

基於字詞內容之適應性對話系統

MAGEN: An Adaptive Conversational System based on Terms

朱育德、張嘉惠

國立中央大學資訊工程學系

Email: peterajchu@db.csie.ncu.edu.tw, chia@csie.ncu.edu.tw

摘要

在資訊科技蓬勃發展的今日，資訊化與多元化時代儼然來臨；眾多線上資訊服務的崛起，整合服務與人機互動介面成為矚目焦點。對話系統是發展長久的一項研究，分支眾多，其中一類即是以系統代理為主要目標，是一種目標導向性的對話系統。

本文系統 MAGEN 是一強調適應性的目標導向對話系統，採用以字詞為基礎，以系統訂定的詞類為輔助，捨棄分類器與文法資訊，做為對話之依據，解決以往系統採用分類器、文法資訊所造成擴充性不足，成長性受限的窘境。在此情況下，MAGEN 將適應性充分發揮在三個方面。首先，在對話領域上，由於採用詞庫與詞類的設計，可在僅變動存放在資料庫中的知識庫即達成領域移轉，低門檻、低成本，讓 MAGEN 易於適用各種對話領域。其次，對話過程中，不同的對象會有不同的詞彙用語，透過線上擴增的機制，系統將可學習這些詞彙，下次使用者再度使用這些詞彙時，系統將可有效辨識，達到適應使用者的對話習慣。最後，系統本身核心相當輕量，對話皆以文字方式進行，無須圖形化介面之輔助，因此可輕易移轉到不同平台裝置之上。

為驗證三項適應性，設計有實驗項目，以不同類別之主題、雙回合的方式驗證適應性的情況，並實作三種應用形態的系統，更突顯實際用途上確實存有其經濟價值。

1. 緒論

在資訊科技蓬勃發展的今日，資訊化與多元化時代儼然來臨；在手持式裝置躍上網路通訊平台的那刻起，漸趨輕巧的機身成了大眾追求之目標，顯示空間逐漸被剝奪遂成不可違之宿命，同時，何以將大量資訊與使用者形成良好互動，亦成工程師的一大考驗。在此時空背景之下，語音對話系統，再度成為眾所討論之焦點。

語音對話系統，顧名思義是透過語音進行對話之系統，基本上可以拆解成：語音辨識 (Speech recognition)，對話系統 (Conversational system) 以及文字朗讀 (Text to speech) 等三部份[15]。當然，語音之中夾帶有文字所缺乏的資訊，如：語調、情緒等，這些或可增加對話系統之準確性，但主要之資訊還是來自於文字訊息本身，因此，雖說語音對話系統包含對話系統，但一般來說是將語音對話系統視為是對話系統的一種型態。

一個能以自然語言進行溝通之電腦系統是為對話系統，不論對話的對象是人抑或是系統。對話系統的發展已超過卅年，其間不計其數的成果，但萬流同源，一個對話系統通常都會有三個元件：對話管理員 (Dialogue Manager)、知識庫 (Knowledge Base) 以及自然語言了解 (Natural Language Understanding, N.L.U.) [1]，因此常有學者透過這些元件作法之差異，來分類一個系統。

分類對話系統的角度眾多，常見的有：輸入/輸出類型、系統目的、知識庫型態、對話管理員、有無人格特質等，然而，在混合型系統日益增多情況下，分類不再是重點，特質才是受到關注的項目，包括：單一作法是否可同時支援多種語言，能否包容未知詞以及知識庫是否具備成長能力；能力愈強大，即代表此一系統的彈性愈佳，能應付較多情況。

對話系統中，有一類的系統即是以資訊查詢、工作委派等系統代理為目的；其對話目標是為達成一需求，對話過程是以收集資訊為主，而在資訊齊全後，發出任務委派，並等候後端代理人回傳結果。由於其對話是針對特定目的而進行，

所以可稱為目標導向之對話，而以進行此一對話模式所設計之對話系統，可稱為：目標導向對話系統。例如近年來開放式代理人架構(Open Agent Architecture, OAA) [6][16]興起，軟體工程界興起了轉譯與整合這些資訊系統的風潮。由於Web Service[9]並不涉及使用者介面，僅透過XML傳遞執行呼叫與執行結果，也因此，對於使用者而言，即便找到所需之服務，使用上依然存在些困難之處，因此以對話形式做為Web Service之人機互動介面，似乎是一個可行的方法。

