

# Instructions for EMNLP 2017 Proceedings

## Anonymous EMNLP submission

## 1 Related Work

### 1.1 Linggle

## 2 Experiments

### 2.1 Feature Collection

To indicate whether an adjective is a central property of a noun, we apply the following features.

#### 2.1.1 Frequency and Collocation

#### 2.1.2 PMI

Pairwise mutual information (or point mutual information), known as PMI (Church and Hanks, 1990; Bouma, 2009), is a widely accepted measure that illustrates the association between two words. The PMI between two words  $x$  and  $y$  can be computed by

$$PMI(x, y) = \log_2 \frac{P(x, y)}{P(x)P(y)} \quad (1)$$

where  $P(x)$  is the probability of word  $x$  to appear,  $P(x, y)$  is the joint probability of  $x$  and  $y$ , namely the probability of the collocation of  $x$  and  $y$ . Based on the statistics, it can be estimated by the frequency of word  $x$  and  $y$ , with the collocation frequency of them. In application, it is usually computed by

$$\begin{aligned} PMI(x, y) &= \log_2 \frac{C(x, y) \cdot N}{C(x)C(y)} \\ &= \log_2 C(x, y) - \log_2 C(x) \\ &\quad - \log_2 C(y) + \log_2 N \end{aligned} \quad (2)$$

where  $N$  is the number of unigrams in total,  $C(x)$  and  $C(y)$  represent the frequency of unigram  $x$  and  $y$ , and  $C(x, y)$  represents the frequency of bigram  $(x, y)$ .

#### 2.1.3 "As ... as" pattern

There is a consensus that the sentence *Something is as A as a/an N* is reasonable if  $A$  is a central

property of  $N$ (*ano*). Thus, the frequency of cooccurrence of an adjective and a noun pair in the pattern of "as adj. as a/an noun." would be a strong feature to detect central properties of nouns. We used Linggle API (Boisson et al., 2013)<sup>1</sup> to extract such patterns, with the queries like "as strong as a/an n." and "as adj. as a/an tiger", and it would return 50 most frequent matched phrases.

#### 2.1.4 Concreteness of words

Turney et al. (2011) proposed a method to compute the concreteness and abstractness of words, and suggested that concreteness of words are strong indicators of metaphors. We replicated the experiments for concreteness computation with the method on a large web scale corpus, and got a list with 75,626 words' concreteness. We also treat concreteness as a feature for central attributes extraction.

## 2.2 Annotation

## 2.3 Future Work

- The concept of central property should not be limited in only adjectives. For example, "have sharp teeth" should be a central property of tigers or lions, so that something could be compared to them in order to indicate such a property.

## References

???? *ano*.

Joanne Boisson, Ting-Hui Kao, Jian-Cheng Wu, Tzu-Hsi Yen, and Jason S Chang. 2013. Linggle: a web-scale linguistic search engine for words in context. In *ACL (Conference System Demonstrations)*. pages 139–144.

<sup>1</sup>add the queries after the url <http://linggle.com/query/>

- Gerlof Bouma. 2009. Normalized (pointwise) mutual information in collocation extraction. In *Proceedings of the Biennial GSCL Conference*. volume 156.
- Kenneth Ward Church and Patrick Hanks. 1990. Word association norms, mutual information, and lexicography. *Computational linguistics* 16(1):22–29.
- Peter D Turney, Yair Neuman, Dan Assaf, and Yohai Cohen. 2011. Literal and metaphorical sense identification through concrete and abstract context. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics, pages 680–690.

150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199