The Semantic Similarity Measures Using Arabic Ontology

مقاييس التشابه الدلالي بإستخدام الانتولوجي العربي

Prepared by

Mohammad Ghandi Aldiery

Supervisor

Prof. Ahmad Kayed

Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of Master of Computer Science

Department of Computer Science

Faculty of Information Technology

Middle East University

Amman-Jordan

January, 2017

Authorization

I, Mohammad Ghandi Aldiery, authorize Middle East University (MEU) to provide copies of my thesis to the concerned libraries, establishments, and institutions upon request.

Name: Mohammad Ghandi Aldiery

Date: 7-2-2017

Signature:

اقرار تفويض

انا محمد غاندي الديري افوض جامعة الشرق الاوسط بتزويد نسخ من رسالتي للمكتبات المعنية، المؤسسات، الهيئات عند طلبها.

الاسم: محمد غاندي الديري

التاريخ: 7-2-2017

التوقيع:

Examination Committee Decision

This is to certify that the thesis entitled "The Semantic Similarity Measures Using Arabic Ontology" was successfully defended and approved on 8-2-2017.

Examination Committee Members	**Signature**
(Chairman of Examination Committee and Supervisor) **Prof. Ahmad K. A. Kayed** *Professor* *Dean Faculty of IT* *Middle East University*	
(Internal Committee Member) **Dr. Ahmad Abu Shariha** *Assistant Professor* *Middle East University*	
(External Committee Member) **Dr. Jehad Al-Saadi** *Assistant Professor* *Arab Open University*	

Acknowledgment

First of all, I would like to express my deepest appreciation to my supervisor Prof. Ahmad Kayed for his guidance and advice through my research.

Special thanks to my parents for all the moral support and the amazing chances they've given me over the years.

I would like to thank my sisters and brothers for their continuous support.

Also special thanks to my friend Dr. Mohammad Alnababteh who has always been my greatest inspiration.

Finally, I would like to express my gratitude to my lovely wife for her support, encouragement and patience.

Table of Contents

List of Figures:

List of Tables:

List of Abbreviations:

ANLP	Arabic Natural Language Processing
AWN	Arabic WordNet
AWSS	Arabic Word Semantic Similarity
HSM	High Similarity of Meaning
IC	Information Content
IR	Information Retrieval
JCN	Jiang & Conrath
LCH	Leacock & Chodorow
LCS	Least Common Subsumer
LSA	Latent Semantic Analysis
LSM	Low Similarity of Meaning
MSA	Modern Standard Arabic
MSE	Mean Square Error
NLP	Natural Language Processing
PR	Passage Retrieval
QE	Query Expansion
RES	Resnik
WN	WordNet
WuP	Wu and Palmer.

The Semantic Similarity Measures on Arabic Ontology

Prepared by

Mohammad Ghandi Aldiery

Supervisor

Prof. Ahmad AL-Kayed

Abstract

The semantic similarity measures have been used in many applications including information retrieval and natural language processing. There are many measures that use a lexical database such as WordNet to calculate the similarity between English concepts. However, few researches have been studied semantic similarity measures using Arabic WordNet.

The traditional semantic similarity measures were classified into four categories: path-based measures, information content-based, feature-based measures, and hybrid measures. Several measures from different categories have been applied on Arabic WordNet to which measure has the best performance using Arabic WordNet. Human benchmark has been used to evaluate the performance of these measures over Arabic WordNet.

Experimental results show that the WuP measure has achieved the minimum mean square error (MSE) with value of (1.64%), and highest value of correlation coefficient with human ratings (0.92). These results indicate that WuP measure has the best performance on Arabic WordNet compared to other measures. Also, the results show that PATH measure has the worst performance.

This thesis proposed a new semantic similarity measure using the taxonomy of Arabic WordNet. The new measure takes three factors into account: depth of concepts in Arabic WordNet tree, distance between two compared concepts and information content of the least common concept that subsumed two compared concepts. The weight of these factors can be adapted manually. However, several experiments have been conducted to find the best weight that achieves the minimum MSE. In order to evaluate the new measure, the Arabic dataset that used previously to evaluate the measures has been used to test the new measure. Then, the results of applying new measure over Arabic WordNet have been compared with the results of the other measures. However, the results showed that the new measure has achieved the highest correlation coefficient with human ratings (0.96), furthermore, the new measure has obtained a very good value of MSE (1.89%) compared with the other measures.

Keywords: Ontology, Arabic ontology, WordNet, Arabic WordNet, Semantic Similarity, Similarity Measures.

مقاييس التشابه الدلالي على الانتولوجي العربي

إعداد: محمد غاندي الديري

إشراف: الأستاذ الدكتور أحمد الكايد

المُلخص

تم إستخدام مقاييس التشابه الدلالي بين الكلمات في عدة تطبيقات، منها استرجاع المعلومات و معالجة اللغات الطبيعية. هنالك العديد من هذه المقاييس التي تستخدم المعجم الالكتروني (WordNet) لحساب نسبة التشابه بين المفاهيم باللغة الانكليزية. أبحاث قليلة جدا قامت بدراسة مقاييس التشابه الدلالي بإستخدام المعجم الالكتروني العربي (Arabic WordNet).

صُنفت مقاييس التشابه الدلالي التقليدية إلى أربعة فئات : المقاييس المبنية على المسار، المقاييس المبنية على محتوى المعلومات، المقاييس المبنية على الخصائص وأخيرا المقاييس الهجينة. قامت هذه الرسالة بدراسة تطبيق هذه المقاييس على المعجم الالكتروني العربي (Arabic WordNet). عدد من المقاييس من فئات مختلفة تم تطبيقها على (Arabic WordNet) وذلك لتقييم فعاليتها. تم إستخدام معيار بشري لتقييم فعالية هذه المقاييس التي تستخدم (Arabic WordNet).

أظهرت نتائج التجارب أن المقياس (WuP) حقق أقل نسبة متوسط مربع الخطأ (MSE) بقيمة (1.64%) و أعلى قيمة إرتباط مع التقييم البشري (0.92). هذا يدل على أن المقياس (WuP) حصل على أفضل فعالية عند تطبيقه بإستخدام الوردنت العربي مقارنة بالمقاييس الأخرى. كما و أظهرت النتائج أن المقياس (PATH) حصل على أسوأ فعالية.

اقترحت هذه الرسالة مقياس تشابه دلالي جديد بإستخدام (Arabic WordNet). المقياس الجديد يأخذ بعين الاعتبار ثلاثة عوامل : عمق المفاهيم في الشجرة الدلالية,طول المسافة بين المفهومين المقارن بينهما و المحتوى المعلوماتي لأقرب مفهوم مشترك يندرج تحته المفهومين المقارن بينهما. الوزن النسبي لهذه العوامل يُعدل يدويا للحصول على الوزن النسبي المناسب. عدة تجارب أجريت لإيجاد أفضل وزن نسبي ليحقق أقل نسبة خطأ. من أجل تقييم المقياس الجديد تم استخدام نفس المعيار البشري المستخدم سابقا لتقييم المقاييس الأخرى. تمت مقارنة نتائج تطبيق المقياس الجديد بإستخام عينة الكلمات العربية على الوردنت العربي مع

النتائج التي حققتها المقاييس الأخرى. أظهرت نتائج التجارب أن المقياس الجديد حقق أعلى معامل ارتباط مع التقييم البشري بقيمة (0.96) وحصل على نسبة متوسط مربع خطأ قريبة جدا من المقياس (WuP) بقيمة (1.89%).

الكلمات المفتاحية: الانتولوجي، الانتولوجي العربي، الوردنت، الوردنت العربي، التشابه الدلالي، مقاييس التشابه.

CHAPTER ONE

Introduction

1.1 Introduction:

Rapid growth of developing traditional Arabic Natural Language Processing (ANLP) and Arabic Information Retrieval applications created the needs to explore well defined semantic similarity measures over Arabic representational vocabulary known as Arabic Ontology. Semantics is acquired by mapping an input text, as words and short texts into an ontology at which these words are getting their semantics by their relation represented in that ontology. To enable the discovery of such relation, several semantic similarity measures have been proposed in the literature.

The semantic measures have been proposed to compute the similarity between a pair of concepts in the structured model of the ontology (Slimani, 2013). Then, these measures have been used to discover the similarity between words in a free text in order to support Natural Language Processing (NLP) and Information Retrieval (IR) applications. Many researchers have studied semantic similarity measures over English ontologies. However, there is lack of researches that focus on Arabic ontology. The interest of the improvement of how to find relevant information in a language other than English is growing, specifically on the collections of information written in Arabic (Elberrichi & Abidi, 2012). Developing new semantic similarity measures over Arabic ontology will improve finding relevant information in Arabic language

1.1.1 Arabic Language

The Arabic language is very rich and complex language, handling Arabic language in NLP and IR field is hard task. The Arabic language considered as a free order with rich morphology. The Arabic letters are written from right to left (Attia, 2008). These letters take different forms based on their location in the word. Diacritics are written above or below the letters to represent the desired sound and to give a word the desired meaning. Also Arabic words show a complex internal structure, where words often incorporate affixes that mark grammatical inflections and diacritics to express different parts of speech (Faaza, James, Zuhair, & Keeley, 2012).

1.1.2 Ontology

Gruber defined ontology as "an explicit specification of a conceptualization" (Gruber, 1993). It is a model for describing the concepts and relationships between them in a hierarchical way. Ontology provides a standardized vocabulary for representing entities in the domain. Ontologies can be classified in their purpose as: general purpose ontologies and domain specific ontologies. Many researches are using ontologies as knowledge resources to measure the semantic similarity between words (Jiang et al., 2013).

1.1.3 WordNet

WordNet is the product of a research project at Princeton University (Miller, 1998). According to Meng, Huang, & Gu (2013) WordNet is a large lexical database of English. Nouns, verbs, adverbs and adjectives in WordNet are organized by set of semantic relations into synonym sets (synsets), which represent one concept. Examples of semantic relations used by WordNet are synonymy, autonomy, hyponymy, member, similar, domain and cause and so on. These relations represented as a hierarchy structure, which makes it a useful tool for computational linguistics and natural language processing (Miller, 1990). WordNet is used by many researchers to measure the semantic similarity or relatedness between a pair of concepts, since it organizes nouns and verbs into hierarchy way.

1.1.4 Arabic WordNet

Black, Elkateb, Rodriguez, and Alkhalifa (2006) developed Arabic WordNet (AWN) which is a lexical resource for Modern Standard Arabic (MSA) following the development process of Princeton WordNet for English.

AWN enables translation on the lexical level to English and dozens of other languages (Elkateb, 2006). AWN 2.0 was released in January of 2008; it contains 9,698 concepts, corresponding to 21,813 MSA words, and 6 different relation types, totaling 143,715 links. A later version of AWN, 2.0.1, is also released and contains 11,269 synsets, corresponding to 23,841 words, and 22 link types, totaling 161,705 links. AWN synsets belong to one of 5 parts of speech: noun (6,438), verb (2,536), adjective (456), adjective satellite (158), and adverb (110) (Cavalli-Sforza, 2013). AWN used in many Arabic Natural Language Processing (ANLP) and Arabic Information Retrieval applications to find common characteristics between concepts. This research will be based on AWN to implement the semantic measures and calculate similarity score between concepts.

1.1.5 Measures of Semantic Similarity and Relatedness

Measures of similarity calculate how much two concepts are alike, based on information obtained from hieratical taxonomy. For example, an automobile might be considered more similar to a boat than a tree, if automobile and boat share vehicle as a common ancestor in the taxonomy (Pederson et al., 2004). Semantic relatedness measures find how much two concepts are related to each other. Measures of relatedness are automatic methods that attempt to emulate human judgments of relatedness (Pedersen, Patwardhan, & Michelizzi, 2007). This research will study and analyze the existing semantics similarity measures; these measures will be called traditional semantics similarity measures.

According to literature, traditional semantics similarity measures can be grouped into four categories: path-based measures, information content-based measures, feature-based measures and hybrid measures.

1.1.6 Arabic Word Semantic Similarity

Few semantic similarity measures have proposed specifically for Arabic. Almarsoomi, O'Shea, Bandar, & Crockett (2013) proposed new algorithm for measuring the semantic similarity of Arabic word pairs. Arabic word semantic similarity (AWSS) method proposed by Almarsoomi, et al. calculated similarity between concepts using information sources extracted from AWN, which are length and depth. They used a previously developed Arabic word benchmark dataset (Fazza et al., 2012) to evaluate AWSS measure by calculating word similarity on an Arabic word set with human judgments. The authors state that the experimental evaluation indicates that the Arabic measure is performing well. It has achieved a correlation value of 0.894 compared with the average value of human participants of 0.893 on evaluation dataset (Almarsoomi et al., 2013).

AWSS approach based on **Li** path-based measure (Li, Bandar, & McLean, 2003), this measure used the same formula to find the similarity between two concepts, but AWSS measure used new method to find depths and lengths of concepts. However, AWSS measure does not take into account information content based measures. In this research AWSS measure will be applied along with traditional semantic similarity measures and compare its performance with these measures.

1.2 Problem Statement

There are several semantic similarity measures that have been used to measure and quantify how much two concepts are alike. However, these measures have been tested, verified and compared in English language, using WordNet (WN). Few concerns have been given to study the impacts of traditional semantic similarity measures on Arabic language, embodied in Arabic WordNet (AWN). This research aims at studying the traditional semantic similarity measures over AWN and their applicability on Arabic-related applications. Having semantic measures for Arabic language will support many Arabic-based natural language processing applications.

Problem will be accomplished by answering the following questions:

1. Which traditional semantic similarity measures can be used on AWN?
2. What is the difference between the structure of WN and the structure of AWN?
3. Which traditional semantic similarity measure has the best performance using AWN?

1.3 Methodology

This research will be combination between descriptive and quantitative methodology. This research methodology will be based on building several experiments to find the best traditional semantic similarity measures using Arabic WordNet. The methodology will include the following main steps:

1. Applying several semantic similarity measures using Arabic dataset over AWN.
2. Evaluating the applied semantic similarity measures to find best measures over AWN.
3. Propose new semantic similarity measure

4. New measure evaluation.

1.4 Objectives

The main objectives of this research are to:

- Apply seven traditional semantic similarity measures from various categories over AWN.
- Find out the appropriate semantic similarity measures that could be applied on AWN.
- Evaluate the performance of the traditional semantic similarity measures that applied on AWN.
- Propose new enhanced semantic similarity measure to obtain good performance over AWN.

1.5 Contribution

Very few researchers have studied the possibility of applying traditional semantic similarity measures on Arabic ontology. This research has applied several semantic similarity measures over AWN. This research contributes to investigating the possibility of applying traditional semantic similarity measures on AWN. Another contribution of this research is to find new adapted semantic similarity measure for AWN.

1.6 Motivation

As Arabic language spoken researchers, it's our responsibility to gain attention to this interesting and rich language. Online Arabic content is increasing rapidly, which makes developing tools and applications to handle processing of Arabic natural language very necessary. Semantic similarity measures are important part to several applications in fields such as artificial intelligence, and natural language processing and linguistics. Many semantic similarity measures have been proposed to measure the semantic similarity over English ontologies, but there is a shortage and lack of researches in measuring semantic similarity using Arabic ontology (Almarsoomi et al., 2013). These reasons motivate this research to study the applicability of applying these measures over Arabic ontology to support Arabic based applications.

1.7 Significance of the Study

This study will be a significant endeavor in finding adapted similarity measures on Arabic ontology. This research will also be beneficial to researchers in Arabic natural language processing and Arabic information retrieval field when they employ these measures in their study. Moreover, this research will provide recommendations on how to evaluate traditional semantic measures over Arabic WordNet.

1.8 Organization of the Thesis

This thesis includes five chapters, and references. The following part explains a brief description for each chapter:

Chapter 2 discusses a theoretical background and literature as follows: classifications of traditional semantic measures, Arabic ontologies, comparison between AWN and WN, and using AWN as knowledge base.

Chapter 3 introduces the methodology of this research. The research methodology has the following main steps: semantic similarity measures selection, applying the semantic similarity measures, gathering the results of all measures. This chapter also presents new hybrid measure

.Chapter 4 explains the experimental results of applying the measures on AWN. The process of evaluation all measures will be discussed in details.

Chapter 5 presents conclusion of this thesis and future work.

CHAPTER TWO

Literature Review & Related Works

Overview

This chapter introduces a theoretical background and literature that relates to this research. Literature review will be divided into four parts: first part discusses the traditional similarity measures that have been proposed and their classifications. Second part discusses the Arabic ontologies that have been proposed. Third part describes the utilization of Arabic WordNet as knowledge base. Fourth part conducts a comparison between Arabic WordNet and WordNet.

2.1 Traditional Similarity Measures

Traditional similarity measures can be classified into four categories: path-based measures, information content-based measures, feature-based measures and hybrid measures.

2.1.1 Path-based Measures

This group of measures relies on the lengths and depths of concepts that extracted from knowledge resource such as WN ontology.

Rada et al (1989) considered as pioneers in using distances between pair of concepts to measure the similarity between them. In their work knowledge based taxonomy viewed as a graph, concepts represented as nodes and relation between concepts represented as edges. This measure uses edge counting method to find the shortest path between two concepts. Therefore, the shortest path length used to calculate the similarity score between concepts (Rada et al, 1989).

Wu & Palmer (1994) introduced a measure of semantic similarity based on both depths and lengths in the taxonomy (Wu & Palmer, 1994). WuP measure takes into account the length between concepts, *C1* and *C2,* as well as the length between the *LCS* and the root of the taxonomy in which the concepts located.

$$sim_{wp}(C1,C2)= \frac{2*depth(LCS(C1,C2))}{depth(C1)+ depth(C2)} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (2.1)$$

Where *depth(C)* is the depth of the synset *C* using edge counting in the taxonomy, *LCS(C1,C2)* is the least common subsumer of *C1* and *C2*. *depth(LCS(C1,C2))* is the length between *LCS* of *C1* and *C2* and the root of taxonomy. If *LCS(C1,C2)* is the root of taxonomy, then *depth(LCS(C1,C2))=1*.

The disadvantage of this method, that two pairs with the same LCS and same lengths of shortest path will have the same similarity. Since we can find LCS in AWN, this measure is simply implemented using AWN.

Leacock & Chodorow's measure is also based on depths and lengths which are information sources in the taxonomy, taking into account the maximum depth of taxonomy and the length between *c1* and *c2* (Leacock & Chodorow, 1998).

$$Sim_{LCH}(C1,C2)= -log \frac{len(C1,C2)}{2*deep\_max} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (2.2)$$

Where the *deep_max* is depth of c_i in the taxonomy. Current version of WN has nine noun taxonomies and the maximum depth is 20. Current version of AWN has also nine noun taxonomies and the maximum depth is 15. The maximum depths of the taxonomies changes considerably. It is clear that this measure can easily implemented in AWN since the information sources that this measure uses are available in AWN.

Li et al (2003) proposed new approach in finding semantic similarity score between word pairs by using multiple information sources in the taxonomy, which are shortest distance between two compared words, and the depth of least common subsumer in the taxonomy, therefore this measure combine the length and depth as follows:

$$sim(c1,c2) = e^{-\alpha * len(c1,c2)} \frac{e^{\beta * depth(lcs(c1,c2))} - e^{-\beta * depth(lcs(c1,c2))}}{e^{\beta * depth(lcs(c1,c2))} + e^{-\beta * depth(lcs(c1,c2))}} \ldots\ldots\ldots\ldots\ldots (2.3)$$

Where parameter α and β need to be adapted manually for good performance. The optimal parameters are $\alpha = 0.2$ and β=0.6.

PATH measure is simple method that uses only one information source which is path length distance between concepts. PATH measure has been introduced to work with semantic taxonomy nets. The distance between concepts is found by counting the node (Michelizzi, 2005). Similarity score between two concepts is calculated as:

$$sim_{PATH}(c1,c2) = \frac{1}{dist_{node}(c1,c2)} \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots (2.4)$$

Where $dist_{node}(c1,c2)$ is the distance between concepts c1 and concept c2 using node counting.

Slimani et al (2006) presented an extension of WuP similarity measure. This measure have been introduced to overcome the following disadvantage of WuP measure: in some cases, the similarity of two concepts in the ontology contained in the neighborhood exceeds the similarity value of two concepts contained in the same hierarchy. According to the author the main objective of the proposed measure is to obtain realistic results for concepts not located in the same way (Slimani et al., 2006).

Noted that all path-based measures rely only on the distances between concepts, and the weight of the concept itself is not taken into account.