本篇論文目的在建立一名為MAGEN之適應性目標導向對話系統。我們希望MAGEN是個高可塑性的系統，在目標導向為前提下，可應用於各種領域，可學習住各種詞彙，可移植到各種平台，深度適應各種情境，而不輕易被時代洪流所淹沒。為此，知識不能以固定的方式儲存，其儲存方式必須能支援系統在運作過程中，從線上所習得的新知識。

以OAA架構為例，Web Service提供的服務眾多而多變，橫跨各種領域，一個對話系統要全然適切其中，除非該對話系統沒有領域限制，亦即具備各領域之詞彙與知識，否則使用者的意圖將無法理解；而若將領域固定，雖是能簡化難度，但同時也大幅降低一個對話系統的存在價值，亦違背我們對適應性的期望。是故，MAGEN在可支援的對話領域上採取折衷之道：易於增添知識領域，藉此，讓MAGEN可以隨需求輕易增加對話領域的知識，而另類的突破領域限制；為此，一個可變動的知識庫更是絕對必要之因素。

自然語言分析器是建構對話系統的根本與基礎，唯有解析來源資訊才能進一步產生回應，而且是即時的回應；過長的回應時間將會讓使用者感到不耐，進而對此系統卻步。因此，選擇適當且適合的自然語言分析技術也是一大重點所在。

2. 相關研究

對話系統的發展已超過卅年[23]，發展出之系統繁雜，焦點與技術各具千秋。概而觀之，可將這些研究分成三種類別：技術導向、擬人導向，以及目標導向等。技術導向的研究並不在提出一個對話系統，而是提出新的技術讓現有之對話系統能夠更強化；從早期的語意分析器(Semantic Parser)如CKIP system、中

期的機械學習法一直到近來的多重對話序等，這些技術並非對話系統專屬，卻是影響深遠。

擬人導向之系統，是以擬人為目的，著重在知識推論以及人格創造。早期，60年代，著名的 ELIZA[22]以專家系統為基礎，使用「知識網絡」達成知識推論與對答，深切影響日後人機對談的發展。1991年，人工智慧領域的 Loebner Prize (www.loebner.net/Prizef/loebner-prize.html) 正式開辦；透過杜林測試 (Turing Test)，尋求最接近真人的系統，十數年來，總有一定數量的系統參加，顯見這部份的發展仍有一定的空間。2005年的冠軍系統，George (<http://chat.jabberwacky.com>)，由 Rollo Carpenter 所打造，是目前最接近真人的系統；在無法完全了解使用者意圖的情況下，仍能以似是而非的回答與使用者互動，也因此獲得各方好評，不過若是使用者一直無法得到正確的回應，終會失去對話的興趣。

與 ELIZA 同一時期另一個系統則是 PARRY[8]；與 ELIZA 不同，PARRY 捨棄知識了解，僅以關鍵字詞來進行解析，搭配「詞庫集合」，尋找適當字詞，並透過程式組成回應句，講求的是快速反應以及對話進行的流暢度。在擬人導向的發展上，PARRY 的影響力遠不如 ELIZA，卻在事隔廿年之後，深刻影響目標導向類的對話系統。

目標導向的對話是以一個目的為目標，而對話內容即是朝向這個目標而前進，因此常結合多媒體影音、代理人、資訊系統等而成，換言之，對話系統成了一種新的操作介面。既然是以介面為目的，深度知識的推論成了非必要之元素，反而了解使用者之目的與整體的流暢性，才是首重之任務，PARRY 的作法在此處，正式發酵。近年來資訊產業的發展注入新思維，以服務為出發點，以體驗為操作介面，軟體強調親合力與智慧性，加上智慧代理人與手持式系統的風行，引發對話系統大舉朝向目標導向發展。

在分類這些具對話目標的系統上，通常依據「對話控制方式」；J. F. Allen 認為可將系統由簡而難歸納成 Finite-state Script、Frame-based、Sets of Contexts、Plan-based Models 以及 Agent-based Models 等五種類型[2]；McTear 則是進一步

整理統併成三類[17]：有限狀態基礎（Finite State- based），語意框架基礎（Frame-based）以及混合主控（Mixed-initiative）。這些方法並無所謂絕對優劣之分，僅有特質上的差異；依不同的應用情境，選用適當的技術才是重點所在。

