Available online at www.sciencedirect.com
ScienceDirect
Knowledge-Based Systems 21 (2008) 52–61
Knowledge-Based SYSTEMS
www.elsevier.com/locate/knosys

Self-associated concept mapping for representation, elicitation and inference of knowledge

W.M. Wang, C.F. Cheung *, W.B. Lee, S.K. Kwok

Department of Industrial and Systems Engineering, The Hong Kong Polytechnic University, Hung Hum, Kowloon, Hong Kong

Received 18 March 2006; received in revised form 1 November 2006; accepted 16 November 2006
Available online 11 December 2006

Abstract

Concept maps have been widely put to educational uses. They possess a number of appealing features which make them a promising tool for teaching, learning, evaluation, and curriculum planning. This paper presents self-associated concept mapping (SACM) which extends the use of concept mapping by proposing the idea of self-construction and automatic problem solving to traditional concept maps. The SACM can be automatically constructed and dynamic updated. A Constrained Fuzzy Spreading Activation (CFSA) model is proposed to SACM for supporting rapid and automatic decisions. With the successful development of the SACM, the capability of Knowledge-based systems (KBS) can be enhanced. The concept and operational feasibility of the SACM is realized through a case study in a consultancy business. The theoretical results are found to agree well with the experimental results.

Keywords: Knowledge representation; Self-associated concept maps; Concept mapping; Knowledge management; Knowledge-based systems

1. Introduction

Cognitive psychology stated that people do not learn by memorizing, instead, they learn by summarizing, relating, and organizing concepts into their cognitive structures [1]. New knowledge is assimilated into their cognitive structures through construction and not merely acquired [30]. Based on this learning theory, a method known as concept maps was developed. Concept maps are widely used as a means of visualizing one's inner cognitive structures. Concept maps require users to identify, graphically display, and link key concepts by organizing and analyzing information. They make the structure of knowledge visually explicit and conceptually coherent. There are numerous applications for concept maps including communication [25,34,40], teaching [2,12,22,31], assessing users understanding [6,37,38], curriculum design [7,28], planning [11,36], etc. Numerous of tools and commercial software have been developed including construction tools of concept maps in different education and business settings [9,14,16,17,20,21,27,39], tools for handheld devices supporting mobile learning [3], tools for navigation and discovery concept map in a repository [26] and so on. Moreover, several researchers have developed different methodologies to extend the usage of concept map [5,6,17,18,24].

In the past, concept maps are manually constructed by the users and the usages of concept maps are focused on educational purpose. Although a number of concept mapping tools are available, the construction and interpretation of concept maps are still heavily relied on human being. Traditional maps are static after the development process, which require human interventions for any later changes of the maps. The construction of the maps is difficult, time consuming and expensive. Moreover, the interpretation of concept maps rely on human who is not suitable tool for computational inference.

On contrary to the current trend of the development of concept maps, this paper attempts to give the idea of

* Corresponding author. Tel.: +852 27667905; fax: +852 23625267.
E-mail address: mfbenny@inet.polyu.edu.hk (C.F. Cheung).

0950-7051/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.knosys.2006.11.015

concept maps with self-construction ability and automatic problem-solving ability. The extended concept map is called Self-Associated Concept Map (SACM). SACMs can be automatically constructed and dynamically updated from a knowledge repository with structural historical records. A Constrained Fuzzy Spreading Activation (CFSA) model is incorporated in the SACM which enables the decision supporting function for providing rapid and automatic decisions. With SACM, the capability of Knowledge-based systems (KBS) can be enhanced and extended. The paper starts by reviewing the related work and then describes the proposed SACM. The capability of the SACM is realized by a case study in a consultancy business. The results indicate that the proposed idea of SACM is well suited for KBS with real-world data.

2. A comparison between Traditional Concept Mapping and Self-Associated Concept Mapping

2.1. Traditional concept maps

Concept map has its root from its relationship to memory and learning theory. Semantic memory theory believes that knowledge is stored in a network format where concepts are connected to each other [8]. The more tightly interconnected the knowledge representation, the more likely it is that a person will recall information at the appropriate time. As a result, a network representation can be used to show the integration of different concepts. The theory has resulted in different terms being used to describe concept maps including semantic networks [13,15] and knowledge maps [19].

In 1984, Novak proposed concept map to represent knowledge [29–31]. It is an instructional method that integrates new information into an old knowledge structure. It promotes conceptual understanding by displaying meaningful patterns of ideas. Knowledge is graphically displayed as a network of nodes and links. A concept map consists of sets of propositions. Each proposition is made up of a pair of nodes and a link connecting them. The labeling of nodes contains the concepts. The labeling of the links provides information about the nature of the relationships. Cross-links sometimes appear to show the connections between and among concepts, create an interdisciplinary space for inquiry and learning, or provide examples for clarifying the meaning of a given concept. Concept maps are varying on the basis of an individual area of interest and style. Fig. 1 shows an example of concept map. There are propositions in the concept map: (Concept Map consists of Concepts), (Concept Map consists of Relations), (Concepts denoted by Nodes), (Relations denoted by Links).

Several research studies have developed different methodologies to extend the usage of concept map. Lin et al. [24] introduce a concept map focusing on the propositions with weights, which is named "weighted concept map". Chen et al. [5] proposed an extended concept maps called attributed concept maps (ACM). ACM associates its

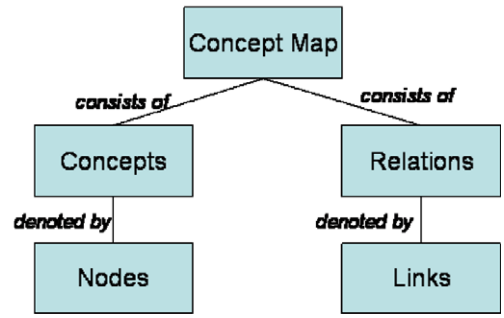

Fig. 1. An example of concept map.

concept nodes and relation links with attribute values which indicate the relative significance of concepts and relationships in knowledge representation. A Two-Phase Concept Map Construction (TP-CMC) algorithm is proposed by Sue et al. [35] to automatically construct a concept map of a course by historical testing records. They apply Fuzzy Set Theory to transform the numeric testing records of learners into symbolic, apply Education Theory to further refine it, and apply Data Mining approach to find its grade fuzzy association rules. Then, they use multiple rule types to further analyze the mined rules and a heuristic algorithm is proposed to automatically construct the concept map according to the results of the analysis.

