Applying Concept Map Mining in English Reading

Tzung Ruei Tsai
National Yunlin University of Science & Technology
g9923742@yuntech.edu.tw
Chung-Chian Hsu
National Yunlin University of Science & Technology
hsucc@yuntech.edu.tw
Yu-Fen Yang
National Yunlin University of Science & Technology
yangy@yuntech.edu.tw
Hui-Chin Yeh
National Yunlin University of Science & Technology
hyeh@yuntech.edu.tw

Abstract

For many years, evidences have shown that more-proficient students incorporate strategies into their reading. Constructing concept map is one of the helpful strategies in enhancing reading comprehension. The purpose of this study is to automatically extract concept maps from the texts to support students to grasp the main ideas in reading. Three phases were conducted in the mining process. The first phase was to extract keywords from each paragraph based on the following two principles: (1) The most important concept should appear many times in a paragraph as well as the other paragraphs in a text; (2) The most important concept should appear in the title and the topic sentence of each paragraph. The second phase was to build the relationships between keywords based on co-occurrence frequency. The final phase was to organize the keywords in three layers of the concept map as the inner first layer indicates the central idea of a text, the second layer presents the main idea of each paragraph, and the outer third layer indicates the supporting ideas in each paragraph. To evaluate accuracy of concept extraction, we compare the result of manual extraction with the extraction result of our approach. From the experiment result, performance criterion (PC) average scores of our maps and the best students' are 81% and 83%, respectively. The average PC score of students' maps is 68.5%. The results of the study show that the automatically extracted concept maps are better than those constructed by eight master students.

Keywords: concept map, computer-assisted learning, e-learning

Applying Concept Map Mining in English Reading

I. INTRODUCTION

Reading is an important and essential skill to acquire knowledge and exchange information. Many studies have shown that more-proficient students incorporate strategies into their reading (Chen et al. 2010; Cline et al. 2010; Huang et al. 2012; Liu et al. 2010). Constructing concept map is one of the helpful strategies which help students understand the text structure and offer them an effective way to find out the important concepts in an article. The concept map is a diagram which shows the relationships among concepts in a text. Students use it not only to understand the central idea in a text, but also to increase the retrieval and memory of knowledge(Friend 2000).

Constructing concept map is, however, very time consuming and labor cost. The previous studies commonly use corpora to analyze a large amount of documents to construct a complex concept map in a specific domain such as physics, algebra, or chemistry (Taşkin et al. 2011; Tseng et al. 2010). Some limitations are found in using the corpora. First, the concept map includes too many concepts in a map and becomes difficult and inconvenient for teachers and students to use in classroom (Anderson-Inman et al. 1993; Reader et al. 1994). Second, students cannot obtain immediate and individualized feedback if the concept map is shown on paper-based materials.

The purpose of this study is to propose a novel approach to construct a concept map from an article according to English writing principles and the frequency counts from the text automatically. In this study, we define a three level concept map. Each level represents the different important degree. The first level represents the main idea of the article. The second level represents the main idea of the paragraph. The third level represents the supporting ideas of the paragraph. For constructing concept map, first, we extract concepts from the article according to the following two principles: (1) The most important concept should appear many times in a paragraph as well as the other paragraphs in a text; (2) The most important concept should appear in the title and the topic sentence of each paragraph. Moreover, we calculate the relation between concepts based on their co-occurrences. Thus, readers can understand not only important the concepts, but also the organization of the article by this system-generated map. Specifically, this study was undertaken in order to create a novel reading learning system to assist students in their reading processes.

The remainder of this paper is organized as follows. Section 2 starts with a brief review of concept map including its construction and application. Section 3 describes the details of the novel approach. Section 4 describes the evaluation procedure and the results obtained. Finally, simulation results are provided in Section 5.

II. LITERATURE REVIEW

In this section, we briefly review tools and studies in the areas related to concept map. Section2.1 introduces concept map mining. Section 2.2 introduces concept map application. Section2.3 introduces automatic construction of concept map.

2.1 Concept map

In recent years, network graph to represent or organize knowledge has been applied in various fields. Concept map can be considered as one of the network graphs, in which concepts and relationship between concepts are represented by points and links, respectively (Novak 1990; Novak et al. 1984). This kind of concept map develops from the central to the external and it includes three levels (Fisher et al. 1991). The first-level word is the center (main concept) of the concept map. The second-level words are sub-concepts around the center. Finally, third-level words are remaining words after removing unimportant words. In general, this kind of concept map has two representations. Figure 1 shows the representations of concept maps. Until now, various concept mapping strategies have been proposed and applied in various domains (Bruillard et al. 2000; Gordon 2000).

