Literature Review: *Metaphor detection and interpretation in NLP*

by

田展铭

1730026102

Supervisor: *Weifeng Su*

BNU-HKBU

UNITED INTERNATIONAL COLLEGE

*May*, 2020

Contents

[Abstract 3](#_Toc41848117)

[1 Introduction 3](#_Toc41848118)

[1.1 Background 3](#_Toc41848119)

[1.2 Basic linguistic theory in metaphor 3](#_Toc41848120)

[1.3 Problems for solving 4](#_Toc41848121)

[2 Related work 5](#_Toc41848122)

[2.1 Metaphor detection 5](#_Toc41848123)

[2.1.1 Word level based metaphor detection 5](#_Toc41848124)

[2.1.2 Neural network based metaphor detection 7](#_Toc41848125)

[2.2 Metaphor interpretation 13](#_Toc41848126)

[2.2.1 Widely used knowledge base 13](#_Toc41848127)

[2.2.2 Knowledge based Metaphor interpretation 14](#_Toc41848128)

[2.2.3 Cooperation theory based Metaphor interpretation 14](#_Toc41848129)

[3 Widely used dataset 15](#_Toc41848130)

[3.1 VUA 15](#_Toc41848131)

[3.2 MOH-X 15](#_Toc41848132)

[3.3 TroFi 16](#_Toc41848133)

[4 Project outlines 16](#_Toc41848134)

[4.1 Metaphor detection 16](#_Toc41848135)

[4.2 Metaphor interpretation 16](#_Toc41848136)

[Reference 17](#_Toc41848137)

# Abstract

The ubiquity of metaphor in our daily life makes the computational realization of metaphor identification and interpretation become a crucial segment in the NLP translation field. This literature review provides the basic theory of metaphor computation, discusses the related work of the state-of-the-art models for metaphor detection and interpretation, introduces three widely used datasets and states our project outline for future work.

# 1 Introduction

## 1.1 Background

Metaphor expressions appear pervasively in daily life as well as in many literary works. According to relevant statistics research, metaphorical usage occurs on average once in every three sentences in typical corpora. People use metaphor frequently since metaphor expression furnish vivid explanations for abstract experience and perception. Meantime, our cognitive system and ways of thinking are based on metaphors aiming to understand our world better. The cognitive and rhetorical phenomena make metaphors active in all kinds of natural language communication scenarios, while a large number of implicit metaphorical expressions have become a problematic task in natural language processing. The computational realization of metaphor identification and interpretation therefore plays a crucial role in the NLP translation field.

## 1.2 Basic linguistic theory in metaphor

Traditional metaphor theory asserts that metaphor is merely a rhetorical method, and only provides an ordinary describing function for expression. In terms of Aristotle's poetics, metaphor is the similarity between objects. This theory emphasizes the feature differences between literal and metaphorical words.

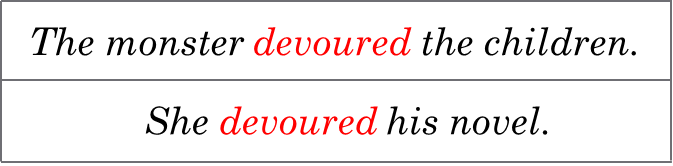
 As the development of cognitive science, in light of modern linguistic theories, Lakoff and Johnson [1] assert the conceptual metaphor theory, arguing that metaphor is systematic reasoning from a concrete conceptual source domain to an abstract target domain. Normally, we use a concrete expression to describe the abstract theory. For instance (Figure 1), the metaphorical sentence "She devoured his novels" utilizes the verb "devour" to describe a sense of greedy reading rather than literally swallowing a novel.

Figure 1 "devoured" in the first sentence is literal usage, and in the second sentence is metaphorical usage, the "devoured" is the source word in the second sentence.