Recently, more and more researches are applying concept maps on knowledge management. Concept mapping is provided as a knowledge management tool so that concepts can be captured, queried, and perhaps most importantly connections discovered and reasoned about [14]. Current research work is focusing on the construction of theoretical frameworks and design of human–machine interfaces (e.g. [19,26]). They provide tools for creating custom templates, publishing the maps as web pages, associating documents and URLs with concepts and some query and search capabilities. However, this is insufficient for performing such knowledge management activities since most of the work is still relied on human being. It seeks for a tool with the ability to automatically discover implicit connections, generate new maps, maintain evolution of maps, reasoning, and inferencing.

In order to support the automatic inference of concept map, spreading activation (SA) is adopted. Similar to concept maps, the SA model also has its roots from its relationship with human memory [33]. It has often been associated with semantic networks. During spreading, the activation input of a node in the network is calculated based on the following equation:

$$I_j = \sum_i O_i w_{ij} \tag{1}$$

where I_j is the total input of node j, O_i is the output of node i connected to node j, and w_{ij} is a weight associated to the link connecting node i to node j.

After the input value of a node has been computed, the activation level of the node is determined by a function of the input:

$$A_j = f(I_j) \tag{2}$$

where A_j is the activation level of node j, f is the activation function, and I_j is the input of node j.

The output of the node, O_j, is usually its activation level, A_j. The output value of the node is fired to all nodes connected to it. Hence, the activation spreads pulse after pulse until a termination condition has been met.

The most salient fault of pure SA is that the activation tends to quickly spread over the entire network [32]. The shortcoming can be partially overcome by the implementation of rules to control the activation. This new model is called Constrained Spreading Activation (CSA). Some common constraints are distance constraint, fan-out constraint, path constraint and activation constraint [10].

In this paper, the authors attempt to propose a new extended concept map: Self-associated Concept Map (SACM). It has a knowledge representation which is similar to the ACM mentioned above. Contrasting to manually constructed ACM, SACM can be automatically constructed and dynamically updated from a knowledge repository with structural historical records. On the other hand, the automatic construction method of SACM is different to TP-CMC that mentioned above. TP-CMC applies fuzzy set theory on measuring the grading of the historical records, while SACM applies fuzzy set theory on dividing the concepts within the historical records which increase the ability of inference. For knowledge inference, a new model named Constrained Fuzzy Spreading Activation (CFSA) is proposed. It integrates fuzzy logic and CSA so as to provide more precise, rapid and automatic solutions.

2.2. Self-Associated Concept Mapping

2.2.1. Knowledge representation

The graphical representation provides insights for describing the relationships among different knowledge concepts. A SACM is represented by a simple graph with nodes and edges. The nodes represent concepts relevant to a given domain and the association relationships between them are depicted by directed edges. An example of SACM is shown in Fig. 2. The importance of the concepts and the associations between different concepts are indicated by the depth of color i.e. darker color indicates higher importance. (The detail symbolic representation of knowledge representation, knowledge elicitation, and knowledge inference is put in the appendix for interested readers.)

2.2.2. Knowledge elicitation

With the advanced development of computer technology and Knowledge-based system (KBS) in the recent decade, organizations are able to record the working activities of each worker at a dramatically lower cost. Some KBSs have been developed to serve this purpose. The knowledge of knowledge workers can be assimilated and stored in a structured format into the knowledge repositories of KBSs when they use the KBSs for performing their daily work [4]. This enables the offer of vast new mines of information on individual working knowledge automatically and objectively. By following the learning theory, individuals' abilities to work depend on whether they have an appropriate concept map of working. In the proposed approach, the concept map of each individual is constructed based on the information in the knowledge repositories of the KBS and it enables dynamic update by adding new records to the knowledge repository.

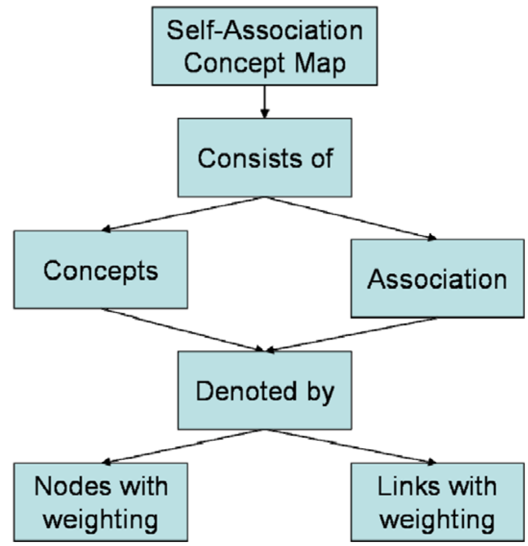

Fig. 2. An example of SACM.

The elicitation algorithm consists of 3 steps:

Step 1: Constructing a temporary SACM based on the inputs of the new record.

- Distinct concepts are extracted from the new record for the construction of a set of concept nodes and the degree of importance of concept is assigned.
- If the value of fields of the new record is symbolic, a concept node is added and its corresponding degree of importance of the added concept is assigned to be 1.
- If the value of fields of the new record is numeric, more than one concept nodes may be added according to the value. For example, if the value is "Quantity (0.7 High, 0.2 Medium, 0 Low)", then two new concept nodes: "Quantity: High" and "Quantity: Medium" are added. The corresponding degree of importance of the added concepts "Quantity: High" and "Quantity: Medium" are 0.7 and 0.2, respectively.
- Assign the degree of importance of relation for each pair of concepts.

Step 2: Combining the temporary SACM with the original SACM.

- The nodes and relations of the temporary SACM are matched with that of the original

SACM. If there is any missing concept existed, the degree of importance of that concept and the degrees of importance of that concept's associated relations are assigned to be 0.
- The degrees of importance of the concepts and the degrees of relation among concepts of the original SACM are adjusted based on that of the temporary SACM.

