IJAMML 1:1 (2014) 55-81 October 2014
ISSN: 2394-2258
Available at http://scientificadvances.co.in

COMPARATIVE STUDY OF SEMANTIC MERGER SOLUTIONS OF ENGLISH MODAL VERB BY ATTRIBUTE PARTIAL ORDER STRUCTURE APPROACH AND FUZZY C-MEANS CLUSTER

Jianping Yu[a*], Xiamei Yuan[a], Wenxue Hong[b], Shaoxiong Li[b] and Deming Mei[c]

[a]College of Foreign Studies, Yanshan University, No. 438 Hebei Street, Qinhuangdao, 066004, Hebei, P. R. China

[b]Institute of Electrical Engineering, Yanshan University, No. 438 Hebei Street, Qinhuangdao, 066004, Hebei, P. R. China

[c]College of International Programs, Shanghai International Studies University, No. 410 Dong Ti Yu Hui Road, Shanghai 200083, P. R. China

---

Abstract

In this article, the attribute partial order structure approach and fuzzy c-means cluster are applied into the solution of semantic merger of English modal verb in order to compare the performance of the two approaches. English modal verb *should* is chosen as the target word for semantic merger solution. First, two models for word sense disambiguation of English modal *should* are established by the two approaches, respectively. The accuracies of word sense disambiguation reach 91% for the attribute partial order structure approach and 90% for the fuzzy c-means cluster. Then, the semantic mergers of *should* are solved based on the two models of word sense disambiguation, respectively. Finally, the performance and accuracies by the two approaches are compared. The comparative result shows that, the

*Corresponding author.

E-mail address: yjp@ysu.edu.cn (Jianping Yu).

2010 Mathematics Subject Classification: 03A65.
Submitted by Jianqiang Gao.
Received August 26, 2014

fuzzy c-means cluster is easier in use than the attribute partial order structure approach, however, the attribute partial order structure approach performs much better than fuzzy c-means cluster in the solution of semantic merger of English modal verb.

Keywords: semantic merger, solution, english modal verb, attribute partial order structure approach, fuzzy c-means cluster.

Nomenclatures

Symbol	Description
WSD	word sense disambiguation.
RT	root meaning of modal verb.
EP	epistemic meaning of modal verb.
EM	semantic merger.
K	a formal context.
O	a set of objects.
A	a set of attributes.
I	a set of relation between objects and attributes.
o	an object.
a	an attribute.
B	extent of a concept.
C	intent of a concept.
FCM	fuzzy c-means.
n	dimension of vectors.
X	sample set.
c	number of clusters.
x_j	the j-th sample.
v_i	the i-th cluster center v_i.
u_{ji}	degree of sample x_j *belongs to cluster center*.
U	fuzzy partition matrix.
J_m	objective function.
d_{ij}	Euclidean distance between x_j and v_i.

Symbol	Description
m	fuzzifier parameter.
MI_i	the i-th mutual information.
w_1, w_2	w_1 is the target word for WSD; w_2 is the word related to w_1.
$P(w_1, w_2)$	probability of co-occurrence of w_1 and w_2.
$P(w_1)$	probability of w_1.
f_k	the k-th syntactic co-occurrence feature.
o_j	the j-th object.
a_i	the i-th attribute, a is corresponding to m.
APOSD	attribute partial order structure diagram.
S	similarity of the attribute patterns of two objects.
X_i	attribute pattern.
S_i	the i-th similarity.

1. Introduction

Semantic merger is semantic convergence of two senses of a word in a context. When the two meanings involved in a word are not in certain contexts mutually exclusive, i.e., they are mutually compatible in a reading of a passage; the word is said to exhibit semantic merger. Semantic merger is regarded as a special case of ambiguity in which a word yields two interpretations closely converged and hard to figure out one from the other. Semantic merger causes troubles for natural language understanding and processing, therefore, it has been a significant and urgent issue to solve semantic mergers, especially to solve the semantic merger in the semantically complex words, such as English modal verbs and prepositions. However, few studies have been found. Everaere et al. [1] investigated the standard unanimity condition for preference aggregation in the setting of propositional merging. Qin [2] studied semantic mergers of English prepositional phrases by a cognitive-corpus approach. Thomas [3] studies vowel mergers in spoken language from the perspective of sociolinguistics. These studies have been concentrated on the description and definition of mergers and investigations of the mechanism of different kinds of mergers. Almost no studies have been found on solutions of semantic mergers.

Attribute partial order structure approach and fuzzy c-means cluster are two approaches, which can be used to solve semantic mergers. The attribute partial order structure approach is based on the principle of partial order in the theory of formal concept analysis [4-5]. Formal concept analysis is a branch of applied mathematics based on the mathematization of concept and conceptual hierarchy. It grows out of the mathematical order theory, in particular, the theory of complete lattices and it can be used for the conceptual unfolding of data contexts. It presents graphical methods for representing conceptual systems that have proved themselves in communicating knowledge. It thereby allows us to mathematically represent, analyze, and construct conceptual structures. Hong et al. [6] proposed the approach of generation of attribute partial order structure diagram. The approach has been applied in word sense disambiguation, data mining, knowledge discovery, and pattern recognition [7-9]. Fuzzy c-means clustering is a data clustering method. The main idea of the method is that each sample belongs to a certain cluster in a certain membership grade. The highest grade of membership determines the cluster a sample belongs to. It maps the data in multiple dimensional spaces onto different cluster sets. Fuzzy c-means clustering has been widely used in different fields, such as in image change detection [10], segmentation for brain MRI image [11], identification of partial discharge location [12], feature recognition [13], information retrieval [14], and principal component averaging fusion [15], etc. However, which of the two approaches performs better in solving semantic mergers? How does one of them perform better than the other? Making clear of these questions is very important for the studies of semantic merger solution and for natural language processing.

