

# A Study of KIND Metaphor and Simile Annotation based on Parsing and ConceptNet

Meng-Hsien Shih

Yu-Siang Shen

Siaw-Fong Chung

Heng-Chia Liao

National Chengchi University

April 23, 2020

# Overview

- 1 Introduction to KIND Metaphor and Simile
- 2 Methodology
- 3 Semi-automatic Metaphor Annotation with Manual Correction

# Table of Contents

- 1 Introduction to KIND Metaphor and Simile
- 2 Methodology
- 3 Semi-automatic Metaphor Annotation with Manual Correction

# Introduction to KIND Metaphor (隱喻) and Simile (明喻)

- KIND Metaphor
- Pattern: X is a KIND of Y (X 是一種 Y)
  - ① In a literal expression, the pattern is used to explain X by its analogy to Y  
豆腐是一種營養好吃的 食物。  
'Dofu is a kind of nutrient and delicious food.'
  - ② In a metaphorical expression, the pattern is used to signal the metaphorical reading  
完全自由的市場只是一種 神話。  
'Totally free market is only a kind of myth.'
- Simile expressions are marked with simile words such as 像 or 一般 'like' in the sentence, while metaphor expressions do not.

# Research Questions

In this corpus-based metaphor study, we begin with two special types of metaphorical expression (with regular patterns), the KIND metaphor and simile, so that we can extract mapped concepts based on the patterns.

In this paper, we will examine the role of concepts in metaphor identification:

- 1 How to automatically capture the mapped nominal concepts in KIND metaphor and simile expressions?
- 2 Can the similarity distance between two concepts facilitate KIND metaphor annotation/identification)?

# Previous Metaphor Identification Procedure (MIP)

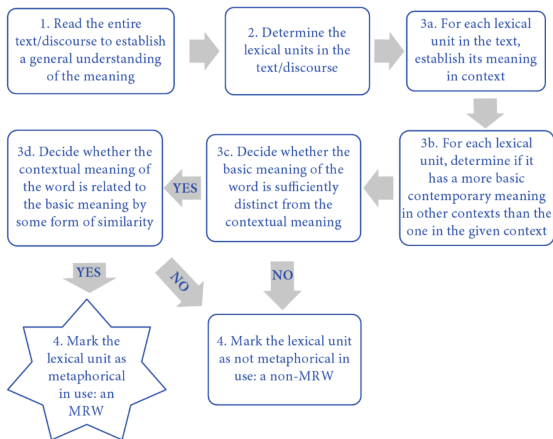
Krennmayr and Steen (2017), Pragglejaz Group (2007):

1. Read the entire text/discourse to establish a general understanding of the meaning.
2. Determine the **lexical units** in the text/discourse
- 3a. For each **lexical unit** in the text, establish its meaning in context, i.e. how it applies to an entity, relation or attribute in the situation evoked by the text (contextual meaning). Take into account what comes before and after the **lexical unit**.
- 3b. For each **lexical unit**, determine if it has a more **basic contemporary meaning** in other contexts than the one in the given context. For our purposes, **basic meanings** tend to be:
  - more concrete; what they evoke is easier to imagine, see, hear, feel, smell, and taste.
  - related to bodily action.
  - more precise (as opposed to vague)
  - historically older.

**Basic meanings** are not necessarily the most frequent meanings of the **lexical unit**.
- 3c. If the **lexical unit** has a more basic current/contemporary meaning in other **contexts** than the given context, decide whether the **contextual meaning** contrasts with the **basic meaning** but can be understood in comparison with it.
4. If yes, mark the **lexical unit** as metaphorical.

# Illustration of the MIP

(Source: Nacey, 2013, p. 79)



# Some issue of the MIP

- MIP is designed for **metaphorical sense** identification.
- MIP adopts a bottom-up approach that does not assume that “related conceptual metaphors guide linguistic metaphor identification” (Krennmayr & Steen, 2017)



# Table of Contents

- 1 Introduction to KIND Metaphor and Simile
- 2 Methodology
- 3 Semi-automatic Metaphor Annotation with Manual Correction

# Data with Metaphor Annotation (Exp. 1)

One annotator identified 19 metaphorical sentences out of 144 是一種 'is a kind of' sentences from the Academia Sinica Balanced Corpus 4.0:

顏色是一種可以直接影響心靈的力量 (隱喻)

豆腐是一種營養好吃的食物 (無)

'Dofu is a kind of nutrient and delicious food.'