Step 3: The parameters of the combined SACM is adjusted and the combined SACM is normalized.

2.3. Knowledge inference

A new model named Constrained Fuzzy Spreading Activation (CFSA) which integrates fuzzy logic and CSA is introduced for knowledge inference. It consists of 3 steps:

Step 1: The enquiry/problem is converted into SACM format
- This step is the same as step 1 of knowledge elicitation.

Step 2: The activation level of each node of the SACM is computed
- The nodes and relations of enquired SACM are matched with that of SACM that used for performing knowledge inference. If there is any missing concept existed, the degree of importance of that concept and the degrees of importance of that concept's associated relations are assigned to be 0.
- The activation level of each node of the SACM is computed based on the degree of importance of the nodes and relations.

Step 3: Generation of the result
- The concept nodes of the result field are extracted and their corresponding activation levels are calculated.
- If the extracted concept node is a fuzzy value, the activation level is computed by defuzzifiion by the centre of gravity (COG) method.
- If the extracted concept node is symbolic, the concept node with the highest activation level is selected to be the result.

2.4. An application example

In this section, an application example is used to illustrate the proposed methodology. This is a simple example to show how a SACM is constructed and how a SACM infers quantitative prediction. Table 1 shows a simplified knowledge repository that stores 2 quotation records. Each record includes a Case ID, Product Type, Diameter and Unit Price. Case ID provides a unique index of the records.

Table 1
A simplified knowledge repository

Case ID	Product type	Diameter	Unit Price
1	Prototype Lens	7	1400
2	Mould Insert	12	2000

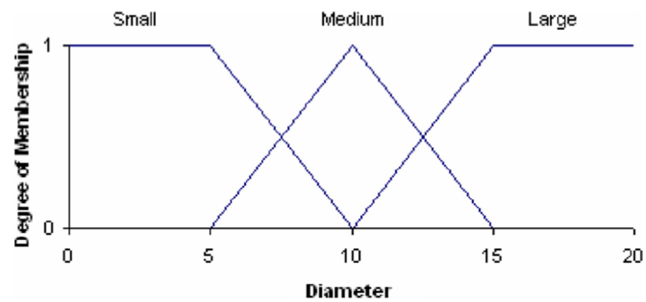

Fig. 3. The membership functions of Diameter.

Fig. 4. The membership functions of Unit Price.

Product Type contains symbolic values. Diameter and Unit Price contain numerical values. The membership functions of Diameter and Unit Price are shown in Figs. 3 and 4, respectively. In this example, Diameter is represented by 3 fuzzy regions which are Small, Medium and Large, respectively. The Unit Price is represented by 3 fuzzy regions which are Low, Medium and High, respectively.

In order to construct a new SACM, all distinct concepts from the first record are firstly extracted for the construction of a set of concept nodes ***C***, the set of degrees of importance of concept ***F*** are assigned, and the set of degrees of importance of relations ***L*** for each pair of concepts are assigned, based on the step 1 of knowledge elicitation that described in Section 2.2.2. The results are shown

Table 2
Temporary results of SACM that assimilates 1 record

i	C_i	F_i	L_{ij}	C_1	C_2	C_3	C_4	C_5
1	Product Type: Prototype Lens	1	C_1	0	0.6	0.4	0.2	0.8
2	Diameter: Small	0.6	C_2	0.6	0	0.4	0.2	0.6
3	Diameter: Medium	0.4	C_3	0.4	0.4	0	0.2	0.4
4	Unit Price: Low	0.2	C_4	0.2	0.2	0.2	0	0.2
5	Unit Price: Medium	0.8	C_5	0.8	0.6	0.4	0.2	0

Table 3
Results of SACM that assimilates 1 record

i	C_i	F_i	L_{ij}	C_1	C_2	C_3	C_4	C_5
1	Product Type: Prototype Lens	1	C_1	0	0.75	0.5	0.25	1
2	Diameter: Small	0.6	C_2	0.75	0	0.5	0.25	0.75
3	Diameter: Medium	0.4	C_3	0.5	0.5	0	0.25	0.5
4	Unit Price: Low	0.2	C_4	0.25	0.25	0.25	0	0.25
5	Unit Price: Medium	0.8	C_5	1	0.75	0.5	0.25	0

$\boldsymbol{P} = (F_{\max} = 1, L_{\max} = 0.8, N = 1)$.

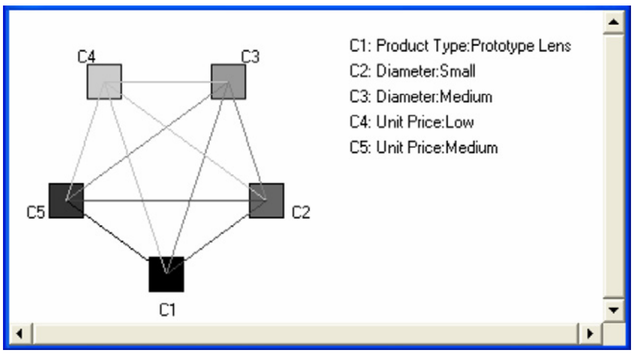

Fig. 5. SACM that assimilates 1 record.

in Table 2. In step 2, since it is the first record to be assimilated, there is no original SACM (i.e. $N = 0$). Thus, $\boldsymbol{F}$ and $\boldsymbol{L}$ are kept the same. In step 3, $\boldsymbol{F}$ and $\boldsymbol{L}$ are normalized, and the parameters of $\boldsymbol{P} = (F_{\max}, L_{\max}, N)$ are assigned. The results are shown in Table 3 and the graphical representation is depicted in Fig. 5.