There is another modern linguistic theory in metaphor derives from Steen et al. [2], namely Metaphor Identification Procedure (MIP): a metaphor is identified if the literal meaning of a word contrasts with the meaning that word takes in this context. Like the previous example, the word "devour" is unusual in the context of 'novel'. A novel cannot be devoured. The contextual meaning contrasts with the literal meaning of the word. Additionally, another similar metaphor theory is SPV. The intuition of the theory is that metaphoricity is identified by detecting the incongruity between a target word and its context.

Moreover, Gentner [3] 's structural mapping theory proposed that the process of structural mapping is the analogous transfer, which seeks for structural similarity by matching different objects, and then extracts the relationship in the source problem through schema induction to solve the problem of the target domain. It is concluded that the process of metaphor mapping is essentially the matching process of the relationship structure between the source domain and the target domain.

In the 1990s, Fauconnier and Turner extended the conceptual metaphor theory and proposed a conceptual integration model [4]. Besides, other theories like basic metaphor theory and characteristic significance imbalance theory have laid a solid theoretical foundation for the metaphorical computing task.

## 1.3 Problems for solving

The metaphor computation task mainly focuses on two aspects. One is how to use a computational method to detect and classify metaphorical expressions in a typical corpus. The other task is to let the computer understand and metaphorical meaning and then interpret it into the most suitable demotic word through word context.

Metaphor detection can be seen as a classification task, which is to identify a target work if it is a metaphor expression. Various approaches have been implemented by word embedding level to sentence level. When a sentence contains words such as "as" and "like", it is readily to be detected. Whereas without such words, the difficulty of detection highly rises. In word level, machine learning methods, including SVM and random forest or word embedding methods like GloVe are developed based on the concreteness difference of the target word and source word. By measuring the cosine similarity of the components in the phrase, metaphor can be effectively detected.

Metaphor detection can also be seen as a sequence annotation task, which is labelling the target word in a sentence. In sentence level, the neural network method has been widely applied for the NLP field for semantic modelling and context modelling. Beginning with the pretraining work embedding vectors, to multi-level processing algorithms, also along with LSTM or CNN method in deep learning or large-scale corpus language models like ELMo and BERT, novel methods are continually lifting the performance of metaphor detection result.

The task for metaphor interpretation aims to let the model understand the implicit literal meaning of metaphorical expression. How to precisely replaced metaphorical expression has been a problematic issue for long. Human beings generally understand metaphor through association with knowledge, providing ideas for the construction of a metaphor interpretation model. However, many difficulties and problems are waiting to overcome, such as considering culture and including novel knowledge base.

# 2 Related work

## 2.1 Metaphor detection

### 2.1.1 Word level based metaphor detection

#### 2.1.1.1 Abstractness for metaphor detection

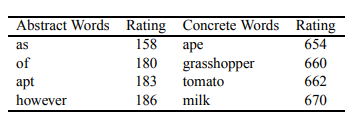
In light of the conceptual metaphor theory, Turney et al. [5] first implemented a method for identifying metaphors. Their approach to the problem of distinguishing literal and metaphorical senses is based on an algorithm for calculating the degree of abstractness of words by MRC Psycholinguistic Database (Figure 2). For instance, time is a word with a high relative abstract attribute. It therefore has a high value in using in metaphor.

Figure 2 examples for abstract words and concrete words in MRC database

Turney constructs the set of abstract words and concrete words. Then calculate the abstractness in terms of the semantic distance between the target words and the pre-designed word set. Finally, the model produces the possibility of if a word is used in metaphor or literal hinge on the word abstractness.

This model provides a new way of thinking for metaphor detection. However, the consideration of just abstractness is not a complete solution to the problem.

#### 2.1.1.2 Metaphor Detection with broader consideration on word

In 2014, Tsvetkov et al. [6] provide a broader view of metaphor detection on different languages besides English. They design three main feature categories:

1. Abstractness and imageability: Abstractness is always concerned as the critical feature for metaphor. Imageability is not a redundant property. The majority of abstract words are hard to visualize.
2. Supersenses: Supersenses are coarse semantic categories from WordNet. It classifies nouns and verbs into 45 classes. For instance, noun.animal, noun.body. This model also uses 13 classes adjective supersenses for addition.
3. Vector space word representation: The model employ a 64-dimensional vector-space word representation where synonyms have similar vectors.

