, http://dx.doi.org/10.1016/j.knosys.2016.07.030.*

Chin, J. L. (2011). Women and Leadership: Transforming Visions and Current Contexts. *Forum on Public Policy: A Journal of the Oxford Round Table*, (2), 1–12.

Damljanovic, D. (2008). Natural Language Interfaces to Ontologies: usability and performance (Transfer report) (2008).

Devendra Singh Chaplot, Pushpak Bhattacharyya, and Ashwin Paranjape (2015) Unsupervised Word Sense Dis-ambiguation Using Markov Random Field and Dependency Parser. *In proceedings of the AAAI, pp. 2217-2223.*

Dongsuk O, Sunjae Kwon, Kyungsun Kim, Youngjoong Ko (2018). Word Sense Disambiguation Based on Word Similarity Calculation Using Word Vector

Representation from a Knowledge-Based Graph, *Proceedings of the 27th International Conference on Computational Linguistics, Association for Computational Linguistics, 2018, 2704–2714.*

Dirk Weissenborn, Leonhard Hennig, Feiyu Xu and Hans Uszkoreit. (2015) Multi-objective optimization for the joint disambiguation of nouns and named entities. *In proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing, pp. 596-605.*

D. Xiong, M. Zhang. (2014) A sense-based translation model for statistical machine translation, *in: Proceedings of the 52th Annual Meeting of the Association for Computational Linguistics, ACL '14, 2014, pp. 1459–1469. http://dx.doi.org/10.1016/j.euroneuro.2010.01.001.*

Eneko Agirre, Aitor Soroa, Mark Stevenson (2010). Graph-based Word Sense Disambiguation of biomedical Documents, *Bioinformatics, Volume 26, Issue 22, 15 November 2010, Pages 2889–2896.*

Fellbaum, C. WordNet and wordnets (2005). *In Encyclopedia of Language and Linguistics, 2nd ed.; Elsevier: Oxford, UK, 2005; pp. 665–670.*

Fensel, D. (2001). Ontologies: *Silver Bullet for Knowledge Management and Electronic Commerce.*

https://wordnet.princeton.edu/

https://babelnet.org/

Janez Brank , Marko Grobelnik , Dunja Mladenić (2005) A survey of ontology evaluation techniques, *In Proceedings of the Conference on Data Mining and Data Warehouses (SiKDD 2005.*

Manasi Kulkarni; Suneeta Sane (2011) An ontology Clarification Tool for WSD, 2011 *3rd International Conference on Electronics Computer Technology , IEEE, DOI: 10.1109/ICECTECH.2011.5941608.*

Meysam Arab; Mansoor Zolghadri Jahromi; Seyed Mostafa Fakhrahmad (2016). A Graph-based Approach to Word Sense Disambiguation. An Unsupervised Method Based on Semantic Relatedness, 2016 *24th Iranian Conference on Electrical Engineering (ICEE), IEEE, DOI : 10.1109/IranianCEE.2016.7585527.*

M. Sanderson,(1994) "Word Sense Disambiguation and Information Retrieval", *Proceedings of the 17th Annual International ACM SIGIR conference on Research and Development in Information Retrieval, SIGIR'94, Dublin, Ireland, Springer, New York, July 03-06, pp 142-151.*

Omar Al-Harbi, Shaidah Jusoh, Norita Md Norwawi (2011). Lexical Disambiguation in Natural Language Question (NLQs), *IJCSI International Journal of Computer Science Issues, Vol. 8, Issue 4, No 2, July 2011 (143-150).*

Oren Melamud, Jacob Goldberger, and Ido Dagan (2016) context2vec: Learning generic context embedding with bidirectional lstm. *In proceedings of the 20th SIGNLL Conference on Computational Natural Language Learn-ing, pp. 51-61.*

Pedersen, Ted et al. (2005) "Maximizing Semantic Relatedness to Perform Word Sense Disambiguation." (2005).

R. A. Kadir, R. A. Yauri, A. Azman (2018). Semantic Ambiguos Query Formulation using Statistical Linguistic Technique, *Malaysian Journal of Computer Science), [S.l.], p. 48-56, dec. 2018. ISSN 0127-9084.*

R. Navigli and P. Velardi. (2005) "Structural semantic interconnections: a knowledge-based approach to word sense disambiguation," *in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 27, no.*

Satanjeev Banerjee and Ted Pedersen (2003) Extended gloss overlaps as a measure of semantic relatedness. *In proceedings of the IJCAI, pp. 805–810.*

S. C. Yee, T. N. Hwee, C. David. (2007) "Word Sense Disambiguation Improves Statistical Machine Translation", *Proceedings of the 45th Annual Meeting of the Association of Computational Linguistics, Prague, Czech Republic, June 2007, pp. 33–40.*

Sin-Jae Kang (2003), Corpus-based Ontology Learning for Word Sense Disambiguation. *Proceedings of the 17th Pacific Asia Conference on Language, Information and Computation, PACLIC 2003. 399-407.*

Sm Zobaed, Md Enamul Haque, Md Fazle Rabby, Mohsen Amini Salehi (2021). SensPick: Sense Picking for Word Sense Disambiguation, *16th IEEE International Conference on Semantic Computing, ICSC'2021.*

Translation. In: Locke, *W.N., Boothe, A.D. (eds.) Machine Translation of Languages, pp. 15–23. MIT Press, Cambridge.*

Weaver, W (1949). Translation. In: Locke, *W.N., Boothe, A.D. (eds.) Machine Translation of Languages, pp. 15–23. MIT Press, Cambridge.*

Z. Zhong, H. Ng. (2012) Word sense disambiguation improves information retrieval, in: *Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics, ACL '12, 2012, pp. 273–282.*