LANGUAGE AND LINGUISTICS 9.2:395-413, 2008
2008-0-009-002-000256-1

Knowledge Representation for Comparative Constructions in Extended-HowNet*

Shu-Ling Huang, Yueh-Yin Shih, and Keh-Jiann Chen
Academia Sinica

In this paper, we propose an approach for studying the semantic representations of comparison words and comparative constructions based on the framework of E-HowNet. Our objective is to determine how the semantic composition mechanism works. The proposed method establishes a mapping between grammatical structures and fine-grained event structures for comparative constructions. We encode the event structures of comparative constructions as part of the representations of comparison words. The fine-grained semantic roles are adopted from FrameNet. A semantic composition mechanism is then developed to unify word sense representations under syntactic constraints. Our ultimate goal is to achieve fully automatic semantic composition.

Key words: knowledge representation, Extended-HowNet, comparative constructions

1. Introduction

In recent years, lexical knowledge representation has become a major research topic in the natural language processing field, because it helps bridge the gap between string processing and conceptual processing. In this paper, we try to represent the senses of comparison words and comparative constructions under the framework of E-HowNet (K.-J. Chen et al. 2005) and determine how the semantic composition mechanism works. E-HowNet, a frame-based entity-relation knowledge representation model evolved from HowNet, is an on-line lexical knowledge base. It extends the word sense definition mechanism of HowNet and uses WordNet (Fellbaum 1998) synsets as vocabulary to describe concepts. In principle, the qualia structure is the major feature of a nominal-type concept (Pustejovsky 1995) and event frames are for eventive concepts (Fillmore, *FrameNet*). For example, the concept of "vintage" is defined as:

* This research was supported in part by the National Science Council under a Center Excellence Grant NSC 94-2752-E-001-001-PAE and Grant NSC 94-2213-E-001-006.

(1) vintage|佳釀
def: {drinks|飲品:
qualification={nice|良好},
telic={addict|嗜好: patient={~}}}.

In the E-HowNet framework, concepts are linked by conventional taxonomic relations, such as synonymy, hypernymy, antonymy, and meronymy, as well as by their attribute features. In the above example, "drinks" is a hypernym of "vintage"; "qualification" is a feature of qualia structures that links the attribute "nice" to the concept "drinks"; and "telic" denotes the purpose and function of "vintage", which connects the event "addict" to "drinks".

In real implementations, we have found that the senses of content words and function words are different; therefore, they should have different representations. Since function words usually mark certain grammatical or semantic functions, they denote relational senses and have few or no content senses. Content words, on the other hand, have implicit or no relational senses. In E-HowNet, the senses of function words are represented by semantic roles/relations (Y.-J. Chen et al. 2005). For example, "because" is defined as shown in (2), which links two events x and y:

(2) because|因爲
def: reason(x)={y}.

Representations of comparison words are even more complicated. Take the comparison word "bi 比" as an example. Its sense is related to a comparative event frame, which describes many conceptual roles involving comparative relations. Some conceptual roles are omitted in surface sentences. Because of the uniqueness and complexity of comparative constructions, representations of the senses of comparison words can be considered as a key indicator of a well defined knowledge representation system. We consider comparative event frames in detail in §2.

In general, a sentence with comparative sense usually contains comparison words, and is therefore called a *comparative sentence*. The sense of a comparative sentence denotes an event frame that compares two items and expresses which of the two is more precise in terms of the attributes and by how much. These core elements comprise a comparative construction (CC) (Li & Thompson 1981). Although sentence patterns are different, almost every language uses the same event roles, i.e. comparative items, attributes, and variances, to express comparative event structures. The Berkeley FrameNet project integrated event roles into a framework called *Evaluative_comparison* to describe a comparison state. FrameNet, which is a semantic approach, forms the basis

of our system of knowledge representation for comparison words.

The remainder of the paper is structured as follows. In §2, we introduce FrameNet's analysis as the background of comparative construction. Section 3 describes the method of applying comparison frames to interpret comparative sentences. Formal representation of different comparison words and sentences is considered in §4. Then, in §5, we summarize our work and present our conclusions.

