*International Journal of Innovative Computing, Information and Control*
Volume **12**, Number **2**, April **2016**

ICIC International ©2016 ISSN 1349-4198
pp. 491–502

# RESTRICTION OF TIME, ASPECT AND VOICE ON SENSES OF ENGLISH MODAL VERBS

Jianping Yu[1,*], Yuting Zhang[1] and Wenxue Hong[2]

[1]The School of Foreign Studies
[2]Institute of Electrical Engineering
Yanshan University
No. 438, Hebei Street, Qinhuangdao 066004, P. R. China
*Corresponding author: yjp@ysu.edu.cn

Received August 2015; revised December 2015

Abstract. *Time, aspect and voice are 3 grammatical categories and they have restrictions on the senses of English modal verbs. How the senses of English modal verbs are restricted by the 3 grammatical categories has been a tough issue in both linguistic study and natural language processing. In this article, the restriction of time, aspect and voice on the senses of English modal verbs is investigated by the approach of attribute partial-ordered structure diagram. Firstly, the theoretical foundation of this study is described. Secondly, the model for word sense disambiguation (WSD) of the English modal verb "must" is built with the accuracy of 93.63%. Thirdly, the restriction of time, aspect and voice on the senses of "must" is investigated from difference angles. It is found that time, aspect and voice attributes are mainly functioning classification in the WSD of "must". Even though they do not contribute much to the accuracies of WSD of "must", they do have some restrictions to the senses of modal verbs. Past time only co-occurs with "must"(necessary) and "must"(possibility). Prograssive and perfect aspect restrict "must" to the sense of "possibility". Passive voice tends to co-occur with "must"(necessary) and "must"(permission). These findings provide significant evidence for both the semantic study and the rule extraction for WSD of the English modal verbs.*
**Keywords:** English modal verb, Attribute partial-ordered structue diagram, Time, Aspect, Voice

1. **Introduction.** English modal verbs have been a complex semantic system, and the senses of modal verbs are influenced by many factors. This has brought a lot of trouble to the linguistic study and the study of natural language processing. Many researchers have studied English modal verbs from different aspects. Ji and Liang [1] made a multi-dimensional diachronic study of the modal meanings and syntactic features of *must*. Liu and Peng [2] studied the interactive restriction of co-occurrence of the Chinese modality and aspect. Zhang [3] used the "dynamic evolutionary model" in cognitive grammar to explain the senses of the English modal verbs. Xu [4] discussed the modality of English tenses from the aspect of systematical functional grammar. Yu et al. [5-9] studied word sense disambiguation (WSD) of the English modal verbs by neural network, naïve Bayesian model, support vector machine and the approach of attribute partial-ordered structure diagram. They have also conducted data mining and knowledge discovery of the interactive relations among the linguistic features.

As we can see from the previous studies, the studies from linguistics field have mainly focused on the understanding of the senses and the diachronic semantic change of the senses of modal verbs. The studies from natural language processing have mainly focused on the WSD and knowledge discovery of English modal verbs. And the knowledge discovery has mainly based on semantic features and syntactic features, such as animate subject, agentive verb, stative verb and negation. Time, aspect and voice are 3 main grammatical categories and they have restrictions on the senses of modal verbs. Making clear how time, aspect and voice restrict the senses of English modal verbs has been a tough and significant issue in both linguistic studies and studies in natural language processing. However, few studies have been found on this issue. Therefore, in this article, this issue is studied by an approach named attribute partial-ordered structure diagram. The approach of attribute partial-ordered structure diagram is proposed by Hong et al. [10] based on the theory of formal concept analysis. In this approach, a hierarchical structure diagram is constructed based on the degrees of covering of attributes over objects, features of attributes and the similarity of objects. It can gather the objects with common features together to show the generality of the dataset, and separate the objects with exclusive attributes apart from others to show the specificity of the dataset. The approach has been used in different applications, such as word sense disambiguation [11-13], knowledge discovery in medicine and semantic studies [8,14,15], and pattern recognition [9], and it has been proven the most effective approach in word sense disambiguation and knowledge discovery [8]. Therefore, the approach of attribute partial-ordered structure diagram is chosen and used in this study to find the answers to the above mentioned issue and provide valuable evidence for the linguistic studies and the studies of natural language processing of English modal verbs. Some nomenclatures used in this study are given as the following.

**Nomenclatures**

| symbol | description | symbol | description |
|---|---|---|---|
| WSD | word sense disambiguation | $P(w_1, w_2)$ | probability of co-occurrence of $w_1$ and $w_2$ |
| $K$ | a formal context | $P(w_1)$ | probability of $w_1$ |
| $G$ | a set of objects | $\mathrm{MI}_i$ | the $i$th mutual information |
| $M$ | a set of attributes | $j(s)$ | the $j$th object with sense of $s$; it is corresponding to $u$ in Definition 2.1. |
| $I$ | a set of relation between objects and attributes | $a_i$ | the $i$th attribute, $a$ is corresponding to $m$ in Definition 2.1. |
| $g$ | an object | APOSD | attribute partial order structure diagram |
| $A$ | extent of a concept | $S$ | similarity of the attribute patterns of two objects |
| $B$ | intent of a concept | $X_i$ | attribute pattern |
| $w_1$, $w_2$ | $w_1$ is the target word for WSD; $w_2$ is the adjacent word to $w_1$ | $S_i$ | the $i$th similarity |

The rest of this article is organized as follows. In Section 2, some theoretical foundation of formal context and attributes are introduced. Section 3 explains the procedures of building the model for word sense disambiguation of *must*. In Section 4, the restriction of time, aspect and voice on the senses of the English modal verb *must* is analyzed. Section 5 comes to the conclusions.

## 2. Theoretical Foundation.

This study is based on the following theoretical foundation of formal context [16] and attributes [17].