Path-based measures depend on two information sources in the taxonomy which are length of the path between synsets and the depth of concepts in the taxonomy. In this approach distance between synsets in the taxonomy quantifies the similarity score (Michelizzi, 2005). The more the distance between synsets, the less similar they are. AWN followed the development process of English WordNet and can be used as a graph by path-based measures to compute the similarity between Arabic concepts. In figure 2.1 **أم** (mother) synset is closer to **والدان** (parent) than it is to **قريب** (relative), and therefore it is more similar to **والدان** (parent) than **قريب** (relative). The distance between two synsets can be calculated using either edge counting or node counting. In edge counting, the distance between two synsets is measured by counting the number of links between two synsets. In node counting the distance between two synsets is calculated by counting the number of nodes along the shortest path between the two synsets. For example in figure 2.1 the distance between **أم** (mother) and **والدان** (parent) is one using edge counting, and two using node counting. Depth of synset is the path length between the synset itself and the root of taxonomy. The depth can be also calculated either by edge counting or node counting.

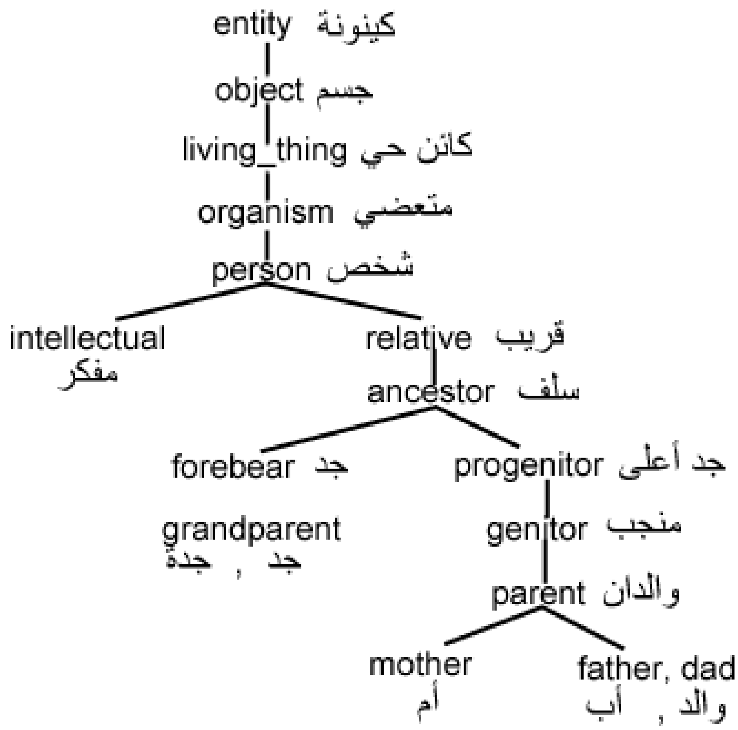

Figure 2.1: A fragment of is-a relation in AWN

A subsumer is a shared parent of two synsets. The least common subsumer (LCS) of two synsets is the most shared parent that subsumed the two synsets. For example in figure 2.1, the LCS of both **أم** (mother) and **مفكر** (intellectual) is **شخص** (person).

2.1.2 Information Content-based Measures

This family can be grouped into two groups; first group is corpus-dependent information content measures. These measures uses statistical analysis extracted from corpus to computes the similarity value. Second group is corpus-independent information content, unlike the first group, this group doesn't rely on the corpus, and instead, these measures use information sources extracted from WN ontology.

2.1.2.1 Corpus-dependent Measures

Resnik (1995) proposed information content corpus based similarity measure, based on the notion of information content. It assumes that the similarity between two concepts is calculated by finding how much shared information is between them. Therefore, the more common information between concepts, the more similar they are. In this method the ontology used to find the instances of concepts, then corpus is used to obtain the frequencies of concepts. According to author this measure is the first to combine ontology and a corpus together (Resnik, 1995).

$$sim_{Res}(c1, c2) = -log\, P(\, LCS(c1, c2)\,) \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (2.5)$$

Resnik's method computes the IC through calculating the probabilities of concepts occurring in the corpus.

$$\boldsymbol{IC(c) = -\,log\, p(c)} \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (2.6)$$

Where *P(c)* is the probability that a randomly selected word in a corpus is an instance of concept *c*. For a given concept, each observed noun is either a member of that concept with probability P(c) not a member of that concept with probability *1-P(c)*. The probability of root node in the taxonomy is the maximum value, *P (root) =1*. The lower a node in hierarchy, the lower its probability. Probability of a concept was estimated as:

$$p(c) = \frac{freq(c)}{N} \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (2.7)$$

The drawback of Resnik semantic similarity measure is that all pairs of synsets with the same LCS will have the same similarity score.

Jiang & Conrath (1997) proposed new method to find a semantic distance between concepts based on information content of compared words and most common subsume. However, this distance converted to represents the similarity score. Like resnik measure, this method used a corpus in addition to a hierarchal taxonomy (Jiang & Conrath, 1997).

$$Dist_{jcn}(c1, c2) = IC(c1) + IC(c2) - 2 * IC(LCS(c1, c2)) \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots (2.8)$$

Semantic similarity is the opposite of the distance:

$$\frac{1}{Dist_{jcn}(c1,c2)} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (2.9)$$

Lin (1998) proposed another information content method, but unlike resnik approach, it doesn't take only the information content of the most shared subsumer into a account, but it takes into an account the information content of two compared concepts. This method assumes that the information content weight of compared concepts should be considered to measure the similarity score (Lin, 1998). The similarity between c1 and c2 is calculated by the ratio between the amount of information needed to state the commonality of c1 and c2 and the information needed to fully describe what c1 and c2 are.

$$sim_{Lin}(c1, c2) = \frac{2\, logP(LCS(c1,c2))}{logP(c1)+logP(c2)} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (2.10)$$

The IC value of LCS is less than or equal to the IC of both concepts *c1* and *c2*, therefore the values of this measure are vary between 1 and 0. As noted from formula (11) if the IC of LCS is zero, then the similarity score is zero and the score is zero if both concepts *c1* and *c2* are zero.

All the above information content-based measures are corpus based, the IC of concepts are calculated using corpus. To apply these measures over AWN, Arabic corpus with diacritics is needed to count the same Arabic words forms as the same concepts, because Arabic words with the same form and without diacritics may have different meaning, for example Arabic word form "رجل" may has two different meaning "رَجُل " (man) and "رِجْل " (leg). In Arabic language the same word form with the same meaning may have different diacritics in another context. Therefore calculating the IC of Arabic concepts using Arabic corpus is hard to implement due to the ambiguity problem.

2.1.2.2 Corpus-independent Measures

Seco (2004) used WordNet as a statistical resource instead of using a corpus to obtain information content (IC) value of concepts. This measure assumes that the more the concept has hyponyms, the more abstract it is. Therefore, the concepts with many children hold less information than concepts that are leaves. Since the root node has the largest numbers of hyponyms, then it is the least informative. Thus, leaf concepts located at the bottom of the tree have the maximum information content value (Seco, 2004). The IC of root node is zero, and the IC of leaf is one. The IC value of a given concept can be calculated as follows:

$$IC(c) = 1 - \frac{log(hypo(c)+1)}{log(node\_max)} \quad \text{......................} (2.11)$$

Sánchez (2011) introduced another corpus independent measure to compute the IC value of concepts. This method takes taxonomical leaves into an account to determining the generality value of concepts, the more the concepts has leaves the more specific it is. Therefore, the group of leaves subsumed in a concept, is fair enough to define its scope. According to author, this method compared with corpora dependent-based methods and obtained better correlation with human benchmark (Sánchez, 2011).

$$IC(c) = -\log\left(\frac{\frac{|leaves(c)|}{|subsumers(c)|}+1}{\max\_leaves+1}\right) \qquad \ldots\ldots\ldots\ldots\ldots\ldots (2.12)$$

Meng et al (2012) presented new corpora–independent method relies on nodes' topology in WordNet. This method takes into an account the depth of concept itself, number of hyponyms and the depth of each hyponym subsumed by that concept. It based on the assumption the topology structure and design of nodes in the taxonomy affects the IC value of concepts (Meng et al, 2012). The authors developed new method (Res_Meng) to calculate the semantic similarity between a pair of concepts based on Resnik approach. Res_Meng measure computes the similarity score by finding the IC value of the LCS.

$$sim_{Res\_Meng}(C1, C2) = \frac{log(depth(LCS(C1,C2)))}{log(deep\_max)} \qquad \ldots\ldots\ldots\ldots\ldots\ldots (2.13)$$

2.1.3 Feature-based Measures

Feature-based similarity measures have been proposed to find how much concepts are related to each other. Unlike the above semantic similarity measures, these measures use different information sources, which are glosses and relations.

Tversky's measure takes into consideration the properties of the concepts to calculate the similarity between two compared concepts in the taxonomy. Information sources, such as path length and information content of concepts are ignored in this measure. Each concept in the taxonomy has a description that contains a set of words represent the features of concept. Shared features between concepts increase the similarity between them. Non-common features between concepts decrease the similarity between them (Tversky, 1977).

Lesk's measure counts overlapping words in glosses of two compared words to find relatedness between them (Lesk, 1987). This measure is based on idea that the more compared concepts have common words in two glosses, the more related they are. Both the number of words in the overlaps and the length of the overlaps are taken into account when calculating semantic relatedness score. The relation functions between synsets are used to determine which glosses are to be compared. Each relation functions pair creates a score, the total relatedness score is the sum of the scores for each pair of relation function. The score for one pair of relation function calculated as follow:

$$pairscore = \sum_{i}^{\#overlaps} length^{2}\,(overlapi) \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{(2.14)}$$

The total relatedness score is the sum of each of these pair's scores:

$$Relatedness(s1, s2) = \sum_{i}^{\#pair\ scores} pair\ score_{j} \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{(2.15)}$$

Lesk's measure has limitations over AWN, due to the very few number of glosses in AWN (Zouaghi et al., 2011). It is possible to attach glosses to the concepts since current version of AWN is open source.

Another feature-based measure that uses glosses to find the similarity between two concepts was proposed by Patwardhan (2003). This measure based on context vectors that combines the glosses content of concepts in the taxonomy with statistical information extracted from the corpus. One advantage of this measure over the Lesk's measure is that the vector method is not limited to finding the same matches between glosses. According to the author, this measure does not rely on the topology of any particular ontology (Patwardhan, 2003).

Zouaghi et al (2011) modified Lesk algorithm, using the different semantic similarity measures to find the similarity relatedness between two concepts in AWN. They replaced the original measure of Lesk by five semantic similarity measures which are used to find the gloss that corresponds to the correct sense of the ambiguous word. The authors developed this method to solve the problem of missing glosses in AWN (Zouaghi et al., 2011).

2.1.4 Hybrid Measures

Hybrid measures are based on the idea of combining multiple methods from the above measures. There are hybrid measures that use both information content and path length of concepts to compute the similarity between two compared concepts. This measure uses the following information sources to calculate the similarity: IC of concepts, lengths between concepts, max depth in the taxonomy and weight factors which can be adapted manually:

$$sim_{zhou}(c_1, c_2) = 1 - k\left(\frac{\log(len(c_1, c_2) + 1)}{\log(2 * (deep\_\max - 1))}\right) - (1 - k) * ((IC(c_1) + IC(c_2) - 2 * IC(lso(c_1, c_2)) / 2) \quad \dots (2.16)$$

Where parameter k needs to be adapted manually.

The advantage of this measure is that the weight of concept itself has been distinguished.

Hybrid semantic similarity measures are applicable on AWN, except the measures that used feature-based measures.

As noted from the above discussion most of semantic similarity measures are applicable. However some measures have limitations over AWN such as feature-based measures and corpus-dependent information content measures. Table 2-3 illustrates the applicability of traditional semantic similarity measures over AWN.

Table 2-3: Applicability of traditional semantic measures on AWN

Measures	Applicable on AWN	Reasons	Type
Path	Yes	Path information source available in AWN	Path-based
WuP	Yes	It depends on length and depth information sources which are available in AWN	
LCH	Yes	count of edges between and log smoothing	
Li	Yes	non-linear function of the shortest path and depth of lso	
Resnik	Not yet	Problem in finding Arabic word frequency with diacritics, and data sparse problem.	IC corpus-dependent
Lin	No yet	Problem in finding Arabic word frequency with diacritics, and data sparse problem.	
Res_Meng	Yes	It depends on depth of LCS and max depth in AWN	IC corpus-independent
Lesk	Has limitations	Glosses does not available in current version of AWN	Feature-based measures
Zhou	Yes	It combines two applicable measures	Hybrid measures

2.2 Arabic Ontologies

Several Arabic ontologies have been developed for supporting Arabic natural language processing. Arabic ontologies are very important for measuring the similarity between Arabic concepts. The most well-known Arabic ontology is Arabic WordNet, section 2.4 will discuss Arabic WordNet ontology in details.

Al-Yahya et al (2010) proposed a computational model for describing Arabic concepts using ontologies. The model has been built using data that obtained from Holy Quran. The new model can easily be extended and linked to other ontologies such as SUMO. The model has been implemented on the Arabic language vocabulary related to "Time" vocabulary in the Holy Quran. According to the authors, Results of the evaluation show that the model is able of describing word semantics in a way that can support semantic analysis of Arabic words and several useful applications (Al-Yahya et al., 2010).

Jarrar presented a methodology for developing a formal Arabic ontology. The proposed work has been taken into an account the semantic relations between concepts instead of words. Unlike WordNet, the proposed Arabic ontology focuses on actual properties of concepts. Jarrar emphasizes that building the Arabic ontology and creating Arabic content should be based on ontological principles (Jarrar, 2011).

Mazari et al (2012) proposed an approach of automatic construction on an Arabic linguistic ontology using statistical techniques to extract entities of ontology from Arabic corpus. The author used "repeated segment" technique to determine the related items that represent main concepts in the domain. They also used "co-occurrence" of extracted concepts to define relations between these concepts in the ontology. To accomplish extraction process the authors used previously prepared Arabic corpus that has been collected from Arabic books and articles (Mazari et al., 2012).

Ishkewy et al (2014) presented an Arabic lexical ontology called Azhary. Like AWN it classifies Arabic words into sets of synsets. Azhary contains 26,195 words, grouped into 13,328 synsets. This ontology has been built a number of relations between words such as synonym, hypernym, hyponym, antonym, holonym and association relations. Authors depend on the Holy Quran to create the seed words the relations between Arabic words have been built manually by using well-known dictionaries. According to the authors, Azhary ontology has larger words and relations between words than AWN (Ishkewy et al., 2014).

2.3 Using AWN as a Knowledge Base

Many researches are using English WordNet as knowledge base to extract useful information, which is used in NLP and IR. In other hand, a few research used AWN as knowledge base.

Elberrichi & Abidi (2012) used Arabic WordNet (AWN) as a lexical and semantic resource for Arabic texts categorization. In their experimental work, they used AWN as a tool map Arabic terms to concepts, and to find similar words (synsets) and representing them in one concept. According to the author, this is the first study that used AWN in Arabic texts classification (Elberrichi & Abidi, 2012).

Abouenour, et al. (2012) presented core modules of a new Arabic question answering system called IDRAAQ. These modules aim at enhancing the quality of obtained passages with respect to a given question. Arabic WordNet in this work is also used as semantic resource, to obtain the semantic relations for given words. Unlike the traditional semantic measures, this work doesn't use AWN to extract information sources such as path and depth of Arabic concepts (Abouenour et al., 2012).

Imam, et al. (2013) introduced Ontology-based Summarization System for Arabic Documents (OSSAD), Domain knowledge is extracted from an Arabic corpus and represented by topic related concepts and the lexical relations among them. The user's query is first expanded by using the Arabic WordNet and then by adding the domain-specific knowledge base to the expansion (Imam et al., 2013).

Abderrahim, et al. (2013) implemented the method of semantic indexing of the documents and query for the information retrieval where are use Arabic WordNet as a semantic resource to exploring the impact of passage from an indexation based on single words to an indexation based on concepts (Abderrahim et al., 2013).

Almarsoomi, et al. (2013) introduced the first semantic similarity measure that has been proposed for Arabic word pairs, they used two information sources in the taxonomy, which are the path distances between concepts and depth (Almarsoomi et al., 2013). They have used knowledge-based approach to calculate the similarity between two Arabic concepts using the latest version of AWN. Their approach has been based on the assumption that similarity score of word pairs increases if the depth of LCS increases as we go deeper in a hierarchical taxonomy. This method extracts shared and distinct properties from AWN to compare two Arabic words. In order to test their work, a previously Arabic dataset benchmark is used (Fazza et al., 2012). The semantic similarity score between two Arabic words over AWN is calculated using the following formula:

$$sim\ (w1, w2) = e\hat{}\,(- \alpha * l) * \tanh\ (\beta * d) \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (2.17)$$

Where α and β are the length and depth factors respectively, obtained at $\alpha = 0.162$ and $\beta = 0.234$.

In this thesis AWSS measure will be applied, to compare its performance with the other semantic similarity measures.

2.4 Comparison between WordNet and Arabic WordNet

This section will conduct a comparison between WordNet (WN) and Arabic WordNet (AWN). Knowing the difference between WN and AWN will help in study the applicability of traditional semantic similarity measures over AWN.

Unlike traditional dictionaries, WordNet is organized words by meaning, rather than word forms, words in close proximity are semantically similar. WordNet considered as a useful knowledge based tool for several semantic similarity measures and used in many natural language processing applications (Miller, 1990). Word senses in WordNet are organized into synonym sets or synsets. A word sense is a given meaning of a word. For example, in figure 2-1 shows that the word cord has four meanings, as a noun. Word sense can be represented as a string by using the word form. This string followed by single letter to represent the part of speech, then followed by a sense number, as shown in figure 2-1, the part of speech letter is **n** for nouns, **v** for verbs, **a** for adjectives, and **r** for adverbs, for example cord#n#3 represents the third sense.

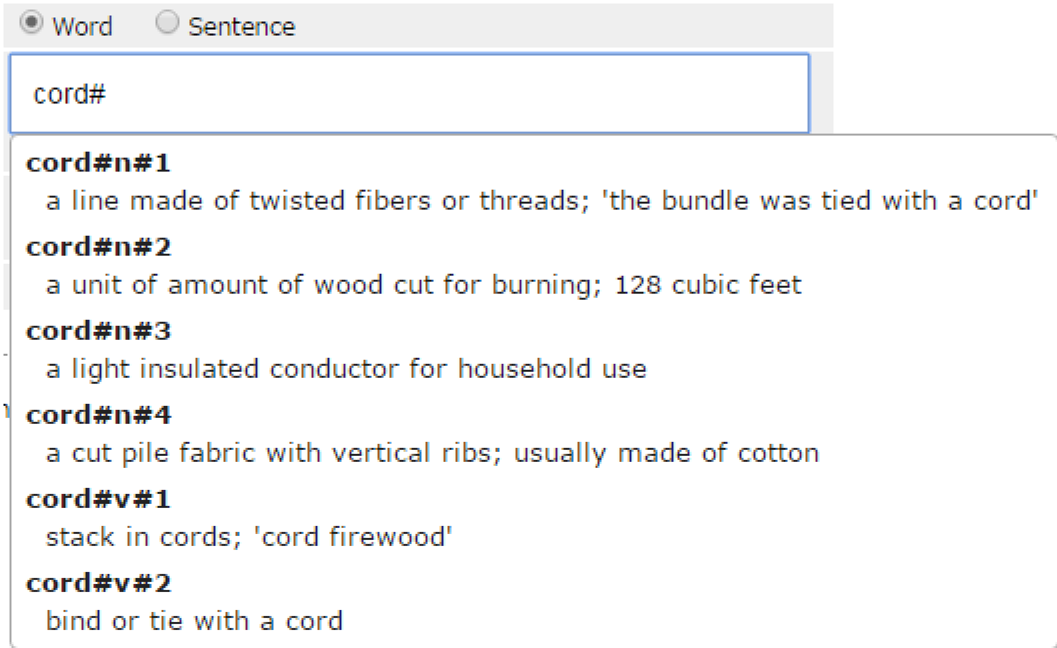

Figure 2-1: Senses of cord in WS4J online tool

WordNet is organized by semantic relations between synsets. Some examples of semantic relations are the synonymy, hypernym, hyponym and meronym relations (Meng et al., 2013). Synonymy is one concept that is expressed by several different word forms that have the same similar meaning, for example {hit, beat, strike} represented as a synonymy (synset). Hypernym relation represent is-a relationship between word meanings, hypernym is general concept for the synset that subsumed by it. Hyponym is the opposite hypernym, which represents the instance of general synset, for example a car is a kind of vehicle. The meronymic relation is a has-a relation and can be used to construct a part hierarchy, for example a finger is part of a hand.

Latest version of WordNet is 2.1[1] for windows, released in March 2005; Version 3.0 for Unix/Linux/Solaris/etc. was released in December, 2006. Table 2-1 illustrates some statistics about WordNet 2.1.

[1] https://wordnet.princeton.edu/wordnet/download/

Table 2-1: Statistics of synsets in WordNet 2.1

Part of speech	Word Forms	Synsets containing word forms
Noun	111,798	82,115
Verb	11,529	13,767
Adjective	21,479	18,156
Adverb	4,481	3,621
Total	155,287	117,659

Arabic WordNet (AWN) is Arabic semantic knowledge source based on the structure and contents of the Princeton WordNet (PWN) and mapped directly onto PWN 2.0 and EuroWordNet (EWN). Most of the synsets of AWN should be linked to English WN, and the structure of AWN hierarchy followed the same WN topology (Elkateb, 2006).