To make classification, the model uses a random forest classifier to consider the three features simultaneously. The classifier learns from independent subsamples of the training data. For input, each tree classifier gives the probability of each label, and those probabilities are averaged to the ensemble. Additionally, it usually performs logistic regression rather than linear results. If this probability is above a threshold, the relation will be viewed as metaphoric; otherwise, it is literal.

This model considers the word level on a higher aspect and combines them with a random tree method. However, the restriction on the input form of the phrase (Subject-Verb-Object or Adjective-Noun) limits the view for further progress.

#### 2.1.1.3 Word Embedding and WordNet for metaphor detection

In 2018, Mao et al. [7] propose an unsupervised learning method that can directly identify and interprets metaphors at word-level without any preprocessing. Their model is based on the Continuous Bag of Words (CBOW) and Skip-gram.

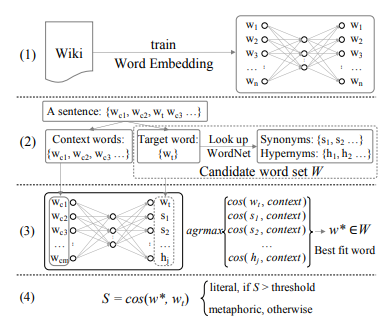
CBOW means given the context of a word, predict its center word, which has the highest probability. Skip-gram is the reverse of CBOW aiming to maximize the probability of predicting each context word given a center word.

Figure 3: word embedding based model structure

In Figure 3, it shows the structure of the entire model. First, input training word embeddings based on Wikipedia dumps for obtaining input and output word vectors. Afterwards, for a sentence, separate its target word and its context word. Furthermore, set Target words set for all the possible senses. Also, eliminate auxiliary verbs since these words contain less contextual meaning. Look up to the WordNet to get the Synonyms and Hypernyms of the Target words set and augment to Candidate word set W. In the next step, identify the best-fit word for the replacement of the target words given its context. Eventually, compute the cosine similarity between the target words the best-fit word. For S, if S is over a threshold, it will be considered as literal; otherwise, it is metaphoric.

This model makes use of the word vectors and its spatial relationship to provide a new way of thinking and directly contribute to the integration of metaphor recognition and metaphor understanding.

### 2.1.2 Neural network based metaphor detection

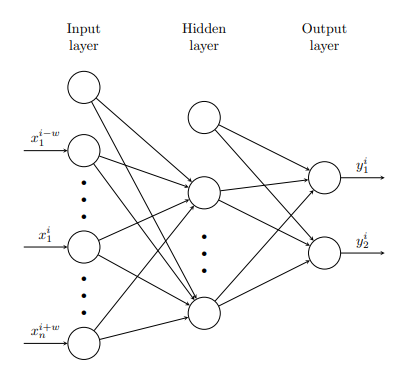
In sentence level, researchers find that metaphorical usage appears in a specific context. Without context, it is difficult to find the differences between metaphorical and literal expression. Therefore, constructing the interaction relation between target word and context, then discovering the connection between the two becomes a new field of thinking. The neural network method has been widely applied for the NLP field for semantic modelling and context modelling.

The main idea of the current neural network based method is to construct a semantic model for the target words and to the context of the target words. If the target word has a big difference with the context of semantics, the target word then can be considered as a metaphor; otherwise, it is not a metaphor. Many neural network models rely on the use of various semantic coding method to achieve different performance.

#### 2.1.2.1 The initial neural network implementation on metaphor detection

In 2016, Do Dinh and Gurevych [8] first implemented a neural network method on the metaphor detection task. The metaphor detection is treated as a tagging problem without any manual labelling. They experiment with the multilayer perceptrons with connected feedforward neural network.

