

# 分析故事內容並生成對話模板用於聊天機器人

倪士恩 (Shih-An Ni)  
Chaoyang University of  
Technology  
Taichung, Taiwan,  
R.O.C  
edai1204@gmail.com

廖韋民 (Wei-Min Liao)  
Chaoyang University of  
Technology  
Taichung, Taiwan,  
R.O.C  
mine851227@gmail.com

廖家浚 (Liao, Jia-Jun)  
Chaoyang University of  
Technology  
Taichung, Taiwan,  
R.O.C  
x222516x@gmail.com

吳世弘 (Shih-Hung  
Wu)\*  
Chaoyang University of  
Technology  
Taichung, Taiwan,  
R.O.C

\*Contact author:  
shwu@cyut.edu.tw

## 摘要

服務機器人可以向需要機器人陪伴的各種用戶講故事，如兒童、學生和老人。而我們希望機器人除了講故事之外，在講完故事後還可以透過語音對話與用戶共同討論故事的內容。但是使用人工編輯模板太耗費成本，因此在本文中，我們報告如何使用資訊擷取與人工智慧標記語言的技術合成為對話的模板以達到自動化的目標，並用以建構一個可以使服務機器人與用戶聊故事內容的系統。

**關鍵字**— 資訊擷取、對話模板、聊天機器人 (Chatbot)、人工智慧標記語言 (AIML)

## I. 簡介

現今使用聊天機器人提供人們各種服務是一種趨勢，像是線上客服、教育服務，甚至是作為一個保姆。我們的目標是聊天機器人提供講故事及之後與孩童聊這個故事的服務。這樣的服務目前因人工編輯對話腳本的成本過高，無法廣泛的實現。因此我們設計了一連串自動化的機制。我們相信資訊擷取的技術可以為幫助系統從故事中找到重要的資訊[9,10,11]，而生成系統則可以有效的運用這些資訊去建構聊天機器人的對話腳本。

我們系統輸入的是孩子們的寓言故事，而輸出則是在 AIML[1]中相對應的對話腳本。這個腳本可以使用 AIML 的引擎去驅動，以達到與用戶對話的目的。

有許多以前的作品研究了對話管理的結構[4][5]或是系統該如何使用情感與用戶進行交流[2]。我們採用了資訊擷取技術來擷取在故事中重要的元素，像是主角、重要的角色、重要的時間點、地點還有一些句子。

辨識出文字內容的情緒是過去幾年很熱門的議題。早期的研究是注重在於辨認文字內所表達的是正面或是負面的情緒。而在較為近期的研究中，則是對於情感有了更多元且嚴謹的分類，像是在生理學中對於情緒的六種基本分類，而這六種分別是憤怒、厭惡、恐懼、快樂、悲傷和驚訝。這些情緒都可以通過文字中對於情感的描述識別出來。而當我們的系統偵測到了故事中對於情感的描述時，我們的系統即可產生出對話來與使用者進行討論，像是探討為何情緒會被激起，或是分享相同的情緒。

## II. 方法

我們為了這個研究，開發了一個類比的轉換過程[12]。為了設計一個這樣的自動系統，我們先手動的編輯了幾個對話的腳本，並試著設計一個系統去模擬執行整個專案執行的過程。手動的過程可以概括成以下步驟：第一步，我們先讀取整個故事並找出在故事中重要的名詞。而判斷重要名詞的關鍵是當你在和別人討論這個故事時，會提到的人物或是字詞，然而這是一個屬於人類直覺上的選擇。第二步，我們藉由選擇一些特定的動詞，去導出故事中重要的動作。特定的動詞除了動詞本身之外，如果在這個動詞前方擁有可以搭配的副詞，我們也會將其一併提取。在第三步，我們會標示在故事中這些重要的人物於故事中發生的情緒，並且根據情緒編入對話的腳本內。

