DOI: 10.24996/ijs.2022.63.6.35

ISSN: 0067-2904

Short Answers Assessment Approach based on Semantic Network

Sumaya H. Mijbel[1]*, Ahmed T. Sadiq[2]
[1] *Education College, Mustansiriyah University, Iraq*
[2] *Computer Science Department, University of Technology, Iraq*

Received: 25/6/2021 Accepted: 25/9/2021 Published: 30/6/2022

Abstract

Finding similarities in texts is important in many areas such as information retrieval, automated article scoring, and short answer categorization. Evaluating short answers is not an easy task due to differences in natural language. Methods for calculating the similarity between texts depend on semantic or grammatical aspects. This paper discusses a method for evaluating short answers using semantic networks to represent the typical (correct) answer and students' answers. The semantic network of nodes and relationships represents the text (answers). Moreover, grammatical aspects are found by measuring the similarity of parts of speech between the answers. In addition, finding hierarchical relationships between nodes in networks. The similarity is then calculated, and students' answers are evaluated. The best results are for weights (α = 0.1, = 0.6, γ = 0.3) = 1.82 from 5, giving more weight in nodes similarity by 6, least similarity by relationships by 3, and least similarity by parts of speech by 1.

Keywords: Semantic networks, short answers, evaluation, assessment, hierarchical relationships.

شبكة دلالية تعتمد على نهج تقييم الإجابات القصيرة

سميه حسام الدين مجبل[1]*، احمد طارق صادق[2]

[1]كلية التربية، جامعة المستنصرية، العراق

[2]قسم علوم الحاسوب، الجامعة التكنولوجية، العراق

الخلاصة

يعد العثور على أوجه التشابه في النصوص أمرًا مهمًا في العديد من المجالات مثل استرجاع المعلومات وتسجيل المقالات الآلي وتصنيف الإجابات القصيرة. تقييم الإجابات القصيرة ليس بالمهمة السهلة بسبب الاختلافات في اللغة الطبيعية. تعتمد طرق حساب التشابه بين النصوص على الجوانب الدلالية أو النحوية. تناقش هذه الورقة طريقة لتقييم الإجابات القصيرة باستخدام الشبكات الدلالية لتمثيل الإجابة النموذجية (الصحيحة) وإجابات الطلاب. تمثل الشبكة الدلالية للعقد والعلاقات النص (الإجابات). علاوة على ذلك، الجوانب النحوية من خلال قياس تشابه أجزاء الكلام بين الإجابات. بالإضافة إلى إيجاد علاقات هرمية بين العقد في الشبكات، ثم يتم حساب التشابه وتقييم إجابات الطلاب. أفضل النتائج هي للأوزان(α = 0.1 ، = 0.6، γ = 0.3) = 1.82 من 5، مما يعطي وزنًا أكبر في تشابه العقد بمقدار 6 ، وأقل تشابه من خلال العلاقات بمقدار 3 ، وأقل تشابه حسب أجزاء الكلام بمقدار 1 .

*Email: smsm142.1988@gmail.com

1. Introduction

Defining the similarity in text is one of the critical jobs in many text applications, concerned with text similarity, or text-related [1], such as grading articles, short answer evaluation, text summarization [2], machine translation, text classification [3], and more. Discovering similarities between words is a key part of the text-similarity process, used as the initial stage for finding similarities between sentences, paragraphs, and documents [4]. Prior to commencing text similarity check, the word pre-processing organizes the data in a way that is convenient for storage and operations such as tokenization, normalization, remove stop words, and stemming [5].

Short answer questions are open-ended questions for students to create an answer. An idea commonly asked on basic tests and comprehension (cognitive levels and comprehension) of a topic before asking more in-depth evaluation questions about the topic.