As in Figure 4, this neural network incorporates an input layer, several hidden layers and an output layer. The output layer utilizes a softmax function to calculate the prediction result. Here is an instance of one hidden layer neural network of multilayer perceptron, the model use input for a token which positions at I concatenate word embedding vectors to . Here is the given window for the context size. Additional features enrich the input vector by concatenation. The output layer utilizes logistic regression to calculate the and , which are the probability for metaphorical and literal usage. This work initially explored the application of neural networks on metaphor detection and somehow required to do more extension on the model.

Figure 4: multilayer perceptron with one hidden layer

#### 2.1.2.2 Neural Metaphor Detection in Context

Gao et al. [9] intend to resolve two problems: given a target verb and classify it whether the word is metaphorical or literal and detect all the metaphorical words in a sentence. The model utilizes a standard architecture based on bi-directional LSTMs (BiLSTM) augments with contextualized word embedding performing surprisingly well on the tasks.

For both sequence labelling and classification encoding, the model uses to represent each token by pre-trained word embedding from GloVe (Global Vectors for Word Representation) and the concatenation of ELMo (Embeddings from Language Model) vector . The combination of the two effectively reduces the word sense disambiguation.

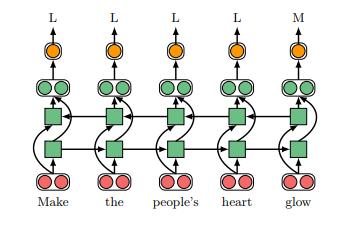
The first task refers to the Sequence Labeling Model. As shown in Figure 5, the model set the word representation to BiLSTM as input. Afterwards, the BiLSTM produces a result of contextualized representation for each token. The model then predicts a label for each word from by using a feedforward neural network.

Figure 5: A sequence labelling model for metaphor detection

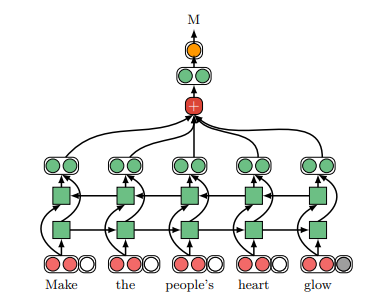
Furthermore, the classification model (Figure 6) of metaphor detection also concatenates an index embedding which indicating whether if it is the target word. Therefore, the word representation as input also produces a contextualized representation to . When computing the weight for each token , the model also utilizes the attention layer. Then compute the result as the weighted sum of the output of BiLSTM, here and are the learning parameters.

Figure 6: A classification model for metaphor detection

This model proves that sentence with context may provide abundant information. The BiLSTM may better capture the crucial information in the sentence and output a better performance.

#### 2.1.2.3 Neural Metaphor Detecting with CNN-LSTM Model

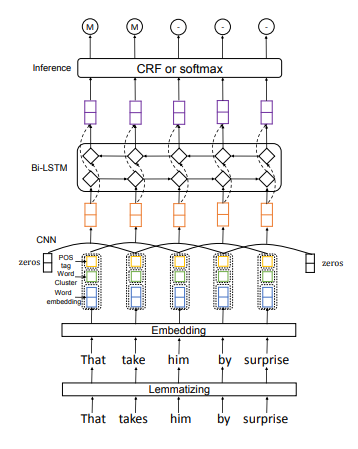
On the 2018 VUA metaphor detection sharing task, an extensive range of neural network methods are implemented containing RNN, LSTM, CNN or the combination of different neural network models. Wu et al. [10] performs the best result in the sharing task, which utilize the pre-trained word vector for input, then use the combination of CNN and BiLSTM for processing. CNN is to capture the local context information; BiLSTM is to capture the global context information. Finally, full-connected CRF or softmax layer is for prediction. This model will be detailed introduced from bottom to top in Figure 7.