Therefore, in this article, both the attribute partial order structure approach and fuzzy c-means cluster are used for semantic merger solution of English modal verb *should* in order to make a comparison between them.

The layout of the article is as follows. Section 1 introduces the research background and the motivation of the study. Section 2 explains semantic merger of English modal verb *should*. Section 3 describes the processes of WSD of English modal verb *should* by attribute partial order structure approach and by fuzzy c-means cluster. Section 4 makes a comparison of the performance in semantic merger solution by the two approaches. Finally, Section 5 draws the conclusion of the study.

2. Semantic Merger of English Modal Verb *Should*

Generally, the senses of English modal verb *should* are categorized into 2 categories [16]: root *should* (RT*should*) and epistemic *should* (EP*should*).

$$Should\begin{cases}\text{RT}\textit{should} - \text{express a weak sense of moral obligation, duty; offer} \\ \qquad\qquad \text{advice or describe correct procedure.} \\ \text{EP}\textit{should} - \text{express a less confident assumption or weak inference.}\end{cases}$$

For instance,

(1) Every citizen *should* obey the law (RT*should*-obligation).

(2) The unemployed *should* be made to do something (RT*should*-advice).

(3) Several potential limitations *should* be considered (RT*should*-correct procedure).

(4) The report is based on many investigations, so it *should* be reliable (EP*should*-less confident assumption).

(5) With his talent and experience, he *should* do well for himself (EP*should*-weak inference).

However, merger sometimes occurs between the root meaning RT*should* (weak obligation) and epistemic meaning EP*should* (weak inference), as shown in Figure 1.

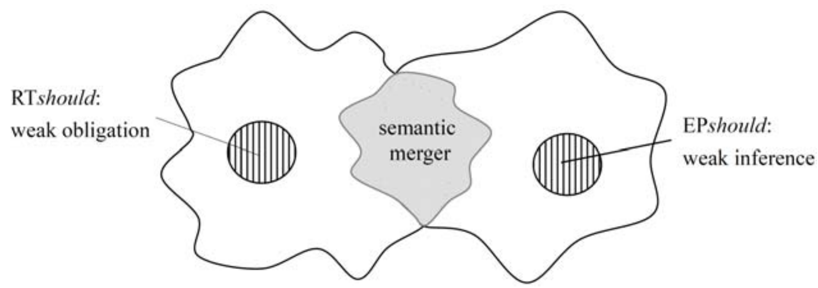


Figure 1. Merger between RT*should* and EP*should*.

For instance,

A: Spring Snow is a good beer.

B: Is it?

A: Well it *should* be at that price.

In this instance, it is not clear whether the speaker is referring to the brewer's obligation to produce good beer (weak obligation), or whether he is making a logical inference according to the price (weak inference). In this case, semantic merger of *should* occurs.

Semantic merger is one of the three categories of indeterminacy: gradience, ambiguity, and merger [16]. Semantic merger occurs when the context fails to exclude one of the possible meanings.

Semantic merger is a feature of natural languages and it may cause troubles in natural language understanding and information processing. A computer needs to receive the exact sense of a modal verb to process the information and to achieve the correct information. Therefore, it is important to recognize semantic merger, and it is more important to solve semantic merger. To solve semantic mergers is to discover the writer's or speaker's intended meanings of the semantically merged words in order to improve the quality of natural language processing and information processing. It may also improve the quality of automatic sense tagging of the merged words and the speed and quality of machine translation, text classification and information retrieval.

3. Solution of Semantic Merger of English Modal Verb by Attribute Partial Order Structure Approach

The approach of attribute partial order structure is based on the following theoretical descriptions of formal context [5].

Definition 1. A formal context $K = (O, A, I)$ consists of two sets O and A, and a relation I between O and A. The elements of O are called the objects and the elements of A are called the attributes of the context. I represents the relation between an object o and an attribute a, written as oIa or $(o, a) \in I$.

Definition 2. Let $K = (O, A, I)$ be a formal context, for a set $B \subseteq O$, $f(B) = \{a \in A | (o, a) \in I, \forall o \in B\}$. Correspondingly, for a set $C \subseteq A$, define $h(C) = \{o \in O | (o, a) \in I, \forall a \in C\}$. A formal concept is a pair (B, C) with $B \subseteq O$, $C \subseteq A$, $f(B) = C$ and $h(C) = B$. B is called the extent of the concept and C is called the intent of the concept.

Definition 3. A binary relation I on a set A is called an order relation, if it satisfies the following conditions for all elements $x, y, z \in A$:

(1) xRx (reflexity).

(2) xRy and $x \neq y \rightarrow$ not yRx (antisymmetry).

(3) xRy and $yRz \rightarrow xRz$ (transitivity).