AWN mapped with the Suggested Upper Merged Ontology (SUMO), which is a formal ontology of about 1000 concepts and 4000 axioms and 750 rules. It is provided in a first order logic language called Standard Upper Ontology Knowledge Interchange format (SUO-KIF) (Pease, 2000). SUMO has been mapped by WordNet of 100,000 noun, verb, adjective and adverb senses (Almarsoomi et al., 2013).

AWN is built in two phases by first building a core WordNet around the most general concepts called base concepts (Vossen 1998), these base concepts encoded as synsets in AWN, other Arabic specific concepts are added and translated manually to the relative synset. Second phase is to extend the core WN downward to the lower level of concepts in the hierarchy (Elkateb, 2006).

The database structure of AWN contains four entity types: item, word, form and link. Items are the synsets, each item has unique identifier and brief description called gloss. A word entity is a word sense. A form entity contains lexical information. A link represents the relation between synsets, examples of relation type are, related_to, has_hyponym, verb_group, has_holo_member and has_derived.

In AWN few Arabic synsets have a translated gloss attached; latest version of the AWN browser comes with an integrated automatic Arabic gloss generator. The generation process works by first obtaining an unglossed Arabic synset and then trying to describe this synset in terms of its surrounding synsets in the tree hierarchy, but the glosses does not really exist in the database. Figure 2-2 shows empty gloss values in the AWN xml file.

```
"" POS="r" source="" gloss="" authorshipid="6522" />
>
 POS="n" source="" gloss="" authorshipid="6523" />
>
 POS="n" source="" gloss="" authorshipid="6524" />
>
" POS="s" source="" gloss="" authorshipid="6525" />
>
" POS="a" source="" gloss="" authorshipid="6526" />
.
" POS="a" source="" gloss="" authorshipid="6527" />
>
" POS="a" source="" gloss="" authorshipid="6528" />
.
" POS="n" source="" gloss="" authorshipid="6529" />
.
" POS="n" source="" gloss="" authorshipid="6530" />
.
" POS="n" source="" gloss="" authorshipid="6531" />
.
" POS="n" source="" gloss="" authorshipid="6532" />
.
" POS="n" source="" gloss="" authorshipid="6533" />
```


Figure 2-2: Empty glosses in xml of AWN database

Latest version of AWN, 2.0.1, contains 11,269 synsets, corresponding to 23,841 words, and 22 link types, totaling 161,705 links. AWN synsets belong to one of 5 parts of speech: noun (6,438), verb (2,536), adjective (456), adjective satellite (158), and adverb (110). Table 2-2 illustrates some statistics about WN and AWN.

Table 2-2: Statistics of synsets in WN and AWN

	WN	AWN
Noun synsets	82,115	6,438
Verb synsests	13,767	2,536
Adjective synsets	18,156	614
Adverb synsets	3,621	110
Total synsets	117,659	9,698

From the information above, AWN followed the structure of WN and has same topology in organizing synsets in the hieratical taxonomy. As illustrated in table 2-2 AWN has few numbers of synsets, 11,269 synsets considered as a small number for rich language such as Arabic. As shown in table 2-2 the total number of synsets in AWN is much less than total number of synsets in WN. The difference of total synsets between AWN and WN will make the distances between concepts in WN longer than distances in AWN. Figure 2-3 compares length path to the root of the word **bus** (حافلة) in AWN and WN tree. Figure 2-3 shows that the depth (length to the root) of the word حافلة (bus) in AWN is 8 and the depth of the word **bus** in WN is 13. Many Arabic words are not found in AWN like فرن (Stove), ساحر (Magician),تل (Hill) and مستشفى (hospital).

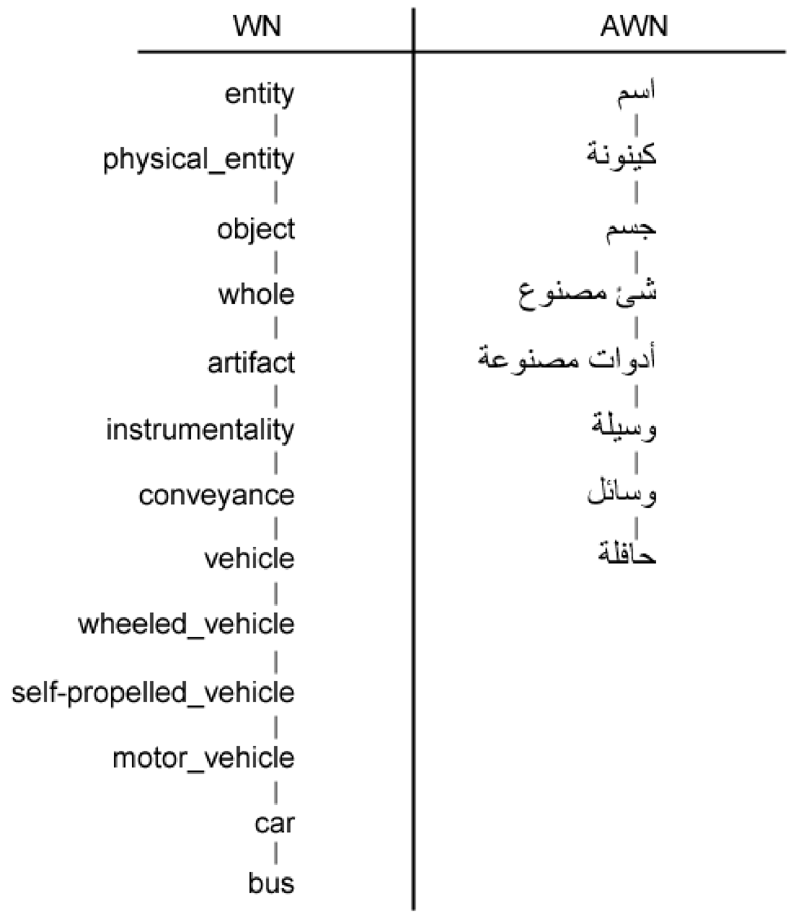

Figure 2-3: depth of the word bus in AWN and WN

2.5 Arabic Dataset Benchmark Used

In this thesis Arabic dataset benchmark called AWSS has been used. This dataset was created by Faza et al (2012), the Arabic dataset uses the same procedures which followed in creating English dataset benchmarks for semantic similarity. The most two common benchmark datasets are Rubenstein & Goodenough R&G (1965) and Miller & Charles (M&C) (1991). To the best of our knowledge there is no Arabic benchmark datasets for semantic similarity except AWSS by Fazza et al (2012).

The AWSS benchmark dataset was prepared mainly in two steps, first, determine the Arabic word pairs set, second, specify human similarity rate for word pairs. The AWSS creators fundamentally used the dataset of Rubenstein & Goodenough R&G (1965). Fazza et al (2012) created a list of Arabic word pairs contains 70 item. They follow the same steps of R&G (1965). 27 Arabic categories were created and employed to select the stimulus Arabic word pairs and to promote the best possible semantic representation. Arabic categories were created based on Rubenstein & Goodenough method, the list of English words in the R&G experiment contains 48 nouns from 22 different categories. In AWSS another five categories added to expand 22 categories to be 27 categories. The 48 English noun pairs from R&G list have been used to create the 22 Arabic categories after translated into Arabic language using English-Arabic dictionary and checked their accuracy from professional translator and fluent lecturers, the categories specified based on the definition of the selected pairs (Rubenstein & Goodenough, 1965). After the 22 categories specified, new 5 categories are added, the added categories relevant to Arabic life style. After that, the first two nouns from each category are selected to generate 56 stimulus Arabic words (Fazza et al 2012).

The 56 noun pairs are divided into two columns, 28 nouns in each column. A sample of 22 Arabic native speakers from 5 different Arabic countries was chosen to generate two sets of Arabic noun pairs ranging from high similarity of meaning (HSM) to medium similarity of meaning (MSM) and low similarity. the participant asked to write 28 Arabic noun pairs which have high similarity from the list by selecting one noun from **Column** A and other from Column B, and write 32 pairs have medium similarity by the same procedure of selecting high similarity pairs. The participants while selecting can choose the same word more than one time without duplicating the pairs. After the list has been processed the final list was contains 57 Arabic noun pairs. Then 13 Arabic noun pairs from low similarity were randomly selected by Faza et al (2012). In order to get list from 70 Arabic word pairs which covered high to low similarity, this list called AWSS. Table 2-3 shows AWSS list.

Table 2-3: AWSS dataset benchmark (Fazza et al 2012)

	Word Pairs		Human Ratings	أزواج الكلمات			Word Pairs		Human Ratings	أزواج الكلمات	
1	Coast	Endorsement	0.03	تصديق	ساحل	36	Slave	Lad	1.77	فتى	عبد
2	Noon	String	0.03	خيط	ظهر	37	Journey	Bus	1.83	باص	رحلة
3	Cushion	Diamond	0.06	الماس	مسند	38	Girl	Odalisque	1.96	جارية	فتاة
4	Gem	Pillow	0.07	مخدة	جوهرة	39	Feast	Fasting	1.96	صيام	عيد
5	Stove	Walk	0.07	مشي	موقد	40	Coach	Means	2.07	وسيلة	حافلة
6	Cord	Midday	0.08	ظهيرة	حبل	41	Brother	Lad	2.15	فتى	أخ
7	Signature	String	0.08	خيط	توقيع	42	Sage	Sheikh	2.26	شيخ	حكيم
8	Boy	Endorsement	0.12	تصديق	صبي	43	Girl	Sister	2.38	أخت	فتاة
9	Boy	Midday	0.16	ظهيرة	صبي	44	Hill	Mountain	2.60	جبل	تل
10	Slave	Vegetable	0.16	خضار	عبد	45	Hen	Pigeon	2.61	حمامة	دجاجة
11	Smile	Village	0.18	قرية	إبتسامة	46	Master	Sheikh	2.66	شيخ	سيد
12	Smile	Pigeon	0.20	حمامة	إبتسامة	47	Food	Vegetable	2.78	خضار	طعام
13	Wizard	Infirmary	0.22	مشفى	ساحر	48	Slave	Odalisque	2.84	جارية	عبد
14	Noon	Fasting	0.29	صيام	ظهر	49	Run	Walk	3.01	مشي	جري
15	Hill	Pigeon	0.33	حمامة	تل	50	Brother	Sister	3.08	أخت	أخ
16	Countryside	Laugh	0.34	ضحك	ريف	51	Cord	String	3.09	خيط	حبل
17	Glass	Diamond	0.36	الماس	كأس	52	Forest	Woodland	3.14	أحراش	غابة
18	Glass	Fasting	0.38	صيام	كأس	53	Sage	Thinker	3.30	مفكر	حكيم
19	Cord	Mountain	0.54	جبل	حبل	54	Gem	Diamond	3.38	الماس	جوهرة
20	Hospital	Grave	0.83	قبر	مستشفى	55	Cushion	Pillow	3.38	مخدة	مسند
21	Forest	Shore	0.86	شاطئ	غابة	56	Journey	Travel	3.39	سفر	رحلة
22	Gem	Young woman	0.87	شابة	جوهرة	57	Countryside	Village	3.41	قرية	ريف
23	Sepulcher	Sheikh	0.89	شيخ	ضريح	58	Smile	Laugh	3.48	ضحك	إبتسامة
24	Tool	Pillow	0.99	مخدة	اداة	59	Stove	Oven	3.55	فرن	موقد
25	Coast	Mountain	1.06	جبل	ساحل	60	Coast	Shore	3.56	شاطئ	ساحل
26	Run	Shore	1.13	شاطئ	جري	61	Signature	Endorsement	3.58	تصديق	توقيع
27	Hill	Woodland	1.19	أحراش	تل	62	Tool	Means	3.68	وسيلة	اداة
28	Countryside	Vegetable	1.24	خضار	ريف	63	Noon	Midday	3.70	ظهيرة	ظهر
29	Tool	Tumbler	1.32	قدح	اداة	64	Boy	Lad	3.71	فتى	صبي
30	Master	Thinker	1.36	مفكر	سيد	65	Girl	Young woman	3.74	شابة	فتاة
31	Feast	Laugh	1.36	ضحك	عيد	66	Sepulcher	Grave	3.75	قبر	ضريح
32	Hen	Oven	1.44	فرن	دجاجة	67	Wizard	Magician	3.76	مشعوذ	ساحر
33	Journey	Shore	1.47	شاطئ	رحلة	68	Coach	Bus	3.80	باص	حافلة
34	Coach	Travel	1.60	سفر	حافلة	69	Glass	Tumbler	3.82	قدح	كأس
35	Food	Oven	1.76	فرن	طعام	70	Hospital	Infirmary	3.91	مشفى	مستشفى

Another 60 participants from different Arabic countries who had not taken part in generating Arabic word pairs were asked to rank the set of 70 Arabic word pairs previously collected. The participants were requested to rate each word pair based on how similar they were in meaning from 0.0 to 4.0 (Fazza et al 2012). In this work, the human rating is divided by four to convert the rating from [0-4] range to [0-1].

In this thesis AWSS benchmark dataset has been chosen for various reasons as follows: first reason is that the Arabic word pairs were created carefully. Second, this benchmark was based on R&G dataset, which is the most influential word dataset for English. The original Arabic dataset contains 24 low similarity, 24 medium similarity and 22 high similarity word pairs. Due to absence of some words in AWN and technical issue in the tool that we used, only 40 word pairs are taken. Sub dataset in this experiment contains 12 word pairs low similarity, 13 word pairs medium similarity and 15 high similarity word pairs.

2.6 Tools Used

Several tools have been used in this research, these tools used for two purposes, first purpose was to study and analyze the structure of both WN and AWN, the second purpose was to applying the semantic measures over AWN

2.6.1 WordNet 2.1 Browser

Provides a window-based interface for browsing the WordNet database, allowing synsets and relations to be displayed as formatted text. For each search word, different searches are available based on syntactic category and information available in the database. Figure 2-4 shows the user interface WordNet 2.1 Browser.

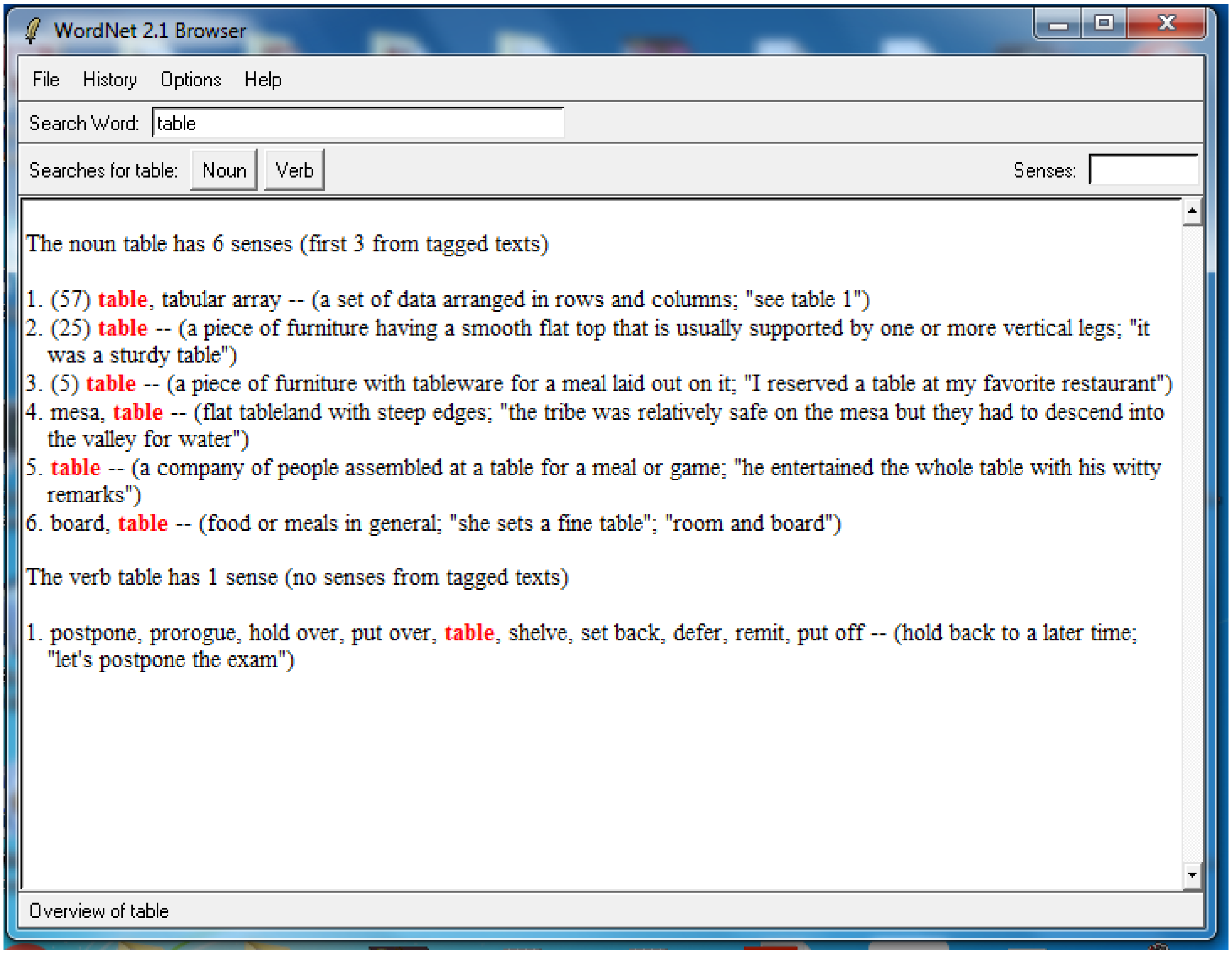

Figure 2-4: WordNet 2.1 Browser user interface

2.6.2 Arabic WordNet Browser

Arabic WordNet browser[1] provides easy interface to search and browse Arabic concepts. The main features of the AWN Browser are as follows:

1. Browsing the AWN: AWN browser represents Arabic concepts in tree. Selecting items from the tree causes English synonyms and gloss to be displayed, as well as Arabic translations if they exist.

[1] http://globalwordnet.org/arabic-wordnet/awn-browser/

2. Searching for Arabic concepts in the AWN: AWN browser supports search for Arabic concepts. Arabic searches may be carried out using either words (entered with or without diacritics) or roots..

3. Updating Arabic data: The AWN Browser has an open source database stored locally, but it provides facilities to update this database automatically from online server.

Figure 2-5 shows the interface of Arabic WordNet browser.

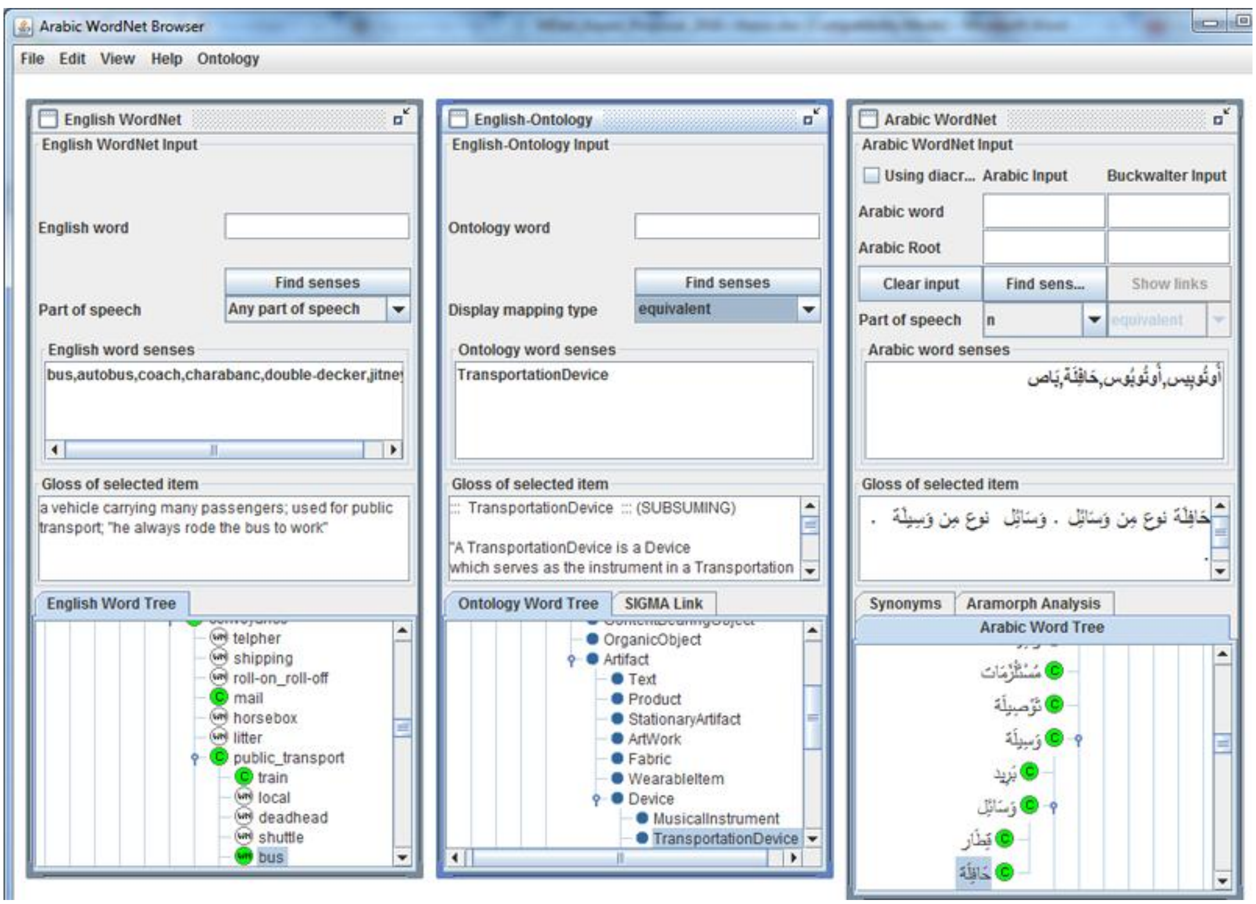

Figure 2-5: Arabic WordNet browser interface

2.6.3 WordNet Similarity for Java (WS4J)

WS4J is an online tool used to measure semantic similarity between two concepts or between two sentences over WordNet. It uses several measures to calculate the similarity scores between words. WS4J tool provides useful information about how each measure calculates the similarity score. Figure 2-6 shows WS4J online demo interface.