To illustrate how to assimilate new records to the existing SACM, the following illustration is considered. Step 1 is similar to the previous section, the set of concept nodes $\boldsymbol{C}$ of the new record, the set of degrees of importance of concept $\boldsymbol{F}$ of the new record, and the set of degrees of importance of relations $\boldsymbol{L}$ for each pair of concepts of the new record are assigned (i.e. the second record). The results are shown in Table 4. In step 2, the nodes and relations of the current SACM is matched with that of the previous SACM. For any identified missing concept, the degree of importance of that concept and the degrees of importance of that concept's associated relations are assigned to 0. Then $\boldsymbol{F}$ and $\boldsymbol{L}$ of the 2 SACMs are combined by Eqs. (A.2) and (A.4). The result is shown in Table 5. $M = (\boldsymbol{C}, \boldsymbol{F}, \boldsymbol{L}, \boldsymbol{P})$ is the previous SACM that assimilates record 1 while $M' = (\boldsymbol{C}', \boldsymbol{F}', \boldsymbol{L}', \boldsymbol{P}')$ is the current SACM that assimilates record 1. Hence, $M'' = (\boldsymbol{C}'', \boldsymbol{F}'', \boldsymbol{L}'', \boldsymbol{P}'')$ is the combined SACM. In step 3, the combined SACM is normalized, and the parameters are assigned. The results are shown in Table 6 and the graphical representation is shown in Fig. 6.

Table 4
Temporary results of SACM that assimilates the second record

i	C_i	F_i	L_{ij}	C_1	C_2	C_3	C_4
1	Product Type: Mould Insert	1	C_1	0	0.6	0.4	1
2	Diameter: Medium	0.6	C_2	0.6	0	0.4	0.6
3	Diameter: Large	0.4	C_3	0.4	0.4	0	0.4
4	Unit Price: Medium	1	C_4	1	0.6	0.4	0

Table 5
Temporary results of SACM that assimilates 2 records

i	C_i	F_i	F_i'	F_i''
1	Product Type: Prototype Lens	1.0	0.0	0.5
2	Diameter: Small	0.6	0.0	0.3
3	Diameter: Medium	0.4	0.6	0.5
4	Unit Price: Low	0.2	0.0	0.1
5	Unit Price: Medium	0.8	1.0	0.9
6	Product Type: Mould Insert	0.0	1.0	0.5
7	Diameter: Large	0.0	0.4	0.2

L_{ij}	C_1 C_1' C_1''	C_2 C_2' C_2''	C_3 C_3' C_3''	C_4 C_4' C_4''	C_5 C_5' C_5''	C_6 C_6' C_6''	C_7 C_7' C_7''
C_1	0	0.75	0.5	0.25	1	0	0
C_1'	0	0	0	0	0	0	0
C_1''	0	0.3	0.2	0.1	0.4	0	0
C_2	0.75	0	0.5	0.25	0.75	0	0
C_2'	0	0	0	0	0	0	0
C_2''	0.3	0	0.2	0.1	0.3	0	0
C_3	0.5	0.5	0	0.25	0.5	0	0
C_3'	0	0	0	0	0.6	0.6	0.4
C_3''	0.2	0.2	0	0.1	0.5	0.3	0.2
C_4	0.25	0.25	0.25	0	0.25	0	0
C_4'	0	0	0	0	0	0	0
C_4''	0.1	0.1	0.1	0	0.1	0	0
C_5	1	0.75	0.5	0.25	0	0	0
C_5'	0	0	0.6	0	0	1	0.4
C_5''	0.4	0.3	0.5	0.1	0	0.5	0.2
C_6	0	0	0	0	0	0	0
C_6'	0	0	0.6	0	1	0	0.4
C_6''	0	0	0.3	0	0.5	0	0.2
C_7	0	0	0	0	0	0	0
C_7'	0	0	0.4	0	0.4	0.4	0
C_7''	0	0	0.2	0	0.2	0.2	0

To illustrate how to infer an enquiry to perform quantitative prediction by SACM, the following illustration is considered. It is assumed that there is an enquiry of requesting the Unit Price of {Product Type: Prototype Lens, Diameter: 9}, and the SACM that constructed above is used as the inference SACM. Based on the steps of knowledge inference in Section 2.3, the enquiry is firstly converted into SACM format. The results are shown in Table 7. In step 2, the enquiry SACM is matched with the SACM that inferencing the SACM. The activation level of each concept is then computed by Eq. (A.10). The results are shown in Table 8. In step 3, the concepts of the result field (i.e. Unit Price) and their corresponding activation levels are extracted (i.e. Unit Price:Low:0.21, Unit Price:Medium:0.93). The consequent membership functions at the level of corresponding activation levels are then clipped and aggregated as shown in Fig. 7. The aggregated fuzzy set is then defuzzified by Eq. (A.11), the result is 1816. Thus, the suggested Unit Price of the enquiry is 1816.

Table 6
Results of SACM that assimilates 2 records

i	C_i	F_i	L_{ij}	C_1	C_2	C_3	C_4	C_5	C_6	C_7
1	Product Type: Prototype Lens	0.56	C_1	0	0.6	0.4	0.2	0.8	0	0
2	Diameter: Small	0.33	C_2	0.6	0	0.4	0.2	0.6	0	0
3	Diameter: Medium	0.56	C_3	0.4	0.4	0	0.2	1	0.6	0.4
4	Unit Price: Low	0.11	C_4	0.2	0.2	0.2	0	0.2	0	0
5	Unit Price: Medium	1	C_5	0.8	0.6	1	0.2	0	1	0.4
6	Product Type: Mould Insert	0.56	C_6	0	0	0.6	0	1	0	0.4
7	Diameter: Large	0.22	C_7	0	0	0.4	0	0.4	0.4	0

$\boldsymbol{P} = (F_{\max} = 0.9, L_{\max} = 0.5, N = 2)$.

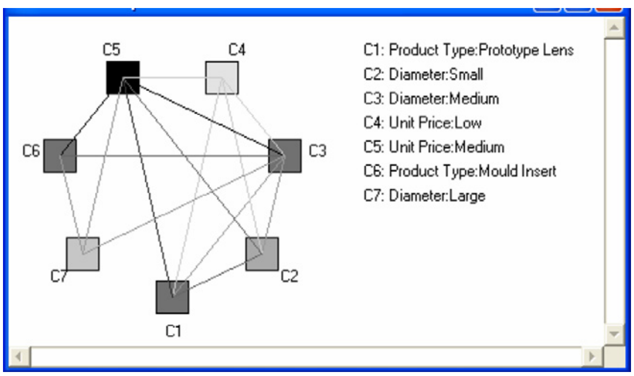

Fig. 6. SACM that assimilates 2 records.