2. Background

Most research on comparative constructions (CC) focuses on their syntactic structure and the syntactic category for comparison words. There are two important questions: (1) Are CC clauses or complements that are linked by comparison words? (2) How can the omitted roles of CC be recovered? (Li & Thompson 1981, Hong 1991, Hsing 2003). However, most approaches discuss issues from a syntactic point of view and focus less on semantic representation. The Berkeley FrameNet project adopted a semantic approach to analyze CC. It lists and describes ten frame elements (i.e. semantic roles) in a complex frame structure called *Evaluative comparison*. The core frame elements are: (a) Attributes, which mark constituents that indicate how Item-1 and Item-2 contrast with each other; (b) Profiled_Attribute, which is a value on a scale that is compared with a Standard_Item or Standard_Attribute; (c) Profiled_Item, which is the grammatically more prominent item among the items compared and is taken as the subject; (d) Standard_Attribute, which is used when the standard is a specific value on a scale; (e) Standard_Item, which is the grammatically less prominent element, i.e. the object or an oblique reference to an object. There are also five non-core elements: (f) Comparison_set, which includes the Standard_Item, is used to evaluate the Profiled_Item; (g) Degree, which indicates how close the Profiled_Item and the Standard_Item are to each other on the scale evoked by the Attribute; (h) Manner. (i) Place, which is the location where the Profiled_Item rivals the Standard_Item; (j) Time, is the time at which the Profiled_Item rivals the Standard_Item. Based on the FrameNet representation cited above, we can analyze sentences (3), (4) as follows:

(3) American caviar now rivals Russia's. (in quality)

Profiled_Item	American caviar
Standard_Item	Russian caviar
(implied) Attribute	quality
Time	now

(4) The current price matches the price last year.
Profiled_Attribute the current price
Standard_Attribute the price last year

No matter how different the syntax and sentence patterns are, almost every language shares the above-mentioned elements.

3. Methodology

Having defined the schema for the comparison framework, we can use the framework to determine the meaning (i.e. construction meaning) of a comparative sentence. To do this, we first derive information about parts-of-speech, grammatical structures and course-grained semantic roles from the sentence through an automatic segmenting and parsing process as exemplified in (5), (6). Then a mapping from course-grained semantic structure to fine-grained semantic frame of FrameNet has to be established.

(5) Theme[Np,S]+Contrast[Pp[比]]+Head[V]+Quantity[DM]
(6) E.g. 張三 比 李四 高 三公分
Zhangsan bi Lisi gao san gong fen
'Zhangsan is three centimeters taller than Lisi.'

Here, the Theme should be mapped to the Profiled_Item, the fine-grained semantic role of FrameNet; the Contrast is the Standard_Item; the Head verb is the Attribute_Value, and the Quantity is the Degree.

Therefore we devise a mapping table to connect the grammatical functions and fine-grained semantic roles. The table acts as a bridge that transfers the parsing result to the fine-grained semantic roles of the comparison frame to help the machine find the corresponding constituents in the comparative sentences. The mapping table between the fine-grained semantic roles and the argument structure is shown in Table 1.

Table 1: Mapping table for the fine-grained semantic roles

Fine-grained Semantic Roles	**Thematic Roles**	**Grammatical Functions**
Profiled_Item + (Profiled_Attribute)	Theme; Experiencer	Subject
Standard_Item + (Standard_Attribute) Comparison_set	Contrast	Object[PP[bi]]
Attribute_Value	Head	Verb
Degree	Quantity; Degree	Complement
Manner	Manner	Adjunct (Manner)
Place	Location	Adjunct (Location)
Time	Time	Adjunct (Time)

These semantic roles are linked in the taxonomy of semantic roles for events in E-HowNet (see Figure 1). We discuss the operation of the mechanism in the following sections.