Definition 4. If (B_1, C_1) and (B_2, C_2) are concepts of a context, (B_1, C_1) is called a subconcept of (B_2, C_2), provided that $B_1 \subseteq B_2$ (which is equivalent to $C_2 \subseteq C_1$). In this case, (B_2, C_2) is a super concept of (B_1, C_1), and we write $(B_1, C_1) \leq (B_2, C_2)$. The relation $\leq$ is called the hierarchical order of the concepts.

Definition 5. Let $K = (O, A, I)$ be a formal context, if for any objects $o_1, o_2 \in O$ from $f(o_1) = f(o_2)$, it always follows that $o_1 = o_2$ and correspondingly, $h(a_1) = h(a_2)$ implies $a_1 = a_2$ for all $a_1, a_2 \in A$ the context $K = (O, A, I)$ is called clarified.

The key of the approach of attribute partial order structure is the construction of the formal context and generation of attribute partial order structure diagram. The diagram can be generated by using an APOSD Tool and it can be used as the model for WSD and rule extraction.

The solution of semantic merger goes through the following steps:

3.1. Data preparation

In this study, a corpus of 3.5 million words is constituted including materials from spoken and written genres: film subtitles, academic papers, literary works, legal documents, speeches, interviews, news reports, and academic forums published during 2008-2012. First, three senses of *should* in the corpus are tagged manually by three symbols: RT(root meaning), EP(epistemic meaning), and ME(semantic merger), respectively, and then the occurrences of them are counted by the Concordance Tool of Wordsmith 4.0. The statistic result is shown in Table 1.

Table 1. Statistical result of different senses of *should*

Should	RT*should*	EP*should*	ME*should*
3103	2076	943	80

Three data sets are prepared. One is the training set and one is the testing sets for disambiguating RT*should* from EP*should*, each contains 50 samples of RT*should* and 50 samples of EP*should*, respectively. The third is the semantic merger data set containing 80 samples of ME*should*. The samples are selected randomly from the tagged corpus.

3.2. Feature selection for WSD of *should*

Some semantic features and syntactic features are extracted from the context for the WSD of *should*. The mutual information (*MI*s) of *should* and the adjacent words are used as the semantic feature and the syntactic co-occurrence patterns are used as syntactic features.

Semantic features	**Syntactic features**
(1) *MI* (subject, RT*should*)	(5) Animate subject
(2) *MI* (subject, EP*should*)	(6) *should* + active agentive verb
(3) *MI* (RT*should*, main verb)	(7) first person
(4) *MI* (EP*should*, main verb)	(8) stative verb

The 4 *MI*s are calculated according to formula (1) [17].

$$MI(w_1, w_2) = \log \frac{P(w_1, w_2)}{P(w_1)P(w_2)}, \tag{1}$$

where w_1 and w_2 are two words, $MI(w_1, w_2)$ embodies the degree of relevance of w_1 and w_2: If $MI(w_1, w_2) > 0$, it means that w_1 and w_2 are relevant. If $MI(w_1, w_2) = 0$, it means that w_1 and w_2 are independent. If $MI(w_1, w_2) < 0$, it means that w_1 and w_2 are compensate. And logical values are given to the syntactic features of the samples. If a sample has the syntactic feature, it is given a value of 1; otherwise, it is given a value of 0. Finally, the training data set is obtained, as shown in Table 2. The testing data set is also obtained in this way.

Table 2. Training set of *should*

o_j	MI_1	MI_2	MI_3	MI_4	f_5	f_6	f_7	f_8
o_1	1.37	0.99	2.23	– 1	1	1	1	0
o_2	1.37	0.99	1.31	0.81	1	1	1	0
o_3	1.37	0.99	1.82	–1	1	1	1	0
o_4	1.37	0.99	0.92	– 1	1	1	1	0
o_5	1.37	0.99	1.49	– 1	1	1	1	0
o_6	0.86	0.93	2.80	– 1	1	1	0	0
o_7	1.37	0.99	1.27	– 1	1	1	1	0
o_8	0.97	1.14	1.39	1.98	1	0	0	1
o_9	1.37	0.99	1.39	0.99	1	1	1	0
o_{10}	1.37	0.99	1.30	0.64	1	0	1	1
o_{11}	0.41	– 1	1.39	1.98	0	0	0	1
o_{12}	0.75	0.80	1.85	– 1	1	1	1	0
o_{13}	0.98	0.74	0.95	0.99	1	1	0	0
o_{14}	0.98	0.74	1.57	1.61	1	1	0	0
o_{15}	1.37	0.99	1.21	0.95	1	1	1	0
...	...	...	...	...	...	...	...	...
o_{86}	– 1	0.71	– 1	2.97	0	0	0	1
o_{87}	– 1	2.12	1.39	1.98	0	0	0	1
o_{88}	0.51	1.03	– 1	1.54	0	0	0	1
o_{89}	1.27	1.62	– 1	2.57	1	1	0	0
o_{90}	– 1	1.51	– 1	1.45	0	1	0	0
o_{91}	– 1	1.48	1.39	1.98	0	0	0	1
o_{92}	1.26	1.08	1.39	1.98	0	0	0	1
o_{93}	0.94	1.29	1.19	1.24	0	0	0	1