The answer to the short questions is in the student's own style. The questions may be about a specific topic or case study. You can refer to the student's educational materials or other sources if necessary, but without copying from the original text. The student expresses them briefly and related to the requested question. Such as identifying, explaining, discussing, comparing, analyzing, and others. [6]

Automatic grading of short answers, known as the task of evaluating answers based on natural language, is automatically used by computation methods and machine learning algorithms. The development of smart learning systems and the importance of evaluation as a major factor in the learning process increases the need for a system that has great flexibility to evaluate tests based on texts. The student's answer compared to the ideal answer and points scored based on their similarity. Correlation and semantic similarity measures can also be used for this goal[7].

The semantic network representation provides the construction of knowledge. In a network, the pieces of knowledge are grouped together into clear semantic groups. It is a natural way of mapping knowledge. A network representation provides a graphical representation of the aspects of knowledge, the relationship between them, and their properties [8].

In this paper, an approach is proposed to evaluate short answers by comparing them to the standard answer using semantic networks and hierarchical relationships in nodes between networks. Initially, word processing is performed, followed by building semantic networks for each answer, then calculating similarity by finding similar nodes and similar relationships between the networks.

This paper is divided as follows: the second part shows the related work, the third part is semantic networks, the fourth part is the proposed approach, and the fifth part is the Experimental Results.

2. Related Work

Evaluating students' answers is one of the important and rapidly developing topics in e-learning. Therefore, many researchers have proposed methods for evaluating answers or vignettes, which differ according to the techniques, used in text processing and score calculation.

In [9], they proposed combining MaLSTM with sense vectors produced from SemSpace, a synchronization-based meaning fusion method that takes advantage of WordNet, to develop an approach to Automatic Classification of Short Answers (ASAG). As inputs to the parallel LSTM structure, simultaneous representations of student answers and reference answers are introduced, in the hidden layer. These two representation vectors are turned into sentence representations, and the vector similarity of these two representation vectors is computed with the data similarity in the output layer. In most of the dataset files, Pearson's (r) value>0.95 was gained.

The proposed approach was tested using the Mohler ASAG dataset and results were obtained in terms of Pearson (r) correlation and RMSE. Also, the proposed approach was tested as a case study using a specific dataset1 (CU-NLP) created from the exam of the "Natural Language Processing" course in the Computer Engineering Department of Cukurova University.

In [10] they proposed a recording system for short answers, constructed on skip thinking vectors. They introduced a simple SAG model called Ans2vec, which depended on changing the answers (student and model) into two vectors and measuring their similarities. Vector Skipping Thought - somehow - word2vec for sentences. When word2vec tries to predict surrounding words from certain words in a sentence, the thought skip vector extends that idea to sentences; it predicts surrounding sentences from a given sentence. Transcendental thinking vectors use the decoder model to first encode a sentence into a vector and then decode that representation into the surrounding sentences. Thought vectors do not consider the order of both words and sentences. Ans2vec achieves the best Pearson correlation value (0.63).

In [11], for this project, they proposed using BERT (Bi-Directional Encoding Representations of Transformers) as a tool to help teachers categorize automated short answers (ASAGs) rather than replace human judgment in high-risk scenarios. A second opinion is required to confirm the validity and consistency of judgment.

They suggested that to simplify the training process and provide data with a smaller percentage of the training data (for each question approximately 70 answers for each student), a compressed version of the BERT model was used called Bert-bas. they were able to achieve sufficiently high reliability among assessors for introductory human assistance with classification tasks.

The language model used by BERT, which represents bi-directional key encryption, in [12] was introduced as a revolutionary linguistic representation model. Many different tasks can be configured, such as ASAG, by adding a single layer to the deep neural network used. For this study a compressed version of the model was used. BERT’s original. The best model achieved a testing accuracy of 0.760 and a Cohen’s Kappa statistic of 0.684.

In [13], they suggested an approach to represent reference and student answers with a graph and to represent concepts using built-in vectors, learned from the knowledge graph directly. The indirect relationships between concepts, the vectors encoded, and a knowledge graph is created by extracting the concepts and their superficial relationships from the reference answers. After that, train the Tensor Neural Network (TNT). The idea is that when trained (TNT), the concept vectors encode the relationships between the entities/concepts in the knowledge graph. The more entities that share the same neighbors and relationships with the entities, the greater the similarity between the vector representations. The result shows that Aug2IP performed best with an average accuracy of 0.644.