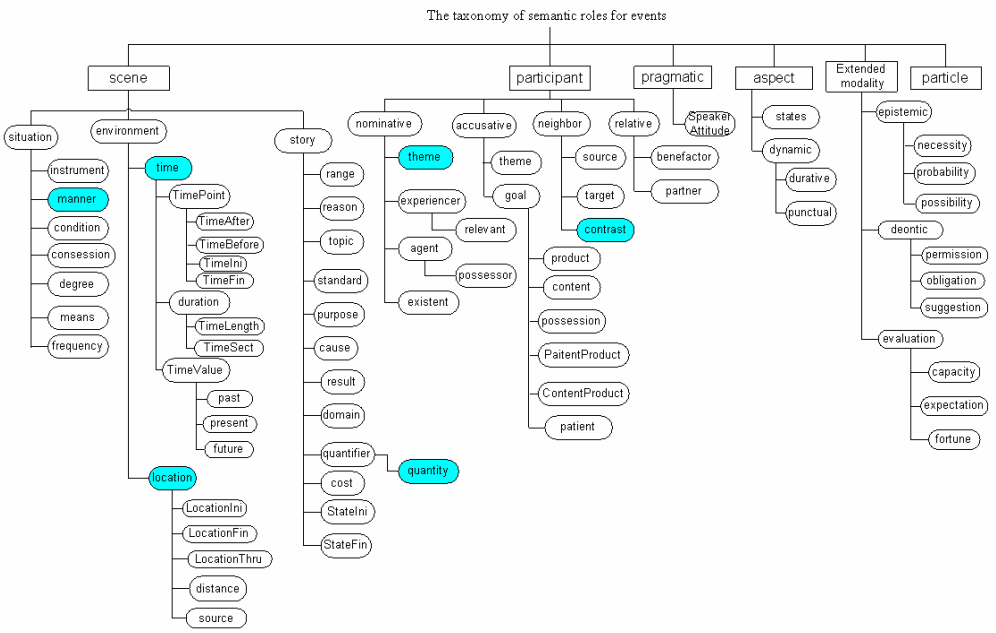

Figure 1: The taxonomy of semantic roles for events

3.1 Filling semantic gaps

Some semantic roles are frequently omitted from surface sentences. Because of the semantic parallels between the Profiled_Attribute and the Standard_Attribute, the recovery of omitted semantic roles follows the default rules below.

Default Rules:

(a) The Attribute is implied by the sentential head of a comparative construction.
For instance, in (6), the attribute role is omitted because it is implied by the head adjective "tall|高". The ontology of E-HowNet[1] shows that "tall|高" is the value of the attribute of "height|身高". Thus, the omitted role of the Attribute can be inferred.

(b) If the Profiled_Attribute is not mentioned in the surface sentence, the inferred Attribute will generate default values for both the Profiled_Attribute and the Standard_Attribute.

(c) The Standard_Attribute is equal to the Profiled_Attribute, unless otherwise specified.
For instance, in example (10), the Standard_Attribute is omitted. By default, the Standard_Attribute is equal to the Profiled_Attribute, i.e. "fly|飛".

(d) If no Standard_Item is mentioned, the average standard is taken as the Standard_Item.

4. Knowledge representation for comparative constructions and comparison words

Although there is only one case frame (argument structure), it can be realized by several different comparative constructions whose surface structures vary. We classify comparative constructions into three groups. First, we discuss the comparative constructions with keywords of "bi 比" and other comparison words, such as "more 較", "inferior 不如", etc. Second, we note the degree adverbs, which imply the comparative sense. Then, we consider the concepts with implicit comparative relations.

4.1 Sense representation for "bi 比"

Unlike other comparison words such as "taller 高過", "bigger 大於", etc., "bi 比" does not convey the meaning of the comparative attribute. However it is one of the major keywords that indicate comparative constructions. We have to represent the sense of "bi" by a comparative event frame (argument structure), which was discussed earlier. Accordingly, "bi" can be defined as shown in (7):

[1] The knowledge about attribute-values is taken from the original HowNet.

(7) def: contrast={} in the course-grained event frame of {AttributeValue: theme={}, contrast={}, quantity (or degree)={}, manner={}, location={}, time={}}

The above definition says that the word "bi 比" has the relational sense of "contrast" and the prepositional phrase of "bi" is playing the semantic role of contrast in a comparative frame.

In practice, the constituents of the grammatical roles in the following sentences are assigned semantic roles according to their syntactic structures.