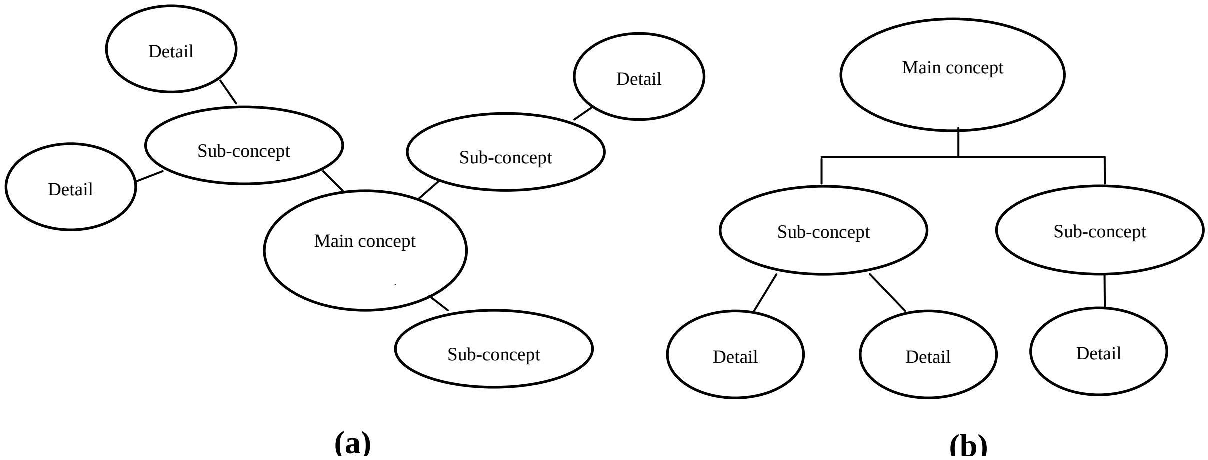

Figure 1.Representations of the concept map

2.2 Concept map applications

In educational domains, concept maps have received a lot of attention and become more and more popular. It has been applied to a wide variety of educational tasks including teaching(Horton et al. 1993; Stewart et al. 1979), learning(Kinchin 2000), evaluation(Chang 2007; Markham et al. 1994; Stoddart et al. 2000; Tsai et al. 2001). Moreover, there are many experiments showing positive impact in education.

Two typical uses of concept map for reading learning are demanding learners to map the concepts as well as establish links among the concepts by themselves and then the expert concept map is provided to assist them in scaffolding their learning. With development of computer-assisted tools of concept map, more and more instructors like to use computer-assisted tools to combine the above two approaches to further facilitate learning.

However, with ever-increasing reading materials, even though the Cmap tool(Clariana et al. 2004) , a free Web-based concept mapping system, had already incorporated the above combined features, it is still time consuming to produce the concept map from a large volume of natural language texts.

2.3 Automatic construction of concept map

How to create a concept map automatically becomes an important research issue in recent studies. In Valerio et al. (2006), they evaluated part of speech effectiveness of concept extraction and their result suggested that performance of using nouns phrases is better than others. In addition, stemming can further enhance the performance. In Valerio et al. (2008), they propose a series concept map extraction steps but they did not perform the quality assessment of the concept maps.

Chen et.al(2008) proposed a way to concept maps by mining a large amount of academic papers. They use the key words listed in the academic articles as concepts. Moreover, they defined the relations with the following four assumptions. First, each keyword listed in a research article represents one essential concept. Second, if two keywords appear in one research article, it implies relation exists between these two key words. Third, The higher the frequency of occurrence of two key words appearing in one sentence, the higher relation would be between them. Fourth, the shorter the “distance” between two keywords in one sentence, the higher the relation would be between them.

Instead of mining concept map from texts, Tseng et al.(2007) proposed a Two-Phase Concept Map Construction (TC-CMC) approach to automatically construct the concept map from learners’ historical testing records. The first phase is to transform numeric data to symbolic ones by Fuzzy Set Theory. The second phase is to analyze the mined association rules and transform them into prerequisite relationship among concepts for constructing concept map.

III. METHOD

In this section, we introduce the details of concepts extraction and relationship identification between concepts. First, an overview of automated concept map mining is presented in Section 3.1. Next, data preprocess are introduced in Section 3.2. Section 3.3 introduces concept identification. Section 3.4 introduces relationship identification. Section 3.5 discusses the integration of the concepts.

3.1 Automated concept map mining in articles

To organize important information in an article is an important step in reading process for learners. Therefore, we propose a concept map to assist learners in rapidly understanding the key ideas in the article.