**Definition 2.1.** *A formal context* $K = (G, M, I)$ *consists of two sets* $G$ *and* $M$ *and a relation* $I$ *between* $G$ *and* $M$. *The elements of* $G$ *are called the objects and the elements of* $M$ *are called the attributes of the context.* $I$ *represents the relation between an object* $u$ *and an attribute* $m$, *written as* $gIm$ *or* $(g, m) \in I$.

**Definition 2.2.** *Let* $K = (G, M, I)$ *be a formal context, for a set* $A \subseteq G$, $f(A) = \{m \in M | (g, m) \in I, \forall g \in A\}$. *Correspondingly, for a set* $B \subseteq M$, *define* $g(B) = \{g \in G | (g, m) \in I, \forall m \in B\}$. *A formal concept is a pair* $(A, B)$ *with* $A \subseteq G$, $B \subseteq M$, $f(A) = B$ *and* $g(B) = A$. $A$ *is called the extent of the concept and* $B$ *is called the intent of the concept.*

**Definition 2.3.** *A binary relation* $I$ *on a set* $M$ *is called an order relation, if it satisfies the following conditions for all elements* $x, y, z \in M$:

*1) reflexity:* $xRx$

*2) antisymmetry:* $xRy$ *and* $x \neq y \rightarrow$ *not* $yRx$

*3) transitivity:* $xRy$ *and* $yRz \rightarrow xRz$

**Definition 2.4.** *If* $(A_1, B_1)$ *and* $(A_2, B_2)$ *are concepts of a context,* $(A_1, B_1)$ *is called a subconcept of* $(A_2, B_2)$, *provided that* $A_1 \subseteq A_2$ *(which is equivalent to* $B_2 \subseteq B_1$*). In this case,* $(A_2, B_2)$ *is a superconcept of* $(A_1, B_1)$, *and we write* $(A_1, B_1) \leq (A_2, B_2)$. *The relation* $\leq$ *is called the hierarchical order of the concepts.*

**Definition 2.5.** *Let* $K = (G, M, I)$ *be a formal context, if for any objects* $g_1, g_2 \in G$ *from* $f(g_1) = f(g_2)$, *it always follows that* $g_1 = g_2$ *and correspondingly,* $g(m_1) = g(m_2)$ *implies* $m_1 = m_2$ *for all* $m_1, m_2 \in M$, *the context* $K = (G, M, I)$ *is called clarified context.*

**Definition 2.6.** *Let* $K = (G, M, I)$ *be a formal context* $m \in M$. *If attribute* $m$ *satisfies the following condition:* $\{g(m) | m \in M\} = U$, *then,* $m$ *is called a maximum common attribute.*

**Definition 2.7.** *Let* $m_0$, $m_1$, $m_2$, $\ldots$, $m_k$ *be the intents of some concepts, if* $g(m_i) \subset g(m_0)$, *where* $i = 1, 2, \ldots, k$, $k \geq 2$, *then* $m_0$ *is a common attribute of the concepts corresponding to* $m_i$.

**Definition 2.8.** *Let* $K = (G, M, I)$ *be a formal context,* $m \in M$. *If attribute* $m$ *satisfies* $|g(m)| = 1$, *then* $m$ *is an exclusive attribute. Here,* $|g(m)| = 1$ *is the length of the object set, i.e., the number of the objects that* $m$ *corresponds to.*

3. **Building of the Model for Word Sense Disambiguation of *Must*.** The investigation of the restriction of time, aspect and voice on the senses of a modal verb is based on the model of WSD of that modal verb. Therefore, the building of a model for WSD is necessary. In this study, the English modal verb *must* is chosen as the target word for WSD. Based on 4 well-known English dictionaries, namely Merriam-Webster's Advanced Learner's English Dictionary [18], Cambridge Advanced Learner's Dictionary [19], Longman Dictionary of Contemporary English [20], and The Concise Oxford English Dictionary [21], 4 modal senses of *must* including the senses of necessity, possibility, insistence and permission are determined for word sense disambiguation, as shown in Table 1.

3.1. **Data preparation.** A corpus of 2.16 million words is built, which is composed of different genres, such as law, literature work, news report, academic paper, interview, speech, movie subtitle, science fiction, popular science books, scientific forum, book review and introduction to products. The materials are evenly extracted from each genre. The 4 senses of *must* in Table 1 are tagged with NEC*must-must*(necessary), POS*mus-must*(possibility),

TABLE 1. The 4 main senses of *must*

| Senses of *must* | Definition | Example |
|---|---|---|
| necessary | something is required of necessary by someone, an organization, a rule or a law | One *must* eat to live. |
| possibility | something is very likely to happen | She *must* think I am a fool. |
| insistence | someone has a definite intension to do something in the future, expressing insistence or emphasizing a statement | I *must* ask you to leave. |
| permission | to tell someone that they are allowed to do or not allowed to do something | You *must* not smoke. |

INS*must*-*must*(insistence) and PER*must*-*must*(permission), respectively. A hundred and fifty-seven sample sentences are extracted from the corpus according to the occurrence of the samples in the 4 senses. Among the 157 samples, 51 are for NEC*must*, 51 for POS*must*, 35 for INS*must* and 20 for PER*must*.

3.2. **Feature exaction for WSD of *must*.** Since this study tries to discover how the time, aspect and voice factors restrict the sense of English modal verbs, some semantic features, such as mutual information (MI) between the target word and the adjacent words, and syntactic features, such as the co-occurred syntactic features of time, aspect and voice are extracted, as shown in Table 2.