(8) 我 比 他 高 一個 頭
wo bi ta gao yige tou
'I am taller than him by a head.'
Surface structure: theme[NP]+contrast[PP[比]]+Head[V]+quantity
Parsing result: {tall|高:
theme={I|我},
contrast={he|他},
quantity={one head|一個頭}}.

(9) 我 數學 比 你 多 五分
wo shuxue bi ni duo wufen
'My math is five points more than yours.'
Surface structure: theme[NP]+contrast[PP[比]]+Head[V]+quantity
Parsing result:{more|多:
theme={my math|我數學},
contrast={you|你},
quantity={five points|五分}}.

In fact, both sentences (8) and (9) contain ellipses and omit the senses. Since the restoration process is based on the Theory of Coercion, it changes the compositional semantics, but leaves the syntactic type unaltered. A key procedure is the selection of the right syntactic function and mapping it to the correct semantic role. In the following discussion, every sentence (x) is assigned a fine-grained semantic role and restored to its complete event structure (x') according to the mapping Table 1. Both sentences (8) and (9) contain ellipses and sense omissions. The full interpretation of the sentences is shown in (8′) and (9′).

(8′) 我 的 身高 比 他 的 身高 高 一個 頭
wo de shengao bi ta de shengao gao yige tou
'My height is one head taller than his height.'
def:{tall|高:
Profiled_Item={I|我},
Profiled_Attribute={height|身高},
Standard_Item={he|他},
Standard_Attribute={height|身高},
Degree={one head|一個頭}}.

(9′) 我 的 數學 分數 比 你 的 數學 分數 多 五分
wo de shuxue fenshu bi ni de shuxue fenshu duo wufen
'My score in math is five points more than your score in math.'
def:{more|多:
Profiled_Item={I|我},
Profiled_Attribute={score of math|數學分數},
Standard_Item={you|你},
Standard_Attribute={score of math|數學分數},
Degree={five points|五分}}.

In example (8), although the Profiled_Attribute "height|身高" is not explicitly expressed in the surface sentence, the mapping Table 1 and the representation of "bi" (7) indicate that the role of Profiled_Attribute is missing. Furthermore the head adjective "tall|高" implies that the Profiled_Attribute is "height|身高", since "tall|高" indicates an attribute value of the attribute of "height|身高" described in HowNet. In contrast, in (9) the subject "my math|我數學" is interpreted as the Profiled_Item + the Profiled_Attribute. Therefore the Profiled_Item is "I|我" and the Profiled_Attribute is "score in math|數學分數". However, the word "score" is missing in sentence (9), and we cannot rely on the head adjective "more|多" to infer the attribute "score". Instead, the quantity "five points|五分" is the clue to inferring that the Profiled_Attribute is the "score in math", since it is a value for the attribute "score". Therefore it fills the gap that the Profiled_Attribute is the "score in math|數學分數".

Instead of being entities, the structure of the Profiled_Item + the Profiled_Attribute can be events, as shown by the following sentence:

(10) 我 飛 得 比 獵鷹 慢
wo fei de bi lieying man
'I fly slower than an eagle.'

In example (10), the surface form of the comparative items is represented by events, not entities. In this case, the restoration process has to recover the complete Standard_Attribute before deriving the full representation of the sentence. By default the attribute of the Standard_Attribute is equal to the attribute of the Profiled_Attribute.

Surface structure: theme[S]+contrast[PP[比]]+Head[V]
Parsing results: {slow|慢:
theme={fly|飛: agent={I|我}},
contrast={eagle|獵鷹}}.
Restore ellipsis: {slow|慢:
theme={fly|飛: agent={I|我}},
contrast={fly|飛: agent={eagle|獵鷹}}}.

In the above representation, instead of adjusting the frame structure for "bi," we add another layer under the semantic roles "theme" and "contrast" to show the event structure "I fly" and "eagle flies." After recovering the attribute "speed" based on its value "slow", the full interpretation of (10) becomes:

(10′) 我 飛 的 速度 比 獵鷹 飛 的 速度 慢
wo fei de sudu bi lieying fei de sudu man
'My speed of flying is slower than an eagle's speed of flying.'
def:{slow|慢:
Profiled_Attribute={fly|飛:
Profile_Item={I|我}},
Standard_Attribute={fly|飛:
Standard_Item={eagle|獵鷹}},
Attribute={speed|速度}}.

