

生活小幫手

專題編號：104-CSIE-S025
執行期限：103 年第 1 學期至 104 年第 1 學期
指導教授：郭忠義
專題參與人員：
101820331 陳亮宇
101820324 林哲琛

一、摘要

現在科技發達，許多資訊都可以在網路上搜尋到，但是網路上搜尋引擎的回覆通常都是不帶感情，並且很難直接找出使用者真正想要的資訊。本為專題為解決此問題，利用中研院剖析系統分析使用者的問題產生的結果，產生語意剖析樹(Parse Tree)，再使用自然語言處理中的問答系統解析出使用者需求，並加以回應。

關鍵詞：剖析樹(Parse Tree)、自然語言處理(Natural Language Process)、問答系統(Question Answering)

二、緣由與目的

搜尋引擎提供我們生活上許多的幫助，但搜尋引擎的回應不帶任何感情，也較少能直接提供使用者真正需要的資訊。因應此問題，使用者在使用搜尋的時候，不會只得到冷冰的回應，讓使用者感受到像個真人在回答問題，並且也會由使用者輸入的基本資料，做出對使用者最貼切的回應。

三、研究範圍

本專題研究的主題為自然語言處理，並著重在下列三項：

- (1) 語意剖析：主要用來分解句子，並將所有語詞加入標籤以解釋此字的詞性，並將各詞語的關係串成剖析樹結構。
- (2) 問答系統：以口語化的問題做輸入，回傳高精準度的結果和明確的答案。
- (3) 剖析樹：將結果依照結構、相對關係及優先順序，串成一個樹狀結構。

四、使用技術方法

由 Android 系統的 TextField 接受使用者問題，例如：最近的便利商店在哪裡？，利用 Socket 傳至中研院剖析系統，系統在剖析後會使用 XML 格式回傳結果，得到的剖析結果再由 Servlet 建立剖析樹物件，利用節點串接起父子節點和手足節點，再由程式邏輯依照剖析樹分析出使用者需求，例如：便利商店和哪裡就是關鍵的詞語，知道使用者正在找有關便利商店，並依此產生回應，例如：最近的是 711，在 XX 路 XX 號。

本專題分析的做法，大致上是從剖析樹找出動詞片語中的動詞以分析使用者想要的動作，再從名詞片語中找出關鍵詞，例如：捷運站(地方詞)、哪裡(位置詞)，來確認使用者需求的目標，再從資料庫依這些訊息調出指令到 Servlet 執行。

五、工具說明

中研院剖析系統是由中研院開發，輸入句子後，會將句子分割並標記文法後傳回，傳回結果大致如下：

S(NP(N: 台新|Head:N: 金控)|Nd: 1 2 月 3 日 |ADV: 將 |Head:Vt: 召開 |NP(N: 股東|Head:N: 臨時會)|VP(Head:Vt: 進行|NP(NP(Head:N: 董監)|Head:Nv: 改選))#)。 (PERIODCATEGORY)

結果之詞類標記大致如下：

- (一) S: 句子
- (二) NP: 名詞片語
- (三) VP: 動詞片語
- (四) N: 名詞

- (五) ADV:副詞
- (六) Nd:時間詞
- (七) Vt:動詞
- (八) Nv:名物化動詞

利用這些資料，找尋我們需要的關鍵字，例如：動詞表示使用者需求的動作、名詞表示使用者需要的目標和結果內容，將這些關鍵字轉換成特定的詞，例如：電影院、地方、時間、天氣，送到資料庫比對已調閱出相關的回應主題和相關查詢，最後將結果用預設的詞組合成句子，例如：「板南線 永寧方向」轉換成「搭乘板南線往永寧方向就可以抵達」回傳。

六、架構流程

用戶端必須使用 Android 系統，並且要在有網路的環境下使用，對中研院端使用 Socket 技術傳送，而對伺服器端則使用 HTTP 的 POST 技術傳送，並且在伺服器回傳 JSON 後，可以運有此結果顯示在對話視窗上。

伺服器端提供 Java EE 伺服器 Tomcat 7 和資料庫系統，資料庫內建有預設用來回答使用者的資料，Servlet 在建立剖析樹物件後，開始依照此剖析樹進行分析，依照結果中的關鍵詞語，到資料庫搜尋，並回傳 JSON 格式的結果。

生活小幫手系統架構

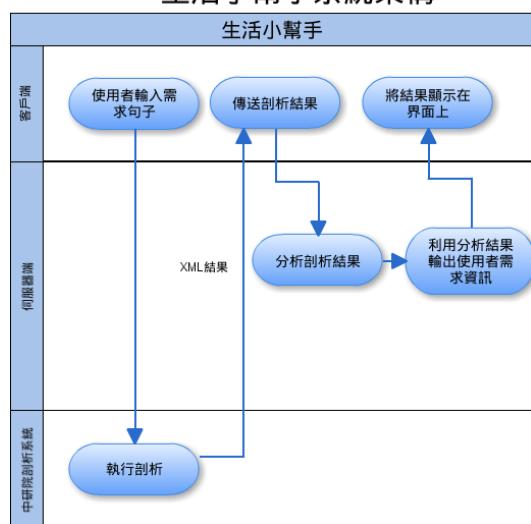


圖 1. 系統運作流程圖

七、預計進行方法

假設使用者輸入：明天的天氣是如何？

一般搜尋引擎可能回應：雨天，20 度。但本系統會回傳：明天有點涼，要記得多穿件外套。明天會下雨，要記得帶雨傘，這類比較人性化的回應。

本專題將利用一些簡單的問題做測輸入，檢驗回傳結果是否真的如我們預期般成功和精準。

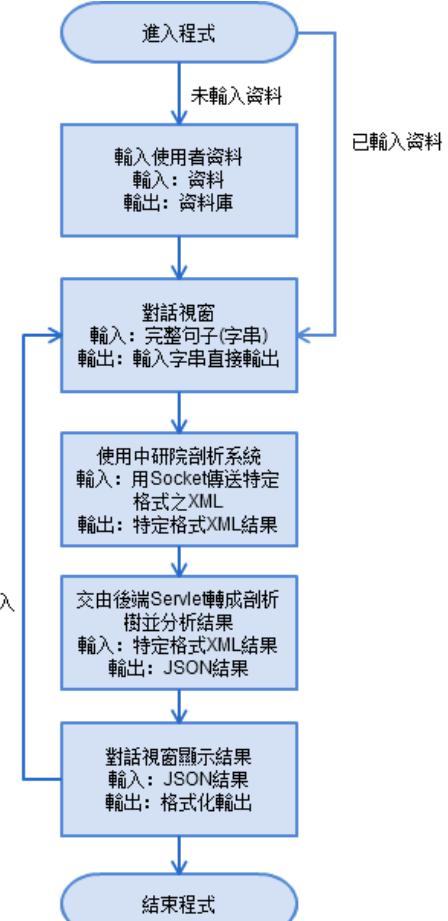


圖 2、系統運作流程

八、參考文獻

- [1]<https://homes.cs.washington.edu/~lsz/papers/fze-kdd14.pdf>
- [2]<http://trec.nist.gov/pubs/trec9/papers/icrjc99a.pdf>
- [3]<http://web.stanford.edu/class/cs224n/handouts/cs224n-QA-2013.pdf>
- [4]<http://www.sinica.edu.tw/as/advisory/journal/14-2/79-82.pdf>
- [5]<https://www.aclweb.org/anthology/O/O03/O03-1013.pdf>