- *有限狀態基礎*

顧名思義，這類型的系統背後都存在有一些事先建立之有限狀態圖，以之引導整個對話的進行；使用者依系統發言做適當回答，通常僅是一個短辭或是單字，故有限狀態基礎之系統不需具備強大的自然語言了解能力，系統建置門檻較低，相對的，過於複雜的對話主題，不易繪製有限狀態圖，故不適以之建構。自動銀行系統[17]是此類型中，相當典型的系統，對話流程完全必須依照事先定義好的有限狀態圖進行。在台灣，這類型系統常見於電話語音系統[12][14]，如：醫院的語音掛號系統等，雖使用者僅透過鍵盤傳送訊息，看似與對話有所差異，但就資訊交換的角度來看，確實是一種對話模式。

- *語意框架基礎*

基於語意分析來產生回應句；在最原始的系統中，對話系統僅對句子進行解析，目的在找出主詞，動詞，受詞等部份，然後透過知識庫與文法的輔助來產生回應句；這類型的系統是由使用者來引導對話，彈性遠比有限狀態基礎之系統來得大。飛利浦自動火車時刻資訊系統 [3]是 1995 年由 Aust 等人製作；這是一個以語意框架為基礎的著名系統，用來查詢火車時刻，使用者描述所要搭乘的日期時間、起點、終點，系統則會回傳查詢的結果。

具意圖萃取之智慧型醫療對話查詢系統[4]是台灣成功大學陳銘軍在 2003 年完成的一套整合多項服務的醫療查詢對話系統，整合的服務有掛號諮詢、科別資訊諮詢以及常見問答集諮詢三項大服務；透過分析語意框架內容以及目的偵測來了解使用者目標，然後透過島嶼演算法的方式半自動建立出醫療概念模型。

- *混合主控*

某個角度上來看，混合主控即是有限狀態和語意框架之結合，同時具備引導

對話的資料結構以及語言分析能力，在此種混合式系統下，問與答的界線不再如同前兩種類型般的清晰，且對話的控制權也不再由任何一方獨佔，而是共享。

混合主控，此一名詞早在 2000 年以前即已提出[7]；當時的定義其實並不明確，僅言及對話主導權並不完全由單一者控制，而此定義可延伸成對話控制權由多個對話系統共享、由多個對話主題共享、由使用者與系統共享、由多種對話控制方法共享……等等。不論何種共享方式，其對話內容就會如同兩位代理人間的對答一般，因此 McTear 將此類系統改稱為代理人基礎 (Agent-based)，然而，此名稱卻易與結合真正代理人之對話系統的名稱相混淆，故至今仍有多篇文章採用原始名稱，本文亦然。此類系統由於自由度甚高，對話彈性也相當受到肯定，所以近年來的對話系統大致以此類為主。

ISIS[18]，2004年香港中文大學，Meng等人提出的系統；這是一套線上股市資訊系統，可以查詢股市資料以及代理下單，能接受中、英文，語音上更含括廣東話。作法上採用文法分析，推論目的以及隱含的資訊；功能委派上採用開放式代理人架構，可和其他現成的代理人接軌，達到功能擴充之目的。對話系統雖然可以擴充，但僅限於知識庫，亦即具有未知詞處理和詞彙成長能力，但文法推論規則和對話目的則是固定不變，在對話領域上的移轉具相當的困難度。

基於統計式語意相依關係之對話語句理解系統[24]是另一個成功大學在 2004，楊茂柱提出的系統；此系統使用一個新的自然語言理解架構，在此架構中利用語意相依關係幫助語意理解，主要精神是考慮到語句中詞與詞之間可能所隱含之語意資訊而不只是考慮到詞的表徵意義，方法為使用語句結構和語意概念資訊當作語意相依分析之依據。此外還加入對話歷史的觀念，考慮對話語意脈絡，幫助對話理解。在理解過程中使用語意相依圖取代語意框架，避免人為介入定義，提高系統移植性。然而，全系統過度依賴歷史對話，反而使人在領域移轉時顯得心有餘而力不足，僅能拚命收集對話紀錄以餵入分類器中，提高準確性；同時，由於系統相當依賴分類模型，導致對話過程中的知識沒有即時進行回饋，彈性上的缺陷暴露無遺。