Let us take another sentence as an example:

(11) 他 比 我 還 喜歡 你
ta bi wo hai xihuan ni
'He likes you more than I.'
Surface structure: theme[NP]+contrast[PP[比]]+degree[還,更]+Head[V]
+goal[NP]

The semantic-role mapping is:

(11′) 他 喜歡 你 的 程度 比 我 喜歡 你 的 程度 還 多
ta xihuan ni de chengdu bi wo xihuan ni de chengdu hai duo
'The degree that he likes you is more than the degree that I like you.'
def:{more|多:
Profiled_Attribute={FondOf|喜歡:
agent=Profile_Item={he|他},
target={you|你}},
Strandard_Attribute={FondOf|喜歡:
agent=Standard_Item={I|我},
target={you|你}},
Attribute ={degree|程度}}.

In this example, the mapping table and the restoration process are more complicated. We add a comparative adjective "more|多" to the sentence as its head. However, this does not add extra information to the original sentence. In fact, the word "hai 還" implies the meaning of "more". And with the event "FondOf", we know the implied profiled attribute is "degree". Although, base on the surface structure of the sentence, the original head is "FondOf", it apparently does not fit the CC requirement that it should be an attribute value. Like "fly" in example (10), we interpret the event structure of (11) in the theme (Profiled_Item+Profile_Attribute) and contrast (Standard_Item+Standard_Attribute) while doing the representation.

In addition, since the negation of "bi" is "bu ru 不如" and "mei you 沒有", we only need to add a logical operator *.NotSo.* in front of the head when we define it, as shown in (12) and (13):

(12) 那 本 書 沒有 這 本 書 好
na ben shu meiyou zhe ben shu hao
'That book is not better than this book.'

(12′) 那 本 書 的 內容 沒有 這 本 書 的 內容 好
na ben shu de neirong meiyou zhe ben shu de neirong hao
'That book's content is not better than this book's content.'
def:{.NotSo.nice|良好:
Profiled_Item={that book|那本書},
Standard_Item={this book|這本書},
Attribute={content|內容}}.

(13) 張 先生 不如 我 跑 得 快
Zhang xiansheng buru wo pao de kuai
'Mr. Chang doesn't run faster than I.'

(13′) 張 先生 跑 的 速度 不如 我 跑 的 速度 快
Zhang xiansheng pao de sudu buru wo pao de sudu kuai
'Mr. Chang's speed of running isn't faster than my speed of running.'
def:{.NotSo.fast|快:
Profiled_Attribute={run|跑:
agent=Profile_Item={Mr.Chang|張先生}},
Standard_Atribute={run|跑:
agent=Standard_Item={I|我}},
Attribute={speed|速度}}.

In example (12), "nice" indicates more than one attribute value of the attributes described in HowNet, so the Profiled_Attribute can be "content", "price", "design", etc. The representation thus depends on the qualia structure of the Profiled_Item "book" to derive the "content" as the Profiled_Attribute.

The Sinica TreeBank contains some difficult examples that share exactly the same argument structure as the basic pattern in (5) without exception. We try to represent them as follows:

(14) 枝葉 比 以前 更 茂盛 了
zhiye bi yiqian geng maosheng le
'Branches and leaves are more abundant than before.'
Parsing result: {exuberant|茂:
theme={branches and leaves|枝葉},
contrast={past|過去},
degree={more|較}}.

(14′) 現在 的 枝葉 比 以前 的 枝葉 更 茂盛 了
xianzai de zhiye bi yiqian de zhiye geng maosheng le
'The branches and leaves now are more abundant than the branches and leaves before.'
def:{exuberant|茂:
Profiled_Item={branches and leaves|枝葉:time={present|當下}},
Standard_Item={branches and leaves|枝葉:time={past|過去}},
Degree={more|較}}.