The mining process is to construct a concept map from an article. In general, a concept map (CM) is defined as a triplet CM={C, R, T}. C is a set of concepts, R is a set of relationships, and T is the topology of concepts. After data pre-process which involves part of speech tagging and stemming, there are three main steps in the CM mining process including concept identification (CI), relationship identification (RI), and integration. The architecture of concept map mining process is depicted in Figure 2.

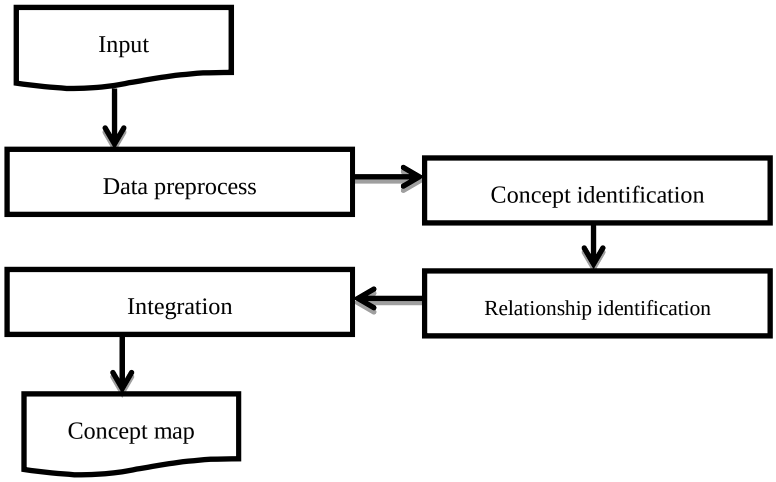

Figure 1. Concept map mining process

3.2 Data preprocess

In the first phase, text preprocessing is conducted. The purpose is to preprocess and clean the original text. Its tasks involves part of speech tagging, stemming, converting capital/small letter, identifying noun phrases and removing stopwords. The processing steps are shown in Figure 3.

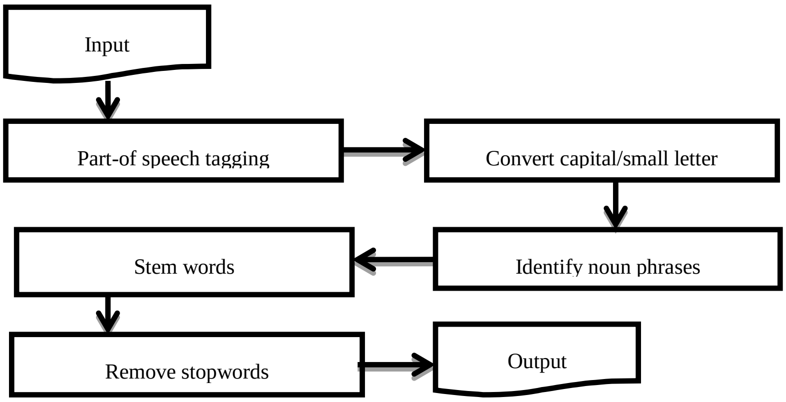

Figure 2. The process of data preprocess

1. Part-of-speech tagging :

   Nouns have more information than other parts of speech in texts. Thus, before we identify concepts, we have to perform part of speech tagging.

2. Convert capital/small letter :

   To avoid word cases to influence the result of concepts extraction. All of the non-proper noun phrases are converted into lower case.

3. Identify noun phrases :

   Noun phrases have more information than simple nouns in texts. Thus, after part of speech tagging, we combine nouns and related adjectives. In this step, we only consider two common noun phrase type: noun (noun noun…) and adjective noun (noun…). Figure 4 shows the algorithm of noun phrase extraction. In Figure 4, the inputs are the target word *w* and source text *t*, the output is the noun phrase *NP* which includes the target word. SetPOS() is to set the part-of-speech of the word to form the noun phrases. POS(*w*) is to get the part-of-speech of the target word *w*. POS(Last *w*) is to get the part-of-speech of the last word *w*.

4. Stem words :

To avoid word forms to influence the result of extraction, we convert all noun forms into singular noun. We use Martin Porter 's Porter Stemming Algorithm (Porter 1980) to perform the task.