In [14], they proposed a semantic-based QA assessment system by an artificial neural network (ANN). The system is also within the closed QA area where users are supposed to be able to access personalized online tests. They discussed two issues related to answer evaluation, i.e., length and paraphrase. They extracted patterns to create the sequence for a given answer. The system has a central filing system that includes reference materials as well as sample answers to questions. These are used as references to match and evaluate the candidate's answer. For each correct answer, a confidence factor for positivity was assigned when the selective pattern required for the candidate's answer matched the model's answer.

3. Semantic Networks

Semantic networks are a graphical representation of the semantic relationships between concepts. In a semantic network, knowledge is expressed by representing concepts with nodes, while directed binary relationships are represented by edges. This type of graph has found use in several subfields of artificial intelligence, including NLP, information retrieval,

and machine translation [15]. All semantic networks share an identifiable graphical representation, which can be used to represent knowledge and concepts [16].
The representation structure of knowledge is done through network representation. In a network, in related semantic groups, the pieces of knowledge are grouped together. Knowledge map representations provide a natural map between the natural language and these networks. Moreover, the network diagram provides a graphic representation of the entities of knowledge, their properties, and their relationships [17]. As shown in Figure (1).
A semantic net is not just information. It also makes it is easier to retrieve relevant facts. For example, all the facts are stored about the existence of "Ahmed", with a direct pointer to a single node representing "Ahmed".
Semantic networks offer a number of benefits regarding knowledge representation:
• They provide a means of clearly defining and representing the facts of the real world, i.e., semantics.
• The way that knowledge is structured in a semantic network reflects the representation of the structure of part of the real world.
• Representing relationships using "is-a" and "is-part of" supports forming hierarchies based on inheritance, which is useful in many applications.
• Hierarchy is useful for hypothetical thinking (For example, an adult's height can be assumed to be 175 cm, but if the person is a basketball player, then it is considered 190 cm).
• Semantic networks are useful for representing events and natural language sentences, the meanings of which can be very precise. Nevertheless, the idea of semantic networks is very general. This causes a problem unless it is clear about the syntax and semantics in each case.
An additional benefit of semantic networks is the inheritance of properties. If the semantic network represents the following bits of knowledge, for example, 'all canaries are yellow' and 'tweety islands', then the network is smart enough to conclude that 'tweety is yellow'. The matcher or retriever implements this conclusion.
Finally, Specific weights are assigned to similarities and measurements at different scales, and then text similarity is obtained by the method of the overlay. Experimental results show that using a tree structure to measure text similarity is feasible and effective [17].
The example below illustrates how a semantic network represents sentences.

The Sentences **The Semantic network**

- Lily is a cat.
- Lily is ginger in color.
- Lily caught a bird.
- John owns Lily.
- The cat sat on the mat.
- Cats like cream.
- A bird is an animal.
- A cat is a mammal.
- All mammals are animals.

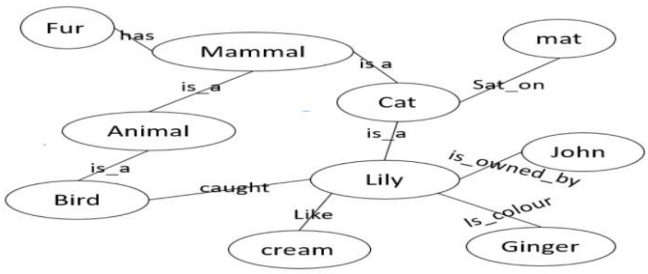

Figure 1- an example of a semantic network

4. Proposed Approach

The proposed approach is characterized by representing the model answer and the student's answer by means of semantic networks. These networks represent knowledge by linking concepts together in semantic relationships, accessing semantic facts, and representing them by a network containing nodes and edges that connect the nodes among themselves. After the semantic networks represent the text, the similarities between the networks are extracted and the student's answer is evaluated. The proposed approach calculates not only the frequency of words between texts but also the similarity of nodes and relationships in semantic networks. Figure 2 shows a block diagram of the proposed approach.