(15)(15′) 寫信 來 訴苦 要 比 當面 罵 他 來 得 安全 有效
xiexin lai suku yao bi dangmian ma ta lai de anquan youxiao
'Complaining about someone in writing is safer and more effective than scolding.'
def: {safe and effective|安全有效:
Profiled_Item={complain|訴苦:
means={write a letter|寫信}},
Standard_Item={ExpressAgainst|譴責:
target={he|他},
manner={overt|公開}}}.

In (14), based on the attribute value "exuberant", we decide that the Profiled_item is "branches and leaves", but not "time". We put the time element in the same layer to describe the Profiled_Item. Because the default time is "now", we then restore it into the sentence. In (15), the difficulty is not restoring the omitted part, but determining whether the Profiled_Item and the Standard_Item are an asymmetric event structure.

Achieving fully automatic semantic composition for the above constructions requires further investigation.

4.2 Knowledge representation for other comparison words

Apart from "bi", there are also some degree adverbs in Mandarin that are frequently used to form comparative constructions, such as "more 更", "too 太", "very 很", etc. Following Lu & Ma's analysis (1999), we classify such adverbs into six types, as shown in Table 2.

Table 2: Degree adverbs that imply comparative constructions

Degree	Words	With/without Standard_Item	Collocated with "bi 比"
High	Much: 很, 挺, 十分, 萬分, 非常, 異常, 太, 極, 極端	–	–
	Most: 最, 最爲, 頂	+	–
	More: 更, 更加, 更爲, 越發, 越加, 愈加, 還$_1$		+
Low	Little: 有點兒, 有些	+	+
	Least: 比較, 較, 較爲, 還$_2$		+
	Less: 稍微, 稍, 稍稍, 多少, 略微, 略略	+	

The adverbs in Table 2 imply a comparative construction in which the Standard_Item is the average standard. Lu & Ma (1999) defined the sentence patterns of the above adverbs as follows:

(16) 相比之下 (比較起來), Profiled_Items+Degree+AttributeValue
Relatively, Profiled_Items+Degree+attribute value

(17) 跟 Standard_Item 相比 (比起 Standard_Item 來), Profiled_Items+Degree+AttributeValue
To compare with Standard_Item, Profiled_Items+Degree+AttributeValue

(18) 在 Comparison_Set 中/上, Profiled_Items+Degree+AttributeValue
Among Comparison_Set, Profiled_Items+Degree+AttributeValue

(19) Profiled_Items+bi+Standard_Item+Degree+AttributeValue

Here, "degree" indicates the degree adverbs listed in Table 2. Among these adverbs, only "much" and "most" never collocate with "bi", but they still imply a comparative construction in which the Standard_Item is the average standard. For instance, we often use the terms "most beautiful" or "very beautiful" to describe a target without mentioning its Standard_Item. This does not mean that the latter does not exist, but—by implication—it is an ordinary looking item. Thus, Lu & Ma add a "bi phrase" in front of each sentence pattern listed above to make the comparative construction clear. Let us take some sentences as examples:

(20) 他 最 討厭 我。
ta zui taoyan wo
'He hates me the most.'

(21) 這件 褲子 較爲 大 些。
zhejian kuzi jiaowei da xie
'This pants is a bit bigger.'

(22) 冬天 時 景色 越發 好。
dongtian shi jingse yuefa hao
'The scenery is better in winter.'

The above comparative sentences use the degree adverbs "most 最", "more 較爲", and "more 越發" instead of "bi 比". However, according to Lu & Ma's analysis, the sentences omit the comparison words shown in (16)-(19). Therefore, we need to restore the omitted comparison phrases in (16)-(19) as well as the comparative frame structure

of the sentences. The comparison phrases in patterns (16)-(19) are similar to “bi”, and can be described as shown in (7). Then, after ellipse restoring, sentences (20)-(22) can be defined as follows:

(20′) 相比之下， 他 最 討厭 我。
xiangbizhixia, ta zui taoyan wo
‘By comparison, he hates me the most.’
def: {more|多:
Profiled_Attribute={disgust|厭惡:
experiencer={he|他},
target={I|我}},
Standard_Attribute={disgust|厭惡:
experiencer={he|他},
target={other person|其他人}},
Degree={most|最},
Attribute={degree|程度}}.