智慧型個人助理[19]是一個建構在PDA之上的對話系統，由Nguyen等人在2005年提出；由於是完全建構在PDA之上，系統所受限制較多，比方說，知識庫與對話目標偵測都必須精簡，同時系統彈性也受到相當嚴峻的挑戰，這也是在手持式系統上建構對話系統的最大問題。此系統採用樣板的作法，透過關鍵字詞比對以及較多次問答，取代複雜的邏輯推理以及機率統計運算，達成速度上的優勢以及彌補知識庫精緻化後的缺陷。整體來說，系統的彈性很低，主要受限於知識庫無法成長，雖有未知詞的處理，但僅限於當下使用，換言之，整個系統完全沒可成長的部份。

2.1 系統比較

表1整理出三種系統和四個評估項目的比較結果。基本上，愈複雜的系統彈性愈好，回應速度愈慢，但請注意，比較是相對性的，實際上，一個對話系統的反應時間都會是在可容忍的範疇之內。對話進行的主導權代表主導對話的走向；在有限狀態基礎和語意框架基礎相當明顯，各為系統與使用者，而在混合起始的系統上則屬於共享。系統必須同時具備有問與答的能力，雖說是共享，卻也不見得是五五平分，通常是系統佔七分，使用者佔三分，換言之，大部分時間仍是由系統在引導整個對話的走向。

由於MAGEN強調適應性，所以這裡不免要歸納整理一下，先前系統在這方面的表現與成果。首先，提到目標導向，直覺必然想到分類器，利用分類器來判斷使用者的目標為何，如同楊的系統[24]一般；然而，分類模型卻會對線上成長造成限制，或許這會是提高辨識率的好主意，但是同時也犧牲了系統的學習性。說到自然語言分析，文法是最常被提及的資訊，ISIS[18]即是立於此點之上，然而產生文法規則並非易事；不論是從歷史對話紀錄萃取文法規則，或是人工自訂，都是偌大工程，尤其是一個嶄新系統要上線時，根本無歷史對話紀錄。在系統平台上的移轉除了要考慮作業系統的差異外，儲存媒體的空間也是一個考驗，例如：要移轉到PDA之上，就不得不考慮知識量與對話流暢性的最大效益比。

表 1：系統比較

	有限狀態基礎	語意框架基礎	混合起始
建構容易度	容易	普通	困難
回應速度	較快	普通	較慢
系統彈性	較低	普通	較高
對話主導權	系統	使用者	共享

3. 系統介紹

本篇系統，MAGEN，是以混合起始為出發點的混合式系統，系統架構如圖 1 所示，包含對話管理員與知識庫；對話管理員是系統的主體，負責對話的進行，由七個元件組成；知識庫是系統的靈魂，以資料庫為儲存媒體，提供對話管理員了解對話以及產生回應的資料來源。

適應性是MAGEN主要強調的重點，為此，我們選擇以資料庫來儲存知識。由前文的歸納可知，若想讓系統的適應性充分發揮，就必須要：避免使用文法資訊、不要依賴對話紀錄、能避免使用分類器就避免、儘可能讓人能夠直接修正知識內容、有限度的知識成長以及簡化運作的複雜度。扣除這些內容，剩下的就僅有字詞本身。以字詞為基礎，知識成長將變得容易，建構出之腳本將是易於閱讀，當然也就能夠由人所訂定，也因此移轉對話領域就輕鬆許多。知識以字詞為基礎，此舉引動自然語言解析器的選擇；顯然，基於文法規則的語意分析器（Semantic Parser）不再符合需求，反而是淺層分析器（Shallow Parser）的斷詞系統會是首選。