5. Remove stopwords :

In general, stopwords, such as “I”, “be”, “at”, “am”, “can”, “you” … etc, appear frequently in a text and do not carry much information. Furthermore, in this study, criterion of concept extraction is mainly relied on frequency. Thus, we have to remove them.

```
GetNounPhrase Algorithm
Symbol Dwfinition:
CNP :candidate noun phrase
NP :the noun phrase including target word
w: the target word
t: source text
S: the set of the part of speech related noun
Input: w, t
Output: NP
Step 1:Initialization
    NP = null
    CNP=null
    S=null
Step 2:Set related part-of-speech
    S SetPOS ( noun, adjective)
Step 3:Getting candidate noun phrase
    While true
    If(POS(w)Contain S)
    CNP= CNP+w
    Else
    If(POS(Last w)Contain S)
    NP =CNP
        Else
    CLEAN CNP
    End while
    Return NP
```

Figure 3. The algorithm of GetNounPhrase

3.3 Concept Identification

In this study, we focus on only expository essays. Concept identification is based on word frequency. Moreover, we assume the articles are well-written and then two principles follow. In general, a well-written article follows two practices. First, there is only one topic in each paragraph. Second, the first sentence of the paragraph is the topic sentence. In addition, important sentences often include more important words (Li et al. 1997). Therefore, one principle is that the most important concept of the article should appear in the important places of the article. In this study, we consider the important places of the article are at the title and the topic sentences. The other principle is that the most important concept of the paragraph should appear many times in the paragraph and fewer times in the other paragraphs.

There are three levels of key concepts in the map. The first-level word is the center (main concept) meaning the main idea of the article. If the title has only one noun phrase, we assign the title noun phrase to be the main idea of the article. Otherwise, according to the first principle, we count the frequency of each noun phrase appearing in the title and the topic sentences. TF_i is occurrence frequency of key concept *i* in the title, TOF_i is occurrence frequency of key concept *i* in the topic sentences and *ParaNum* is the number of paragraphs of the article. Because *TF* is more important than *TOF*, *TOF* is divided by *ParaNum* to reduce its weight. Finally, we sum them up as a measure (*MS*) of key concepts. The noun phrase i' which has the highest score is considered the main idea of the article.

The formula is defined as follows:

$$i' = \text{argmax}\{TF_i + \frac{TOF_i}{ParaNum}\} \tag{1}$$

For example, there are the topic sentences and the title of the article "sharks" in Figure 5. The term "shark" appears in three topic sentences and the title of the article. The MS score of the term "shark" is 1.6. Moreover, because the title has only one noun phrase "shark", it is the main idea of the article in this example.

Sharks

Paragraph 1: People are scared of some types of animals.
Paragraph 2: There are about two hundred and fifty different species of **sharks**.
Paragraph 3**:** **Sharks** have a very good sense of smell.
Paragraph 4: Sometimes when people go swimming they splash in the way that is unnatural for animals in the ocean.
Paragraph 5: In fact, **sharks** should be very scared of people.

Figure 4. The title and topic sentences of article Sharks.

The second-level words are concepts (main ideas of paragraphs) around the center. Each node of the second-level means the main idea of the paragraph. According to the first principle, we count the frequency of each noun phrase appearing in the topic sentence of the paragraph. Next, according to the second principle, the frequency of each noun phrase is divided by its count of occurrence paragraphs where the phrase appears as a paragraph measure (*PMS*) of second-level concepts. We choose noun phrase i' with the highest score excluding central noun phrase from each paragraph. If the scores of two noun phrases are the same, we select the one which appears first under the assumption that an important concept shall appear first. TOF_i is occurrence frequency of key concept *i* in the topic sentence of the paragraph. PF_i is occurrence frequency of *i* in the paragraph. P_i is the count of occurrence paragraphs of *i*. The formula is defined as follows:

$$i' = \text{argmax}\{TOF_i + \frac{PF_i}{P_i}\} \quad (2)$$

For example, in Figure 6, the frequency of the term “people” is four in the first paragraph and occurrence frequency of the term “people” is one in the topic sentence of the paragraph. Moreover, the count of occurrence paragraphs of the term “people” is three. The PF_i of the term “people” is four. The P_i of the term “people” is three. The TOF_i of the term “people” is one. The PMS score is 2.33. After all noun phrases of the first paragraph are calculated, the highest *PMS* score is the term “people”. Therefore, it is main idea of the first paragraph in this example.

Sharks

Paragraph 1: **People** are scared of some types of animals. For example, snakes, bears, spiders and wasps frighten **people** because they have the potential to kill. When **people** go swimming in the ocean, they often get scared of sharks. They think sharks will suddenly appear and attack. But the reason for this fear is mainly due to false information and films about killer sharks. The truth is shark attacks on **people** are very rare.
Paragraph 4: Sometimes when **people** go swimming they splash in the way that is unnatural for animals in the ocean.
Paragraph 5: In fact, sharks should be very scared of **people**.

Figure 5. The first paragraph and topic sentences including ‘people’.