Table 2. (Continued)

o_j	MI_1	MI_2	MI_3	MI_4	f_5	f_6	f_7	f_8
o_{94}	0.23	0.99	1.39	1.98	0	0	0	1
o_{95}	– 1	0.74	1.55	1.27	0	0	0	1
o_{96}	– 1	1.19	1.78	2.42	0	1	0	0
o_{97}	0.23	0.99	1.29	1.16	0	0	0	1
o_{98}	1.12	1.64	1.39	1.98	0	0	0	1
o_{99}	– 1	2.49	0.75	1.39	0	1	0	0
o_{100}	– 1	0.98	1.55	1.27	0	0	0	1

Here in Table 2, $MI_1 - MI$ of the subject and RT*should* (MI(s + RT*should*)); $MI_2 - MI$ of the subject and EP*should* (MI(s + EP*should*)); $MI_3 - MI$ of RT*should* and the main verb (MI(RT*should* + v)); $MI_4 - MI$ of EP*should* and the main verb (EP*should* + v)); $f_k(k = 1, 2, 3, 4)$ are syntactic co-occurrence features; and – 1 denotes the case that there is no co-occurrence of the two words. It is the same for Table 7.

3.3. Vectorization of *MI*s

Most *MI*s in Table 2 are continuous values. They need to be vectorized into bi-values in order to express the direct relation between objects and attributes. The vectorization is based on the scatter diagram of *MI* values. Take MI_1 for instance. The distribution of MI_1 in the training set is shown in the scatter diagram in Figure 2. A way of vectorizing the continuous value into bi-value is to divide the continuous values into ranges. The determination of the dividing point of the 4 *MI*s would directly influence the accuracy of WSD, and the division should facilitate the disambiguation of word senses. In the case of MI_1, the dividing points are set to be 0.5. Then, we get the ranges of $a_1 \leq 0.5$, $a_2 > 0.5$. In the same way, the values of other *MI*s are divided. Then, each range is taken as an attribute. If the *MI* falls in a range, it is given a logical value 1, otherwise it is given 0. The ranges for the *MI*s are listed in Table 3. Finally, the *MI*s are changed into *bi*-values and the formal context of the training set of *should* is formed, as shown in Table 4.

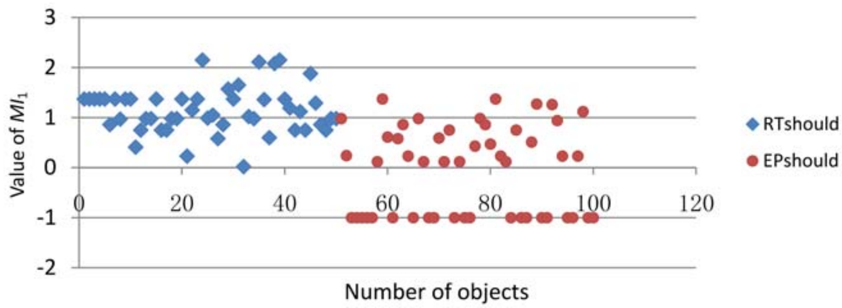


Figure 2. Scatter diagram of MI_1.

Table 3. Ranges of *MI* values

*MI*s	Ranges of *MI* values	*MI*s	Ranges of *MI* values
MI_1(s + RT*should*)	$a_1 \le 0.5$, $a_2 > 0.5$	MI_3(RT*should* + v)	$a_5 \le 1.5$, $a_6 > 1.5$
MI_2(s + EP*should*)	$a_3 \le 1$, $a_4 > 1$	MI_4(EP*should* + v)	$a_7 \le 1.1$, $a_8 > 1.1$

Table 4. Formal context of *should*

o_j \ a_i	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9	a_{10}	a_{11}	a_{12}
o_1		1	1			1	1		1	1	1	
o_2		1	1		1		1		1	1	1	
o_3		1	1			1	1		1	1	1	
o_4		1	1		1		1		1	1	1	
o_5		1	1		1		1		1	1	1	
o_6		1	1			1	1		1	1		
o_7		1	1		1		1		1	1	1	
o_8		1		1	1			1	1			1
o_9		1	1		1		1		1	1	1	
o_{10}		1	1		1		1		1		1	1
o_{11}	1		1		1			1				1
o_{12}		1	1			1	1		1	1	1	
o_{13}		1	1		1		1		1	1		
o_{14}		1	1			1		1	1	1		
o_{15}		1	1		1		1		1	1	1	
$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$
o_{86}	1		1		1			1				1
o_{87}	1			1	1			1				1
o_{88}		1		1	1			1				1
o_{89}		1		1	1			1	1	1		
o_{90}	1			1	1			1		1		
o_{91}	1			1	1			1				1
o_{92}		1		1	1			1				1

Table 4. (Continued)

a_i \ o_j	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9	a_{10}	a_{11}	a_{12}
o_{93}		1		1	1			1				1
o_{94}	1		1	1	1			1				1
o_{96}	1		1			1		1				1
o_{97}	1			1		1		1				
o_{98}	1		1		1			1				1
o_{99}		1		1	1			1				1
o_{100}	1			1	1			1		1		