TABLE 2. Extracted features for WSD of *must*

| Semantic features | Syntactic features |
|---|---|
| MI(s+NEC*must*) | past time |
| MI(NEC*must*+v) | present time |
| MI(s+POS*must*) | future time |
| MI(POS*must*+v) | perfect aspect |
| MI(s+INS*must*) | progressive aspect |
| MI(INS*must*+v) | active voice |
| MI(s+PER*must*) | passive voice |
| MI(PER*must*+v) | negation |

In Table 2, MI(s+NEC*must*) represents the mutual information between the subject and the *must* (necessary) in the sample sentence, and MI(NEC*must*+v) represents the mutual information between the *must* (necessary) and the main verb in the sample sentence. It is the same for the other semantic features. The mutual information is calculated by the following formula [22]:

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

where $w_1$ and $w_2$ are the target word for WSD and the adjacent word, respectively; $P(w_1, w_2)$ is the probability of the co-occurrence of $w_1$ and $w_2$. $P(w_1)$ and $P(w_2)$ are the probability of $w_1$ and $w_2$, respectively. The number of co-occurrences and the occurrences of the words are counted by the Concordance Tool of Wordsmith 4.0 software. $P(w_1, w_2)$ is calculated by Equation (2), and $P(w_1)$ and $P(w_2)$ are calculated by Equation (3). The MIs for WSD of *must* are calculated and listed in Table 3.

$$P(w_1, w_2) = \frac{\text{number of co-occurrence of } w_1, \text{ and } w_2 \text{ in corpus}}{\text{total number of words in corpus}} \tag{2}$$

$$P(w_i) = \frac{\text{number of occurrence of } w_i \text{ in corpus}}{\text{total number of words in corpus}} \tag{3}$$

TABLE 3. MIs for WSD of *must*

| $j(s)$ \ $\mathrm{MI}_i$ | $\mathrm{MI}_1$ | $\mathrm{MI}_2$ | $\mathrm{MI}_3$ | $\mathrm{MI}_4$ | $\mathrm{MI}_5$ | $\mathrm{MI}_6$ | $\mathrm{MI}_7$ | $\mathrm{MI}_8$ |
|---|---|---|---|---|---|---|---|---|
| 1(a) | 1.69 | 2.54 | −1 | −1 | −1 | −1 | −1 | −1 |
| 2(a) | 1.97 | 2.09 | −1 | −1 | −1 | −1 | −1 | −1 |
| 3(a) | 0.86 | 2.17 | 1.32 | −1 | −1 | −1 | 1.08 | −1 |
| 4(a) | 3.48 | 2.64 | −1 | −1 | −1 | −1 | −1 | −1 |
| 5(a) | 2.91 | 1.51 | −1 | 2.04 | −1 | 1.47 | −1 | −1 |
| 6(a) | 1.87 | 1.79 | −1 | −1 | −1 | −1 | −1 | −1 |
| 7(a) | 0.69 | 2.93 | 1.16 | −1 | −1 | −1 | −1 | −1 |
| 8(a) | 2.88 | 1.65 | −1 | −1 | −1 | −1 | −1 | −1 |
| 9(a) | 1.04 | 2.52 | 0.95 | −1 | 1.01 | −1 | −1 | −1 |
| 10(a) | 2.40 | 1.78 | −1 | −1 | −1 | −1 | −1 | 3.71 |
| ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 148(d) | −1 | 1.62 | −1 | −1 | −1 | −1 | 2.83 | 4.73 |
| 149(d) | −1 | −1 | −1 | −1 | −1 | −1 | 3.45 | 3.68 |
| 150(d) | −1 | 1.62 | −1 | −1 | −1 | −1 | 3.50 | 3.68 |
| 151(d) | −1 | 1.62 | −1 | −1 | −1 | −1 | 3.11 | 3.46 |
| 152(d) | 0.56 | 2.01 | 1.27 | −1 | −1 | −1 | 1.49 | 3.46 |
| 153(d) | −1 | 2.01 | −1 | −1 | −1 | −1 | 3.15 | 3.68 |
| 154(d) | −1 | 1.62 | −1 | −1 | −1 | −1 | 3.18 | 3.68 |
| 155(d) | −1 | 1.62 | −1 | −1 | −1 | −1 | 3.13 | 3.68 |
| 156(d) | −1 | 1.62 | −1 | −1 | −1 | −1 | 3.10 | 3.46 |
| 157(d) | −1 | 2.01 | −1 | −1 | −1 | −1 | 3.73 | 4.73 |

Here, $\mathrm{MI}_1$ represents MI(s+NEC*must*); $\mathrm{MI}_2$ – MI(NEC*must*+v); $\mathrm{MI}_3$ – MI(s+POS*must*); $\mathrm{MI}_4$ – MI(POS*must*+v); $\mathrm{MI}_5$ – MI(s+INS*must*); $\mathrm{MI}_6$ – MI(INS*must*+v); $\mathrm{MI}_7$ – MI(s+PER*must*); $\mathrm{MI}_8$ – MI(PER*must*+v). The −1 represents the case of no co-occurrence of $w_1$ and $w_2$.

3.3. **Symbolization of the data set of *must*.** Since some MIs in Table 3 are in continuous values, they need to be transformed into bi-values in order to construct the formal context for WSD. The approach of scattered diagram is used. Firstly, the scattered diagrams for $\mathrm{MI}_i$ are generated, and then, a point is selected for each of $\mathrm{MI}_i$ to nicely separate one class of object from the others. For instance, the scattered diagram for $\mathrm{MI}_2$ is generated as shown in Figure 1.

It can be seen from Figure 1 that the value of 0.36 can nicely separate the first class of objects from the others. Therefore, the values of $\mathrm{MI}_2$ is categorized into two categories: a3-MI(NEC*must*+v) $\leq 0.36$ and a4-MI(NEC*must*+v) $> 0.36$. In this way, the $\mathrm{MI}_i$ are scattered, as shown in Table 4.