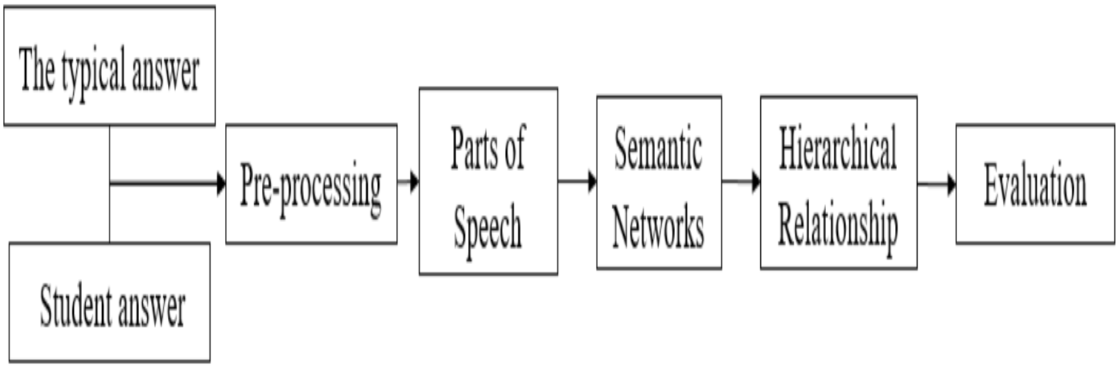

Figure 2- a diagrammatic overview of the proposed evaluation of short answers.

1- The first stage is pre-processing, where the text is processed through tokenization, normalization, and removing stop words.
Tokenization is the division of the text into sentences or words. Normalization is the process of converting the text into a basic (standard) form, reducing randomness in the text, and trying to find a unified pattern for words in the text. Removing stop words from the text documents involves detecting and removing prepositions, supporting nouns, etc., that do not give meaning to the text, e.g., 'I', 'me', 'my', 'myself', 'we', etc. These words are not useful and are removed because they are not considered keywords in text mining applications.
2- The next stage is Parts of Speech (PoS) tagging, which is used in identifying relationships between words and understanding the meaning of sentences. Basic sign groups may contain tags for the most used PoS, e.g., N: noun, V: verb, A: adjective, etc. It is possible to use additional features and distinguish between singular and plural nouns, chronology, verbal conjugations, and others. For languages, similar words may contain dissimilar parts of speech in the text. PoS tags are used to separate the repetition of words after a word is used, such as a noun or a verb. Part of speech tags is also used to search for grammatical and lexical designs. The similarities between the answers (correct, the student's answer) are measured by examining the similarity of nouns, verbs, adjectives, etc.
3- The next stage is building the semantic network. The semantic network has nodes that represent concepts and edges that represent relationships between nodes and labels that denote specific objects and relationships. After mining the text and extracting signs for the parts of speech, words in a sentence are represented as nodes of a network, without repeating the words in the nodes. These nodes represent individuals, events, attributes, etc. To represent the edges, it is necessary to find the link between words from the part of speech, the meaning of the sentence, and the linguistic analysis of the sentence. To clarify this, the sentence consists, for example, of subject, verb, event, description of things, etc., i.e., there is a correlation between the parts of the sentence. Let us consider the following example “John gave the ball to the cat“. There are three facts of natural language grammar, which are subjective: (i) the subject that performs the action, in this case, John, (ii) the objective, i.e., the thing for which the action is being performed, in this case, the ball, and finally, (iii) the dative, i.e., the recipient of the action, in this case, the cat. These different roles of things in a sentence are known as adverbs.

The concepts are John, the ball, and the cat. There is an association between John and the ball and between the ball and the cat (x $\rightarrow$ y), (y z). The network provides a structured representation of the concepts, relationships, and features that occur. In the text, it can be thought of as an approximation of the representation of knowledge covering these concepts.