# Data with Simile Annotation (Exp. 2)

- The other annotator identified 293 simile sentences out of 400 X 一樣的 Y 'Y like X' sentences from the Corpus of Contemporary Taiwanese Mandarin (the written corpus of version 2017)
- 'X 一樣的 Y' is a robust pattern to identify simile and to locate the mapped concepts X and Y in a simile sentences
- 他有一張像 龍 一樣的 臉 (明喻)  
'He has a face like a dragon.'
- 母親給了他和大哥一樣的東西 (無)  
'Mother gave him the thing the same as his brother.'

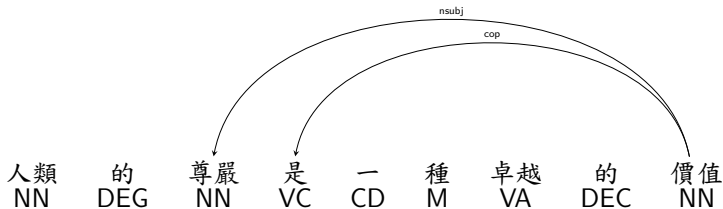
# Table of Contents

- 1 Introduction to KIND Metaphor and Simile
- 2 Methodology
- 3 Semi-automatic Metaphor Annotation with Manual Correction

# Semi-automatic metaphor annotation (Exp. 3)

The following tool and resource are exploited:

- 1 SyntaxNet Dependency Parser to locate the two mapped concepts in a sentence



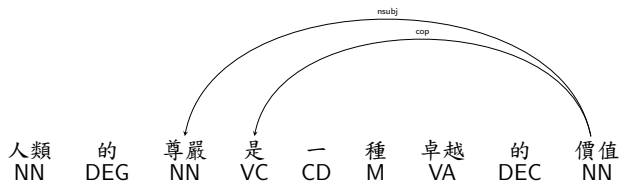
- 2 ConceptNet 5.5

想念 is a type of 思念

思念 is a type of 折磨

# A dependency pattern to capture KIND concepts in copular sentences

- The copular KIND sentence: X ... 是一種 ... Y
- X is the nominal subject (nsubj) of Y in dependency grammar
- 19 out 144 copular sentences were identified as metaphorical



# Example data to rate the metaphoricity

Concept 1	Concept 2	Sentence
市場	神話	完全自由的市場只是一種神話
想像	遊戲	所以我們這一種想像是一種不容易玩的遊戲
聲音	染料	原來，聲音就是一種染料

# Likert rating of the degree of the metaphor usage

5: 完全同意是隱喻

1: 完全不同意是隱喻

	Anno1	Anno2	Anno3	mean_metaphoricity	ConceptDist	ConceptPath	Sentence
1	5	5	4	4.67	5	['市場', '人', '吸血鬼', '傳說', '神話']	完全自由的市場只是一種
2	3	5	5	4.33	3	['想像', '高興', '遊戲']	所以我們這一種想像是一
3	5	5	5	5.00	4	['聲音', '空氣', '衣服', '染料']	原來，聲音就是一種染料
4	5	5	4	4.67	4	['媳婦', '兒子', '金錢', '工具']	這媳婦只是一種襯托的
5	5	5	4	4.67	3	['藝術', '台語', '語言']	電影藝術，是一種世界語
6	4	5	3	4.00	3	['文化', '個', '行動']	文化是一種符號行動。
7	4	5	5	4.67	3	['身體', '大雄', '道具']	我們的身體只是一種做
8	2	4	2	2.67	5	['主義', '共產主義', '人', '法律', '規範']	多元主義是一種取消了所
9	1	1	1	1.00	4	['歌劇', '場', '電影', '藝術']	題目：文學與歌劇。歌劇
10	2	4	2	2.67	3	['信仰', '天使', '力量']	信仰是一種力量，
11	1	1	1	1.00	5	['禪', '/c/en/meditation', '沉思/n', '/c/']	禪不是一種知性的觀念
12	1	1	1	1.00	5	['人格', '人/n', '病人', '虛弱', '病態']	人格異常是一種病態，
13	1	1	1	1.00	99		」空門是一種修行的手段
14	1	1	1	1.00	4	['自我', '人', '上網', '方式']	第二，他的自我超越是一
15	1	1	1	1.00	3	['希望', '幸福', '表現']	「我希望」的希望就是
16	1	1	1	1.00	2	['理論', '假設']	第一、所謂普遍語法理論
17	1	1	1	1.00	4	['辯論', '吵架', '裝死', '方法']	國會的辯論與表決是一種
18	1	1	1	1.00	5	['行為', '愛', '適合', '/c/en/fit', '適應']	近來生物學者認為生物的
19	1	1	1	1.00	4	['血紅素', '補充養分', '進食', '蛋白質']	因此紅血球含有血紅素
20	1	1	1	1.00	1		基本上，不是地理學