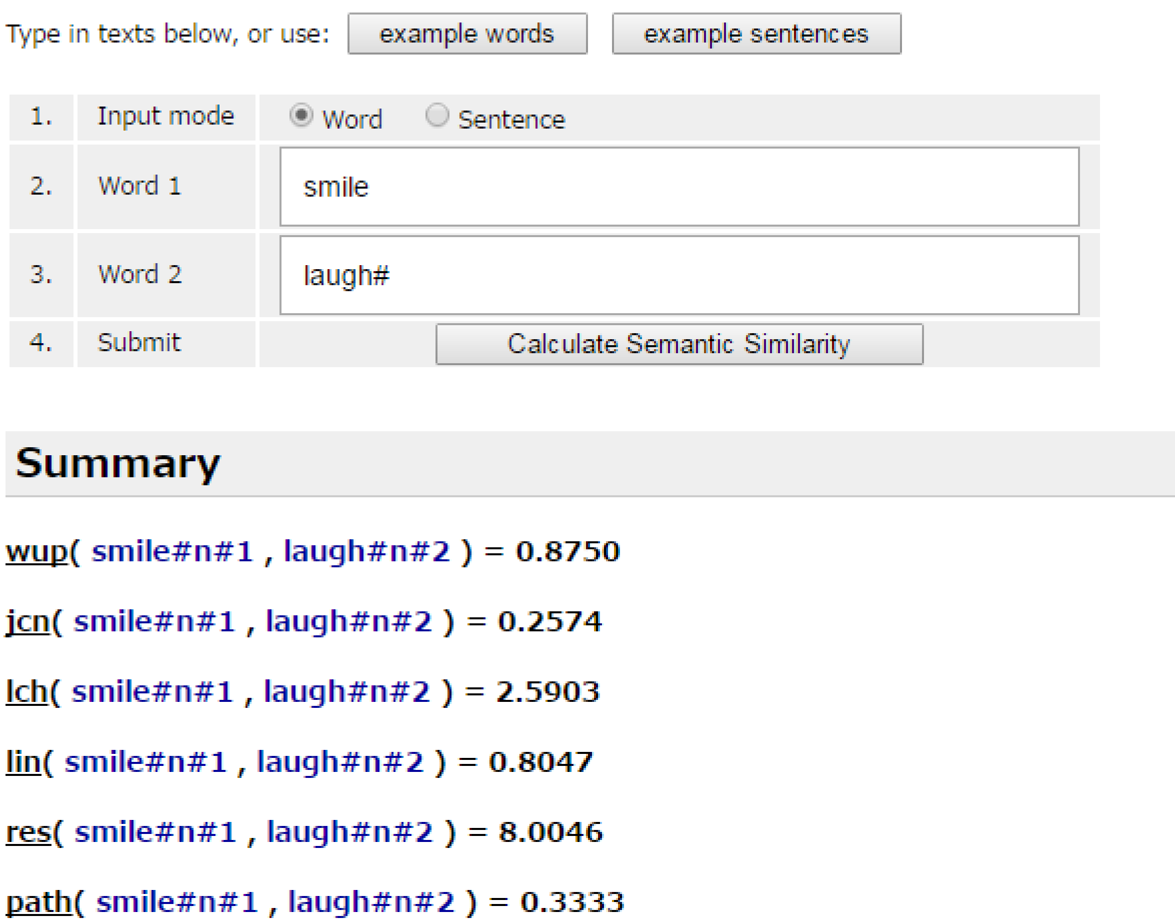

Figure 2-6: WS4J interface

2.6.4 Java API for AWN

Free open source java code that Access XML database for AWN, it has 35 built-in functions. It includes four semantic similarity measures methods: **Path** (Get_word_similirty_edge_counting), **WuP** (Get_word_similirty_WuP), **LCH** (Get_word_similirty_LeakcockChodorow) And **Li** (Get_word_similirty_Li).

2.6.5 NLTK Python Library

Nltk[1] python library is well-known library that provides easy to use interfaces to many lexical resources such as WordNet, along with the built in functions for text processing methods such as classification, tokenization, and similarity calculation. Python 2.7[2] or greater should be installed to run this library. English WN is already installed with NTLK library, to install AWN, database of AWN as xml file should be downloaded[3], and then download and install AWNDatabaseManagement.py[4]. Unfortunately there are no built-in functions to calculate similarity score for AWN, few functions available for AWN like describe and get synsets. Figure 2-7 illustrates how to import AWN in NTLK library.

[1] http://www.nltk.org/
[2] https://www.python.org/downloads
[3] http://nlp.lsi.upc.edu/awn/get_bd.php
[4] http://nlp.lsi.upc.edu/awn/AWNDatabaseManagement.py.gz

test2.py - C:\Python34\test2.py (3.4.4rc1)

File Edit Format Run Options Window Help

```
from AWNDatabaseManagement import wn
from nltk.corpus import wordnet
synsets = wn.get_synsetids_from_word(u"جَمِيل")
w1=wn.get_synsetids_from_word(u"شَاب")
w2=wn.get_synsetids_from_word(u"شَابَة")
wn._words["jamiyl_1"].describe()
```

Figure 2-7: Include AWN in NLTK library

Java AWN API and WS4J will be used in this thesis. Java AWN API is a trusted tool for research and it is accepted from various committee for applying semantic similarity measures on AWN. java AWN API contains implementations of four semantic similarity, WuP, LCH, LI and path. Additionally it gives information sources like number of hyponyms for concepts, depth of the concepts in the taxonomy and path length between concepts. Therefore, in this thesis we apply the four mentioned measures as well as additional measure called Resnik which based on the information provided from AWN.

CHAPTER THREE

Experimental Work & New Proposed Measure

Overview

In this research the well-known semantic similarity measures been applied using Arabic WordNet (AWN) in order to study their performance over AWN. New hybrid semantic similarity measure over AWN has been presented. The results of applying the new measure have been compared with traditional semantic similarity measures in order to evaluate the new measure. This chapter explains in details the main step of the research methodology.

Introduction

The methodology of this research combined the descriptive and quantitative approach. The proposed methodology will use a quantitative research by building several experiments to apply the semantic similarity measures over AWN. The results of running these experiments have been used in the evaluation process in order to find the best measures over AWN. The evaluation process has been based on human benchmark. Thus, the evaluation part has been done by calculating the error which is the difference between the human result and measures results. The results of experiments have been studied to present new measure. The following will illustrate the main steps of the research methodology as shown in figure 3-1:

1. Semantic similarity measures selection.
2. Applying the semantic similarity measures using Arabic dataset over AWN.
3. Gathering the results of all measures and evaluation.
4. Propose new semantic similarity measure.
5. New measure evaluation.

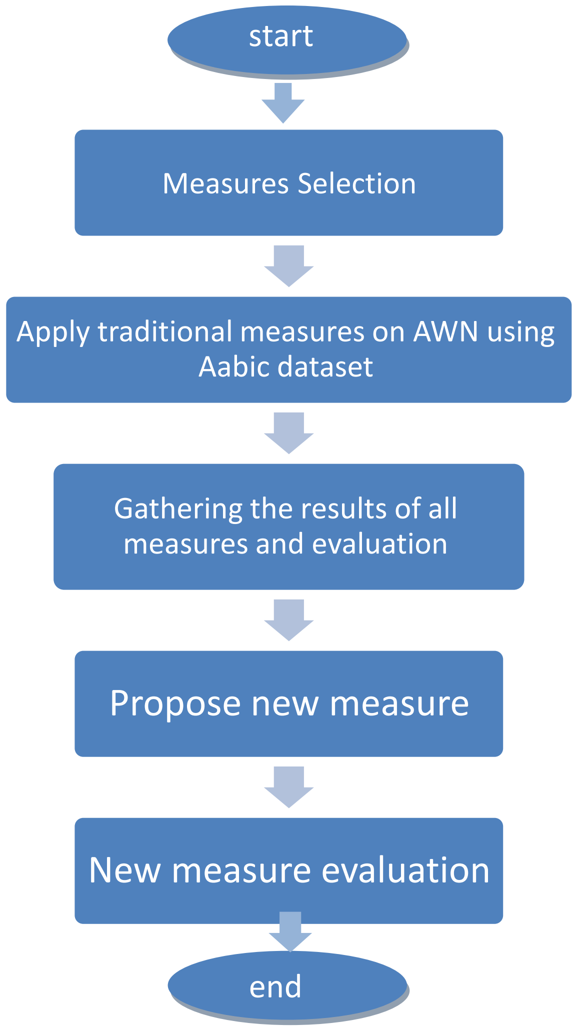

Figure 3-1: Flowchart of the proposed work

The methodology will contain the following steps in details:

3.1 Semantic Similarity Measures Selection

There are many semantic similarity measures based on WN to compute the semantic similarity between two concepts. These measures are divided into four categories, the path-based measures, information content measures, feature-based measures and hybrid measures (Slimani, 2013). In this thesis seven well-known measures from three categories (path-based measures, information content measures and hybrid measure) are selected to study their applicability on AWN. The feature-based measures use the glosses of the concepts which are provided in WN (Meng et al., 2013). However, these glosses are not available in AWN, therefore feature-based measures will not be applied in this thesis. The selected measures in this thesis are:

1. WuP: is path based measure uses the distance between concepts and the depth of the LCS in the taxonomy to compute the semantic similarity.

2. PATH measure: is path-based measure uses the length of the path between concepts to computer the semantic similarity.

3. LCH: is path-based measure uses the length of the path between concepts and the max depth of the taxonomy.

4. Li: is path-based measure uses non-linear equation function based on the length between concepts and the depth of the concepts in the taxonomy.

5. AWSS: is Arabic path-based measure uses LI formula to compute semantic similarity with modification on the depth and length computation to be proper for AWN.

6. Res_Meng.: is node-based measure, also known as information content measure. In this measure we compute the IC using corpus independent method called IC_{meng}.

7. Zhou: is hybrid measure, uses two different measures families, path based measures and information content measures.

The above seven measures consist three path-based measures, two non linear path-based measures and one information content measure. The first three measures are linear path-based measures, and they are selected because they achieve good performance against other measures.

The fourth measure (Li) selected because it is non-linear path based measure, as well as it's the reference measure of AWSS. Fifth measure (AWSS) is selected for experiment because it has been developed especially for AWN, and to compare its result on Arabic dataset against the results of the other five measures. As shown previously the sixth measure is corpus independent measure, there are various corpus dependent measures, but we didn't use them due to the ambiguous and sparse data problem. The seventh measure is selected because it represents hybrid measure category. Table 3-1 illustrates the reasons of selecting each measure.

Table 3-1: reasons of measure selection

Measure	Reason to use
WuP	Uses depth of concepts. Applied to study the effect of concept depth on AWN.
PATH	Uses length of shortest path between concepts. Applied to study the effect of distances between concepts.
LCH	Takes max of depth information source into consideration.
LI	Uses non-linear function.
Res_Meng	Represents information content-based measures and corpus-independent.
AWSS	Developed especially to use AWN. Applied to compare its performance against other measures.
Zhou	Represents hybrid measures

3.2 Applying the Traditional Measures on AWN

In this section we will study the possibility of using the traditional semantic similarity measures on Arabic ontology that are implemented over English ontology and other languages.

The results of this study will give the researchers in Arabic natural language processing good knowledge about the semantic similarity measures that could use in AWN.

The experiments in this section performed according to the following steps:

1- Choosing the proper tools for applying the seven semantic similarity measures over AWN.

2- Handling the AWN in order to be compatible with the selected tool.

3- Applying the seven traditional semantic similarity measures using the selected tool.

4- Extracting the result of implementing the seven semantic similarity measures from the tool.

5- Analyzing and comparing the results of applying the semantic similarity measures over AWN.

3.2.1 Computing the Semantic Similarity Using Java AWN API

In this section the semantic similarity measures will be applied 40 Arabic noun pairs which were selected from AWSS dataset using the java AWN API, and the result for each measure will be described and analyzed. Then, the obtained result from java AWN API for all measures will be compared with human ratings. The process for applying the measures using this tool will as the following steps:

1. Run java AWN API using integrated development environment (IDE).
2. Import the AWN as xml file to the tool.

3. Determine the item-id for all 40 Arabic word pairs.

4. Apply each semantic similarity measure on all Arabic word pairs and write down the result.

To run java source code we need an integrated development environment (IDE) to compile the code and printout the results; we used eclipse IDE, which is mostly used for developing java applications.

In order to run java AWN API tool, we should import the Arabic WordNet (AWN) as xml file. To import the AWN to java AWN API, the path of the AWN xml file should be passed to the tool. Arabic WordNet browser[1] is application available on the internet contains the Arabic WordNet database, the AWN browser gives us the ability to export its database as xml file, figure 3-3 shows how to export xml file from Arabic WordNet browser.

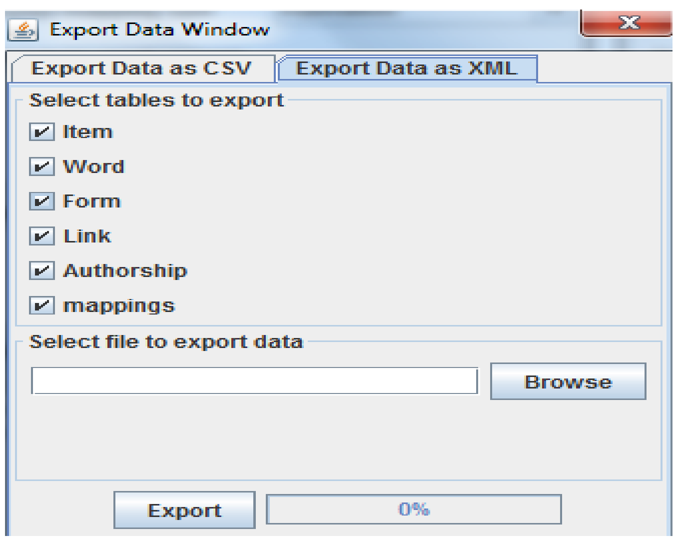

Figure 3-3: Arabic WordNet browser GUI

[1] http://globalwordnet.org/arabic-wordnet/awn-browser/

The exported AWN xml file contains Arabic Synsets, words, forms and links between them. The xml file contains 5 types of nodes, which are:

1- **Item node**: it has information about the synsets, represented by properties like item_id, name, source, offset and gloss, for example synset طبيب (doctor) has synset id "Tabiyb_n1AR" figure 3-4 shows how طبيب (doctor) represented in AWN xml.


```
<item itemid="Tabiyb_n1AR" offset="109380179" lexfile="0" name="طبيب" type="synset" he
<authorship author="sabri" date="20060314" score="0" comment="manchester20060717" cove
<item itemid="TabiyoEiy_a1AR" offset="302845411" lexfile="" name="طبيعي" type="synset'
<authorship author="musa" date="20060621" score="" comment="" covering="1" authorship:
```


Figure 3-4: Item node in AWN xml file

2- **Authorship node:** it has information about author as shown in figure 3-4.

3- **Word node**: it has information about Arabic words, such as synset-id of the word, word value and the word id.

4- **Form node**: it has value, wordid, type and authorshipid as shown in figure 3-5.


```
<word wordid="$axoSiy~ap_1" value="شخصية" synsetid="$axoSiy~ap_n1AR" frequency="1"
<form value="شخص" wordid="$axoSiy~ap_1" type="root" authorshipid="11692" />
```


Figure 3-5: Word and form node in AWN xml file

5- **link node**: it contains the relationship between synsets, examples of relation type are, related_to, has_hyponym, verb_group, has_holo_member and has_derived. Figure 3-6 illustrates how link represented in AWN xml.

```
<link type="related_to" link1="taEal~ama_v1AR" link2="&lt;inojaAz_n2AR"
<authorship author="horacio" date="20080225" score="0" comment="from en
<link type="has_hyponym" link1="taEal~ama_v1AR" link2="&gt;EAd_taEal~vm
<authorship author="horacio" date="20080225" score="0" comment="from en
```

Figure 3-6: Link node in AWN xml

After exporting xml file, the path of the exported xml file should be passed to the java AWN API in order to import the xml file. The tools contain a set of methods and classes to handle it. The first class has been used was AWN class, this class enable us to import the AWN xml file, it takes two parameters, the first parameter is the path of AWN xml file, the second parameter is "true" or "false", to tell the API to remove diacritics (harakat) from the source, "false" parameter should be passed, in our case we need diacritics, so "true" has been passed. The following code shows how to use the class.

AWN aw= new AWN("upc_db.xml",true);

As mentioned above, we have applied the selected semantic similarity measures to all Arabic word pairs in the dataset, this step took a lot of time and effort, because we need to get **synset-id** for all word pairs, this has been done by two steps as follows:

1. We have used AWN browser to get Arabic synonyms with diacritics by typing Arabic concept in Arabic word filed, then choosing proper word sense from the list appeared in Arabic word senses box as shown in figure 3-7, thus Arabic word with diacritics copied to be used in java AWN API tool.

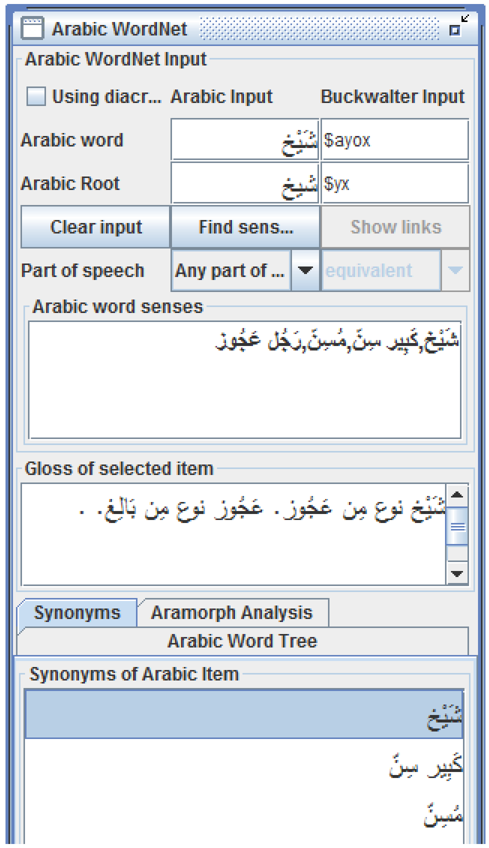

Figure 3-7: Arabic word senses box in AWN browser

2. Arabic word with diacritics have been passed to ***Get_Item_Id_From_Name*** method in java AWN API to get **synset ID** as follows:

```
List<String> ItemID= aw.Get_Item_Id_From_Name("شَيْخ");
System.out.println(ItemID);
```

The above two steps have been repeated for all Arabic noun pairs, all collected **synsets IDs** have have been stored into an excel file as shown in figure 3-8.