故事的內容通常不會只有一句話就影響到整個故事的走向，而是接下來的幾句話或是幾段話都會影響。在人工編輯的腳本裡面，有著很多一問一答的句子對。當用戶在與聊天機器人對話時，可以確保整個聊天的內容都會圍繞在這整個故事裡面。

我們將會在接下來的小節中介紹我們系統的細節。

## III. 系統流程



圖1. 系統簡介

我們的系統分為兩大部分，其中有資訊擷取、對話生成模塊，如圖 1。處理步驟於下面列出。前四個步驟為資訊擷取的部分，最後一個步驟則是生成模塊。

- 文字前處理。
- 使用 CKIP 進行文字分割與 POS 標記。提取所有 Na, Nb 作為關鍵字的候補。
- 計算所有 Na、Nb 的詞頻分數。並找出前數值前五高的作為關鍵字。
- 於故事內搜尋關鍵字，標記出擁有關鍵字的語句。
- 透過將關鍵詞與語句放入對話模板內，生成可以使用的對話模板。

表1. 部分系統中使用的 CKIP 詞類標記<sup>1</sup>

簡易標記	對應的 CKIP 詞類
A	非謂形容詞
D	副詞
Na	普通名詞
Nb	專有名稱
VA	動作不及物動詞
VB	動作類及物動詞
VC	動作及物動詞
VE	動作句賓動詞
VF	動作謂賓動詞

### A. 文字前處理

將文件內的半形標點符號替換為全行的標點符號，並刪除多餘的換行符號，避免分詞系統產生錯誤。

### B. 文字分割與 POS 標記

由於我們的系統是用來處理中文的故事，因此在第一步會先將故事內容分段成為中文的單詞，然後再使用 CKIP 工具包將單詞標記詞性。

CKIP 分詞工具[6]的詞性標注如表 1 所示，值得一提的是，CKIP 的 POS 標詞性工具與其他標詞性工具不同，例如 Penn treebank 中的名詞分類並沒有繼續細分[3]，反之 CKIP 的標詞性工具則有此功能，例如：一般名詞的 Na，專有名詞 Nb，地方名詞 Nc。這些 POS 詞性標記為我們資訊擷取提供不少助力。

### C. 計算詞頻將關鍵字提取

故事的關鍵字是可以藉由詞頻與逆向檔案頻率去加權計算出來的，這個方法已經是被廣泛應用在資訊搜尋和自然語言處理中對於關鍵字的找尋。他可以評估一篇文章中的單詞對於這個文章的重要性。重要性與單詞的頻率是呈正比的，且與文章出現的頻率成反比。也就是說，如果一個術語在這個文章內出現了很多次，代表這個單詞對這個文章來說很重要，但又如果一個單詞他是出現在很多文章裡面，代表這個單詞比起只出現在這個文章內的單詞來說，對這個檔案是不太重要的。在文件  $j$  中單詞的 tfidf 分數是透過以下公式計算而得：

$$tfidf_{i,j} = tf_{i,j} \times idf_i$$

其中  $tf_{i,j}$  是文件  $j$  中單詞  $t_i$  出現的次數除以文件  $j$  的總詞數， $idf_i$  則是總檔案數除以出現有單詞  $j$  的檔案數，公式被定義為：

$$idf_i = \log \frac{|D|}{|\{j: t_i \in d_j\}|}$$

其中  $D$  是檔案總數， $|\{j: t_i \in d_j\}|$  是包含單詞  $t_i$  的檔案總數量。

### D. 關鍵短句選擇

對於關鍵短句的選擇，我們的系統是使用將包含上一個步驟所找出的關鍵字的句子選取出來，再去擷取其中名詞動詞的搭配或是形容詞和名詞的搭配。

### E. 對話腳本生成

我們是將手動的編輯的 AIML 文件去當作任何對話腳本的模板，再透過將選擇出的單詞或短句放入模板中形成新語句來生成 AIML。如此一來便可以為任何的故事生成腳本。