(21′) 在 這些 褲子 中， 這件 褲子 較爲 大 些。
zai zhexie…kuzi zhong, zhejian kuzi jiaowei da xie
‘Among these pants, this pants is a bit bigger.’
def: {big|大:
Profiled_Item={this pants|這件褲子},
Standard_Item={these pants|這些褲子},
Degree={ish|稍},
Attribute={size|尺寸}}.

(22′) 跟 夏天 相比， 冬天 時 景色 越發 好。
gen xiatian xiangbi, dongtian shi jingse yuefa hao
‘Compared with summer, the scenery is better in winter.’
def: {nice|好:
Profiled_Item={winter|冬天},
Standard_Item={summer|夏天},
Attribute={scenery|景色},
Degree={more|較}}.

In this paper, degree adverbs are classified as a type of comparison word because they imply the comparative sense. We think some degree adverbs indicate both the degree and the comparative sense. Based on the comparative event frame, we define these adverbs so that they cover both senses as follows:

(23) def: degree={} in {AttributeValue: theme={}, contrast={}, quantity={}, manner={}, location={}, time={}}

Therefore, it does not matter whether we restore sentences (20)-(22) because they have the same definitions.

"As…as 一樣" and "very…, more…很…更…" are not in Lu & Ma's table, but they can also be analyzed as degree adverbs. Consider the following examples:

(24) 後者 和 前者 一樣 重要。
houzhe he qianzhe yiyang zhongyao
'The latter is as important as the former.'

(24′) 比較起來， 後者 和 前者 一樣 重要。
bijiaoqilai, houzhe he qianzhe yiyang zhongyao
'Relatively, the latter is as important as the former.'
def:{important|重要:
Profiled_Item={the former|前者},
Standard_Item={the latter|後者},
Degree={BeSame|一樣}}.

(25) 北京 的 景山 很 高，我 老家 的 山 更 高。
Beijing de jingshan hen gao, wo laojia de shan geng gao
'Mt. Jing near Peking is very tall, but the mountain near my home town is even taller.'

(25′) 比起 北京 的 景山 來，我 老家 的 山 更 高。
biqi Beijing de jingshan lai, wo laojia de shan geng gao
'Compared with Mt. Jing near Peking, the mountain in my home town is even taller.'
def:{tall|高:
Profiled_Item={mountain|山:
place={HomeTown|老家: owner={I|我}}},
Standard_Item={Mt. Jing|景山:
place={Peking|北京}},
Degree={more|較}}.

4.3 Knowledge representation for concepts with implicit comparative senses

There are a few concepts, especially those that describe human relationships,

which imply a comparative sense. Thus, their representations also contain comparative features. For example, the concept "brother" can be defined as follows:

(26) brother|兄弟
def:{human|人:
gender={male|男},
qualification={BeSame|相同:
theme={parents(X)},
contrast={parents(~)}}}.

To express "a man who has the same parents as another man", we adopt "human|人" as the head, and use the feature "qualification" to connect "human" with the comparative content. The content begins with the comparative adjective "BeSame", and is followed by the theme, i.e. X's parents, and the theme's contrast, the head's (head is denoted by "~") parents. Furthermore, we can infer that the head is older than X by the following comparative structure:

(27) elder brother|哥哥
def:{brother|兄弟:
age={aged|老年:
theme={X},
contrast={~}}}
Expand def: {human|人: gender={male|男},
qualification={BeSame|相同:
theme={parents(X)},
contrast={parents(~)}},
age={aged|老年:
theme={X},
contrast={~}}}.

The following phrase is another example:

(28) father's young friend|世叔
def: {human|人=friend(father(X)):
age={young|青年:
theme={father(X)},
contrast={~}}}.