3.4. WSD and rule extraction for solution of semantic merger of *should*

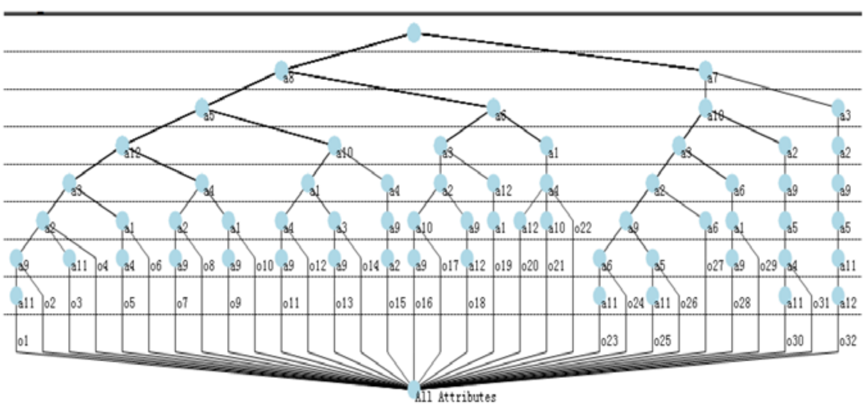
The attribute partial order structure diagram (APOSD) is generated and tested by the Leave-one-out approach [18], and the tested result shows that the WSD accuracy of the model reaches 96%. Then, the repeated attribute patterns in the formal context are deleted in order to generate a clarified attribute partial order structure diagram for rule extraction for the WSD. The clarified formal context is shown in Table 5, and the corresponding APOSD is shown in Figure 3.

Table 5. Clarified formal context of training set

o_j \ a_i	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9	a_{10}	a_{11}	a_{12}
o_1		1	1			1	1		1	1	1	
o_2		1	1		1		1		1	1	1	
o_6		1	1			1	1		1	1		
o_8		1		1	1			1	1			1
o_{10}		1	1		1		1		1		1	1
o_{11}	1		1		1			1				1
o_{13}		1	1		1		1		1	1		
o_{14}		1	1			1		1	1	1		
o_{21}	1		1			1	1			1		
o_{22}		1		1	1		1		1	1	1	
o_{24}		1	1			1		1	1			1
o_{25}		1	1			1	1			1		
o_{31}		1		1	1		1		1	1		
o_{32}	1		1			1	1		1	1		
o_{41}		1	1			1		1		1		
o_{47}		1	1		1			1	1			1
o_{51}		1	1		1			1				1
o_{52}	1			1	1			1				1
o_{57}	1		1		1			1	1	1		
o_{59}		1	1		1			1			1	1
o_{61}	1			1		1		1		1		
o_{67}	1		1		1			1		1		
o_{68}	1			1	1			1	1			1

Table 5. (Continued)

o_j \ a_i	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9	a_{10}	a_{11}	a_{12}
o_{69}	1			1	1			1		1		
o_{72}		1	1		1			1	1		1	1
o_{75}	1			1	1			1	1	1		
o_{88}		1		1	1			1				1
o_{89}		1		1	1			1	1	1		
o_{94}	1		1	1	1			1				1
o_{95}	1		1			1		1				1
o_{96}	1			1		1		1				
o_{100}	1			1		1		1				1



o1:o_{72} o2:o_{47} o3:o_{59} o4:o_{51} o5:o_{94} o6:o_{11} o7:o_8 o8:o_{88} o9:o_{68} o10:o_{52}
o11:o_{75} o12:o_{69} o13:o_{57} o14:o_{67} o15:o_{89} o16:o_{14} o17:o_{41} o18:o_{24} o19:o_{95}
o20:o_{100} o21:o_{61} o22:o_{96} o23:o_1 o24:o_6 o25:o_2 o26:o_{13} o27:o_{25}
o28:o_{32} o29:o_{21} o30:o_{22} o31:o_{31} o32:o_{10}

Figure 3. Clarified attribute partial order structure diagram of the training set.

Table 6. Extracted rules for solution of semantic merger of *should*

Rules for RT*should*	Rules for EP*should*
a_3, a_2, a_9, a_{12}, a_8, a_6	a_3, a_2, a_9, a_{12}, a_8, a_5
a_3, a_2, a_{10}, a_6, a_8	a_3, a_2, a_5, a_8, a_{12}
a_3, a_2, a_9, a_{10}	a_3, a_1, a_8
a_7	a_4, a_8

The rules for disambiguating RT*should* from EP*should* are extracted based on Figure 3 according to the hierarchical distribution of the attributes and the relations between the attributes and the objects, as shown in Table 6. The extracted rules are tested with the testing set, and the accuracy of WSD reaches 91%.

3.5. Solution of semantic mergers of *should*

The extracted rules in Table 6 are used to solute the semantic mergers in the ME*should* data set in Table 7, the corresponding formal context is shown in Table 8, and the final result is shown in Table 9.