Table 7
SACM of the example of enquiry

i	C_i	F_i	L_{ij}	C_1	C_2	C_3
1	Product Type: Prototype Lens	1	C_1	0	0.2	0.8
2	Diameter: Small	0.2	C_2	0.2	0	0.2
3	Diameter: Medium	0.8	C_3	0.8	0.2	0

Table 8
Activation level of each concept of this example

i	C_i	A_i
1	Product Type: Prototype Lens	0.22
2	Diameter: Small	0.51
3	Diameter: Medium	0.25
4	Unit Price: Low	0.21
5	Unit Price: Medium	0.93
6	Product Type: Mould Insert	0.27
7	Diameter: Large	0.18

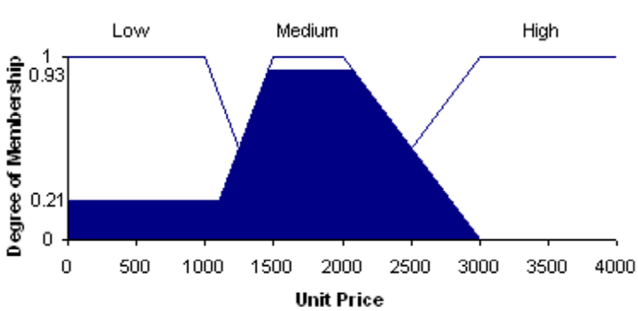

Fig. 7. Aggregated fuzzy set of this example.

Table 9
The structure of each record in the knowledge repository

Case number	
Customer information	Customer ID
Staff information	Staff ID
Quotation date	
Product information	Product type
	Material
	Diameter
	Radius
	Surface type
	Quantity
	Rough blank
	Inspection
Decision	Unit price
	Expected delivery date
	Payment method

3. Experimental verification

To evaluate the effectiveness of the proposed SACM model, it is applied in a consulting company. The company provides various consulting services in design, manufacture and evaluation of surface quality of precision optical components. A KBS has been built in the company [4]. Based on the KBS, the knowledge of experienced consultants can be captured continuously in a structured format such as cases. Table 9 shows the structure of each record in the knowledge repository. The case number is assigned sequentially for providing unique designation of individual case. As different customers may entitle different privileges on the customer services (e.g. discount, shorter delivery time, longer payment term, etc.), the customer information is important to the decision making process. The product information is related to the product type, material, geometry, quantity, requirements of inspection, and preparatory instruction of the machining jobs. The decision is given in form of suggested quotation price, expected delivery date and payment method. When there is any requests for quotation, customer will provide the product information, and then the staff needs to make the decision based on the provided information.

To facilitate the quotation process, a decision support function is developed based on Case-Based Reasoning (CBR) approach as mentioned in [4]. CBR is a problem-solving approach that relies on past and similar cases to find solutions to new problems [23]. It simulates human decision making processes and enables the accumulation of previous experience. One of the important advantages of CBR is its learning capability. Its problem-solving ability improves with the increasing amount of accumulated cases. CBR cycle starts with the codification of customer request as a new case. Then the new case is compared with all the cases in the knowledge. The similarity between 2 cases is determined as follows:

$$\text{Similarity} = \frac{\sum_{j=1}^{m} w_j \text{sim}(v_j^o, v_j^r)}{\sum_{j=1}^{m} w_j} \tag{3}$$

where m is the number of inputs, w_j is the weighting of the jth input, v_j^o and v_j^r are values of the jth inputs and that for the retrieved cases, $\text{sim}(v_j^o, v_j^r)$ is the similarity function for the jth inputs as follows:

For numerical value, the similarity is calculated based on the normalized distance of the feature between two cases:

$$\text{sim}(v_j^o, v_j^r) = 1 - \frac{|v_j^o - v_j^r|}{\max_j - \min_j} \tag{4}$$

where $\max_j$ and $\min_j$ are the maximum and minimum value of the jth input

For symbolic value:

$$\text{sim}(v_j^o, v_j^r) = 1 \text{ if } v_j^o = v_j^r \tag{5}$$

$$\text{sim}(v_j^o, v_j^r) = 0 \text{ if } v_j^o \neq v_j^r \tag{6}$$

The retrieved cases are ranked in descending order according to the similarity. The most similar case is then selected for suggestion.

In this paper, the proposed approach SACM is compared with the CBR approach and a group of human being who are laymen of the domain. The fuzzy membership functions used in SACM and the weightings used in CBR are determined by the expertise of the company. The accuracy of the suggested results is calculated by the following equations:

$$a = (1 - \frac{|v^s - v^r|}{v^r}) \times 100\% \tag{7}$$

where a is the accuracy, v^r is the actual value and, v^s is the suggested value by the model.

Initially, there is only one case is used as the learning record and a case is used as the testing record. After the model/human has answered the same testing record, the actual result of the testing record will be provided. In other words, one new learning record is added to the knowledge base of the model/human (i.e. two learning records in the knowledge base of the model/human). Then, another testing record is provided.

The results of average accuracy of SACM, CBR, and human against the number of learning records are shown in Fig. 8 and Table 10. The results show that the accuracy of CBR is the highest and the accuracy of SACM is similar to the accuracy of human. Based on the correlation analysis, SACM has a higher correlation with the result of laymen of the domain, while CBR has a higher correlation with the actual results (which is provided by the expert of the domain).

CBR has both its strengths and weaknesses; there are several characteristics of using SACM over CBR:

(i) *Smaller data size*. Since CBR needs to store all historical records in terms of cases in the knowledge base, it occupies huge amount of storage space. It constantly requires the maintenance of the case base. For SACM, the amount of data size increases with the number of distinct concepts. It reduces the data size to minimal.
(ii) *Faster speed*. When there is a new enquiry, CBR needs to compare the new case with all the cases in its case base. It is inefficient when the number of cases of the case base is large. For SACM, the computation speed remains stable.
(iii) *Graphical representation*. Graphical representation of SACM provides users a rough image of the domain effectively and efficiently.