Finally, we discard unimportant words and then consider the remaining noun phrases of each paragraph to be supporting concepts. A noun phrase is considered unimportant if its total frequency is smaller than two in the article and it doesn’t appear in the topic sentence of the paragraph. For an important word, authors usually use it more than twice or it appears in the important places of the article according to the first factor. For example, in Figure 7, because the term “type” appears in the topic sentence of the first paragraph, we extract it to be a supporting idea. Moreover, the frequency of the term “ocean” is six, i.e., more than two, in the article. Therefore, we also extract it to be a supporting idea.

Sharks

Paragraph 1: People are scared of some **types** of animals. For example, snakes, bears, spiders and wasps frighten people because they have the potential to kill. When people go swimming in the **ocean**, they often get scared of sharks. They think sharks will suddenly appear and attack. But the reason for this fear is mainly due to false information and films about killer sharks. The truth is shark attacks on people are very rare.
Paragraph 2: Most sharks just live in the **ocean** like other fish.
Paragraph 3: they are often called the cleaners of the **ocean**. For example, a bird with a broken wing that splashes helplessly in the **ocean** may attract a shark and be eaten.
Paragraph 4: Sometimes when people go swimming they splash in the way that is unnatural for animals in the **ocean**. It is only doing what its instincts tell it: to clean the **ocean** of injured animals.

Figure 6. An example of the supporting ideas of the first paragraph.

3.4 Relationship Identification

The idea of relationship identification is based on the co-occurrence frequency between key concepts. We consider that high co-occurrence frequency between two concepts in the paragraph indicates tight relationship. The processing steps are shown in Figure 8.

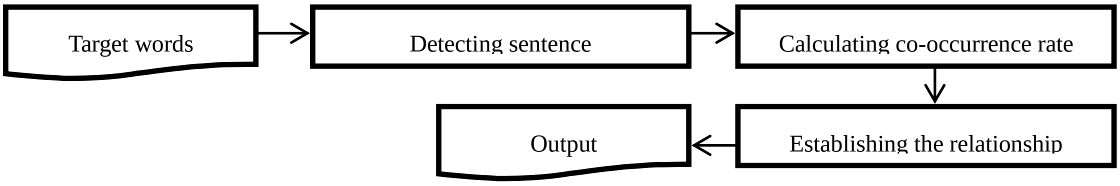

Figure 7. The processing steps of relationship identification

1. Sentence detection

This step is to extract all sentences which include the target word. Then, we use the sentences to determine the relation between concepts. Figure 10 shows the algorithm of sentence extraction. In Figure 10, the inputs are the target word *w* and source text *t*, and the output is the sentence *S* which includes the target word. GetPosition(*w*) is the index of the target word *w* from source text *t*. GetBackwardPeriodPosition (*w*) is the index of the closest period before target word appears. GetForwardPeriodPosition (*w*) is the index of the closest period after target word appears. GetLastWord(*t*) is the last word of text. Figure 9 shows an example of sentence detection.

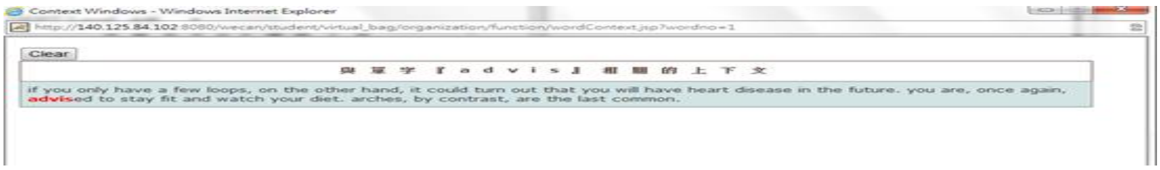

Figure 8. An example of sentence detection

```
GetSentence Algorithm
Input: the target word w
, source text t
Output: the sentence including target word S
Step 1:Initialization
    S = null
    p = getPosition(w)
    startPoint = 0
    endPoint = 0
Step 2:getting sentence starting point
    While true
    if (getPosition(w) != null)then
        if(getBackwardPeriodPosition (w)!=null)then
        startPoint = getBackwardPeriodPosition (w)
        else
        startPoint =0
Step 3:getting sentence end point
    if (getPosition(w) != null)then
        if ( getForwardPeriodPosition (w)!=null)then
        endPoint = getForwardPeriodPosition (w)
        else
        lw=getLastWord(t)
    endPoint = getPosition(lw)
    End while
    S = getSubString(startPoint, endPoint)
    Return S
```

Figure 9. The algorithm of GetSentence

2. Co-occurrence rate calculation

   The step is to count the co-occurrence rate between key concepts by their co-occurrence sentences. First, we calculate the relation between the center and each main idea of paragraphs. Second, for each branch, we calculate the relation between the main idea of the paragraph and each supporting idea of the paragraph. We use the parent concept of them as denominator to calculate co-occurrence rate.