Arabic word1	SynId1	Arabic word2	SynId2
شَيْخ	$ayox_n1AR	ضَرِيح	qabor_n1AR
حَافِلَة	HaAfilap_n1AR	وَسِيلَة	wasiylap_n1AR
سَيِّد	say~id_n1AR	شَيْخ	$ayox_n1AR
طَعام	TaEAm_n1AR	خُضار	xuDaAr_n1AR
مَشْي	ma$oy_n1AR	جَرْي	jaroy_n1AR
صَبِيّ	Sabiy~_n2AR	فَتًى	muraAhiq_n1AR
أَدَاة	>adaAp_n2AR	وَسِيلَة	wasiylap_n1AR
مُفَكِّر	mufak~ir_n1AR	سَيِّد	say~id_n1AR
تَل	rukaAm_n1AR	جَبَل	jabal_n1AR
دَجَاجَة	dajaAjap_n1AR	حَمَامَة	HamaAm_n1AR
حَافِلَة	HaAfilap_n1AR	وَسِيلَة	wasiylap_n2AR
حَكِيم	fayolasuwf_n1AR	مُفَكِّر	mufak~ir_n1AR

Figure 3-8: Arabic word noun pairs with synset IDs

The semantic similarity for all Arabic noun pairs have been computed by Java AWN API tools. As said previously this tool has only 4 measures, namely, **edge counting** (*Get_word_similirty_edge_counting*), **WUP** (*Get_word_similirty_WuP*), **Leakcock and Chodorow** (*Get_word_similirty_LeakcockChodorow*) And **Li** (*Get_word_similirty_Li*). For the two measures (Resnik$_{meng}$ and Zhou), we developed two new methods. Arabic word pairs were already implemented by AWSS measure (Almarsoomi et al., 2013).

The similarity has been computed using the measures methods. To perform that, the **synset ID** for Arabic word pairs should pass to the methods of the measures in java AWN API to return the similarity score between them. For example if we need to find the similarity score between شيخ (Sheikh) and ضريح (Sepulcher), we should pass **synset ID** for both concepts as follows:

```
System.out.println(aw.Get_word_similirty_WuP("$ayox_n1AR","qabor_n1AR"));
System.out.println(aw.Get_word_similirty_Li("$ayox_n1AR","qabor_n1AR",0.2,0.6));
System.out.println(aw.Get_word_similirty_LeakcockChodorow("$ayox_n1AR","qabor_n1AR"));
System.out.println(aw.Get_word_similirty_edge_counting("$ayox_n1AR","qabor_n1AR"));
```

The measures score for similarity between Arabic words pairs printed out in the Eclipse console user interface as shown in figure 3-9.

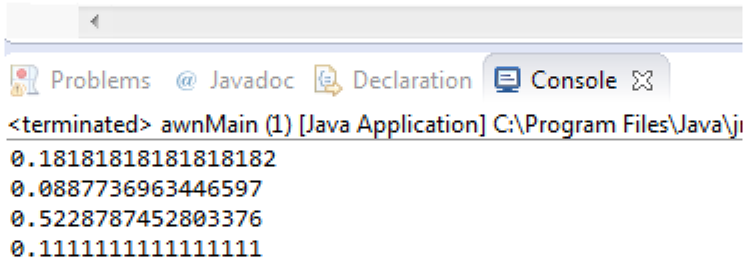

Figure 3-9: Similarity scores in Eclipse console interface

As shown in the example, the semantic similarity measures called using the interface of the method, the WuP measure called by write `Get_word_similirty_WuP("$ayox_n1AR","qabor_n1AR").`

Before demonstrating and analyzing the results of applying the semantic similarity measures on all Arabic noun pairs and because all Arabic noun pairs will translate to English in order to compute the semantic similarity to all of them. The tool that will be used to compute the semantic similarity for English noun pairs will illustrate briefly.

Finding semantic similarity scores for English word pairs will help in evaluation process during the results comparison. The computation of semantic similarity for English word pairs is much easier by using the online tools.

WS4J[1] online tool gives the ability to compute the similarity between English concepts by simply typing the two words then click on calculate semantic similarity button as shown in figure 3-10.

[1] http://ws4jdemo.appspot.com/

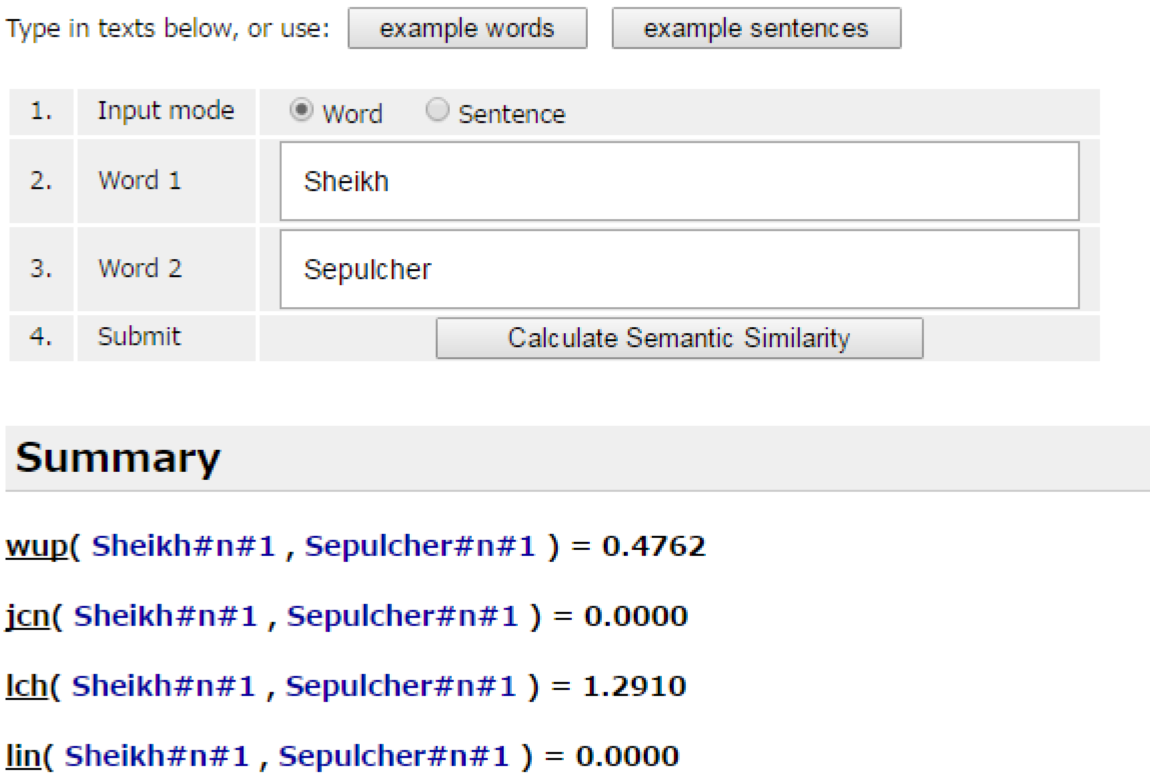

Figure 3-10: interface of WS4J online tool

3.3 Gathering the Results for All Measures

After calculating the similarity score for all Arabic word pairs and English word pairs using the above mentioned techniques, we have gathered the similarity scores values for each measure. The results of applying the selected measures have been collected into seven tables. The collected data in these tables will help in study the performance of the applied measure over AWN as will be shown in chapter 5.

3.4 New Hybrid Measure

This thesis presents new hybrid measure to compute the semantic similarity between a pair of Arabic concepts using Arabic WordNet. As stated in formula (3.1), the proposed measure takes three factors into an account:

1. Depth of concepts in AWN tree, this factor represented by formula (2.1).
2. Distance between two compared concepts, this factor represented by formula (2.4).
3. Information content of LCS, this factor represented by formula (2.13).

$$\boldsymbol{Sim(c1,c2)=tanh}\left(\frac{2*depth(LCS)}{depth(c1)+depth(c2)}+W*\left(\frac{1}{len(c1,c2)}+\frac{log(depth(LCS))}{log(Max-depth)}\right)\right) \quad .....(3.1)$$

Formula (3.1) contains three operands to compute the similarity score, as noted from the above formula, weight value has been given for these operands; the first operand has been multiplied by one, second and third operands have been multiplied by adapted weight ***W.*** In order to find the best value of ***W*** we have conducted several experiments by applying formula (3.1) on the Arabic word pairs that we have used in the previous experiments. Our experiments have been done by adapting ***W*** value to find the lowest MSE. As shown in table 3-2, the lowest MSE value was obtained at $W = 0.5$. We have used hyperbolic function in our formula to normalize the result between 0 and 1.

Table 3-2: adapted vlaues of *W*

W	**MSE**
0.3	0.025409
0.4	0.020182
0.5	**0.018932**
0.6	0.020761
0.7	0.023029
0.9	0.028132

Table 3-3 shows how the lowest MSE value has been obtained when calculating the similarity of Arabic word pairs using the formula (3.1). We have collected the needed information of Arabic word pairs using Java API tool. Then we have computed the similarity scores by applying the collected values on the formula (3.1). For example the similarity between Arabic word pair جري (run) and مشي (walk) can be calculated by applying their information that are located in the table 3-3 on the formula as follows:

***Sim*(مشي, جري)=** $\boldsymbol{tanh}\left(\frac{2*5}{6+6}+\boldsymbol{W}*\left(\frac{1}{2}+\frac{log(5)}{log(Max-depth)}\right)\right)$

Where *Max-depth* in the current version of AWN is 15, and ***W*** =0.5. After substituting the values and doing the calculation in the above formula, we find that the similarity between جري (run) and مشي (walk) is 0.87. The calculated values of **Error** and **Square Error** have been used to calculate the MSE value that helps us in adapting ***W*** value.

Table 3-3: calculating MSE at *W=0.5*

C1	C2	Depth(LCS)	Depth(c1)	Depth(c2)	Len(C1,C2)	Similarity	Human Ratings	Error	Square Error
تصديق	ساحل	0	5	5	-	0	0.01	0.01	0.0001
خيط	ظهر	0	6	5	-	0	0.01	0.01	0.0001
مشي	**موقد**	-	-	-		0	0.01	-	-
ظهيرة	حبل	0	6	6	-	0	0.02	0.02	0.0004
خيط	توقيع	0	4	4	-	0	0.02	0.02	0.0004
تصديق	صبي	0	5	5	-	0	0.03	0.03	0.0009
ظهيرة	صبي	0	6	5	-	0	0.04	0.04	0.0016
قرية	إبتسامة	0	5	8	-	0	0.05	0.05	0.0025
صيام	ظهر	0		6	-	0	0.07	0.07	0.0049
الماس	كأس	1	7	9	14	0.15	0.09	-0.06	0.0036
ضريح	شيخ	1	5	6	9	0.23	0.22	-0.01	0.0001
خضار	ريف	1	6	5	9	0.23	0.31	0.08	0.0064
قدح	أداة	2	4	4	8	0.59	0.33	-0.26	0.0676
عيد	ضحك	1	6	7	11	0.19	0.34	0.15	0.0225
جارية	فتاة	3	5	6	14	0.64	0.49	-0.15	0.0225
صيام	عيد	2	6	9	8	0.41	0.49	0.1	0.01
وسيلة	حافلة	5	7	8	5	0.78	0.52	-0.26	0.0676
شيخ	حكيم	3	7	6	7	0.62	0.56	-0.06	0.0036
أخت	فتاة	3	5	6	5	0.68	0.60	-0.08	0.0064
حمامة	دجاجة	9	11	12	5	0.85	0.65	-0.2	0.04
جبل	**تل**	-	-	-	-	-	0.65	-	-
شيخ	سيد	3	6	6	6	0.65	0.67	0.02	0.0004
خضار	طعام	2	6	4	6	0.54	0.69	0.15	0.0225
جارية	عبد	4	5	4	2	0.82	0.71	-0.11	0.0121
مشي	جري	5	6	6	2	0.87	0.75	-0.12	0.0144
خيط	حبل	4	6	6	4	0.77	0.77	0	0
أحراش	غابة	8	9	9	1	0.94	0.79	-0.15	0.0225
مفكر	حكيم	4	4	6	2	0.86	0.82	-0.04	0.0016
سفر	رحلة	5	5	6	1	0.93	0.84	-0.09	0.3347
ألماس	جوهرة	5	7	5	2	0.88	0.84	-0.04	0.0016
قرية	ريف	4	5	5	1	0.91	0.85	-0.06	0.0036
مخدة	مسند	5	7	8	6	0.82	0.85	0.03	0.0009
ضحك	إبتسامة	5	8	8	6	0.75	0.87	0.12	0.0144
توقيع	تصديق	4	5	5	2	0.91	0.89	-0.02	0.0004
وسيلة	أداة	5	6	7	2	0.86	0.92	0.06	0.0036
ضريح	قبر	5	5	5	1	0.94	0.93	-0.01	0.0001
فتى	صبي	4	5	5	1	0.92	0.93	0.01	0.0001
مشعوذ	**ساحر**	-	-	-	-	-	0.94	-	-
حافلة	باص	8	8	8	1	0.95	0.95	0	0
قدح	كأس	7	9	9	2	0.87	0.95	0.08	0.0064
								MSE=	**0.018932**

CHAPTER FOUR

Experimental Results & Measures Evaluation

Overview

This chapter discusses the results of applying traditional semantic similarity measures over Arabic WordNet. The results have been used to evaluate the performance for all measures. Each measure will be evaluated and compared with the other measures. In this chapter, the new hybrid measure will be evaluated and compared with other measures to study its performance over Arabic WordNet.

4.1 Results of Applying All Measures and Evaluation

The results of applying the selected semantic similarity measures over AWN have been collected into seven tables. The collected data in these tables have been used in measures evaluation process.

Table 4-1 shows the results of WuP measure, the table contains the Arabic word pairs and their translations, the Arabic word pairs have been translated into English word pairs in order to be applied over WN. The results of applying Arabic and English word pairs have been compared to study the differences between AWN and. The table includes **Human Rating** column which contains the human judgment similarity score of the Arabic noun pairs, this score has been used to be compared with computer based result (i.e output of applying WuP measure) , human based score is considered as benchmark to compute the error rate of the computerized semantic similarity measure. Table also contains two columns (**EN**, **AR**) to show the similarity score of WuP for English and Arabic pairs. The last two columns (**Err**, **Sqr Err**) in the table contain the Error which is the difference between the computed similarity score by WuP and human rating score, and the square error to compute the mean square error. Human rating and the results of measures columns have been divided have into three groups, these are: low similarity, medium similarity and high similarity. We have applied the same form and structure of table 4-1 to create tables for other measures, the tables can be found in the appendix.

Evaluation process in this thesis carried out by finding two factors. The first factor is a correlation coefficient between similarity measure score and human rating. The correlation coefficient has been considered to study the strength of relation between human judgment and similarity scores calculated by machine. The stronger the association between human ratings and similarity scores calculated by applied measures, the closer the correlation coefficient will be to one Furthermore, the correlation coefficient between machine similarity scores and human ratings for each group (i.e low, medium and high) has been calculated separately to figure out which group's result has the strongest relation with human ratings. Second factor is mean square error (MSE) of measures results; the smaller value of MSE is the better measure accuracy.

Table 4-1 WuP measure results

NO.	Sim. level	Word Pairs		Arabic word pairs		Human Ratings	EN	AR	Err.	Sqr. Err.
1	Low Similarity	Coast	Endorsement	تصديق	ساحل	***0.01***	0.28	0	0.01	0.0001
2		Noon	String	خيط	ظهر	***0.01***	0.35	0	0.01	0.0001
3		Stove	Walk	**مشي**	**موقد**	***0.01***	**0.16**	**-**	-	-
4		Cord	Midday	ظهيرة	حبل	***0.02***	0.21	0	0.02	0.0004
5		Signature	String	خيط	توقيع	***0.02***	0.23	0	0.02	0.0004
6		Boy	Endorsement	تصديق	صبي	***0.03***	0.23	0	0.03	0.0009
7		Boy	Midday	ظهيرة	صبي	***0.04***	0.28	0	0.04	0.0016
8		Smile	Village	قرية	إبتسامة	***0.05***	0.37	0	0.05	0.0025
9		Noon	Fasting	صيام	ظهر	***0.07***	0.36	0	0.07	0.0049
10		Glass	Diamond	الماس	كأس	***0.09***	0.35	0.12	-0.03	0.0009
11		Sepulcher	Sheikh	ضريح	شيخ	***0.22***	0.47	0.18	0.04	0.0016
12		Countryside	Vegetable	خضار	ريف	***0.31***	0.40	0.18	0.13	0.0169
13	Medium similarity	Tumbler	Tool	قدح	أداة	***0.33***	0.73	0.5	-0.17	0.0289
14		Laugh	Feast	عيد	ضحك	***0.34***	0.40	0.15	0.19	0.0361
15		Girl	Odalisque	جارية	فتاة	***0.49***	0.83	0.54	-0.05	0.0025
16		Feast	Fasting	صيام	عيد	***0.49***	0.5	0.18	0.31	0.0961
17		Coach	Means	وسيلة	حافلة	***0.52***	0.77	0.66	-0.14	0.0196
18		Sage	Sheikh	شيخ	حكيم	***0.56***	0.76	0.46	0.1	0.01
19		Girl	Sister	أخت	فتاة	***0.60***	0.40	0.54	0.06	0.0036
20		Hen	Pigeon	حمامة	دجاجة	***0.65***	0.84	0.78	-0.13	0.0169
21		Hill	Mountain	**جبل**	**تل**	***0.65***	0.85	-	-	-
22		Master	Sheikh	شيخ	سيد	***0.67***	0.90	0.5	0.17	0.0289
23		Food	Vegetable	خضار	طعام	***0.69***	0.85	0.4	0.29	0.0841
24		Slave	Odalisque	جارية	عبد	***0.71***	0.72	0.66	0.05	0.0025
25		Run	Walk	مشي	جري	***0.75***	0.90	0.83	-0.08	0.0064
26	High Similarity	Cord	String	خيط	حبل	***0.77***	0.94	0.66	0.11	0.0121
27		Forest	Woodland	أحراش	غابة	***0.79***	1	0.88	-0.09	0.0081
28		Sage	Thinker	مفكر	حكيم	***0.82***	0.85	0.8	0.02	0.0004
29		Journey	Travel	سفر	رحلة	***0.84***	0.95	0.90	-0.06	0.0036
30		Gem	Diamond	ألماس	جوهرة	***0.84***	0.95	0.83	0.01	0.0001
31		Countryside	Village	قرية	ريف	***0.85***	0.77	0.80	0.05	0.0025
32		Cushion	Pillow	مخدة	مسند	***0.85***	0.94	0.57	0.28	0.0784
33		Smile	Laugh	ضحك	إبتسامة	***0.87***	0.87	0.62	0.25	0.0625
34		Signature	Endorsement	توقيع	تصديق	***0.89***	0.94	0.8	0.09	0.0081
35		Tools	Means	وسيلة	أداة	***0.92***	0.82	0.76	0.16	0.0256
36		Sepulcher	Grave	ضريح	قبر	***0.93***	0.94	1	-0.07	0.0049
37		Boy	Lad	فتى	صبي	***0.93***	0.95	0.88	0.05	0.0025
38		**Wizard**	**Magician**	**مشعوذ**	**ساحر**	***0.94***	1	-	-	-
39		Coach	Bus	حافلة	باص	***0.95***	1	1	-0.05	0.0025
40		Glass	Tumbler	قدح	كأس	***0.95***	0.94	0.77	0.18	0.0324
									MSE=	0.016475676

Table 4-1 shows that WuP measure has obtained a good value of MSE (0.016475). MSE values for each similarity group (i.e. low, medium and high) were calculated separately. MSE value for high similarity group is (0.01740). Low and medium similarity group have the same MSE value (0.0027). These results indicate better performance for WuP in high similarity.

WuP measure has obtained a high value of correlation coefficient (0.94) with human ratings, this means that WuP measure has good linear relation with human rating. Figure 4-1 shows the correlation between human ratings and the scores of WuP measure.

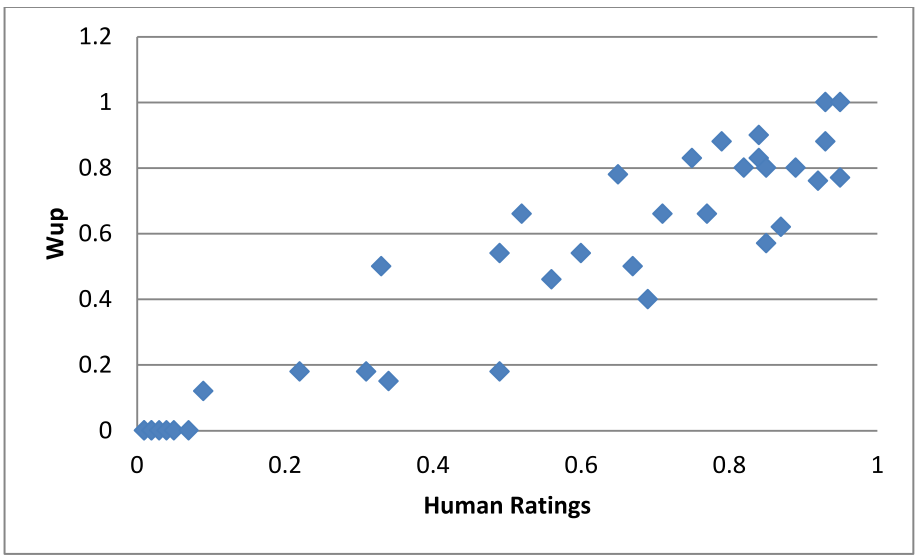

Figure 4-1: The correlation between human ratings and WuP measure scores.