## IV. 實驗

在接下來的這個章節內，我們將舉例說明系統如何從故事中提取預期會使用到的資訊。在實驗時我們找了二十個故事，但在以下我們取一個故事局部來當作範例。

表2. 故事資訊擷取範例

句子	TFIDF 找出的關鍵詞	人力事先找出的關鍵詞
米可奮力抵抗，但微弱的閃光比螢光蟲還小，只能認命了！	米可	米可 奮力抵抗
年輕人將米可帶回家，心想它既是一朵雲，就需要水氣，於是燒了一大鍋熱水，用蒸氣薰米可，米可才悠悠轉醒，慢慢恢復元氣。	米可 雲 年輕人	年輕人 米可
米可冷靜下來，心想現在只有自己能救他了，於是深呼吸，將附近的烏雲吸附過來，自己也膨脹成一顆大氣球，再用盡全身力氣發出此生最明亮的一閃，一道閃電劃破天際直瀉而下，正巧命中一棵枯樹，枯樹瞬間起火燃燒，也將商人前進的路截斷了！	米可 雲 閃電	米可 閃電 劃破天際 枯樹 起火燃燒
此時天空烏雲略散，微亮的星空下仍閃著電光，商人向它揮揮手，閃電似乎明白了，回以數道足以照亮幽暗夜空的明晰亮光。	雲 商人 閃電	商人 閃電

### A. 故事資訊擷取

我們在這個實驗中使用的是閃電報恩這個故事。故事的內容在表三的第一欄，第二欄是我們的系統在通過 TFIDF 分數識別後所選出的關鍵字，第三欄則是在建構

<sup>1</sup> <http://ckipsvr.iis.sinica.edu.tw/>

系統前使用手動識別出的關鍵字。在這個故事中，我們的召回率是 74.1%，而準確率是 71.4%。

## B. AIML 生成

在這個部分，我們將透過將關鍵字插入模板來展示我們是如何為故事生成 AIML。AIML 模板是手動編輯的文件。

### 1) 故事名稱

從文件名稱抓取，將所有在模板內的「STORY」字詞替換為當前故事的名稱。假設程式抓取到的故事名稱為「閃電報恩」，則會被替換為以下形式。

```
<category>
  <pattern>*/</pattern>
  <template>
    <random>
      <li>你想跟我聊聊閃電報恩這個故事嗎?</li>
    </random>
  </template>
</category>
```

### 2) 使用普通名詞關鍵詞的句子

我們的系統會選擇出故事中包含著常用名詞關鍵字的句子並將其變成問句，以詢問使用者是否注意到這個名詞為故事的重要部分。

假設系統選擇出的句子是：年輕人將米可帶回家，心想它既是一朵雲，就需要水氣，於是燒了一大鍋熱水，用蒸氣薰米可，米可才悠悠轉醒，慢慢恢復元氣。

```
<category>
  <pattern>_</pattern>
  <that>是不是年輕人將米可帶回家，心想它既是一朵雲，就需要水氣，於是燒了一大鍋熱水，用蒸氣薰米可，米可才悠悠轉醒，慢慢恢復元氣?</that>
  <template>
    <think><set
      name="want"><star/></set></think>
    <condition name="want">
      <li value="是">很好哦！小朋友看故事很認真！給你一百個讚！那我繼續問囉</li>
      <li value="不是">是這樣嗎？請在認真看一下故事唷！然後我再問你</li>
    </condition>
  </template>
</category>
```

### 3) 詢問使用者如果你是這故事其中的角色，你會怎麼做？

我們系統的模板是像這樣的：如果你是 S 你也會 ADV+VERB 嗎？藉此來詢問使用者是否會與角色做一樣的事情。動作是由系統從故事中提取的動詞或是副詞加動詞來取代。假設主角是「米可」，從故事中抓取到的動詞與副詞搭配為「奮力」「抵抗」，那麼在經過替換後 AIML 會像以下這樣。

```
<category>
  <pattern>_</pattern>
  <that>如果你是米可會不會也奮力抵抗?</that>
  <template>
    <think><set
      name="want"><star/></set></think>
    <condition name="want">
      <li value="會">為什麼你也會奮力抵抗呢?</li>
      <li value="不會">為什麼你不會奮力抵抗?</li>
    </condition>
  </template>
</category>
```