4- There is a list of words (nodes) and a list of word pairs (relationships). To find out the semantic relationships between nodes in networks, hierarchical relationships between words are found, such as synonyms, word definitions, hyponyms, and hypernyms.

Natural language differs in the way speech is expressed for the same meaning, and this is the difficulty of understanding natural language. Therefore, hierarchical relationships are used to understand the true meaning of the use of certain words in a sentence and to find semantic similarities.

For example, let us consider two networks, Network A and Network B. If node X is in network A and is found in network B, then it will be added to the set of matching nodes. If there is no match, then The rest of the relationships will depend on the result of a search for synonyms for node X. Finally, the definition of node X is sought, and if there are nodes in network B that are considered within the definition.

A hyponym is a word with a more specific meaning than a generic term or at a semantic level above the word to which it is applied to. Alternatively, it can be thought of as a term denoting a subclass of a more general category. For instance, “chair” and “table” are hyponyms of “furniture.”

A hypernym is a word with a broad import under which more words are specifically included; these are known as subordinates. Moreover, it can be considered as a word whose meaning includes the meanings of other words. For example, “flower” is a hypernym of “daisy” and “rose”. In another word, hypernyms are common words, while hyponyms (known as subordinates) are parts of more common words.

The semantic relationship between a more definite word (e.g., “daisy” and “rose”) and the more common term (e.g., “flower”) is known as hyponymy or inclusion. After the nodes and similarity relationships are identified, the similarity of the answers is calculated and evaluated.

5- The following equations show how to calculate similarity and evaluate answers:

A number of similarities H, in parts of speech, i.e., verbs, nouns, adjectives, and adverbs, separated in each text, were compared between the answers, and the ratio between similar parts of speech and parts of speech (h) in the typical answer in X exists:

$$\mathbf{X = H / h} \tag{1}$$

The number of similar nodes N, between semantic networks of similar words, is counted in terms of the same word or its synonyms. The ratio between the number of similar nodes and the number of nodes (n) in the first semantic network of the typical answer in Y is calculated:

$$\mathbf{Y = N / n} \tag{2}$$

The edges and relationships between nodes are found by comparing the first and second nodes in both answers, extracting similar edges in E, and finding the ratio of similar relationships with relationships and edges (E) from the model answer in Z.

$$\mathbf{Z = E / e} \tag{3}$$

From Equations (1)-(3), total score of similarity is defined as:

$$\mathbf{Total\_similarity\_score = \alpha * X + \beta * Y + \gamma * Z} \tag{4}$$

Where $\boldsymbol{\alpha + \beta + \gamma = 1}$ **and** $\boldsymbol{0 < \alpha, \beta, \gamma < 1}$.

The total similarity score is computed from the weighted sum of X, Y, and Z similarity scores. The weights α, β, γ represent the contributions of the similarity ratios for parts of speech, nodes, and relationships, respectively.

The sum of the similarity points ranges between 1 and 0. The students' final scores are evaluated from 1 to 5, for example, if the similarity is 0, the final score is 0, and if the similarity is 0.9, the score is 5

The example in the Figure 3 illustrates building the semantic networks, extracting relationships, and hierarchical relationships.