If an object has the semantic feature described in Table 4, it is given a logical value of 1; otherwise, it is given 0. The syntactic features are also given logical values of 1 for having the feature or 0 for not having the feature. By now, altogether 24 attributes are prepared and used in the generation of the formal context for WSD of *must*, as shown in Table 4. Finally, the formal context for WSD of *must* is prepared, as shown in Table 5.

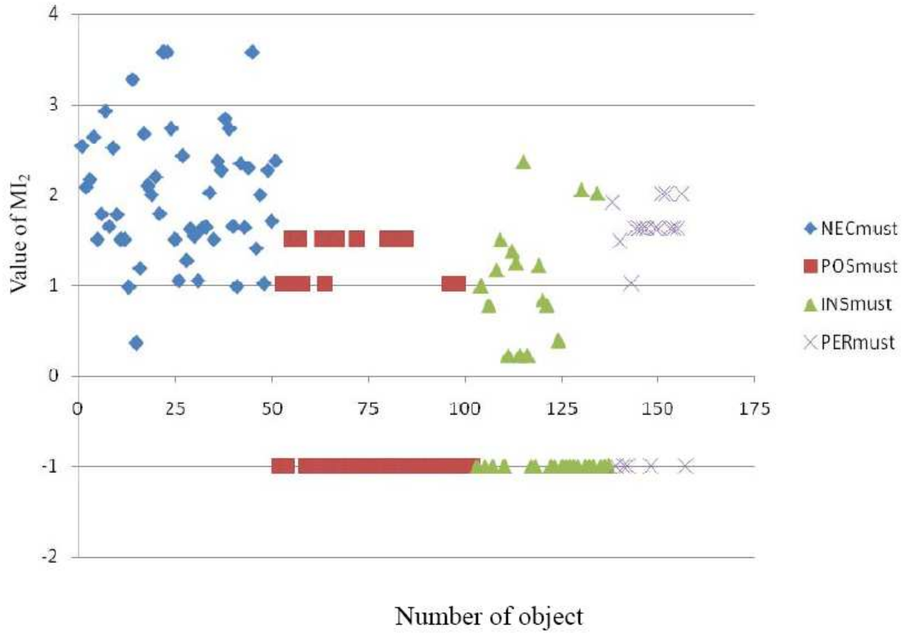



FIGURE 1. Scattered diagram of $MI_2$

TABLE 4. Descriptions of semantic and syntactic features

| | Descriptions | | Descriptions | | Descriptions |
|---|---|---|---|---|---|
| a1 | MI(s+NEC*must*) $\leq -0.05$ | a9 | MI(s+INS*must*) $\leq 0.91$ | a17 | past time |
| a2 | MI(s+NEC*must*) $> -0.05$ | a10 | MI(s+INS*must*) $> 0.91$ | a18 | present time |
| a3 | MI(NEC*must*+v) $\leq 0.36$ | a11 | MI(INS*must*+v) $\leq 1.05$ | a19 | future time |
| a4 | MI(NEC*must*+v) $> 0.36$ | a12 | MI(INS*must*+v) $> 1.05$ | a20 | perfect aspect |
| a5 | MI(s+POS*must*) $\leq 0.3$ | a13 | MI(s+PER*must*) $\leq 0.59$ | a21 | progressive aspect |
| a6 | MI(s+POS*must*) $> 0.3$ | a14 | MI(s+PER*must*) $> 0.59$ | a22 | active voice |
| a7 | MI(POS*must*+v) $\leq 0.41$ | a15 | MI(PER*must*+v) $\leq 1.39$ | a23 | passive voice |
| a8 | MI(POS*must*+v) $> 0.41$ | a16 | MI(PER*must*+v) $> 1.39$ | a24 | negation |
| a1∼a16 – semantic features; a17∼a24 – syntactic features | | | | | |

3.4. **Generation of attribute partial-ordered structure diagram.** By using the Attribute Partial-ordered Structure Diagram Tool [10], an attribute partial-ordered structure diagram is generated as a model for WSD of *must*, as shown in Figure 2. Figure 2 shows the partial ordered and hierarchical structure of the attributes and objects in the formal context in Table 5. The attributes at the higher layers have larger covering of the objects and they have the feature of extent functioning generalization. And the attributes at the lower layers have much less covering of the objects, and they have the feature of intent functioning specification. In Figure 2, a$i$ ($i = 1, 2, 3, \ldots, 24$) represents attributes and o$j$ represents objects. o26:118(c) means the o26 in Figure 2 represents the 118th object in the formal context and the object belongs to the class $c$. The generated model is tested by the approaches of Leave-One-Out Cross Validation [23] and the similarity method [10]. If o1' is the taken out object and its attribute pattern can be described by: $X_1' = \{x_1', x_2', x_3', \ldots, x_i'\}$, and o1 is an object in the APOSD which has similar attribute pattern to o1' and can be described by: $X_1 = \{x_1, x_2, x_3, \ldots, x_m\}$, the similarity of $X_1'$ and $X_1$ is represented by $S$:

$$S = |X_1' \cap X_1| / \max(|X_1'|, |X_1|) \tag{4}$$

Table 5. Formal context of training set for WSD of *must*

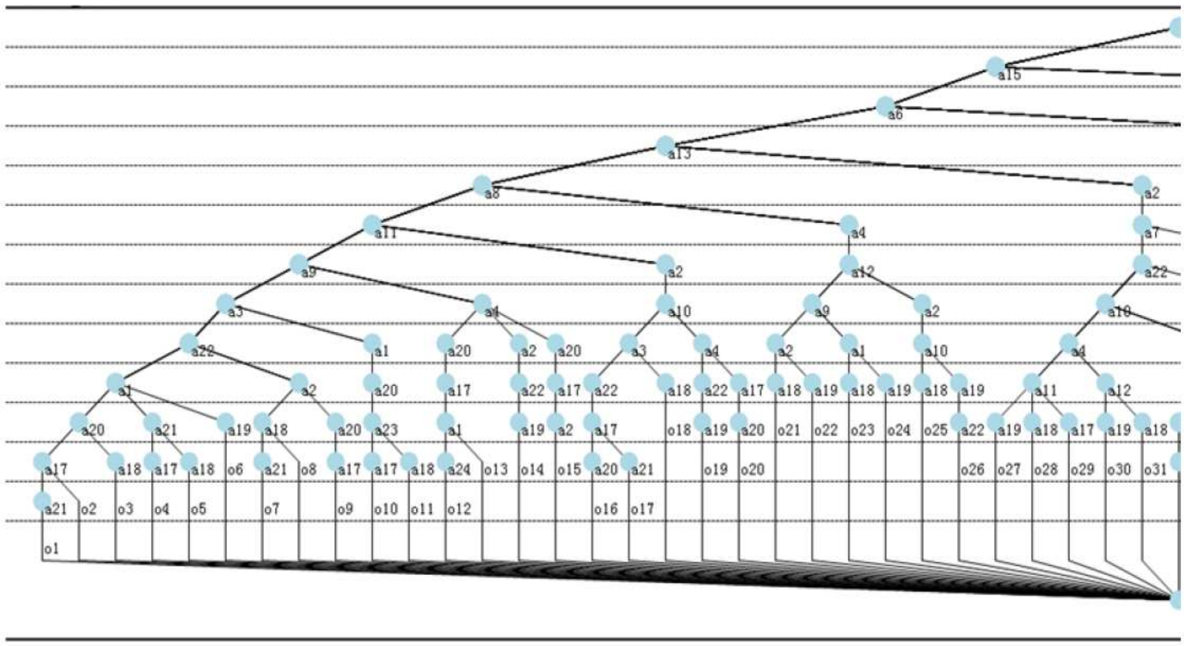
| $j(s)$ \ $ai$ | a1 | a2 | a3 | a4 | a5 | a6 | a7 | a8 | a9 | a10 | a11 | a12 | a13 | a14 | a15 | a16 | a17 | a18 | a19 | a20 | a21 | a22 | a23 | a24 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1(a) | | 1 | | 1 | 1 | | 1 | | 1 | | 1 | | 1 | | 1 | | | | 1 | | | | 1 | |
| 2(a) | | 1 | | 1 | 1 | | 1 | | 1 | | 1 | | 1 | | 1 | | | 1 | | | | 1 | | |
| 3(a) | | 1 | | 1 | | 1 | 1 | | 1 | | 1 | | | 1 | 1 | | | | 1 | | | 1 | | |
| 4(a) | | 1 | | 1 | 1 | | 1 | | 1 | | 1 | | 1 | | 1 | | | | 1 | | | | 1 | |
| 5(a) | | 1 | | 1 | 1 | | | 1 | 1 | | | 1 | 1 | | 1 | | | 1 | | | | | | |
| 6(a) | | 1 | | 1 | 1 | | 1 | | 1 | | 1 | | 1 | | 1 | | | | 1 | | | 1 | | |
| 7(a) | | 1 | | 1 | | 1 | 1 | | 1 | | 1 | | 1 | | 1 | | | | 1 | | | | 1 | |
| 8(a) | | 1 | | 1 | 1 | | 1 | | 1 | | 1 | | 1 | | 1 | | | 1 | | | | 1 | | |
| 9(a) | | 1 | | 1 | | 1 | 1 | | | 1 | 1 | | 1 | | 1 | | | 1 | | | | 1 | | |
| 10(a) | | 1 | | 1 | 1 | | 1 | | 1 | | 1 | | 1 | | | 1 | 1 | | | | | | 1 | |
| … | … | … | … | … | … | … | … | … | … | … | … | … | … | … | … | … | … | … | … | … | … | … | … | … |
| 148(d) | 1 | | 1 | | 1 | | 1 | | 1 | | 1 | | | 1 | | 1 | | | 1 | | | | 1 | 1 |
| 149(d) | 1 | | | 1 | 1 | | 1 | | 1 | | 1 | | | 1 | | 1 | | | 1 | | | | 1 | 1 |
| 150(d) | 1 | | | 1 | 1 | | 1 | | 1 | | 1 | | | 1 | | 1 | | | 1 | | | | 1 | 1 |
| 151(d) | 1 | | | 1 | 1 | | 1 | | 1 | | 1 | | | 1 | | 1 | | | 1 | | | | 1 | 1 |
| 152(d) | | 1 | | 1 | | 1 | 1 | | 1 | | 1 | | | 1 | | 1 | | | 1 | | | | 1 | 1 |
| 153(d) | 1 | | | 1 | 1 | | 1 | | 1 | | 1 | | | 1 | | 1 | | | 1 | | | | 1 | 1 |
| 154(d) | 1 | | | 1 | 1 | | 1 | | 1 | | 1 | | | 1 | | 1 | | | 1 | | | | 1 | 1 |
| 155(d) | 1 | | | 1 | 1 | | 1 | | 1 | | 1 | | | 1 | | 1 | | | 1 | | | | 1 | 1 |
| 156(d) | 1 | | | 1 | 1 | | 1 | | 1 | | 1 | | | 1 | | 1 | | | 1 | | | | 1 | 1 |
| 157(d) | 1 | | 1 | | 1 | | 1 | | 1 | | 1 | | | 1 | | 1 | | | 1 | | | | 1 | 1 |