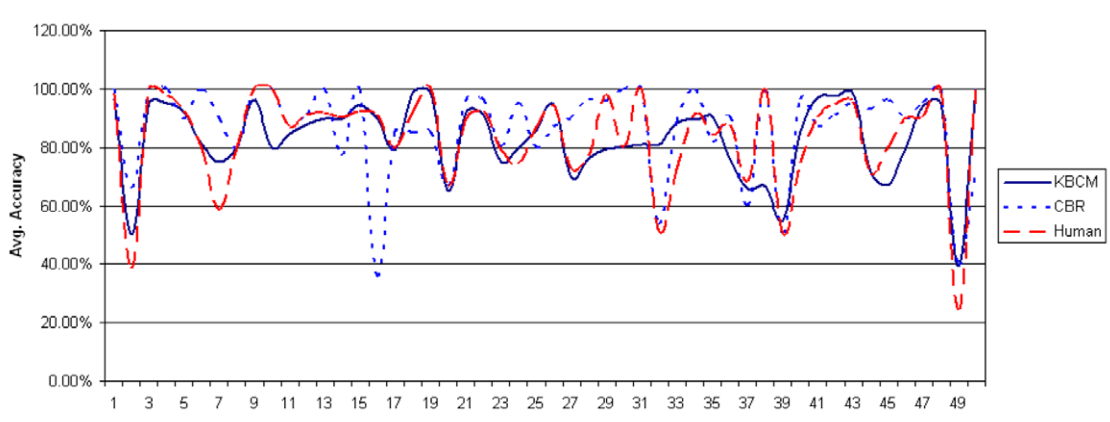

Fig. 8. The results of SACM, CBR and human.

Table 10
Experiment results

Model	Average accuracy (%)	Correlation with human layman	Correlation with human expert
SACM	83.27	0.90	0.35
CBR	87	0.69	0.47
Human layman	83.75	1	0.36
Human expert	100	0.36	1

(iv) *Higher accuracy against human laymen.* From the experiment, it shows that SACM has a higher accuracy and higher correlation with the results which are provided by the laymen of the domain. It shows that it is good at emulating laymen learning.

(v) *Lower accuracy against human experts*. From the experiment, it in interesting to note that CBR has a higher accuracy and higher correlation with the actual results which are provided by the domain expert. It shows that it is good at emulating expert reasoning.

(vi) *Narrower range of choices of algorithms*. CBR covers a very broad range of systems and algorithms. The performance of the model can be fine tuned for different application systems.

4. Conclusions

The performance of KBS relies on its abilities on knowledge elicitation, knowledge representation and knowledge inference. In this paper, a Self-Associated Concept Map (SACM) is proposed for enhancing these abilities. SACM has its origin from concept map. It is extended and enhanced so as to apply to KBS. An elicitation algorithm which embedded with fuzzy set is presented. It provides an automatic solution for the construction and dynamic update of SACM from data without human intervention. The graphical network graph of knowledge representation facilitates the knowledge sharing of different concepts or ideas among users. It has the advantages of simplicity, naturalness, visionless, and clarity. An inference algorithm of SACM which embedded with fuzzy set is built for quantitative predication. It has been evaluated through a case study undertaken in a consultancy business. The results are compared with a widely used problem-solving method, i.e. case-based reasoning (CBR), and human laymen of the consultancy business. The results show that the knowledge inference of SACM has advantages of higher laymen learning capability, smaller data size and faster speed than CBR. It would be useful for simulating human learning activities. Further work will be done on exploring the usage of SACM to other applications and how it may work with more difficult and unstructured knowledge area.

Acknowledgement

The authors express their sincere thanks to the Research Committee of the Hong Kong Polytechnic University for financial support of the research work.

Appendix A. Knowledge representation

A SACM is defined with all necessary notations as follows:

Let $K = [0,1]$,
A SACM is a 3-tuple $(\boldsymbol{C},\boldsymbol{F},\boldsymbol{L},\boldsymbol{P})$ where
$\boldsymbol{C} = (C_1,C_2,\ldots,C_n)$ is a set of n distinct concepts forming the nodes of a SACM
$\boldsymbol{F} = (F_1,F_2,\ldots,F_n)$ is a function that at each C_i associates its degree of importance F_i with $F_i \in K$
$\boldsymbol{L}:(C_i,C_j) \rightarrow L_{ij}$ is a function that a pair of concept (C_i,C_j) associates its degree of importance L_{ij}, with L_{ij} denoting a weighting of directed edge from C_i to C_j, $L_{ij} \in K$ if $i \neq$ j, and $L_{ij} = 0$ if $i = j$, $\boldsymbol{L}$ represents a set of degree of association between all concepts in a SACM.
$\boldsymbol{P} = (F_{\max},L_{\max},N)$ is a set of parameter which facilitates the computation of knowledge elicitation and inference, with $F_{\max}$ and $L_{\max}$ indicate the maximum value of $\boldsymbol{F}$ and $\boldsymbol{L}$ before normalization, respectively, and N indicates the total number of records that have been assimilated to this SACM.

Appendix B. Knowledge elicitation

The elicitation algorithm consists of 3 steps:

Step 1: A temporary SACM is constructed based on the information of the new record.

- Distinct concepts are extracted from the new record for the construction of a set of concept nodes $\boldsymbol{C}$ and the degree of importance of concept F_i for each $C_i \in \boldsymbol{C}$ is assigned by the following conditions;
  Let $\mathbf{S}$ be the set of fields of the knowledge repository, and $\mathbf{V}$ be the set of values of fields of the new record, where V_i corresponds to S_i for $V_i \in \mathbf{V}$ and $S_i \in \mathbf{S}$
  If V_i is a symbolic value, a concept node is added as a text value in the format of S_i:V_i, and the corresponding degree of importance of the added concept is assigned to be 1. For example, if S_i = "Product Type" and V_i = "Mould Insert", then a new concept node "Product Type: Mould Insert" is added to $\boldsymbol{C}$, and the corresponding degree of importance of added concept $F = 1$;
  If V_i is a numerical value, V_i is fuzzified based on the corresponding membership function of S_i, concept node(s) is/are added as text value(s) in

the format of S_i:V_i's belonged membership(s), and the corresponding degree(s) of importance of the added concept(s) is/are assigned to be the corresponding degree(s) of membership. For example, if S_i = "Quantity" and V_i = (0.7 High, 0.2 Medium, 0 Low), then two new concept nodes: "Quantity: High" and "Quantity: Medium" are added to $\boldsymbol{C}$, the corresponding degree of importance of the added concepts $F = 0.7$ of concept "Quantity: High" and $F = 0.2$ of concept "Quantity: Medium".