3. Paragraph relationship establishment

   The final step is to eliminate the influence of the number of sentences in the paragraph. From the previous step, we find the co-occurrence rate between key concepts. Next, we divided the rate by the count of sentences of the paragraph to normalize the relationship score between key concepts. R_{ijk} is the score of relationship between concept i and its child concept j in the k paragraph. OCF_{ijk} is the number of co-occurrence sentences between i and j in the k paragraph. F_{ik} is the frequency of i for each branch in the k paragraph. P_k is the count of sentences in the k paragraph. The formula is defined as follows:

$$R_{ijk} = \frac{OCF_{ijk}}{F_{ik} \times P_k} \tag{3}$$

There is an example in Figure 11. In the article "shark", the main idea of first paragraph is the term "people". Then, we calculate the relationship between people and the supporting idea "type" of the first paragraph. The OCF_{ij1} of the term "people" and the term "type" in the first paragraph is one. Because the term "people" is the parent concept of "type", F_{i1} is four and P_1 is six. Therefore, the relationship score R_{ij1} between people and type is about 0.04.

Sharks

Paragraph 1: **People** are scared of some **types** of animals. For example, snakes, bears, spiders and wasps frighten **people** because they have the potential to kill. When **people** go swimming in the ocean, they often get scared of sharks. They think sharks will suddenly appear and attack. But the reason for this fear is mainly due to false information and films about killer sharks. The truth is shark attacks on **people** are very rare.

Figure 10. An example of calculating of the relationship score.

3.5 Integration

The development of our concept map is from inner to outer. The center is the main idea of the article. Concepts of a branch are extracted from the same paragraph. After the extraction step, we link concepts between different levels to form a complete concept map. Concepts in the same level are not linked. Figure 12 shows an example of a concept map. For the article "Sharks", the central idea is shark. The main ideas of paragraphs are people, different species, bird, animal and fact. The words in the third level are the supporting ideas of each paragraph.

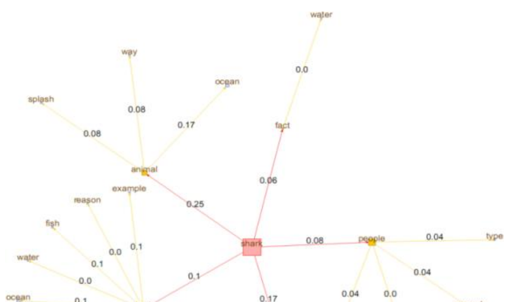

Figure 11. An example of concept map.

IV. EXPERIMENT

The experiment is to evaluate accuracy of extracted concept maps. Section 4.1 introduces data collection. Section 4.2 introduces experiment participants. Section 4.3 introduces the experimental result of comparing the result of manual extraction with the extraction result of our approach. Section 4.4 analyzes experiment error.

4.1 Data collection

In English articles, the structure of expository essay is more systematic than that of other styles. Therefore, the experiment data consisted of eleven expository essays from the book "Reading Comprehension (basic)" published in 2001 by Matthew McGinniss. There are thirty articles in the book. The length of the experimental articles is between 300 and 400 of words.

4.2 The experiment participants

The participants consisted of eight master students and a professor of the Department of Applied Foreign Languages from National Yunlin University of Science and Technology. The students are the comparison objects and the professor assists in identifying the right concepts.

4.3 The experiment result of concept extraction

To evaluate accuracy of concept extraction, we compare the result of manual extraction with the extraction result of our approach. Each article was read by two students and then maps were drawn by them. The first and second level concepts mean the main idea of the article and of the paragraphs, respectively. Moreover, the number of main ideas is fixed. Therefore, we focus on measuring their correctness. We design a performance criterion (PC). For each article, we choose the higher PC score one from experimental objects of the article to be the comparison object. To identify correct maps, we first found out the difference of main ideas of the paragraphs and the article between our map and students' maps. Then, the right concepts are identified by consulting the professor. *MI* is the score of the main idea. If the extracted main idea is correct, the concept map gets a five-point score. *PI* is the score of a paragraph main idea. For each paragraph idea extracted correctly, the map can get a five-point score from the paragraph. *ParaNum* is the number of paragraphs of the article. Finally, we sum up *MI* and the average of *PI* and then the sum is divided by ten. Table 1 shows the experiment result of concept extraction. Average PC scores of our maps and the best students are 81% and 83%, respectively. Our maps are worse than the maps of best experiment objects. However, we found some students can't identify the key ideas of the article accurately. Therefore, on the whole, our maps are better than students. Table 2 shows the comparison result of the article "shark". The criterion is defined as follows:

$$\mathrm{PC} = (MI + PI \,/\, ParaNum) \,/\, 10 \tag{4}$$

Table 1. The experiment result of concept extraction.