Table 6. Extracted rule for WSD of *must*

| Class of senses | Extracted rules |
|---|---|
| Class *a* (NEC*must*) | a2,a4,a7,a11<br>a2,a4,a5,a9,a13<br>a2,a4,a11,a13,a16<br>a2,a4,a11,a13,a22<br>a2,a4,a11,a15,a19<br>a2,a4,a6,a8,a9,a12,a13,a15,a19<br>a2,a4,a6,a8,a10,a12,a14,a15,a19 |
| Class *b* (POS*must*) | a20<br>a21<br>a3,a6,a8,a11,a15<br>a6,a8,a11,a15,a17<br>a1,a6,a8,a9,a13,a15<br>a2,a6,a8,a9,a14,a15<br>a6,a8,a15,a18 |
| Class *c* (INS*must*) | a2,a6,a10,a12,a15,a22 |
| Class *d* (PER*must*) | a11,a14,a16 |

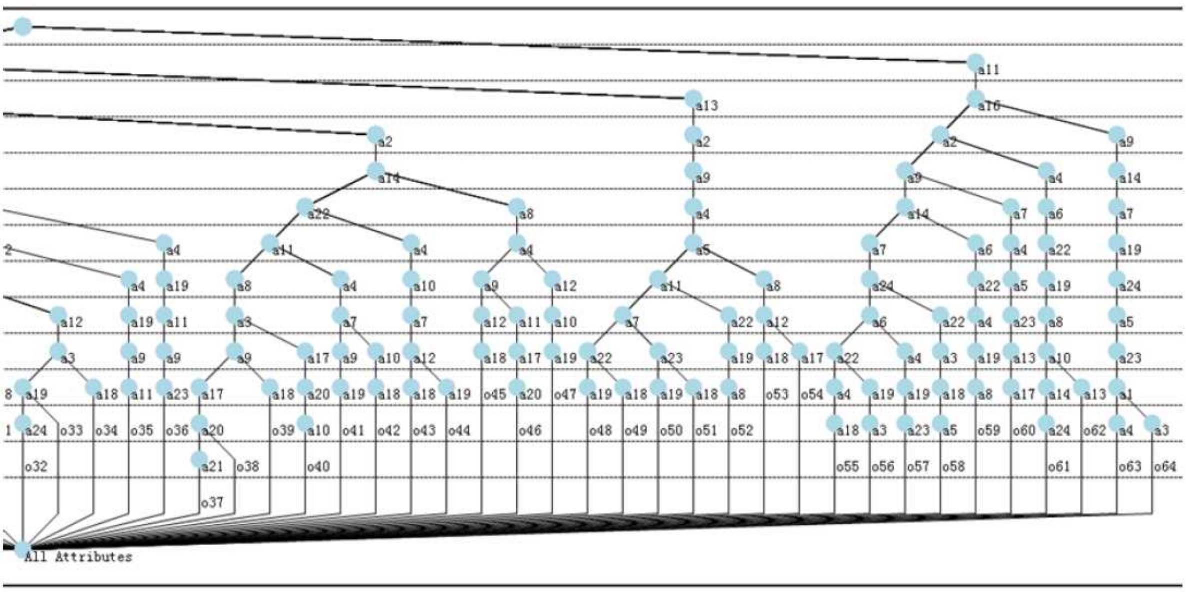
If $S = 1$, then $X_1$' and $X_1$ have the same attribute pattern description and are in the same class. If $S \neq 1$, and $S$ is the maximum of a group of $S_i$, then $X_1$' is most approximate to $X_1$, and they are classified into one class. If $S$ of $X_1$' is the same as $S_i$ of several other objects, then $X_1$' is classified into the class which has the maximum members of objects having co-occurred attributes. If the similarity between $X_1$' and $X_1$ is greater than 2/3, then the two objects are in the same class. In this way, the model is checked and then it is tested by using the testing data set.

The accuracy of WSD of the model reaches 95.54%, which is proven effective and can be used for investigation of restriction of time, aspect and voice on the senses of English modal verb *must*. Based on the clarified model for WSD of *must* in Figure 2, the rules for disambiguating the 4 senses of *must* are extracted, as listed in Table 6.



o1:93(b) o2:59(b) o3:82(b) o4:76(b) o5:87(b) o6:74(b) o7:61(b) o8:68(b) o9:65(b) o10:95(b) o11:52(b)
o12:64(b) o13:98(b) o14:46(a) o15:56(b) o16:83(b) o17:75(b) o18:89(b) o19:28(a) o20:53(b) o21:57(b)
o22:25(a) o23:63(b) o24:72(b) o25:55(b) o26:118(c) o27:41(a) o28:9(a) o29:26(a) o30:123(c) o31:103(c)

(a)




o32:122(c) o33:125(c) o34:104(c) o35:36(a) o36:7(a) o37:54(b) o38:62(b) o39:73(b) o40:92(b) o41:3(a)
o42:33(a) o43:111(c) o44:107(c) o45:67(b) o46:58(b) o47:11(a) o48:6(a) o49:2(a) o50:1(a) o51:22(a)
o52:16(a) o53:5(a) o54:12(a) o55:138(d) o56:141(d) o57:145(d) o58:139(d) o59:140(d) o60:10(a)
o61:143(d) o62:48(a) o63:144(d) o64:142(d)

(b)

FIGURE 2. Clarified APOSD of the training set for WSD of *must*

The extracted rules are tested with the testing data set and the accuracy of WSD reaches 93.63%.

It is noticed that in the extracted rules for the 4 senses, the attributes representing the higher value of MI for each sense occur in the rules for the sense. For instance, a2 and a4 are the higher values of MI(s+NEC*must*) and MI(NEC*must*+v), respectively for the sense of *must* (necessary), a6 and a8 represent the higher value of MI(s+POS*must*) and MI(POS*must*+v), respectively, a10 and a12 are the high values of MI(s+INS*must*) and MI(INS*must*+v), respectively and a14 and a16 represent the high values of MI(s+PER*must*) and MI(PER*must*+v), respectively, and they all occur in the rules for the corresponding sense class. The higher values of MI imply the relevance of *must* and the adjacent words. This suggests that the attributes representing the relevance of *must* and the adjacent words are the main and core attributes in the rules for each sense.