Applying the selected measures using AWN shows that the LCH measure has obtained MSE value of (0.037075). The results show that the LCH measure performs better in low similarity group with MSE value of (0.00231). The LCH measure has the worst performance in high similarity group, due to the highest value of MSE (0.06085) that this measure has achieved.

LCH measure has a good correlation coefficient compared with human ratings (0.89). This indicates a strong relation between LCH measure and human ratings. Less correlation has been scored when compared with LCH measure on WN (0.82). Figure 4-2 shows the correlation between the scores of LCH measure and human ratings.

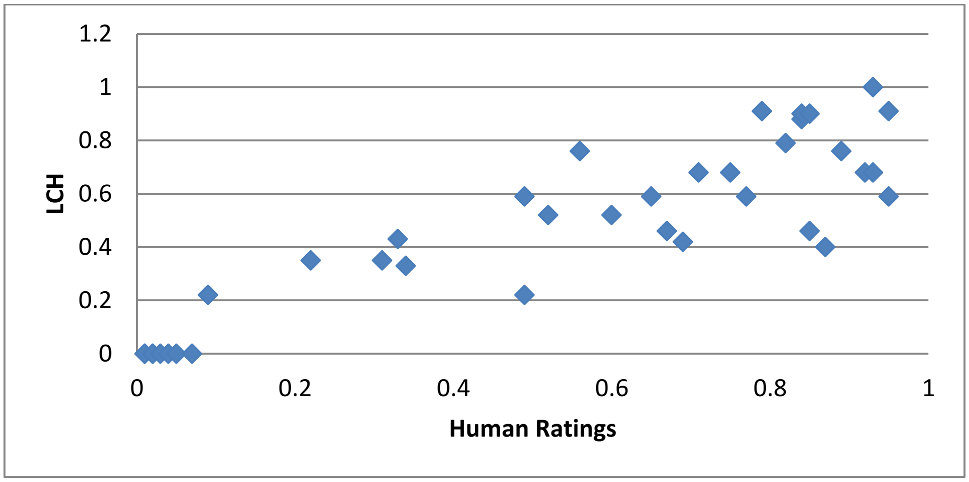

Figure 4-2: The correlation between human ratings and LCH measure

PATH measure has obtained the highest MSE value (0.160383) compared to the MSE values of other measures, which indicates bad performance for PATH measure. Highest MSE value (0.301057) for this measure in high similarity group shows that PATH measure has scored very poor results in high similarity. The correlation coefficient of PATH measure is 0.75. Figure 4-3 shows an empty area between 0.5 and 1. However, this empty area reduces the correlation with human ratings. PATH measure on AWN has scored better value of correlation coefficient compared with PATH measure that has been applied on WN with value of (0.79).

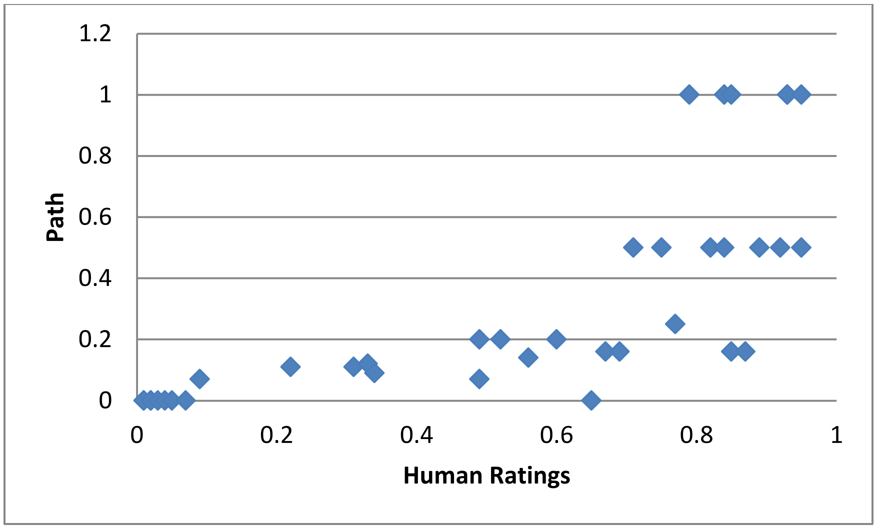

Figure 4-3: The correlation between result of path measure and human ratings.

MSE value for Li's measure is (0.1020513); this high value of error indicates poor performance for this measure. The results show that Li's measure has obtained better scores for low similarity group than scores for medium and high similarity group.

Correlation coefficient of Li's similarity measure using AWN beats the PATH measure with value of (0.84). Li's measure has scored high correlation coefficient with corresponding Li's measure that has been applied over WN with value of (0.95).

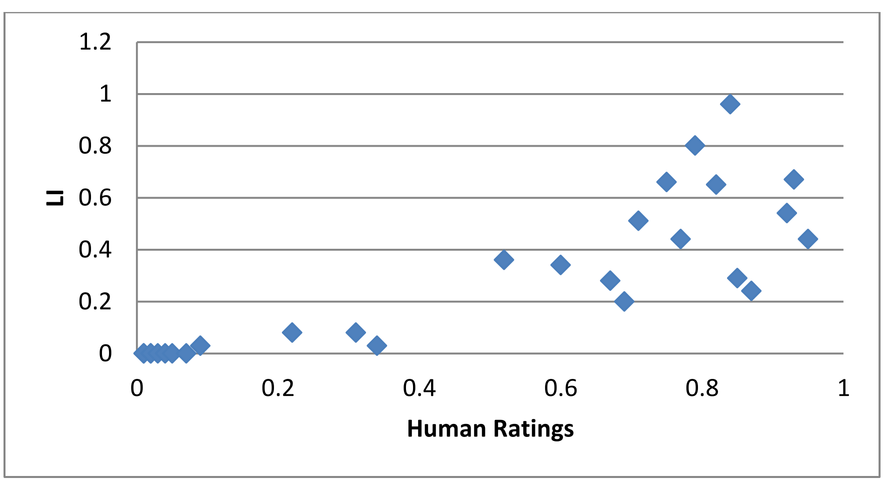

Figure 4-4: The correlation between human ratings and the results of Li's measure

Information content-based measure (Res_Meng) has obtained medium value of MSE (0.077056). Compared to the other measure, this measure has achieved intermediate outcome. This measure performs well in low similarity group by achieving (0.014863) of MSE in low similarity group. However, the results show weakness for this measure in high similarity.

Res_Meng measure has obtained a good correlation coefficient (0.91) with human ratings and comes in third place. Correlation coefficient value between Res_Meng measure over AWN and between Res_Meng measure over WN is 0.82.

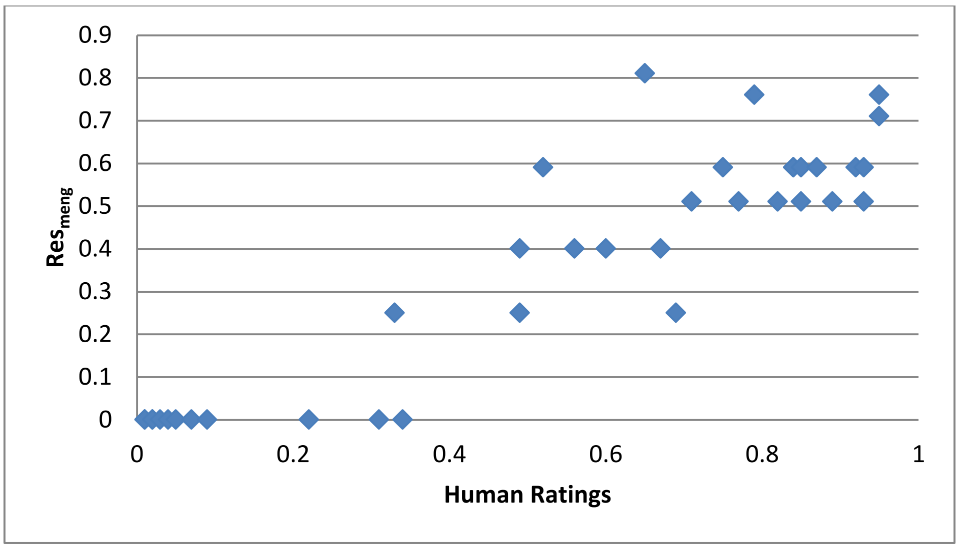

Figure 4-5: Correlation between human ratings and the results of Res_Meng measure

AWSS measure has achieved good MSE score (0.044237). AWSS measure has scored best results in low similarity group and worst results in high similarity.

Human raring correlation with AWSS method (0.88) is very close to LCH correlation coefficient with human scores. Figure 4-6 shows the correlation coefficient between the scores of AWSS measure and the human ratings.

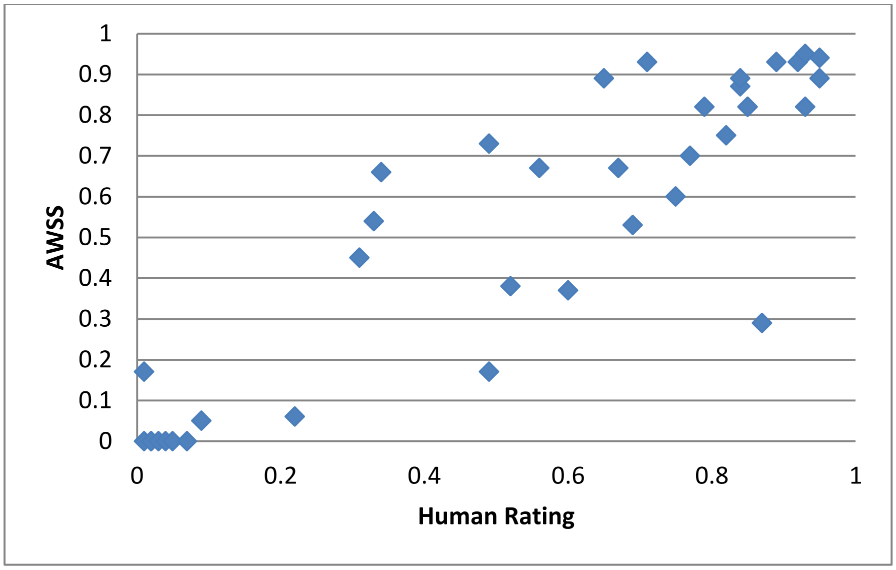

Figure 4-6: The correlation between human ratings and AWSS measure

The MSE value (0.03174) of Zhou measure is very close to MSE of LCH measure. MSE value of (0.07202) in high similarity group indicates the weakness of this measure in high similarity group. However, Zhou measure has achieved better performance in medium and low similarity.

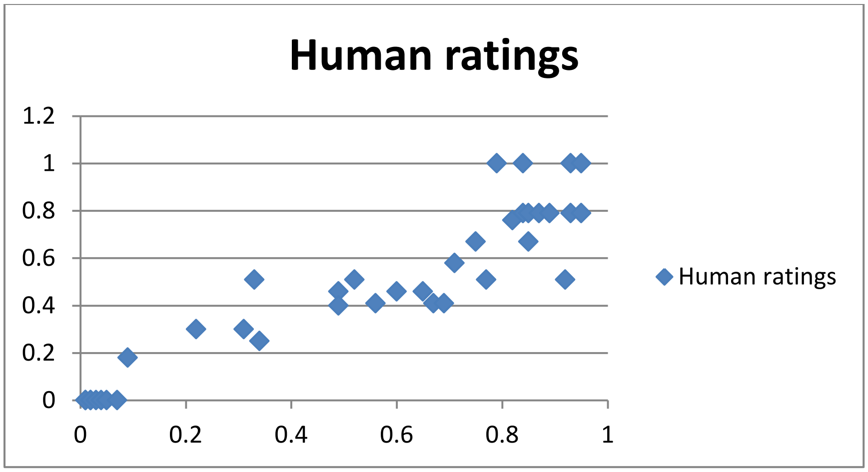

Figure 4-7: The correlation between human rating and Zhou measure

Figure 4-7 shows the correlation coefficient between Zhou measure and human ratings, this measure has a high correlation score after WuP measure (0.92).

4.2 Measures Evaluation

In this section the obtained results from previous experiments have been analyzed to find which measures achieve good performance over AWN. Table 4-8 shows that WuP measure has achieved the highest correlation with human ratings and the lowest value of MSE. Therefore, this indicates that the WuP measure has the best performance in calculating the similarity of Arabic word pairs using AWN compared to the other measures. In other hand, PATH measure has the worst performance, because of the lowest correlation coefficient with human ratings and highest value of MSE that it has achieved. Table 4-8 shows the correlation coefficient between each measure and human ratings, and the MSE values for all measures. Correlation coefficient values multiplied by 10 and MSE values multiplied by 100 to make the comparison between measures easier.

Table 4-8: list of correlation and MSE values for all measures

Measure	Correlation coefficient with human ratings	MSE (%)
WuP	**0.94**	**1.6475**
Res_Meng	0.91	7.7056
LCH	0.89	3.7075
AWSS	0.88	4.4237
Li	0.84	10.205
PATH	**0.75**	**16.038**
Zhou	0.92	3.17432

Figure 4-8 shows that the correlation coefficient values of all measures are almost close to each other. However, the correlation value of WuP measure is the highest, followed by Zhou measure and the correlation coefficient value of path measure is the lowest.

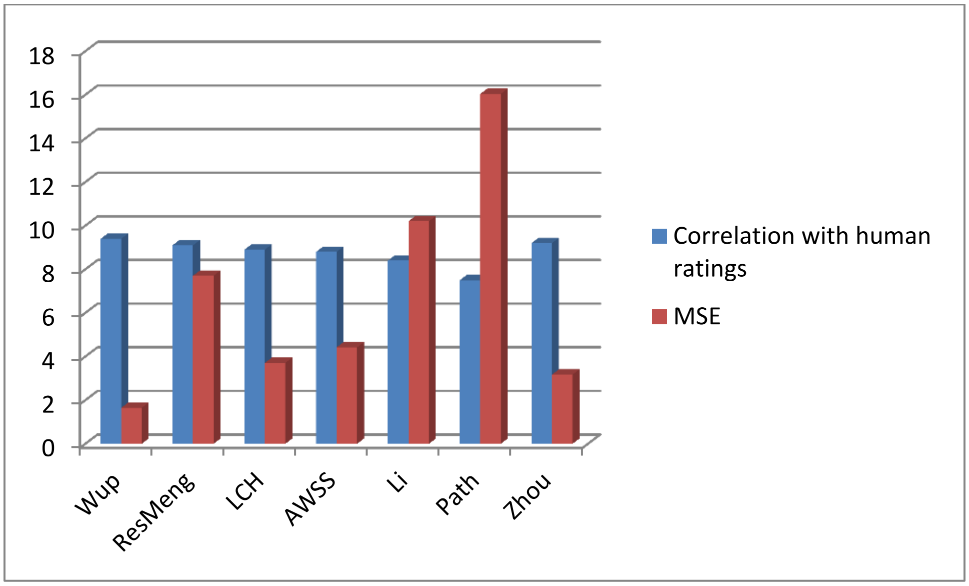

Figure 4-8: The correlation and MSE values for all measures

4.3 New Measure Evaluation

The evaluation of the new measure has been conducted by finding MSE value for this measure, to compare it with the MSE values of other measures. Moreover, finding correlation coefficient with human ratings to compare it with the correlation coefficient values of other measures. Table 4-9 shows that MSE value (1.89 %) of the new measure is close to WuP measure and better than the MSE values of other measures.

The correlation coefficient with human ratings for new measure is very high and beats the correlation values of all measures. The value of correlation coefficient is 0.96; this means that the performance of this measure is very good. Figure 4-9 shows the strong relation between scores of new measure and human ratings.

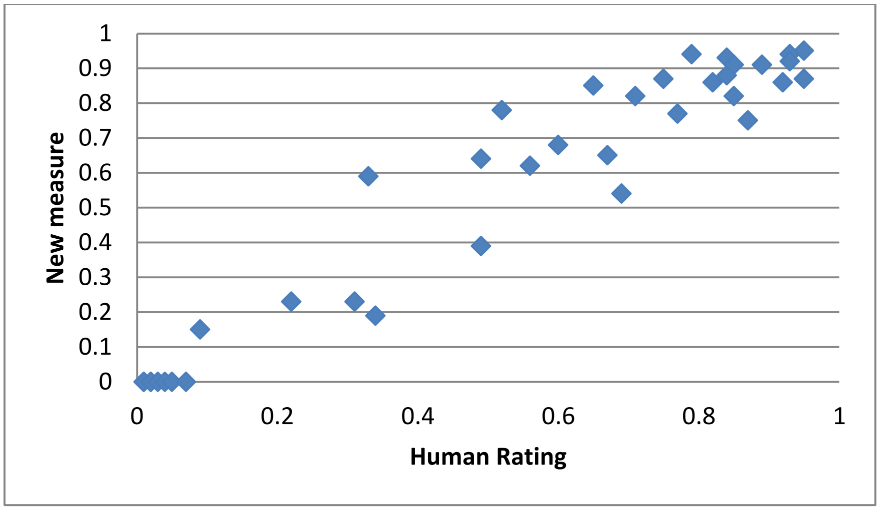

Figure 4-9: The correlation between human rating and new measure

Table 4-9 shows that new measure has the highest value of correlation coefficient with human ratings which indicates the strongest relation with human ratings compared to the other measures.

Table 4-9: list of correlation and MSE values for all measures and new measure

Measure	Correlation coefficient with human ratings	MSE (%)
New measure	**0.96**	1.8932
WuP	0.94	1.6475
Res_Meng	0.91	7.7056
LCH	0.89	3.7075
AWSS	0.88	4.4237
Li	0.84	10.205
PATH	0.75	16.038
Zhou	0.92	3.17432

Figure 4-10 shows that correlation coefficient and MSE values of new measure are very close to the correlation and MSE values of WuP measure. New measure has better relation with human ratings than WuP measure. However, the error in the scores of WuP measure is less than the error in the new measure.

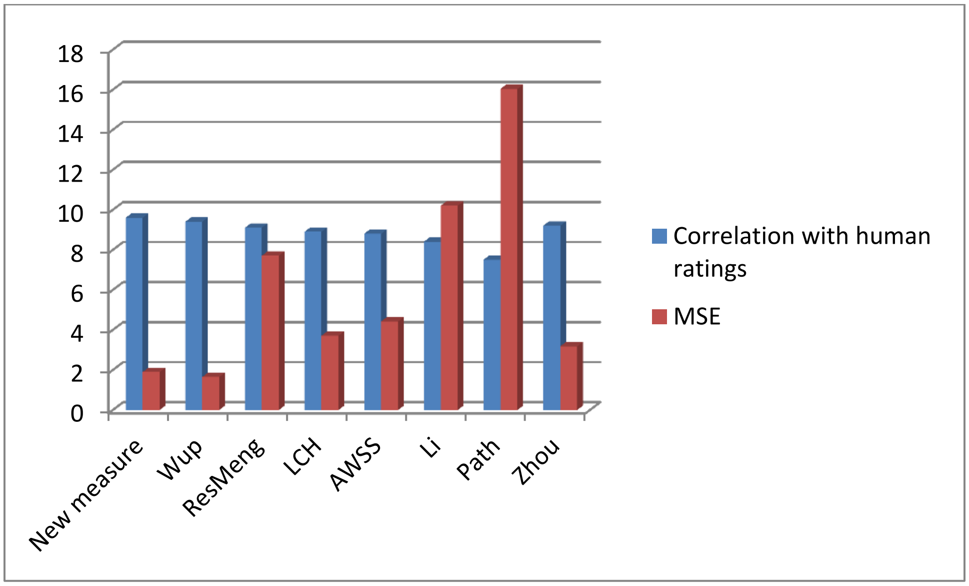

Figure 4-10: comparison between new meaure and all measures

CHAPTER FIVE

Conclusions and Future Work

Overview

This chapter summarizes the work done through this thesis. It discusses the conclusion from the results in the experiments. It also discusses future work.

5.1 Conclusion and Contributions

This thesis has studied the traditional semantic similarity measures over AWN. Then, these measures have been applied using Arabic dataset on AWN. This thesis shows that AWN provides information sources which are: distances between concepts, depths of concepts and information content of concepts. Therefore, these information sources could be used by different categories of measures such as path-based measures, corpus-dependent information content based measures, and hybrid measures to calculate the similarity score between Arabic word pairs.

The thesis shows that AWN has missing information sources such as glosses of concepts. However, some of feature-based measures need these glosses to be applied on AWN. Therefore, Lesk's measure which is well known feature-based measure is hard to be applied on AWN. Furthermore, the corpus-dependent information content-based measures are not applied yet over AWN due to the ambiguity and sparse data problem. However, to avoid these problems, this thesis recommends using corpus-independent information content-based measures.