Table 7. Data set of semantic merger of *should*

o_j	MI_1	MI_2	MI_3	MI_4	f_1	f_2	f_3	f_4
o_1	1.37	0.99	1.07	– 1	1	1	0	0
o_2	0.75	0.80	0.89	1.08	1	1	0	0
o_3	1.37	0.99	– 1	– 1	1	1	0	0
o_4	– 1	– 1	2.32	– 1	0	1	0	0
o_5	1.01	– 1	1.34	0.94	0	1	0	0
o_6	– 1	– 1	1.39	1.98	0	1	0	0
o_7	– 1	– 1	2.49	– 1	0	1	0	0
o_8	0.63	0.67	– 1	– 1	0	1	0	0
o_9	0.86	0.93	– 1	– 1	1	0	0	0
o_{10}	– 0.22	0.60	2.52	– 1	0	1	1	1
…	…	…	…	…	…	…	…	…
o_{71}	0.98	0.74	0.78	1.42	1	0	1	1
o_{72}	1.37	0.99	– 1	– 1	1	0	1	1
o_{73}	0.58	0.69	1.39	1.98	1	0	1	1
o_{74}	– 0.17	– 1	1.39	1.98	1	0	1	1
o_{75}	0.98	0.74	1.61	1.35	1	0	1	1
o_{76}	0.98	0.74	1.13	– 1	1	1	0	0
o_{77}	0.98	0.74	1.11	0.85	1	1	0	0
o_{78}	3.73	3.49	0.42	0.91	1	0	1	1
o_{79}	0.75	0.80	0.78	1.42	1	0	1	1
o_{80}	1.37	0.99	– 1	– 1	1	1	0	0

Table 8. Formal context of ME*should*

a_i \ o_j	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9	a_{10}	a_{11}	a_{12}
o_1		1	1		1		1		1	1		
o_2		1	1		1		1		1	1		
o_3	1	1	1		1		1		1	1		
o_4	1		1			1	1			1		
o_5		1	1		1		1			1		
o_6	1		1		1			1		1		
o_7	1		1			1	1			1		
o_8		1	1		1		1			1		
o_9		1	1		1		1		1			
o_{10}	1		1			1	1			1	1	1
o_{11}	1		1		1			1	1		1	1
o_{12}	1		1		1		1		1	1		
o_{13}		1	1		1		1			1		
o_{14}	1		1		1		1			1		
o_{15}		1	1		1		1			1		
$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$
o_{66}		1	1		1		1			1		
o_{67}	1			1		1	1			1		
o_{68}		1	1			1		1	1		1	1
o_{69}		1	1		1		1		1	1		
o_{70}		1	1		1		1		1		1	1
o_{71}		1	1		1			1	1		1	1

Table 8. (Continued)

a_i / o_j	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9	a_{10}	a_{11}	a_{12}
o_{72}		1	1		1		1		1		1	1
o_{73}		1	1		1			1	1		1	1
o_{74}	1		1		1			1	1		1	1
o_{76}		1	1			1		1	1		1	1
o_{77}		1	1		1		1		1	1		
o_{78}		1	1		1		1		1	1		
o_{79}		1		1	1		1		1		1	1
o_{80}		1	1		1			1	1		1	1

Table 9. Result of solution of semantic mergers of *should* by APOSD approach

Class	Objects
RT*should* (53)	$o_1, o_2, o_3, o_4, o_5, o_7, o_8, o_9, o_{10}, o_{12}, o_{13}, o_{15}, o_{17}, o_{18}, o_{20}, o_{21}, o_{22}, o_{23}, o_{25}, o_{27}, o_{28}, o_{29}, o_{30}, o_{31}, o_{32}, o_{33}, o_{34}, o_{35}, o_{37}, o_{39}, o_{40}, o_{42}, o_{44}, o_{46}, o_{48}, o_{53}, o_{56}, o_{57}, o_{59}, o_{61}, o_{63}, o_{64}, o_{65}, o_{66}, o_{68}, o_{69}, o_{70}, o_{72}, o_{75}, o_{76}, o_{77}, o_{78}, o_{80}$
EP*should* (27)	$o_6, o_{11}, o_{14}, o_{16}, o_{19}, o_{24}, o_{26}, o_{36}, o_{38}, o_{41}, o_{43}, o_{45}, o_{47}, o_{49}, o_{50}, o_{51}, o_{52}, o_{54}, o_{55}, o_{58}, o_{60}, o_{62}, o_{67}, o_{71}, o_{73}, o_{74}, o_{79}$

4. Solution of Semantic Merger of English Modal Verb by Fuzzy C-Means Cluster

A fuzzy c-means (FCM) partitions a set of n-dimensional vectors $\boldsymbol{X} = \{x_1, x_2, \ldots, x_N\}$ into c clusters, where $x_j = \{x_{j1}, x_{j2}, \ldots, x_{jn}\}$ $(j = 1, 2, \ldots, n)$ represents the j-th sample. For the j-th sample x_j and the i-th cluster center v_i, there is a membership degree u_{ji} which indicates to what degree the sample x_j belongs to the cluster center $v_i = \{v_{1i}, v_{2i}, \ldots, v_{ni}\}$ $(i = 1, 2, \ldots, c)$ resulting in a fuzzy partition matrix $\boldsymbol{U} = (u_{ji})_{dxc}$. The FCM algorithm is based on minimizing the objective function J_m $(\boldsymbol{U}, \boldsymbol{V})$, which is defined as

$$J_m(\boldsymbol{U}, \boldsymbol{V}) = \sum_{j=1}^{N}\sum_{i=1}^{c} u_{ji}^m d_{ji}^2, \tag{2}$$

here $\boldsymbol{V} = \{v_1, v_2, \ldots, v_c\}$, and d_{ji} is the Euclidean distance between x_j and the cluster center v_i, which is defined as