It can be seen from the extracted rule that some rules are combination of semantic attributes, and others are combination of semantic and syntactic attributes. It can also be seen that the rules for a certain class tend to be the combination of the attributes of higher values of MIs of the class and the lower values of MIs of the other classes. This is logically reasonable combination because the senses of a modal verb are sensitive to its adjacent words. In most cases, it co-occurs with some words, but seldom or never co-occurs with other words, which means it is relevant to some words but not the others.

4. **Restrictions of Time, Aspect and Voice on the Senses of *Must*.** The restriction of time, aspect and voice is examined from the following three angles: (1) to examine the hierarchical distribution of the attributes representing time, aspect and voice in the APOSD in order to see the role of the attributes in the formation of the senses of *must*; (2) to examine the influence of time, aspect and voice upon the accuracies of WSD of *must* in order to see the effect of the each attribute on the disambiguation; (3) to examine the features of the attributes of time, aspect and voice in order to find their special characteristics and special functions in the WSD of *must*.

4.1. **Hierarchical distribution of time, aspect and voice in APOSD.** The hierarchical distribution of the attributes in the attribute partial-ordered structure diagram for Figure 2 is shown in Figure 3.

It can be seen from Figure 3 that semantic attributes (a1-a16) mostly occur in the upper layers playing the role of extent, and syntactic attributes (a17-a24) mostly occur in the lower layers as intent. This implies that the semantic attributes are mainly functioning generalization and syntactic attributes, i.e., the time, aspect and voice attributes are mainly functioning classification in the WSD of *must*.

4.2. **Influence of time, aspect and voice upon the accuracies of WSD.** The influence of time, aspect and voice upon the accuracies of WSD is investigated through a series of experiments. The experiments are carried out by deleting one or more attributes from the original formal context and then generate a model for WSD, and then test the accuracy of WSD. The difference between the original accuracy and the new accuracy is believed to reflect the influence of the deleted attribute(s) upon the WSD of *must*. The experimental results are shown in Table 7.

Models 2 and 3 in Table 7 show that both semantic attributes and syntactic attributes can be used independently for WSD of *must*, with semantic attributes having a greater influence on the accuracy of WSD of *must* than syntactic attributes. We can see from models 4-15 that time, aspect and voice attributes do not bring much influence to the accuracies of WSD of *must*, no matter they are used jointly or individually. When only

∅

a15a11

a13a2a16

a2a6a9a19

a6a4a9a14a7

a22a4a8a9a7a19a23a2

a10a11a8a9a5a1a4a24a6

a7a8a9a12a11a18a4a10a19a2a5a22

a12a4a11a8a18a9a17a19a10a7a3a20a23a22a6a2

a18a19a11a3a4a9a10a20a7a22a17a12a5a2a13

a3a4a18a19a17a12a7a10a9a23a22a8a20a21a1a2a6a5a14a13

a24a20a21a17a18a4a3a6a22

a21

Figure 3. Distribution of the attributes

Table 7. Experimental results

| Models | Deleted attributes | Errors | Accuracy (%) |
|---|---|---|---|
| 1 | none | 7 | 95.54 |
| 2 | semantic attributes | 34 | 78.34 |
| 3 | syntactic attributes | 8 | 94.90 |
| 4 | all but time, aspect and voice | 53 | 66.24 |
| 5 | time, aspect and voice | 8 | 94.90 |
| 6 | time | 6 | 96.18 |
| 7 | past time | 7 | 95.54 |
| 8 | present time | 8 | 94.90 |
| 9 | future time | 8 | 94.90 |
| 10 | aspect | 7 | 95.54 |
| 11 | perfect aspect | 7 | 95.54 |
| 12 | progressing aspect | 8 | 94.90 |
| 13 | voice | 7 | 95.54 |
| 14 | active voice | 7 | 95.54 |
| 15 | passive voice | 8 | 94.90 |
| 16 | all but time | 67 | 57.32 |
| 17 | all but past time | 126 | 19.75 |
| 18 | all but present time | 130 | 17.20 |
| 19 | all but future time | 125 | 20.38 |
| 20 | all but aspect | 121 | 22.93 |
| 21 | all but perfect aspect | 126 | 19.75 |
| 22 | all but progressing aspect | 149 | 5.1 |
| 23 | all but voice | 104 | 33.76 |
| 24 | all but active voice | 123 | 21.66 |
| 25 | all but passive voice | 138 | 12.10 |

time, aspect and voice attributes are used for WSD of *must*, the accuracy drops to 66.24%, and when the time, aspect and voice attributes are used individually for WSD of *must*, as seen from Models 16-25, time attributes contribute more to the accuracy of WSD of *must* than the aspect and voice attributes, and voice attributes contribute more than aspect attributes.

4.3. **Features of attributes of time, aspect and voice.** Based on Definitions 2.6-2.8 in Section 2 and the formal context in Table 5, we can discover the following knowledge of the features of some attributes: (1) a17 (past time) only occurs in class $a$ (necessary) and class $b$ (possibility), but not in class $c$ (insistence) and class $d$ (permission). While attribute a18 (present time) and a19 (future time) occur in all the 4 classes. This implies that the senses of *must* are sensitive to the time. On the contrary, time has restrictions on the senses of *must*. The senses of "insistence" and "permission" indicate something not happened yet; therefore, they do not co-occur with past time. Attribute a23 (passive voice) does not occur in class $c$ (insistence) because "insistence" is an active action, it does not co-occur with passive voice; (2) It is observed from the formal context in Table 5 that attributes a20 and a21 occur only in class $b$ (possibility); therefore, they are the exclusive attributes of class $b$ (*must*-possibility) and they can be used as the rules for WSD of *must*.