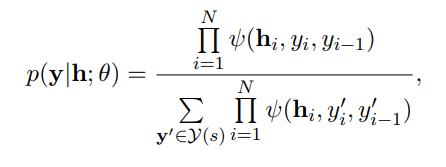
Figure 7: The structure for CNN-LSTM Model

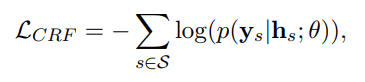
The model initially uses lemmatization to let the verb in the sentence transfer from different forms into one by the following strategy from Klebanov et al. (2016). In this way, the lemmatized word can simplify the semantic information and reduce the word number.

The next layer is a word embedding layer. The sequences of words in the sentence are converted into low-dimension vectors via a lookup table. The weight of words utilized pre-trained word2vec and is fine-tuned during model training. The model contains one-hot POS (Part Of Speech) tags as additional feature augmented into word embedding as well. K-means method is used to resolve the case that similar words may incorporate similar metaphor information.

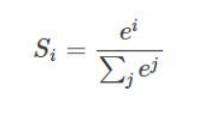
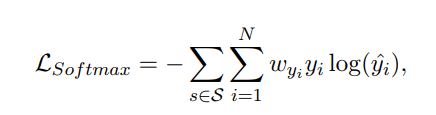
The third layer is the multiple kernel convolutional neural network for capturing local contextual information with different windows size for modelling. Afterwards, the fourth layer is the BiLSTM layer aiming to extract long-range contextual information from the CNN features map. This layer synthesizes the previous and future context information to output the hidden state at time .

The last layer implements two alternatives for inference, which are CRF and softmax. The two alternatives both produce better in the different test cases.

**CRF**: This is for prediction of the labels for each word. Given the matrix of hidden layer , the probability for the output sequence is:

Here is the parameter and is the set of all possible label sequence. And is for the potential function.

W and T mean the linear transformation parameter. Moreover, the CRF loss function is the negative log function for all the training samples. and are for hidden states.

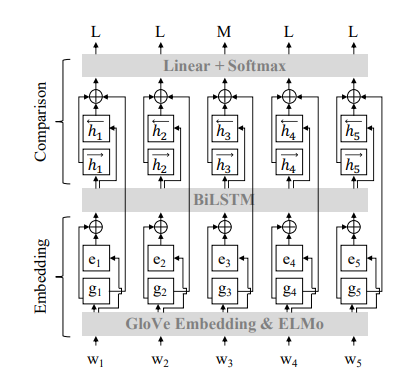
**Softmax**: The loss function for softmax is also a negative log function. is the metaphor label for word and predict score as . The model sets larger loss weight on positive class due to the larger quantity of non-metaphorical words.

This model tactfully uses a combination of the two neural modules to add the necessary features to improve the performance, which is a very effective method. However, the model does not effectively use the metaphor theory to optimize the results.

#### 2.1.2.3 Sequential Metaphor Identification Inspired by Linguistic Theories

As this review mentioned before, the two linguistic theory in metaphor SPV (Selectional Preference Violation) and MIP (Metaphor Identification Procedure) have not been emphasized before. Mao et al. [11] explored the two theory to implementation in 2019. Their model has the best result until now.

Figure 8 shows the implementation on MIP: a metaphor expression is classified by the contrast between a word's contextual and literal meanings. For the basis word encoding, the author continues to use the concatenation of ELMo vector and GloVe word vector. Since the model requires the contextual meaning, BiLSTM that can effectively capture contextual information will be set as the hidden states. The result of each token will be outputted as label probability, which is L (literal) or M (metaphorical). The probability of a label prediction for a target word at position , which is a contextual and literal meaning representation of the target word.

Figure 8: RNN\_HG model framework based on MIP

Here, is the same method as Gao et al., which is concatenation as input for the BiLSTM and get the hidden state . Next is the softmax function, is trained parameters and is the bias.