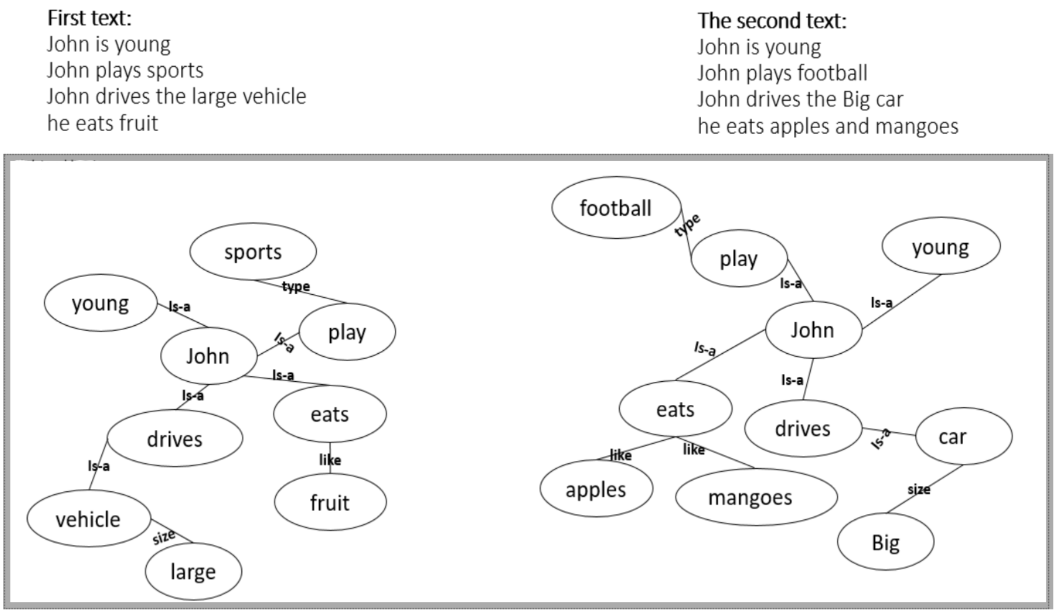

Figure 3- Illustrated Example

The nodes in the two texts are identified as follows:
Node 1= (John, young, plays, sports, drives, large, vehicle, eats, fruit)
Node 2= (John, young, plays, football, drives, big, car, eats, apples, mangoes)
While the relations between words are given below:
Relation 1 =[(John, young)(John, plays)(John, drives,)(John, eats)(plays, sports)…etc.]
Relation 2 =[(John, young)(John, plays)(John, drives,)(John, eats)(plays, football)…etc.]
Next, similar nodes, synonyms, hypernyms and hyponyms are identified:
Similar nodes= (John, young, plays, drives, eats)
Synonyms= (large, big)
Hypernym = (Vehicle is hypernym of car)
Hyponym = (mango is a hyponym of fruit)
As shown above, the semantic networks for two texts were built, the nodes and relationships were extracted, and hierarchical relationships (definition, synonyms, hyponyms, and hyponyms) were found between the networks and added to similar nodes and similar relationships.

5. Experimental Results

5 .1 Dataset description

In our proposed approach, the Mohler data set [18] was used. It is a data set commonly used for measurement purposes in ASAGE studies. The Mohler dataset was constructed using computer science tests. It consists of 12 tests; each test contains 7-8 questions. In the data set, there are 87 questions with a reference answer for each question. 26 to 31 students answered each question. The assessors evaluated the students' answers by comparing each student's answer with the reference answer and judging the similarity. Rating scores were given from 1 to 5. 5 is the most relevant and similar to the reference answer and 1 is not at all similar.

The dataset of questions was generated from introductory computer science assignments with answers provided by a class of undergraduate students. The tasks were administered as part of Data Structures course at the University of North Texas. For each task, student answers were collected via the WebCT online learning environment.
The answers were ranked independently by two human examiners, using an integer scale from 0 (not quite correct) to 5 (perfect answer). Both human referees were graduate students in the Department of Computer Science. Table 1 shows two questions, their correct answers, and three answers for students for each question. With the ratings of two of the human examiners [18] listed in the table.

Table 1-Two sample questions with short answers provided by students and results determined by human judges

Sample questions, correct answers, and student answers	**Grade**
Question: “What is the role of a prototype program in problem solving? “ “Correct answer: To simulate the behavior of portions of the desired software product.”	
Student answer 1: “A prototype program is used in problem solving to collect data for the problem.” Student answer 2:”It simulates the behavior of portions of the desired software product. “ Student answer 3: “To find problem and errors in a program before it is finalized.”	1, 2 5, 5 2, 2
Question: “What are the main advantages associated with object-oriented programming? “ “Correct answer: Abstraction and reusability.”	
Student answer 1:” They make it easier to reuse and adapt previously written code and they separate complex programs into smaller, easier to understand classes.” Student answer 2:” Object oriented programming allows programmers to use an object with classes that can be changed and manipulated while not affecting the entire object at once. “ Student answer 3: “Reusable components, Extensibility, Maintainability, it reduces large problems into smaller more manageable problems.”	5, 4 1, 1 4, 4