The experimental results of applying the traditional semantic similarity measures on AWN found out that WuP measure has the highest correlation value with human ratings. Furthermore, WuP measure has obtained the lowest MSE value compared to the other measures; therefore, this result indicates that the WuP measure has the best performance compared to other measures.

Thus, PATH measure has the worst performance, with lowest correlation with human rating and lowest MSE value.

The thesis presented new hybrid semantic similarity measure using AWN. The new hybrid measure takes three factors into consideration: depth of concepts in AWN taxonomy, length of shortest paths between two compared concepts and information content of the LCS. The weight of these factors can be adapted manually. However, several experiments have been conducted to find the best weight that achieves the minimum MSE. However, the result of applying new measure shows that new measure has obtained the highest correlation value compared with the other measures. Furthermore, it has achieved very good value of MSE compared with the performance of the other measures; the new measure has achieved very good performance.

This research found out that there is a shortage in using AWN as a semantic knowledge base in finding the similarity score between Arabic word pairs, due to the following reason: absence of concepts' glosses, many of Arabic words are missing, and there are not enough links (relations) between Arabic words. Moreover, AWN contains only 9,698 synsets, which considered as a few number for a rich language such Arabic.

5.2 Future Work

As mentioned above, AWN suffers from some shortages; this affects the performance of similarity measures for the Arabic language. Therefore, propose new Arabic ontology to cover the AWN shortages will help to enhance Arabic similarity measures. However, well-structured Arabic ontology with enough relations between concepts, and with suited glosses for all concepts, will make applying feature-based measures possible.

AWN database is open source and designed to be extended, therefore, more Arabic concepts, glosses, and relations could be added, this needs to cooperate with Arabic lexicographer.

Calculating IC value of Arabic concepts using corpora is a challenging task. Developing new methods or tools for obtaining the IC value of Arabic concepts will help researchers to propose new corpus-dependent information content-based measures. In the other hand, IC corpus-independent methods are promising techniques in developing new semantic similarity measures.

This research will open the door to propose new hybrid similarity measures, since the experimental result of this study showed that hybrid measures achieved good performance in computing similarity scores between Arabic word pairs. However, new measures from different categories could be proposed as long as AWN provides the information sources that are needed for these categories.

References

- Abderrahim, M. A., Abderrahim, M. E. A., & Chikh, M. A. (2013). Using Arabic wordnet for semantic indexation in information retrieval system. arXiv preprint arXiv:1306.2499.

- Abouenour, L., Rosso, P., & Bouzoubaa, K. (2012). IDRAAQ: New Arabic question answering system based on query expansion and passage retrieval.

- Al-Khiaty, M. A. R., & Ahmed, M. (2016). UML Class Diagrams: Similarity Aspects and Matching. *Lecture Notes on Software Engineering*, *4*(1).

- Almarsoomi, F. A., O'Shea, J. D., Bandar, Z., & Crockett, K. (2013, October). AWSS: An Algorithm for Measuring Arabic Word Semantic Similarity. In Systems, Man, and Cybernetics (SMC), 2013 IEEE International Conference on (pp. 504-509). IEEE.

- Atoum, I., Otoom, A., & Kulathuramaiyer, N. (2016). A Comprehensive Comparative Study of Word and Sentence Similarity Measures. *International Journal of Computer Applications*, *135(1), 10-17*.

- *Al-Yahya, M., Al-Khalifa, H., Bahanshal, A., Al-Odah, I., & Al-Helwah, N. (2010). An ontological model for representing semantic lexicons: an application on time nouns in the holy Quran. Arabian Journal for Science and Engineering, 35(2), 21.*

- Attia, M. A. (2008). Handling Arabic morphological and syntactic ambiguity within the LFG framework with a view to machine translation (Doctoral dissertation, University of Manchester).

- Black, W., Elkateb, S., Rodriguez, H., Alkhalifa, M., Vossen, P., Pease, A., & Fellbaum, C. (2006, January). Introducing the Arabic wordnet project. *In Proceedings of the third international WordNet conference (pp. 295-300)*.

- Cavalli-Sforza, V., Saddiki, H., Bouzoubaa, K., Abouenour, L., Maamouri, M., & Goshey, E. (2013, May). Bootstrapping a WordNet for an Arabic dialect from other WordNets and dictionary resources. In AICCSA (pp. 1-8).

- Elberrichi, Z., & Abidi, K. (2012). Arabic text categorization: a comparative study of different representation modes. *Int. Arab J. Inf. Technol., 9(5), 465-470.*

- Elkateb, S., Black, W., Rodríguez, H., Alkhalifa, M., Vossen, P., Pease, A., & Fellbaum, C. (2006, May). Building a wordnet for arabic. *In Proceedings of The fifth international conference on Language Resources and Evaluation (LREC 2006)*.

- Faaza, A., James, D., Zuhair, A., & Keeley, A. (2012). Arabic Word Semantic Similarity. *Proceedings of World Academy of Science, Engineering and Technology. No. 70. World Academy of Science, Engineering and Technology*

- Gruber, T. R. (1993). A translation approach to portable ontology specifications. Knowledge acquisition, 5(2), 199-220..

- Imam, I., Nounou, N., Hamouda, A., & Khalek, H. A. A. (2013). An ontology-based summarization system for arabic documents (ossad). International Journal of Computer Applications, 74(17).

- Ishkewy, H., Harb, H., & Farahat, H. (2014). Azhary: An arabic lexical ontology. arXiv preprint arXiv:1411.1999.

- Jiang, R., Gan, M., & Dou, X.,. (2013). From ontology to semantic similarity: calculation of ontology-based semantic similarity. *The Scientific World Journal,2013*.

- Jarrar, M. (2011, April). Building a formal Arabic ontology. In Invited Paper). In proceedings of the Experts Meeting on Arabic Ontologies and Semantic Networks. Alecso, Arab League. Tunis: April (pp. 26-28).

- Leacock, C., & Chodorow, M. (1998). Combining local context and WordNet similarity for word sense identification. WordNet: An electronic lexical database, 49(2), 265-283.

- Lesk, M. (1986, June). Automatic sense disambiguation using machine readable dictionaries: how to tell a pine cone from an ice cream cone. *In Proceedings of the 5th annual international conference on Systems documentation (pp. 24-26). ACM.*

- Li, Y., Bandar, Z. A., & McLean, D. (2003). An approach for measuring semantic similarity between words using multiple information sources. *Knowledge and Data Engineering, IEEE Transactions on, 15(4), 871-882.*

- *Lin, C. Y. (2004). Looking for a Few Good Metrics. Working Notes of NII-NACSIS Test Collection for IR Systems.*

- *Lin, D. (1998, July). An information-theoretic definition of similarity. In ICML (Vol. 98, pp. 296-304).*

- McInnes, B., & Pedersen, T. (2013). Evaluating measures of semantic similarity and relatedness to disambiguate terms in biomedical text. *Journal of Biomedical Informatics, 46(6), 1116-1124.*

- Meng, L., Gu, J., & Zhou, Z. (2012). A new model of information content based on concept's topology for measuring semantic similarity in WordNet. International Journal of Grid and Distributed Computing, 5(3), 81-94.

- 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.

- *Michelizzi, J. (2005). Semantic relatedness applied to all words sense disambiguation (Doctoral dissertation, University of Minnesota).*

- Mihalcea, R., Corley, C., & Strapparava, C. (2006, July). Corpus-based and knowledge-based measures of text semantic similarity. *In AAAI (Vol. 6, pp. 775-780).*

- *Miller, G. A., Beckwith, R., Fellbaum, C., Gross, D., & Miller, K. J. (1990). Introduction to WordNet: An on-line lexical database. International journal of lexicography, 3(4), 235-244.*

- Miller, G., & Fellbaum, C. (1998). Wordnet: An electronic lexical database.

- Mittal, K., & Jain, A. (2015). Word Sense Disambiguation Method Using Semantic Similarity Measures And Owa Operator. *ICTACT Journal on Soft Computing: Special Issue on Soft–computing Theory, Application and Implications in Engineering and Technology, 5(02).*

- Mazari, A. C., Aliane, H., & Alimazighi, Z. (2012). Automatic Construction of Ontology from Arabic Texts. In ICWIT (pp. 193-202).

- Patwardhan, S. (2003). Incorporating dictionary and corpus information into a context vector measure of semantic relatedness (Doctoral dissertation, *University of Minnesota, Duluth).*

- Patwardhan, S. (2003). *Incorporating dictionary and corpus information into a context vector measure of semantic relatedness*. Master's thesis, University of Minnesota Duluth.

- *Pease, A., (2000). Standard Upper Ontology Knowledge Interchange Format. Web document http://suo.ieee.org/suokif. html.*

- Pedersen, T., Pakhomov, S. V., Patwardhan, S., & Chute, C. G. (2007). Measures of semantic similarity and relatedness in the biomedical domain. *Journal of biomedical informatics, 40(3), 288-299*.

- *Pedersen, T., Patwardhan, S., & Michelizzi, J. (2004, May). WordNet:: Similarity: measuring the relatedness of concepts. In Demonstration papers at HLT-NAACL 2004 (pp. 38-41). Association for Computational Linguistics.*

- Resnik, P. (1995). Using information content to evaluate semantic similarity in a taxonomy. In *Proceedings of the 14th International Joint Conference on Artificial Intelligence*, 448–453.

- Rubenstein, H., & Goodenough, J. B. (1965). Contextual correlates of synonymy. Communications of the ACM, 8(10), 627-633.

- Sánchez, D., Batet, M., & Isern, D. (2011). Ontology-based information content computation. *Knowledge-Based Systems*, *24*(2), 297-303.

- *Seco, N., Veale, T., & Hayes, J. (2004, August). An intrinsic information content metric for semantic similarity in WordNet. In ECAI (Vol. 16, p. 1089).*

- Slimani, T. (2013). Description and Evaluation of Semantic Similarity Measures Approaches. *International Journal of Computer Applications, 80(10), 25-33. http://dx.doi.org/10.5120/13897-1851*

- Slimani, T., Yagahlane, B. B., & Mellouli, K. (2006). A new similarity measure based on edge counting. Proceedings of the World Academy of Science, engineering and Technology, 17.

- Turney, P. 2001. Mining the web for synonyms: PMI-IR versus LSA on TOEFL. In Proceedings of the Twelfth European Conference on Machine Learning (ECML-2001).

- Tversky, A. (1977). Features of similarity. Psychological review, 84(4), 327.

- *Vossen, P. (ed.) (1998) EuroWordNet: A Multilingual Database with Lexical Semantic Networks. Dordrecht: Kluwer Academic Publishers.*

- Wu, Z., and Palmer, M. 1994. Verb semantics and lexical selection. *In Proceedings of the Annual Meeting of the Association for Computational Linguistics*

- *Zouaghi, A., L. Merhbene, and M. Zrigui. "Word Sense disambiguation for Arabic language using the variants of the Lesk algorithm." WORLDCOMP 11 (2011): 561-567.*

- *Zhou, Z., Wang, Y., & Gu, J. (2008, November). New model of semantic similarity measuring in wordnet. In Intelligent System and Knowledge Engineering, 2008. ISKE 2008. 3rd International Conference on (Vol. 1, pp. 256-261). IEEE.*

Appendix

LCH measure results

NO.	Sim. level	Word Pairs		Arabic word pairs		Human Ratings	EN	AR	Err.	Sqr. Err.
1	Low Similarity	Coast	Endorsement	تصديق	ساحل	***0.01***	0.43	0	0.01	0.0001
2		Noon	String	خيط	ظهر	***0.01***	0.33	0	0.01	0.0001
3		Stove	Walk	**مشي**	**موقد**	***0.01***	**0.17**	**-**	-	-
4		Cord	Midday	ظهيرة	حبل	***0.02***	0.25	0	0.02	0.0004
5		Signature	String	خيط	توقيع	***0.02***	0.28	0	0.02	0.0004
6		Boy	Endorsement	تصديق	صبي	***0.03***	0.33	0	0.03	0.0009
7		Boy	Midday	ظهيرة	صبي	***0.04***	0.33	0	0.04	0.0016
8		Smile	Village	قرية	إبتسامة	***0.05***	0.35	0	0.05	0.0025
9		Noon	Fasting	صيام	ظهر	***0.07***	0.27	0	0.07	0.0049
10		Glass	Diamond	الماس	كأس	***0.09***	0.40	**0.22**	-0.13	0.0169
11		Sepulcher	Sheikh	ضريح	شيخ	***0.22***	0.35	0.35	-0.13	0.0169
12		Countryside	Vegetable	خضار	ريف	***0.31***	0.33	0.35	-0.04	0.0016
13	Medium similarity	Tumbler	Tool	قدح	أداة	***0.33***	0.52	0.43	-0.1	0.01
14		Laugh	Feast	عيد	ضحك	***0.34***	0.42	0.33	0.01	0.0001
15		Girl	Odalisque	جارية	فتاة	***0.49***	0.57	0.59	-0.1	0.01
16		Feast	Fasting	صيام	عيد	***0.49***	0.33	0.22	0.27	0.0729
17		Coach	Means	وسيلة	حافلة	***0.52***	0.56	0.52	0	0
18		Sage	Sheikh	شيخ	حكيم	***0.56***	0.52	0.76	-0.2	0.04
19		Girl	Sister	أخت	فتاة	***0.60***	0.33	0.52	0.08	0.0064
20		Hen	Pigeon	حمامة	دجاجة	***0.65***	0.57	0.59	0.06	0.0036
21		Hill	Mountain	**جبل**	**تل**	***0.65***	0.70	-	-	-
22		Master	Sheikh	شيخ	سيد	***0.67***	0.70	0.46	0.21	0.0441
23		Food	Vegetable	خضار	طعام	***0.69***	0.70	0.42	0.27	0.0729
24		Slave	Odalisque	جارية	عبد	***0.71***	0.47	0.68	0.03	0.0009
25		Run	Walk	مشي	جري	***0.75***	0.70	0.68	0.07	0.0049
26	High Similarity	Cord	String	خيط	حبل	***0.77***	0.81	0.59	0.18	0.0324
27		Forest	Woodland	أحراش	غابة	***0.79***	0.35	0.91	-0.12	0.0144
28		Sage	Thinker	مفكر	حكيم	***0.82***	0.63	0.79	0.03	0.0009
29		Journey	Travel	سفر	رحلة	***0.84***	0.70	0.88	-0.04	0.3598
30		Gem	Diamond	ألماس	جوهرة	***0.84***	0.83	0.9	-0.06	0.0036
31		Countryside	Village	قرية	ريف	***0.85***	0.55	0.9	-0.05	0.0025
32		Cushion	Pillow	مخدة	مسند	***0.85***	0.70	0.46	0.39	0.1521
33		Smile	Laugh	ضحك	إبتسامة	***0.87***	0.70	0.40	0.47	0.2209
34		Signature	Endorsement	توقيع	تصديق	***0.89***	0.8	0.76	0.13	0.0169
35		Tools	Means	وسيلة	أداة	***0.92***	0.63	0.68	0.24	0.0576
36		Sepulcher	Grave	ضريح	قبر	***0.93***	0.80	0.68	0.25	0.0625
37		Boy	Lad	فتى	صبي	***0.93***	0.79	1	-0.07	0.0049
38		**Wizard**	**Magician**	**مشعوذ**	**ساحر**	***0.94***	0.98	-	-	-
39		Coach	Bus	حافلة	باص	***0.95***	1	0.91	0.04	0.0016
40		Glass	Tumbler	قدح	كأس	***0.95***	0.70	**0.59**	0.36	0.1296
									MSE=	0.031743243

PATH measure results

NO.	Sim. level	Word Pairs		Arabic word pairs		Human Ratings	EN	AR	Err.	Sqr. Err.
1	**Low Similarity**	Coast	Endorsement	تصديق	ساحل	***0.01***	0.12	0	0.01	0.0001
2		Noon	String	خيط	ظهر	***0.01***	0.08	0	0.01	0.0001
3		Stove	Walk	**مشي**	**موقد**	***0.01***	0.04	-	-	-
4		Cord	Midday	ظهيرة	حبل	***0.02***	0.06	0	0.02	0.0004
5		Signature	String	خيط	توقيع	***0.02***	0.07	0	0.02	0.0004
6		Boy	Endorsement	تصديق	صبي	***0.03***	0.09	0	0.03	0.0009
7		Boy	Midday	ظهيرة	صبي	***0.04***	0.06	0	0.04	0.0016
8		Smile	Village	قرية	إبتسامة	***0.05***	0.09	0	0.05	0.0025
9		Noon	Fasting	صيام	ظهر	***0.07***	0.06	0	0.07	0.0049
10		Glass	Diamond	الماس	كأس	***0.09***	0.11	0.07	0.02	0.0004
11		Sepulcher	Sheikh	ضريح	شيخ	***0.22***	0.09	0.11	0.11	0.0121
12		Countryside	Vegetable	خضار	ريف	***0.31***	0.08	0.11	0.2	0.04
13	**Medium similarity**	Tumbler	Tool	قدح	أداة	***0.33***	0,16	0.12	0.21	0.0441
14		Laugh	Feast	عيد	ضحك	***0.34***	0.16	0.09	0.25	0.0625
15		Girl	Odalisque	جارية	فتاة	***0.49***	0.2	0.2	0.29	0.0841
16		Feast	Fasting	صيام	عيد	***0.49***	0.09	0.07	0.42	0.1764
17		Coach	Means	وسيلة	حافلة	***0.52***	0.20	0.2	0.32	0.1024
18		Sage	Sheikh	شيخ	حكيم	***0.56***	0.16	0.14	0.42	0.1764
19		Girl	Sister	أخت	فتاة	***0.60***	0.08	0.2	0.4	0.16
20		Hen	Pigeon	حمامة	دجاجة	***0.65***	0.2	0,2	0.45	0.2025
21		Hill	Mountain	**جبل**	**تل**	***0.65***	0.33	-		
22		Master	Sheikh	شيخ	سيد	***0.67***	0.33	0.16	0.51	0.2601
23		Food	Vegetable	خضار	طعام	***0.69***	0.33	0.16	0.53	0.2809
24		Slave	Odalisque	جارية	عبد	***0.71***	0.14	0.5	0.21	0.0441
25		Run	Walk	مشي	جري	***0.75***	0.33	0.5	0.25	0.0625
26	**High Similarity**	Cord	String	خيط	حبل	***0.77***	0,5	0.25	0.52	0.2704
27		Forest	Woodland	أحراش	غابة	***0.79***	1	1	-0.21	0.0441
28		Sage	Thinker	مفكر	حكيم	***0.82***	0.25	0.5	0.32	0.1024
29		Journey	Travel	سفر	رحلة	***0.84***	0.5	1	-0.16	2.1363
30		Gem	Diamond	ألماس	جوهرة	***0.84***	0.5	0.5	0.34	0.1156
31		Countryside	Village	قرية	ريف	***0.85***	0.2	1	-0.15	0.0225
32		Cushion	Pillow	مخدة	مسند	***0.85***	0.5	0.16	0.69	0.4761
33		Smile	Laugh	ضحك	إبتسامة	***0.87***	0.33	0.16	0.71	0.5041
34		Signature	Endorsement	توقيع	تصديق	***0.89***	0.5	0.5	0.39	0.1521
35		Tools	Means	وسيلة	أداة	***0.92***	0,25	0.5	0.42	0.1764
36		Sepulcher	Grave	ضريح	قبر	***0.93***	0.5	1	-0.07	0.0049
37		Boy	Lad	فتى	صبي	***0.93***	0.5	1	-0.07	0.0049
38		**Wizard**	**Magician**	**مشعوذ**	**ساحر**	***0.94***	1	-	-	-
39		Coach	Bus	حافلة	باص	***0.95***	1	1	-0.05	0.0025
40		Glass	Tumbler	قدح	كأس	***0.95***	0.5	0.5	0.45	0.2025
									MSE=	0.160383784