Article	PC Scores of the worse student	PC Scores of The best student	PC Scores of Our map
Alfred Nobel	75%	100%	75%
Mother Teresa	58%	83%	75%
Jeans	50%	91%	75%
Reduce, Reuse and Recycle	50%	80%	90%
Rollerblading	12.5%	100%	87.5%
Tennis at Wimbledon	63.3%	75%	83%
Shark	50%	70%	70%
Amazon Rainforest	62.5%	75%	75%
Arctic	10%	80%	90%
World cup	60%	60%	80%
Traveling	100%	100%	90%
Average PC Score	54%	83%	81%

Table 2. The comparison result of the article "shark"

Idea	The best student	The worse student	Our approach	Teacher
Main idea of the article	shark	shark	shark	shark
Main idea of the paragraph 1	types	potential	people	scared
Main idea of the paragraph 2	different species	ocean	different species	different species
Main idea of the paragraph 3	good sense	smell	bird	good sense of smell
Main idea of the paragraph 4	animal	people	animal	animal
Main idea of the paragraph 5	Skin	fact	fact	scared
PC score	70%	50%	70%	100%

4.4 Error analysis

In this experiment, the concept map of "Shark" is the worst one. In our map, the main ideas of paragraph 1, 3 and 5 are incorrect. Comparing with the map by the English expert, we can find two main reasons. First, not only noun phrases but also verbs can be considered the key ideas of the article, for instance, the one 'scared' is the third row of Table 2. In present study, our approach considers only noun phrases. Moreover, we only consider two types of noun phrases: noun (noun noun…) and adjective noun (noun…). Other form of noun phrases can be not extracted yet, such as the one 'good sense of smell' in the fifth row.

V. CONCLUSION

In reading, concept map is a useful tool. However, it is very time consuming to manually construct a concept map from an article. In this study, we propose an approach to automatically constructing concept map from the article. The map can be used to assist readers in understanding not only important concepts of the article, but also the relationship between concepts. From the experiment result, PC average scores of our maps and the best students' are 81% and 83%, respectively. The average PC score of students' maps is 68.5%. The performance of our maps is worse than best students'. However, it is better than the average of students'. In other words, under the assumption that the articles are well-written, our maps are useful for learners. In the near future, we will consider more types of noun phrases and verbs to increase accuracy of concept extraction.

ACKNOWLEDGEMENTS

The work of this paper is supported by the National Science Council, Taiwan, under grant NSC 100-2410-H-224-003-MY2 and NSC100-2410-H-224-017.

REFERENCE

1. Anderson-Inman, L., and Zeitz, L. "Computer-based concept mapping: Active studying for active learners.," *Computing Teacher* (21:1) 1993, pp 6–11.
2. Bruillard, E., and Baron, G. "Computer-based concept mapping a cognitive tool for students: A rerview," 2000, pp. 331-338.
3. Chang, S.-N. "Externalizing students'mental models through concept maps. ," *Journal of Biological Education* (41:3) 2007, pp 107–112.
4. Chen, N.-S., Kinshuk, Wei, C.-W., and Chen, H.-J. "Mining e-Learning domain concept map from academic articles," *Computers & Education* (50:3) 2008, pp 1009-1021.
5. Chen, S.-M., and Bai, S.-M. "Using data mining techniques to automatically construct concept maps for adaptive learning systems," *Expert Systems with Applications* (37:6) 2010, pp 4496-4503.
6. Clariana, R. B., and Koul, R. " A computer-based approach for translating text into concept map-like representations.," Paper presented at the first international conference on concept mapping, Pamplona, Spain., 2004.
7. Cline, B. E., Brewster, C. C., and Fell, R. D. "A rule-based system for automatically evaluating student concept maps," *Expert Systems with Applications* (37:3) 2010, pp 2282-2291.
8. Fisher, C., Gleitman, H., and Gleitman, L. R. "On the semantic content of subcategorization frames.," *Cognitive Psychology* (23) 1991, pp 331–392.
9. Friend, R. "Teaching summarization as a content area reading strategy," *Journal of Adolescent & Adult Literacy* (44:4) 2000, pp 320-329.
10. Gordon, J. L. "Creating knowledge maps by exploiting dependent relationships," *Knowledge-Based Systems* (13:2) 2000, pp 71-79.
11. Horton, P. B., McConney, A. A., Gallo, M., Woods, A. L., Senn, G. J., and Hamelin, D. "An investigation of the effectiveness of concept mapping as an instructional tool," *Science Education* (77:1) 1993, pp 95-111.
12. Huang, H.-S., Chiou, C.-C., Chiang, H.-K., Lai, S.-H., Huang, C.-Y., and Chou, Y.-Y. "Effects of multidimensional concept maps on fourth graders' learning in web-based computer course," *Computers & Education* (58:3) 2012, pp 863-873.
13. Kinchin, I. M. "Concept-Mapping Activities To Help Students Understand Photosynthesis--and Teachers Understand Students," *School science review* (82) 2000, p n299.
14. Li, J.-J., and Choi, K.-S. "Corpus-Based Chinese-Korean Abstracting Translation System ", IJCAI, 1997, pp. 972-976.
15. Liu, P.-L., Chen, C.-J., and Chang, Y.-J. "Effects of a computer-assisted concept mapping