Figure 9 shows the implementation of SPV: metaphor expression can be identified by detecting the incongruity between a target word and its context. The model is RNN\_MHCA (Recurrent Neural Network Multi-Head Contextual Attention). The second model compares the target word (BiLSTM hidden state) with context . Notice that context incorporates two parts: left-side and right-side contextual attention representation. means the attention mechanisms on a window of context words.

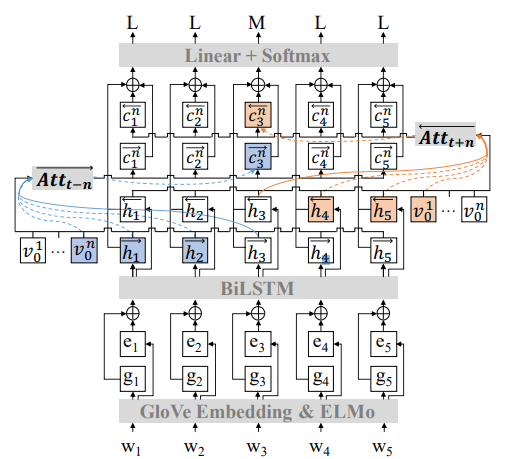
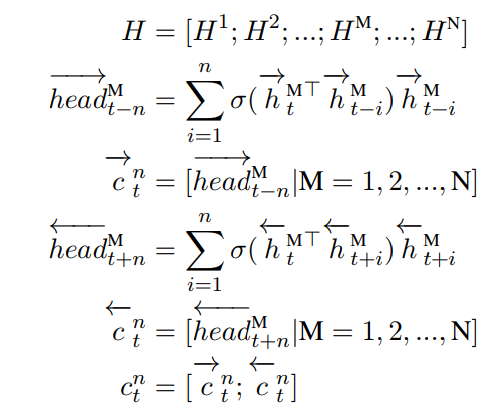


Figure 9: RNN MHCA model framework based on SPV



Where is the hidden state of BiLSTM and is the number of heads. The model employs a window size for contextual words. And is the hidden state for the target word. In this formula, consider the forward contextual information and past contextual information on both directions, effectively capture the information.

This model provides another aspect of thinking, which is linguistic advancement on the model. Previous neural models ignore the significance of different conceptual theory on the linguistic aspect. Therefore, linguistic consideration is a cardinal part for further work.

## 2.2 Metaphor interpretation

### 2.2.1 Widely used knowledge base

For metaphor interpretation, directly use the method like reasoning from context. This way seems possible, but the accuracy cannot assure. Consider the way that human being understands the metaphor. A man will remind of the previous knowledge he learned and extract the attribute in the word. For this, a knowledge base is necessary. Next, some crucial knowledge base for metaphor interpretation will be listed:

**WordNet**

WordNet is widely used in metaphor detection and metaphor interpretation for capturing the semantic information of words. It is an extensive English word database developed at Princeton University, where nouns, verbs, adjectives and adverbs are groups cognitively into synonyms for different concepts.

**VerbNet**

VerNet the biggest online verb word database. Every verb contains the subject description, parameter restriction and framework for semantic or syntax analysis. It is greatly useful for verb metaphor processing.

**FrameNet**

FrameNet is a framework semantics based word knowledge base. The majority of the word inside can be displayed by the semantic framework.

**HowNet**

Hownet is a semantic network based on word concepts, reflecting morpheme inclusion, morpheme equivalence, and domain-specific interactions between word concepts.