這些 AIML 對話腳本可以持續創建並附加到現有的 AIML 內。而我們在其中所使用的<that>標籤可以標記在這之前的上一個問題，這對系統來說是一個短期的記憶，可以藉此達成一個問題的延伸問答，而不是只有一問一答就結束這一個問題。當使用者在這一問一答中完整的問答完並且之後不再有我們所創建的<that>標籤時，聊天機器人的系統會即刻切換聊天的主題。因此聊天機器人可以自由地談論許多不同主題的問題，而不是存在於一定的問題循環內。

### 4) 詢問使用者如果你是這故事其中的角色，你是否會有同樣的情緒？

表 3 顯示了 Robinson 的情緒分類[8]，在這張表中，我們可以看到不同類型的共同情緒和相應的情緒。我們的系統辨識出故事中的各種情緒之後，會詢問使用者如果你是這故事其中的角色，是否會有同樣的情緒。藉由討論故事中的情緒變化，深度理解故事的內涵。

表3. 情緒種類列表[8]

情緒種類	正面情緒	負面情緒
與對象相關	有興趣、好奇、熱情	冷淡、習慣、無聊
	有吸引力、有慾望、欽佩	厭惡
	驚喜	恐慌
對未來的評估	希望、興奮	恐懼、焦慮
與事件相關	感謝	憤怒
	喜悅、興高采烈、勝利、歡騰	悲傷
	忍耐	沮喪、失望
滿意	不滿、煩躁	
自我評價	謙虛	傲慢
社會	慈悲	貪婪、吝嗇
	同情	殘忍
情感	愛	討厭

## V. 討論

這個系統的好處是可以節省編輯故事對話腳本的時間。根據我們的經驗，每一個人編輯一個對話腳本需要至少兩個小時，編輯出的該腳本需涵蓋童話故事中所有主題和動作的簡短對話。這樣的腳本由我們的系統生成則只需要幾秒鐘。因此，人工編輯器可以使用自動生成的腳本作為初始版本，並以更少的時間將更多有趣的對話內容添加到腳本中。假設我們現在有 100 種這樣的故事；最初，編輯基本質量的腳本需要 200 個小時，然而如果在我們的系統的幫助下，可能只需要一天時間就能夠獲得更好的腳本。

## VI. 結論

我們實現了一個只給故事內容資訊就可以與用戶討論任何故事內容的系統。系統可以從故事中提取重要的主題，動作和情感，並將它們放入 AIML 的對話腳本中。AIML 腳本可以由 AIML 引擎執行，用自然語言與用戶互動。系統可以降低編輯聊天機器人對話腳本的成本，並增加聊天機器人的多功能性。

未來，我們可以進一步識別故事中人物的更多情緒狀態，以構建更有趣的對話。有了各種情感理論就可以處理更多情緒的細節，例如 Plutchik 的情緒輪[7]。要實現完整的情感識別系統，需要更多的訓練數據或字典。即使我們平均的召回率和準確率分別如表 4 只有 45.7%和 46.3%，但是對於我們的系統，並沒有必要建立一個具有高召回率和高準確率的完整識別系統，因為我們的系統可以從故事的任何細節中受益，藉以幫助聊天機器人談論故事，根據我們的統計，在產生出來的句子當中，能直接使用的句子占了 64.1304%，經過稍微修正後可使用的占 9.0581%完全無法使用的則只占了 26.8115%。因此，即使召回率低，識別更多情緒也會有所幫助。