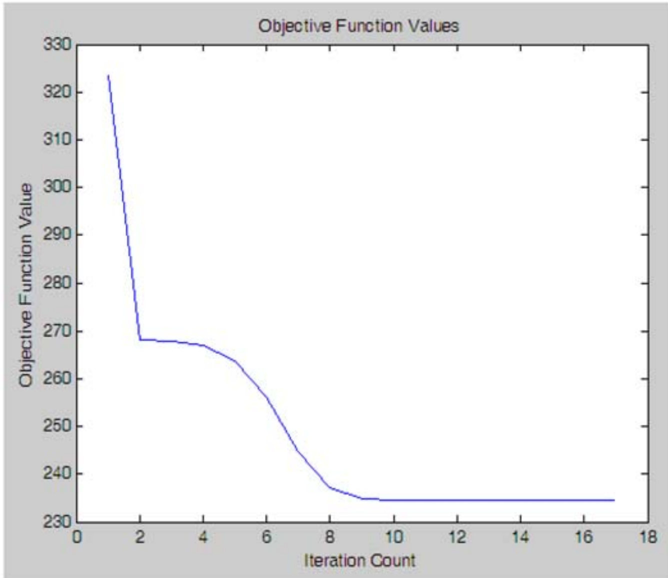
$$d_{ji} = \sqrt{\sum_{p=1}^{N} (v_{pi} - x_{pj})^2}. \tag{3}$$

The exponent m in Equation (2) is the fuzzifier parameter and it defines the fuzziness of the clustering. The u_{ji} and v_i are calculated from the following formulae:

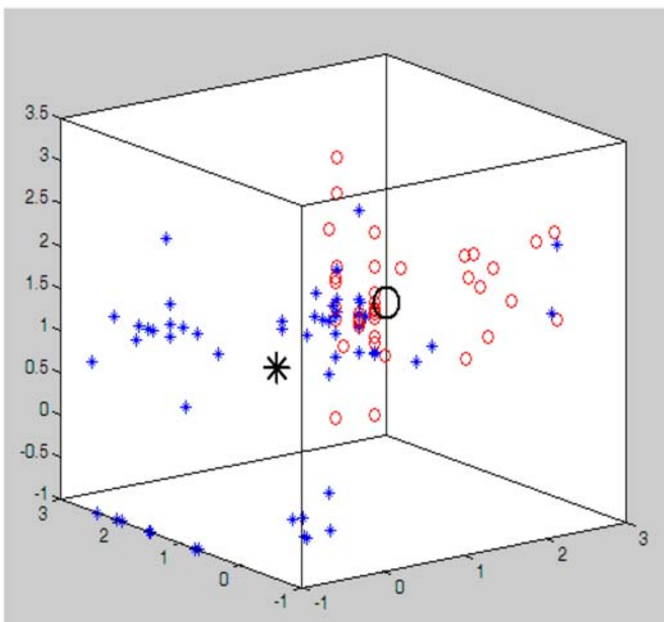
$$u_{ji} = 1 \Big/ \sum_{k=1}^{c} (d_{ji} / d_{jk})^{\frac{2}{m-1}}, \quad \text{here} \quad m \neq 1, \quad v_i = \sum_{j=1}^{N} u_{ji}^m x_j \Big/ \sum_{j=1}^{N} u_{ji}^m. \tag{4}$$

Finally, the FCM algorithm results in a fuzzy partition matrix and cluster centers.

The model for solving the semantic mergers of *should* is established. The number of cluster is set to be $c = 2$. The data in Table 2 are input into the FCM model and run Matlab 7.10. By comparing the degree of membership and the actual senses of the samples, the word senses of *should* are disambiguated. The accuracy of WSD of *should* reaches 90%.



(a) FCM objective function for training set



(b) FCM cluster for training set

Figure 4. Results of WSD by FCM cluster.

Then, the established model is used for the semantic merger solution of *should*. Input the data set in Table 7 into the FCM cluster, and the run Matlab 7.10, the semantic mergers of should are solved. The result is shown in Table 10.

Table 10. Result of solution of semantic mergers of *should* by FCM cluster

Class	Objects
RT*should* (36)	$o_1, o_3, o_4, o_7, o_8, o_9, o_{10}, o_{12}, o_{14}, o_{15}, o_{20}, o_{21}, o_{25}, o_{27}, o_{28}, o_{31}, o_{33}, o_{34}, o_{37}, o_{39}, o_{42}, o_{46}, o_{53}, o_{56}, o_{57}, o_{59}, o_{62}, o_{63}, o_{64}, o_{66}, o_{67}, o_{70}, o_{72}, o_{76}, o_{78}, o_{80}$
EP*should* (44)	$o_2, o_5, o_6, o_{11}, o_{13}, o_{16}, o_{17}, o_{18}, o_{19}, o_{22}, o_{23}, o_{24}, o_{26}, o_{29}, o_{30}, o_{32}, o_{35}, o_{36}, o_{38}, o_{40}, o_{41}, o_{43}, o_{44}, o_{45}, o_{47}, o_{48}, o_{49}, o_{50}, o_{51}, o_{52}, o_{54}, o_{55}, o_{58}, o_{60}, o_{61}, o_{65}, o_{68}, o_{69}, o_{71}, o_{73}, o_{74}, o_{75}, o_{77}, o_{79}$

Compared with the result by the attribute partial order structure approach in Table 9, 72% of the semantic mergers of *should* are correctly solved by the fuzzy c-means cluster.