5.2 Results

To assess students' answers, the proposed approach was applied to 2025 Student Answers. From the Mohler dataset . The algorithm was implemented using the Python programming language version 3.8.5. Four different options for weights, as described in Section 4, and degrees of similarity compared to human assessments, were explored. The four weight configurations were as follows:
(i) Without weights (X, Y, Z),
(ii) A case with weights ($\alpha = 0.1$, $= 0.6$, $\gamma = 0.3$),
(iii) A case with weights ($\alpha = 0.1$, $\beta = 0.5$, $\gamma = 0.4$), and
(iv) A case with weights ($\alpha = 0.2$, $\beta = 0.5$, $\gamma = 0.3$).
To provide a numerical illustration of how Equation 4 works, let us say X equals 0.6, Y equals 0.6 and Z equals 0.5. Then, for weights $\alpha = 0.1$, $= 0.6$, and $= 0.3$, the sum of similarity points is 0.57. This means that the student's final grade is (3).
Figure 4. Shows the error ratio of applied different weight values on Mohler dataset .

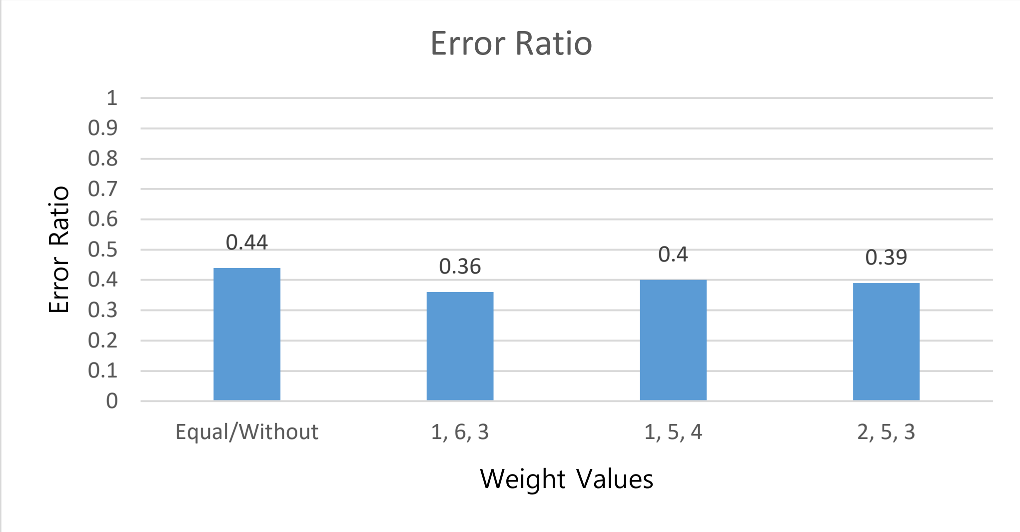

Figure 4- Error Ratio of Applied Different Weight Values of our Approach

The graph shows the error percentage of the results obtained from applying the proposed approach to the Mohler data set. The best results are for weights($\alpha = 0.1$, $\beta = 0.6$, $\gamma = 0.3$), which gives more weight in nodes similarity by 6, least similarity by relationships by 3, and least similarity by parts of speech by 1.

In some cases, we do not get exact similarities between sentences. This is because when a person thinks about two sentences, they don't only use semantic or grammatical similarities but also acquired perspectives in determining the similarity between the two, which cannot be considered by mathematical procedures. In the proposed approach, we use semantic similarity and appropriate weights are given to aspects of grammatical similarity to consider different aspects of sentences. The words were taken into account and linked together in the sentence. Finding similarities between words was profound in terms of taking aspects of the words' synonyms, definitions, word parts.