- Assign the degree of importance of relation L_{ij} for each pair of concepts (C_i,C_j), where $i,j \in n$ and n is the total number of distinct concepts $\boldsymbol{C}$, by the following equation:

$$L_{ij} = \mathrm{Min}(V_i, V_j) \tag{A.1}$$

Step 2: The temporary SACM is combined with the original SACM;

- The nodes and relations of the temporary SACM are matched with that of the original SACM. If there is any missing concept existed, the degree of importance of that concept and the degrees of importance of that concept's associated relations are assigned to be 0.
- $\boldsymbol{L}$ and $\boldsymbol{F}$ of the original SACM are adjusted based on $\boldsymbol{L}$ and $\boldsymbol{F}$ of the temporary SACM by the following equations:
  Let F_i, F'_i, F''_i be the original, temporary and combined degree of importance of C_i, N be the total number of records of the original SACM.

$$\text{For} \quad N > 0, \quad F''_i = \frac{F'_i + NF_iF_{\max}}{N+1} \tag{A.2}$$

$$\text{For} \quad N = 0, \quad F''_i = F'_i \tag{A.3}$$

Let L_{ij}, L'_{ij}, L''_{ij} be the original, temporary and combined degree of relations between C_i and C_j, N be the total number of records of the original SACM.

$$\text{For} \quad N > 0, \quad L''_i = \frac{L'_i + NL_iL_{\max}}{N+1} \tag{A.4}$$

$$\text{For} \quad N = 0, \quad L''_i = L'_i \tag{A.5}$$

Step 3: The parameters of the combined SACM is adjusted and the combined SACM is normalized.

- The parameters of the combined SACM, $\boldsymbol{P} = (F_{\max}, L_{\max}, N)$, is adjusted by the following equations:

$$F_{\max} = \mathrm{Max}(F_1, F_2, \ldots, F_n) \tag{A.6}$$

$$L_{\max} = \mathrm{Max}(L_1, L_2, \ldots, L_m) \tag{A.7}$$

- The total number of record N is increased by 1.
- $\boldsymbol{L}$ and $\boldsymbol{F}$ of the combined SACM is normalized to the range between 1 and 0 by the following equations:

$$\mathrm{Normalize}(\boldsymbol{F}) = \frac{\boldsymbol{F}}{F_{\max}} \tag{A.8}$$

$$\mathrm{Normalize}(\boldsymbol{L}) = \frac{\boldsymbol{L}}{L_{\max}} \tag{A.9}$$

Appendix C. Knowledge inference

A new model named Constrained Fuzzy Spreading Activation (CFSA) which integrates fuzzy logic and CSA is introduced for knowledge inference. It consists of 3 steps:

Step 1: The enquiry/problem is converted into SACM format

- This step is the same as step 1 of knowledge elicitation.

Step 2: The activation level of each node of the SACM is computed which is used for performing knowledge inference

- The nodes and relations of enquired SACM are matched with that of SACM that used for performing knowledge inference. If there is any missing concept existed, the degree of importance of that concept and the degrees of importance of that concept's associated relations are assigned to be 0.
- The activation level of each node of the SACM is computed which is used for performing knowledge inference by the following equations:
  Let $M' = (\boldsymbol{C}',\boldsymbol{F}',\boldsymbol{L}',\boldsymbol{P}')$ be the enquired SACM, and $M = (\boldsymbol{C},\boldsymbol{F},\boldsymbol{L},\boldsymbol{P})$ be the SACM that are used for performing knowledge inference, A_j be the activation level of node j

$$A_j = \sum_i F'_iF_iL_{ij} \tag{A.10}$$

Step 3: Generation of the result

- The concept nodes of the result field are extracted and their corresponding activation levels are based on the following conditions:
  If the extracted concept nodes are fuzzy values, their consequent membership functions at the level of corresponding activation levels of that concept nodes are clipped. Then the clipped membership functions are aggregated into a single fuzzy set. The fuzzy set is then defuzzified by the centre of gravity (COG) method and the result is displayed. The COG of fuzzy set A on the interval a_1 to a_2 with membership function μ_A is given by the following equation:

$$\mathrm{COG}(A) = \frac{\int_{a_1}^{a_2} \mu_A(x)x\mathrm{d}x}{\int_{a_1}^{a_2} \mu_A(x)\mathrm{d}x} \tag{A.11}$$

- If the extracted concept nodes are not fuzzy values, the concept node with the highest activation level is selected to be the result.

References

[1] D.P. Ausubel, The psychology of meaningful verbal learning, Grune and Stratton, New York, 1963.

[2] C. Van Boxtel, J. Van der Linden, G. Kanselaar, Collaborative construction of conceptual understanding: interaction processes and learning outcomes emerging from a concept mapping and a poster task, J. Interact. Learn. Res. 8 (3/4) (1997) 341–361.

[3] T. Chan, M. Sharples, A concept mapping tool for pocket PC computers, in: Proceedings of the IEEE International Workshop on Wireless and Mobile Technologies in Education (WMTE'02), 2002.

[4] C.F. Cheung, W.B. Lee, W.M. Wang, K.F. Chu, S. To, A Multi-perspective Knowledge-based System for Customer Service Management, Expert Syst. Appl. 24 (4) (2003) 455–468.

[5] S.W. Chen, S.C. Lin, K.E. Chang, Attributed Concept Maps: Fuzzy Integration and Fuzzy Matching, IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics 31 (5) (2001) 842–852.