# Inter-Annotator Agreement

After removing 33 KIND metaphor sentence candidates with parsing errors or rating differences larger than 2:

```
kripp.alpha(t(as.matrix(df[,c('Anno.1', 'Anno.2', 'Anno.3')]))
```

Krippendorff's alpha

Subjects = 111

Raters = 3

alpha = 0.645 ( < 0.67 )

# The Investigation of Concept Distance

- Concept distance in this study: the shortest path in ConceptNet
- e.g., 完全自由的市場是一種神話
- In ConceptNet: ‘市場’, ‘人’, ‘吸血鬼’, ‘傳說’, ‘神話’.
- Why shortest path? We want to examine each connection between two mapped concepts (for metaphor explanation)

# The correlation between annotation and concept distance

	Anno1	Anno2	Anno3	mean_metaphoricity	ConceptDist	ConceptPath	Sentence
1	5	5	4	4.67	5	['市場', '人', '吸血鬼', '傳說', '神話']	完全自由的
2	3	5	5	4.33	3	['想像', '高興', '遊戲']	所以我們
3	5	5	5	5.00	4	['聲音', '空氣', '衣服', '染料']	原來，聲
4	5	5	4	4.67	4	['媳婦', '兒子', '金錢', '工具']	這媳婦只
5	5	5	4	4.67	3	['藝術', '台語', '語言']	電影藝術
6	4	5	3	4.00	3	['文化', '個', '行動']	文化是一
7	4	5	5	4.67	3	['身體', '大雄', '道具']	我們的身
8	2	4	2	2.67	5	['主義', '共產主義', '人', '法律', '規範']	多元主義
9	1	1	1	1.00	4	['歌劇', '場', '電影', '藝術']	題目：文
10	2	4	2	2.67	3	['信仰', '天使', '力量']	信仰是一
11	1	1	1	1.00	5	['禪', '/c/en/meditation', '沉思/n', '/c/']	禪不是一
12	1	1	1	1.00	5	['人格', '人/n', '病人', '虛弱', '病態']	人格異常
13	1	1	1	1.00	99		」空門是
14	1	1	1	1.00	4	['自我', '人', '上網', '方式']	第二，他
15	1	1	1	1.00	3	['希望', '幸福', '表現']	「我希望
16	1	1	1	1.00	2	['理論', '假設']	第一、所
17	1	1	1	1.00	4	['辯論', '吵架', '裝死', '方法']	國會的辯
18	1	1	1	1.00	5	['行為', '愛', '適合', '/c/en/fit', '適應']	近來生物
19	1	1	1	1.00	4	['血紅素', '補充養分', '進食', '蛋白質']	因此紅血
20	1	1	1	1.00	4	['理性', '小說', '快樂', '態度']	基本上，
21	1	1	1	1.00	4	['瞎畫', '教堂', '椅子', '裝飾']	瞎畫

# Pearson correlation test

Correlations

		Metaphoricity	ConceptDist
Metaphoricity	Pearson Correlation	1	-.203*
	Sig. (2-tailed)		.033
	N	111	111
ConceptDist	Pearson Correlation	-.203*	1
	Sig. (2-tailed)	.033	
	N	111	111

\*. Correlation is significant at the 0.05 level (2-tailed).

# Results and Discussion

- A quality dependency pattern to capture the two mapped concepts in a copular sentence
- The Krippendorf's alpha shows a tentative inter-annotator agreement (  $0.645 < 0.67$  )
- The annotation guideline could be more strictly followed
- A significant slightly negative correlation between the degree of metaphorical usage and the concept distance (based on ConceptNet)

# Contribution (Take-home message)

- 144 annotated metaphor sentences and 400 simile sentences
- A significant slightly negative correlation between the degree of metaphor usage and the concept distance (based on ConceptNet)

# Limitation and Future Work

- Currently results are still limited to the annotated data
- Need to examine the application to other unannotated sentences in the future

# Questions and Answers

Comments are welcome!