5. **Conclusions.** In this article, the restriction of time, aspect and voice on the senses of English modal verb *must* is investigated by the approach of attribute partial-ordered structure diagram. Based on the theoretical foundation of formal concept analysis and the approach of attribute partial-ordered structure diagram, the model for WSD of English modal verb *must* is built with the accuracie of 93.63%. Based on the model and the formal context of the training set, the following restrictions of time, aspect and voice on the senses of English modal verb *must* are found: (1) time, aspect and voice attributes are mainly playing the role of intent and functioning classification in the WSD of *must*; (2) time, aspect and voice do not contribute much to the accuracies of WSD of *must*, but they do have some restrictions to the senses of *must*. Past time restricts *must* to the sense of either *must*(necessary) or *must*(possibility). Prograssive and perfect aspects are the exclusive attributes of *must*(possibility) and they restrict *must* to the sense of "possibility". Passive voice tends to co-occur with *must*(necessary) and *must*(permission). The results of this study provide significant evidence for both semantic study and the rule extraction for WSD of English modal verbs. In the further study, we will focus on the influences of some covert linguistic features and contextual features upon the senses of the English modal verbs.

**Acknowledgements.** This work is supported by the National Social Science Foundation under Grant No. 12BYY121 and the Social Sciences Foundation of Hebei Province under Grant No. HB15YY002. It is also partially supported by the National Natural Science Foundation of China under Grant No. 61273019. The authors gratefully acknowledge the supports.

## REFERENCES

[1] J. Ji and M. Liang, A multi-dimensional diachronic study of the modal meanings and syntactic features of *must*, *Shangdong Foreign Language Teaching Journal*, vol.26, no.4, pp.40-46, 2014 (in Chinese).

[2] Y. Liu and L. Peng, On the interactive restriction of cooccurrence of modality and aspect, *Journal of Foreign Language*, vol.33, no.5, pp.41-48, 2010 (in Chinese).

[3] C. Zhang, The meaning of cognitive grammar's view on modality, *Foreign Studies*, vol.2, no.1, pp.7-14, 2014.

[4] Y. Xu, Modality of English tenses, *Journal of Xinyang Normal College*, vol.13, no.1, pp.87-90, 2011.
[5] J. Yu, L. An and J. Fu, Word sense disambiguation of English modal verb *must* by neural network, *ICIC Express Letters*, vol.4, no.1, pp.83-88, 2010.
[6] J. Yu, J. Fu and J. Duan, Syntactic feature based word sense disambiguation of English modal verbs by Naïve Bayesian model, *ICIC Express Letters*, vol.4, no.5(B), pp.1817-1822, 2010.
[7] J. Yu, L. Huang and J. Fu, Word sense disambiguation of English modal verbs by support vector machines, *ICIC Express Letters*, vol.4, no.5, pp.1477-1482, 2010.
[8] 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*, vol.9, no.3, pp.1189-1200, 2013.
[9] J. Yu, C. Li, W. Hong, S. Li and D. Mei, A new approach of rules extraction for word sense disambiguation by features of attributes, *Applied Soft Computing*, vol.27, no.1, pp.411-419, 2015.
[10] W. Hong, S. Li, J. Yu and J. Song, A new approach of structural partial-ordered attribute diagram, *ICIC Express Letters, Part B: Applications*, vol.3, no.4, pp.823-830, 2012.
[11] J. Yu, N. Chen, R. Sun, W. Hong and S. Li, Word sense disambiguation and knowledge discovery of English modal verb *can*, *ICIC Express Letters*, vol.7, no.2, pp.577-582, 2013.
[12] J. Fu, J. Yu and H. Liu, An investigation of influence of different subjective factors to WSD of English modal verb *can*, *ICIC Express Letters, Part B: Applications*, vol.6, no.5, pp.1473-1478, 2015.
[13] 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*, vol.9, no.9, pp.3627-3638, 2013.
[14] J. Song, J. Yu, E. Yan, Z. Zhang and W. Hong, Syndrome differentiation of six meridians for warm disease based on structural partial-ordered attribute diagram, *ICIC Express Letters*, vol.7, no.3(B), pp.947-952, 2013.
[15] H. Li and J. Yu, Attribute significance analysis of English modal verb *shall* in word sense disambiguation, *ICIC Express Letters, Part B: Applications*, vol.6, no.5, pp.1287-1294, 2015.
[16] B. Ganter and R. Wille, *Formal Concept Analysis – Mathematical Foundations*, Springer-Verlag, 1996.
[17] W. Hong, J. Mao, J. Yu and J. Song, The complete definitions of attributes and abstract description of attribute features of the formal context, *ICIC Express Letters*, vol.7, no.3(B), pp.997-1003, 2013.
[18] *Merriam-Webster's Advanced Learner's English Dictionary*, 1st Edition, Beijing, Encyclopedia of China Publishing House, 2010 (in Chinese).
[19] *Cambridge Advanced Learner's Dictionary*, 1st Edition, Foreign Language Teaching & Research Press, Beijing, 2008 (in Chinese).
[20] *Longman Dictionary of Contemporary English*, 1st Edition, Foreign Language Teaching & Research Press, Beijing, 2004 (in Chinese).
[21] D. Thompson, *The Concise Oxford English Dictionary*, 1st Edition, Foreign Language Teaching & Research Press, Beijing, 2003 (in Chinese).
[22] C. Zong, *Statistical Natural Language Processing*, Tsinghua University Press, Beijing, 2008.
[23] Z. Bian and X. Zhang, *Pattern Recognition*, Tsinghua University Press, Beijing, 2000.