### 2.2.2 Knowledge based Metaphor interpretation

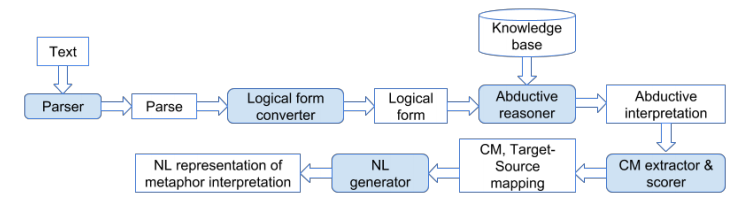
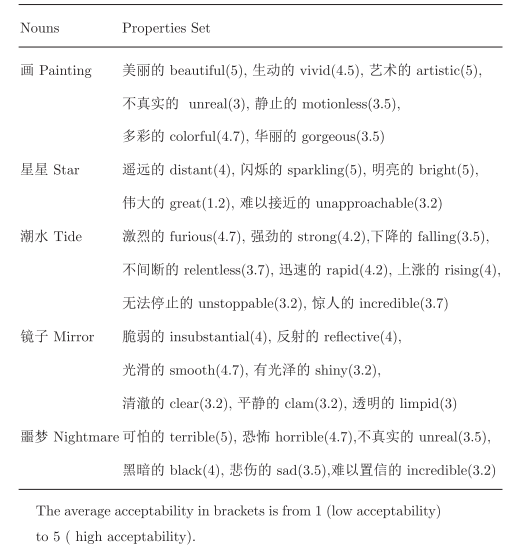
Ovchinnikova et al. [12] constructed a system of multilingual metaphorical understanding based on abductive reasoning (Figure 10). The system first analyzes the fragments containing the metaphorical text and generates the logical expression. The knowledge base is input into the inference system to get the inference result with the highest possibility, which is mapped into the conceptual metaphor. Finally, the conceptual metaphor is translated into logical predicates and translated into natural language expressions.

Figure 10: Abduction-based metaphor interpretation system

Shutova et al. [13] also proposed an unsupervised metaphorical comprehension system, which can change the metaphorical text into non-metaphorical text with the same meaning. The system trains the semantic similarity vector and the two meanings obtained by the UKWac corpus. The system mainly focused on subject-verb and verb-object types of metaphor, and the system accuracy reached 0.52.

### 2.2.3 Cooperation theory based Metaphor interpretation

Su et al. [14] believed that the result of metaphor understanding is a cooperation between the source domain and the target domain, and the focus of metaphor understanding was to extract their similarities. They divide metaphor understanding into two steps: attribute extraction and attribute transfer. The system retrieves the source domain properties from the property database. The system then calculates the relationships between the attributes from the target domain and the source domain. Cosine similarity is used to measure the correlation based on the word vector model. The most relevant attributes are systematically selected as the best interpretation of the metaphor. Eventually, the model gets a decent result.

Figure 11: The properties of several nouns and the calculated acceptability

# 3 Widely used dataset

## 3.1 VUA

VUA (VU Amsterdam Metaphor Corpus) is the largest manually labelled figurative language Corpus among a various field that published in the Metaphor recognition task. It consists of four categories of texts, namely academic texts, novel texts, news texts and dialogue texts. The language contains over 2626 paragraphs, more than 16,000 sentences and about 200,000 words. This corpus is based on MIPVU metaphor recognition rules, and some texts are selected from BNC-baby for annotation.

|  |  |  |
| --- | --- | --- |
| Type | vocabulary | Segments quantity |
| Academic text | 49561 | 16 |
| Dialogue text | 48001 | 24 |
| Novel text | 44892 | 12 |
| News text | 45116 | 63 |
| Total | 187570 | 115 |

Table 1: vocabulary and segments quantity in VU Amsterdam Metaphor Corpus

## 3.2 MOH-X

The MOH-X corpus is a subset of the MOH corpus. The texts in the corpus are all from the WordNet dictionary, and the sentences in the corpus are generally of short length.

## 3.3 TroFi

TroFi Corpus came from The Wall Street Journal, in which 50 verbs were marked, including the literal and metaphorical usage of these words. The packet contains three fields. The first field is used to mark the position of the sample text in the Wall Street Journal; the second field is used to represent the label. It contains three values: L (Literal), N (Nonliteral), and U (Unannotated), and the third field is the sample text.