learning strategy on EFL college students' English reading comprehension," *Computers & Education* (54:2) 2010, pp 436-445.

16. Markham, K. M., Mintzes, J. J., and Jones, M. G. "The concept map as a research and evaluation tool: Further evidence of validity.," *Journal of Research in Science Teaching* (31:1) 1994, pp 91–101.
17. Novak, J. D. "Concept maps and Vee diagrams: Two metacognitive tools to facilitate meaningful learning," *Instructional science* (19:1) 1990, pp 29-52.
18. Novak, J. D., and Gowin, D. B. *Learning how to learn* Cambridge Univ Pr, 1984.
19. Porter, M. F. "An algorithm for suffix stripping," (14:3) 1980, pp 130–137.
20. Reader, W., and Hammond, N. "Computer-based tools to support learning from hypertext: Concept mapping tools and beyond.," *Computers and Education* (12) 1994, pp 99–106.
21. Stewart, J., Vankirk, J., and Rowell, R. "Concept maps: A tool for use in biology teaching.," *American Biology Teacher* (41:3) 1979, pp 171–175.
22. Stoddart, T., Abrams, R., Gasper, E., and Canaday, D. "Concept maps as assessment in science inquiry learning-a report of methodology," *International Journal of Science Education* (22:12) 2000, pp 1221-1246.
23. Taşkin, M., Pepe, H., Taşkin, C., Gevat, C., and Taşkin, H. "The effect of concept maps in teaching sportive technique," *Procedia - Social and Behavioral Sciences* (11:0) 2011, pp 141-144.
24. Tsai, C., Lin, S. S. J., and Yuan, S. "Students' use of web-based concept map testing and strategies for learning," *Journal of Computer Assisted Learning* (17:1) 2001, pp 72-84.
25. Tseng, S.-S., Sue, P.-C., Su, J.-M., Weng, J.-F., and Tsai, W.-N. "A new approach for constructing the concept map," *Computers & Education* (49:3) 2007, pp 691-707.
26. Tseng, Y.-H., Chang, C.-Y., Rundgren, S.-N. C., and Rundgren, C.-J. "Mining concept maps from news stories for measuring civic scientific literacy in media," *Computers & Education* (55:1) 2010, pp 165-177.
27. Valerio, A., and Leake, D. "Jump-starting concept map construction with knowledge extracted from documents.," Citeseer, Paper presented at the second international conference on concept mapping, San José, Costa Rica., 2006.
28. Valerio, A., and Leake, D. " Associating documents to concept maps in context.," Citeseer, Paper presented at the third international conference on concept mapping, , Finland 2008.

運用概念圖於英文閱讀

蔡宗叡
國立雲林科技大學
g9923742@yuntech.edu.tw
許中川
國立雲林科技大學
hsucc@yuntech.edu.tw
楊育芬
國立雲林科技大學
yangy@yuntech.edu.tw
葉惠菁
國立雲林科技大學
hyeh@yuntech.edu.tw

摘要

本研究的目的在於從文章中自動地建構概念圖，協助學生閱讀以及了解文章主旨。建構的過程主要分為以下三個階段，第一階段為抽取主要概念，第二階段為概念間關係的建立。最後階段為依據不同的重要性，把關鍵字分三個層級組織起來，第一層級為中心點，它表示整篇文章最重要的概念，第二層級為每個段落最重要概念，第三層為每段落中其他重要概念。實驗主要與應用外語系學生進行比較，針對不同層級抽取概念，由英語專家評估雙方的正確性，結果為本研究的自動化概念圖的正確率為 81%略低於學生中的最佳結果 83%，但高於學生們的平均結果 68.5%。

關鍵詞：概念圖、電腦輔助學習、電子化學習