5. Summary and conclusion

To build a well defined knowledge representation system, we propose a case frame representation mechanism for Mandarin comparison words, which we define as follows:

def: contrast={} (in {AttributeValue: theme={}, contrast={}, quantity={}, manner={}, location={}, time={}})
def: degree={} (in {AttributeValue: theme={}, contrast={}, quantity={}, manner={}, location={}, time={}})

The frame elements of comparative constructions (i.e. Profiled_Item, Standard_Item, Attribute, Profiled_Attribute, Standard_Attribute, Comparison_Set, Degree, Manner, Place, and Time) are all included in the representations. By building a mapping table, we can connect the parsing result of sentences with the case frame so that a computer can understand the meaning of comparative sentences. Our goal is to establish an automatic composition mechanism that can combine every word sense representation with the correct features of a comparative construction. To achieve this, coercion and gap filling processes are an integral part of the mechanism. They will be addressed more in our future research.

References

Chao, Yuen Ren. 1968. *A Grammar of Spoken Chinese*. Berkeley: University of California Press.

Chen, Keh-Jiann, Shu-Ling Huang, Yueh-Yin Shih, and Yi-Jun Chen. 2005. Extended-HowNet: a representational framework for concepts. *Proceedings of OntoLex 2005—Ontologies and Lexical Resources*, 1-6.

Chen, Yi-Jun, Shu-Ling Huang, Yueh-Yin Shih, and Keh-Jiann Chen. 2005. Semantic representation and definitions for function words in Extended-HowNet. Paper presented at the Workshop on Chinese Lexical Semantics. Xiamen: Xiamen University. (In Chinese)

Fellbaum, Christiane. 1998. *WordNet: An Electronic Lexical Database*. Cambridge: MIT Press. (http://wordnet.princeton.edu/)

Hong, Wei-mei. 1991. *Comparative Structure in Mandarin Chinese.* Hsinchu: National Tsing Hua University MA thesis.

Hsing, Jen-chieh. 2003. *On Comparative Sentences with Bi-marker in Mandarin Chinese*. Hsinchu: National Tsing Hua University MA thesis.

Li, Charles N., and Sandra A. Thompson. 1981. *Mandarin Chinese: A Functional Reference Grammar*. Berkeley: University of California Press.

Liao, Wei-wen. 2005. The comparative construction and the *wh*-movement. *UST Working Papers in Linguistics* (*USTWPL*), Vol. 1, ed. by Yin-Ling Christina Chen, Han-Chun Huang & Barry C.-Y. Yang, 187-200. Hsinchu: Graduate Institute of Linguistics, National Tsing Hua University.

Lu, Jianming, and Zhen Ma. 1999. *Xiandai Hanyu Xuci Sanlun*. Beijing: Language & Culture Press.

Pustejovsky, James. 1995. *The Generative Lexicon*. Cambridge: MIT Press.

Shih, Yueh-Yin, Shu-Ling Huang, Yi-Jun Chen, and Keh-Jiann Chen. 2005. Semantic representation and composition for spatial concepts in Extended-HowNet. *Proceeding of NLPKE 2005*, 735-740.

Website Resources

HowNet: http://www.keenage.com/
FrameNet: http://www.icsi.berkeley.edu/~framenet/
Sinica TreeBank: http://treebank.sinica.edu.tw/

[Received 28 December 2006; revised 19 September 2007; accepted 1 November 2007]

Shu-Ling Huang
Institute of Information Science
Academia Sinica
130, Sec. 2, Academia Road
Taipei 115, Taiwan
josieh@hp.iis.sinica.edu.tw

廣義知網中比較結構的語意表達

黃淑齡　　施悅音　　陳克健
中央研究院

廣義知網是以「個體–關係」爲框架的知識表達模型，我們希望藉著這個表達模型，使詞彙與詞彙間的自動語意合成機制得以建立，從而幫助計算機由了解詞彙進一步發展到了解文意。爲達成此目標，本文以比較詞與比較結構爲例，討論比較詞的分類、定義以及比較句的合成等等問題。比較詞本身並沒有核心語意，而須在一個結構中才能顯現它的意義，因此我們參考框架網 (FrameNet) 所提出的比較事件框架，以其事件框架元素 (frame element) 作爲定義比較詞的特徵。然後提出一個對應表，使比較句成分能一一對應到比較詞的語意特徵中，亦即使句子的表面結構轉換爲一種深層的語意結構，達成計算機理解語言的目標。

關鍵詞：知識表達，廣義知網，比較結構