5. Comparison of Semantic Merger Solution of English Modal Verb *Should* by the two Approaches

In order to know which of the above mentioned approaches performs better in solution of semantic mergers, a comparison is conducted between the two approaches from the perspectives of easiness, accuracy, and effectiveness. The comparative results show that:

(1) The attribute partial order structure diagram approach uses more steps in the process of solution of semantic merger than FCM cluster. The former needs to extract rules for disambiguating RT*should* from EP*should* from the APOSD model first, and then used the rules to solve the semantic merger of *should*; while the latter solves the semantic merger directly from the FCM model.

(2) The accuracies of word sense disambiguation by APOSD approach and by the FCM cluster are about the same (90% to 91%); but the rate of semantic merger solution by the APOSD approach is much higher than the one by the FCM cluster (72% to 100%).

(3) In the attribute partial order structure approach, the model of WSD can also be used for knowledge discovery, such as the interactive relations and hierarchical relations between the word senses and the attributes, and the relations between the attributes, and so on. However, the FCM cluster can only be used for WSD and solution of semantic mergers. Therefore, the attribute partial order structure approach is more effective than the FCM cluster.

6. Conclusion

In this study, both the approach of attribute partial order structure and the fuzzy c-means cluster are used in semantic merger solution of English modal verb *should*. It is proven through comparative analysis that the attribute partial order structure approach performs better than the fuzzy c-means cluster in accuracy of WSD and solution of semantic merger of *should*. In addition, the attribute partial order structure approach can also be used in knowledge discovery, data mining, and pattern recognition because the working principle of attribute partial order follows the basic principle of human being's cognition of classification of natural things. A further study will compare the attribute partial order structure approach with other possible approaches in order to find the best approach for solution of semantic mergers of English modal verbs.

Acknowledgement

This work is supported by the National Social Science Foundation under Grant No. 12BYY121 and by the Humanities and Social Sciences Foundation of the Ministry of Education of China under Grant

No. 12YJA740096. It is also partially supported by National Natural Science Foundation of China under Grant No. 61074130. The authors gratefully acknowledge the supports.

References

[1] P. Everaere, S. Konieczny and P. Marquis, Disjunctive merging: Quota and Gmin merging operators, Artificial Intelligence 174 (2010), 824-849.

[2] H. Qin, Semantic Overlap and Construal of English Prepositional Phrases: A Cognitive-corpus Analysis, Master Degree Thesis, Nanchang University, P. R. China, 2009.

[3] E. R. Thomas, Encyclopedia of Language & Linguistics, Second Edition, 2006.

[4] R. Wille, Restructuring Lattice Theory: An Approach Based on Hierarchies of Concepts, In Ivan Rival (ed.), Ordered Sets, Reidel, Dordrecht-Boston, 1982.

[5] B. Ganter and R. Wille, Formal Concept Analysis, Mathematical Foundations, Berlin/Heidelberg/New York: Springer-Verlag, 1996.

[6] W. Hong, S. Li, J. Yu and J. Song, A new approach of generation of structural partial-ordered attribute diagram, ICIC Express Letters, Part B: Applications 3(4) (2012), 823-830.

[7] J. Yu, W. Hong, S. Li, T. Zhang and J. Song, A new approach of word sense disambiguation and knowledge discovery of English modal verbs by formal concept analysis, International Journal of Innovative Computing, Information and Control 9(3) (2013), 1189-1200.

[8] J. Fu, J. Yu and H. Liu, Data mining of grammatical errors in Chinese postgraduate's academic writing, ICIC Express Letters 7(12) (2013), 3221-3227.

[9] J. Yu, N. Chen, W. Hong, S. Li and T. Zhang, Interactive relations between semantic and syntactic features in word sense disambiguation of semantically complex words, International Journal of Innovative Computing, Information and Control 9(9) (2013), 3627-3638.

[10] W. Ma, L. Jiao, M. Gong and C. Li, Image change detection based on an improved rough fuzzy c-means clustering algorithm, International Journal of Machine Learning and Cybernetics 5(3) (2014), 369-377.

[11] X. Yin, Y. Li and F. Li, Segmentation for brain MRI image based on the fuzzy c-means clustering algorithm, WIT Transactions on Information and Communication Technologies 46(2) (2014), 1177-1182.

[12] J. Hsieh, C. Tai, S. Su and Y. Lin, Identification of partial discharge location using probabilistic neural network and the fuzzy c-means clustering approach, Electric Power Components and Systems 42(1) (2014), 60-69.

[13] D. Zhao, L. Song and J. Yan, Feature recognition method based on fuzzy clustering analysis and its application, Computer Integrated Manufacturing Systems 15(12) (2009), 2417-2423.

[14] X. Fang, On the problems of fuzzy clustering analysis used in information retrieval, Science Technology and Engineering 10(11) (2010), 2272-2276.

[15] R. Vijayarajan and S. Muttan, Fuzzy c-means clustering based principal component averaging fusion, International Journal of Fuzzy Systems 16(2) (2014), 153-159.

[16] J. Coates, The Semantics of the Modal Auxiliaries, London: Routledge Press, 1983.

[17] C. Zong, Statistical Natural Language Processing, Beijing: Tsinghua University Press, 2008.

[18] Z. Bian and X. Zhang, Pattern Recognition, Beijing: Tsinghua University Press, 2000.

■