表4. 關鍵詞的精準率與召回率

故事名稱	精準率	召回率
老上帝還沒有滅亡	0.5	0.2
老頭子做事不會錯	0.428571	0.130435
阿力的學習之旅	0.78125	0.78125
茶壺	0.5	0.75
規矩鴨	0.043478	0.071429
窮女人和她的小金絲鳥	0.176471	0.214286
碗豆上的公主	0.4	0.545455
賣火柴的小女孩	0.870968	0.84375
蕎麥	0.521739	0.631579
螃蟹小不點	0.764706	0.866667
新北風與太陽	0.3	0.136364
跳高者	0.68	0.944444
地下迷宮	0.285714	0.190476

雷公谷	0.6	0.681818
小雁的選擇	0.16	0.222222
小草莓	0.090909	0.071429
小白兔的紅眼睛	0.133333	0.095238
醜小鴨	0.777778	0.823529
一個豆莢里的五粒豆	0.391304	0.346154
閃電報恩	0.740741	0.714286
<b>平均</b>	<b>0.457</b>	<b>0.463</b>

## 參考資料

- [1] A.L.I.C.E. AI Foundation (2005) Artificial Intelligence Markup Language (AIML) Version 1.0.1, [http://www.alicebot.org/TR/2005/WD-aiml/AIML\\_syntax](http://www.alicebot.org/TR/2005/WD-aiml/AIML_syntax)
- [2] Arjan Egges, Xuan Zhang, Sumedha Kshirsagar, Nadia Magnenat-thalmann, Emotional communication with virtual humans, In Proceedings of the 9th International Conference on Multi-Media Modeling, pp. 243-263, 2003.
- [3] Beatrice Santorini. 1990. "Part-of-Speech Tagging Guidelines for the Penn Treebank Project (3rd Revision)", July.
- [4] Churcher, G, Atwell, ES and Souter, C (1997) Dialogue management systems: a survey and overview. University of Leeds, School of Computing Research Report 1997.
- [5] James F. Allen, Donna K. Byron, Myroslava Dzikovska, George Ferguson, Lucian Galescu, and Amanda Stent, Towards Conversational Human-Computer Interaction, AI MAGAZINE, vol. 22, pp.27-37, 2001.K. Elissa, "Title of paper if known," unpublished.
- [6] Ma, Wei-Yun and Keh-Jiann Chen, 2003, "Introduction to CKIP Chinese Word Segmentation System for the First International Chinese Word Segmentation Bakeoff", Proceedings of ACL, Second SIGHAN Workshop on Chinese Language Processing, pp168-171.
- [7] Plutchik, Robert (16 September 1991). "The Emotions". University Press of America. p. 110. Retrieved 16 September 2017 – via Google Books.
- [8] Robinson, D. L. (2009). Brain function, mental experience and personality. The Netherlands Journal of Psychology, 64, 152-167.
- [9] Shih-Hung Wu, Hsien-You Hsieh, Sentence Parsing with Double Sequential Labeling in Traditional Chinese Parsing Task, in Proceedings of the second CIPS-SIGHAN Joint Conference on Chinese Language Processing (CLP 2012), Tianjin, Dec. 20-21, 2012.
- [10] Shih-Hung Wu, Wen-Feng Shih, Che-Cheng Yu, Liang-Pu Chen, and Ping-Che Yang, CYUT-III Short Text Conversation System at NTCIR-13 STC-2 Task, in Proceedings of the 13rd NTCIR Conference on Evaluation of Information Access Technologies, Dec. 5-8, 2017, Tokyo Japan.
- [11] Shih-Hung Wu, Wen-Feng Shih, Liang-Pu Chen and Ping-Che Yang, CYUT Short Text Conversation System for NTCIR-12 STC, in Proceedings of the 12th NTCIR Conference on Evaluation of Information Access Technologies, June 7-10, 2016, Tokyo Japan, pp.541-546.
- [12] S. Wu, L. Chen, P. Yang and T. Ku, "Automatic Dialogue Template Synthesis for Chatbot by Story Information Extraction," 2018 IEEE International Conference on Information Reuse and Integration (IRI), Salt Lake City, UT, 2018, pp. 485-490.