6. Conclusion and future work

The proposed approach finds the similarity between the answers by representing them with a semantic network and the hierarchical relationships between words and finding the similarity of the parts of speech between the answers. More than one weight was adopted to calculate the similarity, and the best of them was ($\alpha = 0.1$, $\beta = 0.6$, $\gamma = 0.3$), which achieved the best results. The data set used was the Mohler data set. For future work, this approach can be applied to essay questions and to represent the answers with a semantic grid.

When the evaluation is between sentences from a human perspective, measurements sometimes differ from arithmetic operations because the person does not rely only on the semantic or lexical aspect, but on the experiences gained, and this needs to develop processes and methods to get the best proportions in the similarity of texts.

References

[1] Palakorn Achananuparp , Xiaohua Hu, and Xiajiong Shen, "*The evaluation of sentence similarity measures", Philadelphia, pa 19104, springer-verlag berlin Heidelberg* 2008.

[2] Dheyaa Abdulameer Mohammed, Nasreen J. Kadhim,"Extractive Multi-Document Summarization Model Based on Different Integrations of Double Similarity Measures". *Iraqi Journal of Science*, vol. 61, no. 6, pp. 1498-1511, 2020

[3] Nur Syafiqah Mohd Nafis, Suryanti Awang." The Evaluation of Accuracy Performance in an Enhanced Embedded Feature Selection for Unstructured Text Classification". *Iraqi Journal of Science*, vol. 61, no. 12, pp. 3397-3407 , 2020

[4] Wael h. gomaa, aly a. fahmy, "A survey of text similarity approaches", Cairo, Egypt, *international journal of computer applications* (0975 – 8887) volume 68– no.13, april 2013.
[5] Dr. s. Vijayarani, Ms. j. Ilamathi, Ms. Nithya, "preprocessing techniques for text mining - an overview", Coimbatore, tamilnadu, India, *international journal of computer science & communication networks*, vol. 5, no. 1, pp. 7-16, 2015.
[6] Puthiaparampil, T., & Rahman, M. M. "Very short answer questions: a viable alternative to multiple choice questions". *BMC medical education*, vol. 20, no. 1, pp. 1-8, 2020.
[7] Mozhdeh Nazari Soleimandarabi, "semantic relatedness for automatic short answer grading", *Islamic Azad university*, Rasht, iran, 2020.
[8] John f. sowa, semantic networks, "The encyclopedia of artificial intelligence", 1992.
[9] Tulu, Cagatay Neftali, Ozge Ozkaya, and Umut Orhan." Automatic Short Answer Grading with SemSpace Sense Vectors and MaLSTM". IEEE Access 9 (2021): 19270-19280.
[10] Wael hassan gomaa and aly aly fahmy," Ans2vec: a scoring system for short answers", Cairo, Egypt, springer nature Switzerland ag 2020.
[11] Aubrey condor, "Exploring automatic short answer grading as a tool to assist in human rating", USA, springer nature Switzerland ag 2020.
[12] Devlin, Jacob, et al. "Bert: Pre-training of deep bidirectional transformers for language understanding." arXiv preprint arXiv:1810.04805 2018.
[13] Dipesh gautam and Vasile rus," Using neural tensor networks for open-ended short answer assessment", USA, springer nature Switzerland ag 2020.
[14] Toshavi patil, "Automatic assessment of descriptive answers for online examination using semantic analysis", pune institute of computer technology, 2019.
[15] Pia-ramona wojtinnekand Stephen pulman, "Building semantic networks from plain text and Wikipedia with application to semantic relatedness and noun compound paraphrasing", *international journal of semantic computing*,2012.
[16] Baydin, Atılım Günes, Ramon López de Mántaras, and Santiago Ontanón. "A semantic network-based evolutionary algorithm for modeling memetic evolution and creativity." Neural and Evolutionary Computing , 2014.
[17] k. r. Chowdhary, "Fundamentals of artificial intelligence", k. r. chowdhary, springer nature India private limited 2020.
[18] Mohler, Michael, Razvan Bunescu, and Rada Mihalcea. Learning to grade short answer questions using semantic similarity measures and dependency graph alignments. Proceedings of the 49th annual meeting of the association for computational linguistics: Human language technologies. 2011.