# 4 Project outlines

Our final year project task incorporates two parts: first is for innovation on metaphor detection, which is identifying the existence of metaphor among various sorts of words. Afterwards, the second task is for metaphor interpretation, which is transferring the metaphor word into the most suitable demotic word through word context.

## 4.1 Metaphor detection

In the first task, in order to raise the performance of the previous work, we initially tend to use a neural network method containing modules like CNN, BiLSTM, CRF or combinations of the above and augment essential detailed process like emotional analysis into it. If possible, consider more on the linguistic part of the metaphor theory for advancement. Then restructure the model based on the previous work.

After constructing our model, we will set our training and test dataset as the three widely used metaphor datasets: VUA, MOH-X, TroFi. We may set the results of the state-of-the-art model like Mao et al. (2019) as the baselines for comparison.

If the output result is less than expectation, we will change the aspect of thinking, think about innovation points, restructure our model, adjust the learning parameters or augment new modules in our model.

## 4.2 Metaphor interpretation

In this task, we tend to follow the previous work and do some innovations. We tend to utilize the cooperation mechanism of the source domain and the target domain. On one side, we may append one of the knowledge bases, let each word contains more knowledge information, including emotion and environment. On the other side, we may add a process after the selection, which is checking if the replaced word in the context is suitable by CBOW.

After constructing our model, we may set our training and test dataset as the same as the Su et al. works. Furthermore, set the results of the state-of-the-art model as the baselines for comparison.

# Reference

[1] L. George, and M. Johnson. Metaphors We Live by, Chicago: University of Chicago Press, 2003.

[2] G. J. Steen, A. G. Dorst, J. B. Herrmann, A. Kaal, T. Krennmayr, and T. Pasma, A method for linguistic metaphor identification: John Benjamins Publishing, 2010.

[3] D. Gentner. Structure-mapping: Cognitive science, 1983, pp. 155-170.

[4] G. Fauconnier, M. Turner. Conceptual integration networks: Cognitive science, 1998, pp. 133-187.

[5] S. Mohammad, E. Shutova, P. Turney. "Metaphor as a medium for emotion: An empirical study", in Proceedings of the Fifth Joint Conference on Lexical and Computational Semantics. 2016, pp. 23-33.

[6] Y. Tsvetkov, L. Boytsov, A. Gershman, E. Nyberg, and C. Dyer, "Metaphor detection with cross-lingual model transfer", in Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics, 2014, pp. 248-258.

[7] R. Mao, C. Lin and F. Guerin, "Word embedding and wordnet based metaphor identification and interpretation", in Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics, 2018, pp. 1222-1231.

[8] E. L. D. Dinh and I. Gurevych, Token-level metaphor detection using neural networks, in Proceedings of the Fourth Workshop on Metaphor in NLP, 2016, pp. 28-33.

[9] G. Gao, E. Choi, Y. Choi and L. Zettlemoyer, "Neural metaphor detection in context", arXiv:1808.09653, 2018.

[10] C. Wu, F. Wu, Y. Chen, S. Wu, Z. Yuan and Y. Huang, "Neural metaphor detecting with CNN-LSTM model", in Proceedings of the Workshop on Figurative Language Processing, pp. 110-114, 2018.

[11] R. Mao, C. Lin and F. Guerin, "End-to-End Sequential Metaphor Identification Inspired by Linguistic Theories", in Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, 2019, pp. 3888-3898.

[12] E. Ovchinnikova, R. Israel, S. Wertheim, V. Zaytsev, N. Montazeri and J. R. Hobbs, "Abductive inference for interpretation of metaphors", in Proceedings of the Second Workshop on Metaphor in NLP, pp. 33-41, 2014.

[13] D. Bollegala and E. Shutova, "Metaphor interpretation using paraphrases extracted from the web", PloS one vol. 8, pp. 1-10, 2013.

[14] C. Su, S. Huang and Y. Chen, "Automatic detection and interpretation of nominal metaphor based on the theory of meaning", Neurocomputing, vol. 219, pp 300-311, 2017.