[6] Hsien-Hsun Chen, Shein-Yung Cheng, J.-S. Heh, Assessing Users' mental knowledge by using structural approach and concept map, machine learning and cybernetics, 2005, in: Proceedings of 2005 International Conference, vol. 4, 18–21 Aug. 2005, pp. 2166–2171.

[7] J.W. Cliburn, Using concept maps to sequence instructional materials, J. Coll. Sci. Teach. (1986) 377–379.

[8] in A. Collins, E.E. Smith (Eds.), Readings in Cognitive Science: A Perspective from Psychology and Artificial Intelligence, Morgan Kaufman, San Mateo, CA, 1988.

[9] ConceptDraw: <http://www.conceptdraw.com/en/products/mindmap/main.php>.

[10] Fabio Crestani, Puay Leng Lee, Searching the web by constrained spreading activation, Info. Process. Manage. 36 (2000) 585–605.

[11] N. Dabbagh, Concept Mapping as a Mindtool for Critical Thinking, J. Comput. Teach. Educ. 17 (2) (2001) 16–24.

[12] D.L. Darmofal, D.H. Soderholm, D.R. Brodeur, Using concept maps and concept questions to enhance conceptual understanding, in: 32nd ASEE/IEEE Frontiers in Education Conference, Boston, MA, 2002.

[13] D. Mularz, M. Lyell, Integrating concept mapping and semantic web technologies for effective knowledge management, dexa, pp. 449–453, Database and Expert Systems Applications, in: 15th International Workshop on (DEXA'04), 2004.

[14] Embeddable-Maps: <http://ksi.cpsc.ucalgary.ca/articles/ConceptMaps>.

[15] K.M. Fisher, Semantic networking: The new kid on the block, J. Res. Sci. Teach. 27 (10) (1990) 1001–1018.

[16] GetSmart: <http://rocky.dlib.vt.edu/∼rshen/ConceptMap/GetSmart.htm>.

[17] E. Gouli, A. Gogoulou, K. Papanikolaou, M. Grigoriadou, COMPASS: an adaptive web-based concept map assessment tool, in: Proceedings of the First International Conference on Concept Mapping, 2004.

[18] E. Gouli, A. Gogoulou, K. Papanikolaou, M. Grigoriadou, Evaluating learner's knowledge level on concept mapping tasks, in: Proceedings of the Fifth IEEE International Conference on Advanced Learning Technologies (ICALT'05), 2005.

[19] H.E. Herl et al., Reliability and validity of a computer-based knowledge mapping system to measure content understanding, Comput. Hum. Behav. 15 (3–4) (1999) 315–333.

[20] IHMC Concept Maps: <http://cmap.ihmc.us>.

[21] Inspiration: <http://www.inspiration.com/home.cfm>.

[22] D.H. Jonassen et al., Concept maps as cognitive learning and assessment tools, J. Interactive Learn. Res. 8 (1997) 3/4.

[23] J.L. Kolodner, Cased-Based Reasoning, Morgan Kaufman, San Mateo, 1993, ISBN 1-55860-237-2.

[24] Shui-Cheng Lin, Kuo-En Chang, Yao_Ting Sung, Gwo-Dong Chen, A new structural knowledge assessment based on weighted concept, in: Proceeding of the International Conference on Computers in Education (ICCE'02), 2002, pp. 679–680.

[25] C.V. Lloyd, The elaboration of concepts in three biology textbooks: facilitating student learning, J. Res. Sci. Teach. 27 (10) (1990) 1019–1032.

[26] J.H. McClellan, L.D. Harvel, R. Velmurugan, M. Borkar, C. Scheibe, CNT: concept-map based navigation and discovery in a repository of learning content, Frontiers in Education, 2004. FIE 2004, in: 34th Annual, 2004, vol. 2, 2004, pp. F1F-13–18.

[27] MindGenius: <http://www.ygnius.com>.

[28] E.M.J.C. Moen, K.TH. Boersma, The significance of concept mapping for education and curriculum development, J. Interactive Learn. Res. 8 (1997) 487–502.

[29] J.D. Novak, Learning, Creating, and Using Knowledge: Concept Maps As Facilitative Tools in Schools and Corporations, Lawrence Erlbaum Assoc, 1998.

[30] J.D. Novak, D.B. Gowin, Learning How to Learn, Cambridge University Press, Cambridge (UK), 1984.

[31] J.D. Novak, Concept Mapping: a useful tool for science education, J. Res. Sci. Teach. 27 (10) (1990) 937–949.

[32] S. Preece, A spreading activation model for Information Retrieval. Ph.D. thesis, University of Illinois, Urbana-Champaign, USA, 1981.

[33] D. Rumelhart, D. Norman, Representation in memory. Technical report, Department of Psychology and Institute of Cognitive Science, UCSD La Jolla, USA, 1983.

[34] R.W. Schvaneveldt (Ed.), Pathfinder Associative Networks: Studies in Knowledge Organization, Ablex, Norwood, NJ, 1990.

[35] Pei-Chi Sue, Jui-Feng Weng, Jun-Ming Su, Shian-Shyong Tseng, A new approach for constructing the concept map, in: Proceedings of the IEEE International Conference on Advanced Learning Technologies (ICALT'04), 2004.

[36] W.M.K. Trochim, An introduction to concept mapping for planning and evaluation, Eval. Program Plann. 12 (1989) l–16.

[37] J. Turns, C.J. Atman, R. Adams, Concept Maps for Engineering Education: a Cognitively Motivated Tool Supporting Varied Assessment Functions, IEEE Trans. Edu. 43 (2) (2000) 164–173.

[38] J.M.T. Walker, P.H. King, Concept mapping as a form of student assessment and instruction, in: Proc. 2002 ASEE Annual Conf. and Exp., Montreal, Canada, 2002.

[39] WebMap: <http://ksi.cpsc.ucalgary.ca/articles/www/www4wm>.

[40] M. Willerman, R.A.M. Harg, The concept map as an advance organizer, J. Res. Sci. Teach. 28 (8) (1991) 705–711.