LI measure results

NO.	Sim. level	Word Pairs		Arabic word pairs		Human Ratings	EN	AR	Err.	Sqr. Err.
1	**Low Similarity**	Coast	Endorsement	تصديق	ساحل	***0.01***	*0.09*	0	*0.01*	*0.0001*
2		Noon	String	خيط	ظهر	***0.01***	*0.09*	0	*0.01*	*0.0001*
3		Stove	Walk	**مشي**	**موقد**	***0.01***	*0.12*	-	-	-
4		Cord	Midday	ظهيرة	حبل	***0.02***	*0.09*	0	*0.02*	*0.0004*
5		Signature	String	خيط	توقيع	***0.02***	*0.16*	0	*0.02*	*0.0004*
6		Boy	Endorsement	تصديق	صبي	***0.03***	*0.16*	0	*0.03*	*0.0009*
7		Boy	Midday	ظهيرة	صبي	***0.04***	*0.18*	0	*0.04*	*0.0016*
8		Smile	Village	قرية	إبتسامة	***0.05***	*0.11*	0	*0.05*	*0.0025*
9		Noon	Fasting	صيام	ظهر	***0.07***	*0.14*	0	*0.07*	*0.0049*
10		Glass	Diamond	الماس	كأس	***0.09***	*0.09*	0.03	*0.06*	*0.0036*
11		Sepulcher	Sheikh	ضريح	شيخ	***0.22***	*0.18*	0.08	*0.14*	*0.0196*
12		Countryside	Vegetable	خضار	ريف	***0.31***	*0.2*	0.08	*0.23*	*0.0529*
13	**Medium similarity**	Tumbler	Tool	قدح	أداة	***0.33***	*0.25*	0.19	*0.14*	*0.0196*
14		Laugh	Feast	عيد	ضحك	***0.34***	*0.18*	0.03	*0.31*	*0.0961*
15		Girl	Odalisque	جارية	فتاة	***0.49***	*0.26*	0.34	*0.15*	*0.0225*
16		Feast	Fasting	صيام	عيد	***0.49***	*0.40*	0.03	*0.46*	*0.2116*
17		Coach	Means	وسيلة	حافلة	***0.52***	*0.80*	0.36	*0.16*	*0.0256*
18		Sage	Sheikh	شيخ	حكيم	***0.56***	*0.66*	0.65	*-0.09*	*0.0081*
19		Girl	Sister	أخت	فتاة	***0.60***	*0.76*	0.34	*0.26*	*0.0676*
20		Hen	Pigeon	حمامة	دجاجة	***0.65***	*0.80*	0.36	*0.29*	*0.0841*
21		Hill	Mountain	**جبل**	**تل**	***0.65***	*0.82.*	-	-	-
22		Master	Sheikh	شيخ	سيد	***0.67***	*0.76*	0.28	*0.39*	*0.1521*
23		Food	Vegetable	خضار	طعام	***0.69***	*0.85*	0.20	*0.49*	*0.2401*
24		Slave	Odalisque	جارية	عبد	***0.71***	*0.87*	0.51	*0.2*	*0.04*
25		Run	Walk	مشي	جري	***0.75***	*0.90*	0.66	*0.09*	*0.0081*
26	**High Similarity**	Cord	String	خيط	حبل	***0.77***	*0.85*	0.44	*0.33*	*0.1089*
27		Forest	Woodland	أحراش	غابة	***0.79***	*0.96*	0.80	*-0.01*	*0.0001*
28		Sage	Thinker	مفكر	حكيم	***0.82***	*0.92*	0.65	*0.17*	*0.0289*
29		Journey	Travel	سفر	رحلة	***0.84***	*0.96*	0.96	*-0.12*	*1.2004*
30		Gem	Diamond	ألماس	جوهرة	***0.84***	*0.95*	0.66	*0.18*	*0.0324*
31		Countryside	Village	قرية	ريف	***0.85***	*0.93*	0.65	*0.2*	*0.04*
32		Cushion	Pillow	مخدة	مسند	***0.85***	*0.91*	0.29	*0.56*	*0.3136*
33		Smile	Laugh	ضحك	إبتسامة	***0.87***	*0.95*	0.24	*0.63*	*0.3969*
34		Signature	Endorsement	توقيع	تصديق	***0.89***	*0.90*	0.65	*0.24*	*0.0576*
35		Tools	Means	وسيلة	أداة	***0.92***	*0.94*	0.54	*0.38*	*0.1444*
36		Sepulcher	Grave	ضريح	قبر	***0.93***	*0.96*	0.69	*0.24*	*0.0576*
37		Boy	Lad	فتى	صبي	***0.93***	*0.94*	0.67	*0.26*	*0.0676*
38		**Wizard**	**Magician**	**مشعوذ**	**ساحر**	***0.94***	*0.94*	-	-	-
39		Coach	Bus	حافلة	باص	***0.95***	*0.96*	0.88	*0.07*	*0.0049*
40		Glass	Tumbler	قدح	كأس	***0.95***	*0.89*	0.44	*0.51*	*0.2601*
									MSE=	0.102051351

Res_Meng measure results

NO.	Sim. level	Word Pairs		Arabic word pairs		Human Ratings	EN	AR	Err.	Sqr. Err.
1	Low Similarity	Coast	Endorsement	تصديق	ساحل	*0.01*	0.23	0	0.01	0.0001
2		Noon	String	خيط	ظهر	*0.01*	0.36	0	0.01	0.0001
3		Stove	Walk	**مشي**	**موقد**	*0.01*	0.23	-	-	-
4		Cord	Midday	ظهيرة	حبل	*0.02*	0.31	0	0.02	0.0004
5		Signature	String	خيط	توقيع	*0.02*	0.20	0	0.02	0.0004
6		Boy	Endorsement	تصديق	صبي	*0.03*	0.23	0	0.03	0.0009
7		Boy	Midday	ظهيرة	صبي	*0.04*	0.25	0	0.04	0.0016
8		Smile	Village	قرية	إبتسامة	*0.05*	0.36	0	0.05	0.0025
9		Noon	Fasting	صيام	ظهر	*0.07*	0.46	0	0.07	0.0049
10		Glass	Diamond	الماس	كأس	*0.09*	0.59	0	0.09	0.0081
11		Sepulcher	Sheikh	ضريح	شيخ	*0.22*	0.53	0	0.22	0.0484
12		Countryside	Vegetable	خضار	ريف	*0.31*	0.46	0	0.31	0.0961
13	Medium similarity	Tumbler	Tool	قدح	أداة	*0.33*	0.64	0.25	0.08	0.0064
14		Laugh	Feast	عيد	ضحك	*0.34*	0.36	0	0.34	0.1156
15		Girl	Odalisque	جارية	فتاة	*0.49*	0.76	0.25	0.24	0.0576
16		Feast	Fasting	صيام	عيد	*0.49*	0.25	0.40	0.09	0.0081
17		Coach	Means	وسيلة	حافلة	*0.52*	0.64	0.59	-0.07	0.0049
18		Sage	Sheikh	شيخ	حكيم	*0.56*	0.53	0.40	0.16	0.0256
19		Girl	Sister	أخت	فتاة	*0.60*	0.46	0.40	0.2	0.04
20		Hen	Pigeon	حمامة	دجاجة	*0.65*	0.76	0.81	-0.16	0.0256
21		Hill	Mountain	**جبل**	**تل**	*0.65*	0.59	-	-	-
22		Master	Sheikh	شيخ	سيد	*0.67*	0.73	0.40	0.27	0.0729
23		Food	Vegetable	خضار	طعام	*0.69*	0.59	0.25	0.44	0.1936
24		Slave	Odalisque	جارية	عبد	*0.71*	0.69	0.51	0.2	0.04
25		Run	Walk	مشي	جري	*0.75*	0.76	0.59	0.16	0.0256
26	High Similarity	Cord	String	خيط	حبل	*0.77*	0.69	0.51	0.26	0.0676
27		Forest	Woodland	أحراش	غابة	*0.79*	0.64	0.76	0.03	0.0009
28		Sage	Thinker	مفكر	حكيم	*0.82*	0.73	0.51	0.31	0.0961
29		Journey	Travel	سفر	رحلة	*0.84*	0.76	0.59	0.25	0.944
30		Gem	Diamond	ألماس	جوهرة	*0.84*	0.81	0.59	0.25	0.0625
31		Countryside	Village	قرية	ريف	*0.85*	0.51	0.51	0.34	0.1156
32		Cushion	Pillow	مخدة	مسند	*0.85*	0.69	0.59	0.26	0.0676
33		Smile	Laugh	ضحك	إبتسامة	*0.87*	0.64	0.59	0.28	0.0784
34		Signature	Endorsement	توقيع	تصديق	*0.89*	0.76	0.51	0.38	0.1444
35		Tools	Means	وسيلة	أداة	*0.92*	0.76	0.59	0.33	0.1089
36		Sepulcher	Grave	ضريح	قبر	*0.93*	0.76	0.59	0.34	0.1156
37		Boy	Lad	فتى	صبي	*0.93*	0.76	0.51	0.42	0.1764
38		**Wizard**	**Magician**	**مشعوذ**	**ساحر**	*0.94*	0.76	-	-	-
39		Coach	Bus	حافلة	باص	*0.95*	0.76	0.76	0.19	0.0361
40		Glass	Tumbler	قدح	كأس	*0.95*	0.73	0.71	0.24	0.0576
									MSE=	0.07705675

AWSS measure results

NO.	Sim. level	Word Pairs		Arabic word pairs		Human Ratings	EN	AR	Err.	Sqr. Err.
1	**Low Similarity**	Coast	Endorsement	تصديق	ساحل	***0.01***	-	0	0.01	0.0001
2		Noon	String	خيط	ظهر	***0.01***	-	0.17	-0.16	0.0256
3		Stove	Walk	**مشي**	**موقد**	***0.01***	-	-	-	-
4		Cord	Midday	ظهيرة	حبل	***0.02***	-	0	0.02	0.0004
5		Signature	String	خيط	توقيع	***0.02***	-	0	0.02	0.0004
6		Boy	Endorsement	تصديق	صبي	***0.03***	-	0	0.03	0.0009
7		Boy	Midday	ظهيرة	صبي	***0.04***	-	0	0.04	0.0016
8		Smile	Village	قرية	إبتسامة	***0.05***	-	0	0.05	0.0025
9		Noon	Fasting	صيام	ظهر	***0.07***	-	0	0.07	0.0049
10		Glass	Diamond	الماس	كأس	***0.09***	-	0.05	0.04	0.0016
11		Sepulcher	Sheikh	ضريح	شيخ	***0.22***	-	0.06	0.16	0.0256
12		Countryside	Vegetable	خضار	ريف	***0.31***	-	0.45	-0.14	0.0196
13	**Medium similarity**	Tumbler	Tool	قدح	أداة	***0.33***	-	0.54	-0.21	0.0441
14		Laugh	Feast	عيد	ضحك	***0.34***	-	0.66	-0.32	0.1024
15		Girl	Odalisque	جارية	فتاة	***0.49***	-	0.73	-0.24	0.0576
16		Feast	Fasting	صيام	عيد	***0.49***	-	0.17	0.32	0.1024
17		Coach	Means	وسيلة	حافلة	***0.52***	-	0.38	0.14	0.0196
18		Sage	Sheikh	شيخ	حكيم	***0.56***	-	0.67	-0.11	0.0121
19		Girl	Sister	أخت	فتاة	***0.60***	-	0.37	0.23	0.0529
20		Hen	Pigeon	حمامة	دجاجة	***0.65***	-	0.89	-0.24	0.0576
21		Hill	Mountain	**جبل**	**تل**	***0.65***	-	-	-	-
22		Master	Sheikh	شيخ	سيد	***0.67***	-	0.67	0	0
23		Food	Vegetable	خضار	طعام	***0.69***	-	0.53	0.16	0.0256
24		Slave	Odalisque	جارية	عبد	***0.71***	-	0.93	-0.22	0.0484
25		Run	Walk	مشي	جري	***0.75***	-	0.60	0.15	0.0225
26	**High Similarity**	Cord	String	خيط	حبل	***0.77***	-	0.70	0.07	0.0049
27		Forest	Woodland	أحراش	غابة	***0.79***	-	0.82	-0.03	0.0009
28		Sage	Thinker	مفكر	حكيم	***0.82***	-	0.75	0.07	0.0049
29		Journey	Travel	سفر	رحلة	***0.84***	-	0.87	-0.03	0.6391
30		Gem	Diamond	ألماس	جوهرة	***0.84***	-	0.89	-0.05	0.0025
31		Countryside	Village	قرية	ريف	***0.85***	-	0.82	0.03	0.0009
32		Cushion	Pillow	مخدة	مسند	***0.85***	-	0.82	0.03	0.0009
33		Smile	Laugh	ضحك	إبتسامة	***0.87***	-	0.29	0.58	0.3364
34		Signature	Endorsement	توقيع	تصديق	***0.89***	-	0.93	-0.04	0.0016
35		Tools	Means	وسيلة	أداة	***0.92***	-	0.93	-0.01	0.0001
36		Sepulcher	Grave	ضريح	قبر	***0.93***	-	0.82	0.11	0.0121
37		Boy	Lad	فتى	صبي	***0.93***	-	0.95	-0.02	0.0004
38		**Wizard**	**Magician**	**مشعوذ**	**ساحر**	***0.94***	-	-	-	-
39		Coach	Bus	حافلة	باص	***0.95***	-	0.94	0.01	0.0001
40		Glass	Tumbler	قدح	كأس	***0.95***	-	0.89	0.06	0.0036
									MSE=	0.044237

ZHOU measure results

NO.	Sim. level	Word Pairs		Arabic word pairs		Human Ratings	EN	AR	Err.	Sqr. Err.
1	Low Similarity	Coast	Endorsement	تصديق	ساحل	***0.01***	-	0	0.01	0.0001
2		Noon	String	خيط	ظهر	***0.01***	-	0	-0.16	0.0256
3		Stove	Walk	**مشي**	**موقد**	***0.01***	-	-	-	-
4		Cord	Midday	ظهيرة	حبل	***0.02***	-	0	0.02	0.0004
5		Signature	String	خيط	توقيع	***0.02***	-	0	0.02	0.0004
6		Boy	Endorsement	تصديق	صبي	***0.03***	-	0	0.03	0.0009
7		Boy	Midday	ظهيرة	صبي	***0.04***	-	0	0.04	0.0016
8		Smile	Village	قرية	إبتسامة	***0.05***	-	0	0.05	0.0025
9		Noon	Fasting	صيام	ظهر	***0.07***	-	0	0.07	0.0049
10		Glass	Diamond	الماس	كأس	***0.09***	-	0.18	-0.09	0.0081
11		Sepulcher	Sheikh	ضريح	شيخ	***0.22***	-	0.30	-0.08	0.0064
12		Countryside	Vegetable	خضار	ريف	***0.31***	-	0.30	0.01	0.0001
13	Medium similarity	Tumbler	Tool	قدح	أداة	***0.33***	-	0.51	-0.18	0.0324
14		Laugh	Feast	عيد	ضحك	***0.34***	-	0.25	0.09	0.0081
15		Girl	Odalisque	جارية	فتاة	***0.49***	-	0.46	0.03	0.0009
16		Feast	Fasting	صيام	عيد	***0.49***	-	0.40	0.09	0.0081
17		Coach	Means	وسيلة	حافلة	***0.52***	-	0.51	0.01	0.0001
18		Sage	Sheikh	شيخ	حكيم	***0.56***	-	0.41	0.15	0.0225
19		Girl	Sister	أخت	فتاة	***0.60***	-	0.46	0.14	0.0196
20		Hen	Pigeon	حمامة	دجاجة	***0.65***	-	0.46	0.19	0.0361
21		Hill	Mountain	**جبل**	**تل**	***0.65***	-	-	-	-
22		Master	Sheikh	شيخ	سيد	***0.67***	-	0.41	0.26	0.0676
23		Food	Vegetable	خضار	طعام	***0.69***	-	0.41	0.28	0.0784
24		Slave	Odalisque	جارية	عبد	***0.71***	-	0.58	0.13	0.0169
25		Run	Walk	مشي	جري	***0.75***	-	0.67	0.08	0.0064
26	High Similarity	Cord	String	خيط	حبل	***0.77***	-	0.51	0.26	0.0676
27		Forest	Woodland	أحراش	غابة	***0.79***	-	1	-0.21	0.0441
28		Sage	Thinker	مفكر	حكيم	***0.82***	-	0.76	0.06	0.0036
29		Journey	Travel	سفر	رحلة	***0.84***	-	0.79	0.05	0.4379
30		Gem	Diamond	ألماس	جوهرة	***0.84***	-	1	-0.16	0.0256
31		Countryside	Village	قرية	ريف	***0.85***	-	0.67	0.18	0.0324
32		Cushion	Pillow	مخدة	مسند	***0.85***	-	0.79	0.06	0.0036
33		Smile	Laugh	ضحك	إبتسامة	***0.87***	-	0.79	0.08	0.0064
34		Signature	Endorsement	توقيع	تصديق	***0.89***	-	0.79	0.1	0.01
35		Tools	Means	وسيلة	أداة	***0.92***	-	0.51	0.41	0.1681
36		Sepulcher	Grave	ضريح	قبر	***0.93***	-	1	-0.07	0.0049
37		Boy	Lad	فتى	صبي	***0.93***	-	0.79	0.14	0.0196
38		**Wizard**	**Magician**	**مشعوذ**	**ساحر**	***0.94***	-	-	-	-
39		Coach	Bus	حافلة	باص	***0.95***	-	1	-0.05	0.0025
40		Glass	Tumbler	قدح	كأس	***0.95***	-	0.79	0.16	0.0256
									MSE=	0.031743243

New hybrid measure results

NO.	Sim. level	Word Pairs		Arabic word pairs		Human Ratings	EN	AR	Err.	Sqr. Err.
1	Low Similarity	Coast	Endorsement	ساحل	تصديق	***0.01***	-	0	0.01	0.0001
2		Noon	String	ظهر	خيط	***0.01***	-	0	0.01	0.0001
3		Stove	Walk	**موقد**	**مشي**	***0.01***	-	-	-	-
4		Cord	Midday	حبل	ظهيرة	***0.02***	-	0	0.02	0.0004
5		Signature	String	توقيع	خيط	***0.02***	-	0	0.02	0.0004
6		Boy	Endorsement	صبي	تصديق	***0.03***	-	0	0.03	0.0009
7		Boy	Midday	صبي	ظهيرة	***0.04***	-	0	0.04	0.0016
8		Smile	Village	إبتسامة	قرية	***0.05***	-	0	0.05	0.0025
9		Noon	Fasting	ظهر	صيام	***0.07***	-	0	0.07	0.0049
10		Glass	Diamond	كأس	الماس	***0.09***	-	0.15	-0.06	0.0036
11		Sepulcher	Sheikh	شيخ	ضريح	***0.22***	-	0.23	-0.01	0.0001
12		Countryside	Vegetable	ريف	خضار	***0.31***	-	0.23	0.08	0.0064
13	Medium similarity	Tumbler	Tool	أداة	قدح	***0.33***	-	0.59	-0.26	0.0676
14		Laugh	Feast	ضحك	عيد	***0.34***	-	0.19	0.15	0.0225
15		Girl	Odalisque	فتاة	جارية	***0.49***	-	0.64	-0.15	0.0225
16		Feast	Fasting	عيد	صيام	***0.49***	-	0.39	0.1	0.01
17		Coach	Means	حافلة	وسيلة	***0.52***	-	0.78	-0.26	0.0676
18		Sage	Sheikh	حكيم	شيخ	***0.56***	-	0.62	-0.06	0.0036
19		Girl	Sister	فتاة	أخت	***0.60***	-	0.68	-0.08	0.0064
20		Hen	Pigeon	دجاجة	حمامة	***0.65***	-	0.85	-0.2	0.04
21		Hill	Mountain	**تل**	**جبل**	***0.65***	-	-	-	-
22		Master	Sheikh	سيد	شيخ	***0.67***	-	0.65	0.02	0.0004
23		Food	Vegetable	طعام	خضار	***0.69***	-	0.54	0.15	0.0225
24		Slave	Odalisque	عبد	جارية	***0.71***	-	0.82	-0.11	0.0121
25		Run	Walk	جري	مشي	***0.75***	-	0.87	-0.12	0.0144
26	High Similarity	Cord	String	حبل	خيط	***0.77***	-	0.77	0	0
27		Forest	Woodland	غابة	أحراش	***0.79***	-	0.94	-0.15	0.0225
28		Sage	Thinker	حكيم	مفكر	***0.82***	-	0.86	-0.04	0.0016
29		Journey	Travel	رحلة	سفر	***0.84***	-	0.93	-0.09	0.3347
30		Gem	Diamond	جوهرة	ألماس	***0.84***	-	0.88	-0.04	0.0016
31		Countryside	Village	ريف	قرية	***0.85***	-	0.91	-0.06	0.0036
32		Cushion	Pillow	مسند	مخدة	***0.85***	-	0.82	0.03	0.0009
33		Smile	Laugh	إبتسامة	ضحك	***0.87***	-	0.75	0.12	0.0144
34		Signature	Endorsement	تصديق	توقيع	***0.89***	-	0.91	-0.02	0.0004
35		Tools	Means	أداة	وسيلة	***0.92***	-	0.86	0.06	0.0036
36		Sepulcher	Grave	قبر	ضريح	***0.93***	-	0.94	-0.01	0.0001
37		Boy	Lad	صبي	فتى	***0.93***	-	0.92	0.01	0.0001
38		**Wizard**	**Magician**	**ساحر**	**مشعوذ**	***0.94***	-	-	-	-
39		Coach	Bus	باص	حافلة	***0.95***	-	0.95	0	0
40		Glass	Tumbler	كأس	قدح	***0.95***	-	0.87	0.08	0.0064
									MSE=	0.018932