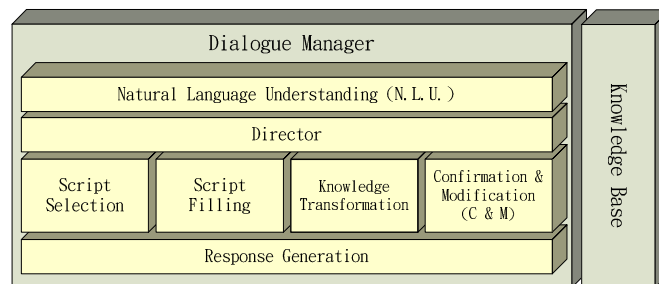


圖 1：MAGEN 系統架構圖

淺層分析器僅對字面做解釋，而忽略背後文法資訊；相較於語意分析器（Semantic Parser），使用斷詞系統不僅喪失文法相關資訊，並且詞彙散亂，但請注意，這是針對正規文句而言；一般口語並非如同撰文一般注重句法結構，尤以今日網路次文化的盛行，遣詞用語愈趨雜亂無章，在此情況下，額外考慮文法資訊，分析效果不見得會比較好，但較為費時卻是可見的。採用斷詞，捨棄可能錯誤的文法資訊，僅將句子視為字詞堆砌，專注於文字內容之上，反是上上之選。同時，因為斷詞不在乎文法，故當句子中出現多語言夾雜出現，這種令各對話系統棘手的問題亦能一併解決。[11][12]

僅以斷詞來解析自然語言顯然是不足，系統是難以理解世上多變複雜的字詞各代表何種意義，有鑑於此，我們設計了「系統詞類」（TermType）來輔助系統了解各種字詞所代表的概念為何，就如同WordNet裡的Concepts一般，唯詞類是客製化的存在，並非固定不動的各種類別，雖然看似貧乏，卻會是最適切對話領域的存在，讓知識庫不至於過度龐大，除了加快系統反應時間，同時也有效節省知識儲存空間。

添加了系統詞類會不會造成腳本的撰寫與詞彙的學習產生困擾？詞類是詞的概念，比之文法，定義詞的概念應是更加簡單，而有了概念的輔助，相信腳本的撰寫只會更清晰，不會造成太大的困擾才是。因此，建構出來的腳本是相當人性化，不僅易於閱讀，更能輕易異動之；這關鍵的特性，使得領域移轉成了簡單的事情，僅需建構腳本即可，而且MAGEN具有線上成長的特性，無需事先準備過多的字詞知識即可順利運作。

線上成長儼然是目前對話系統備受考驗的地方，而MAGEN也具備此一能力。MAGEN會在對話過程中偵測出新的字詞，並且透過使用者給予定義，更新知識庫，達到學習新字詞的目的；然而，僅如此的話，對話將會相當令人煩悶，尤其在系統尚未習得夠多字詞時。有鑒於此，我們不只設計有學習新字詞的功能，更搭配有推論機制，讓系統先行猜測此未知詞的可能意義，並透過隨後的對話與腳本之記載來確認此推測的正確性；此一作法將可有效降低對話的繁雜度。

最後，為加快系統回應速度，除了實作上的最佳化外，邏輯設計上也有加速的空間。以往的對話系統會經常對於使用者輸入之句子進行目標偵測（Conversational Act Detection）[20]。MAGEN完全不偵測使用者該次輸入目的為何，而採用預設立場的方式，直接認定當次輸入之資訊應當為何；由於，系統與使用者的互動進程皆是透過此一資訊進行，因此，若使用者沒有刻意提供系統錯誤資訊，該預設之狀態不致會錯得離譜。如此作法固然會侷限對話的自由度，無法讓多個主題交互進行，但我們認為「一次一件事」的這種限制，並非太過，屬合理範疇；然此一限制，卻能讓系統執行速度與對話精準度大幅提升，利弊得失，淺顯可見。倒是使用者可以一次將多個資訊告之MAGEN，來省對話回合數。

3.1 知識庫

系統進行對話的依據，以資料庫儲存，可分成用來辨識對話內容的詞庫與引導對話進行的腳本。腳本其實就是對話目的之紀錄，之前的系統需要辨識使用者的需求，而MAGEN就是要辨識對話目的，選擇正確的腳本，以引導對話進行，並指示系統該收集何種資訊，相當於有限狀態圖和樣板的角色。

使用資料庫儲存之知識庫具有成長的能力，能夠在系統運作中，透過對答或是操作介面進行成長，達到契合不同使用環境。初始時，這些資料須依靠人工建立少量部份，然而資料量之多寡並不影響系統的能力，僅會使對答流暢性較低，例如：過多未知的字詞就會造成系統不停的進行未知詞概念的確認，但是仍能完成對話目標。知識庫包含詞庫、腳本兩個部份，以下分別介紹如下：

- 詞庫

詞庫預設有十六個系統詞類(Termtyp)，其名稱與範例如表 2 所示；每個詞都會對應到一個或多個類別，例如：再見<結束>，或中央大學<地點/組織>。若詞彙難以歸類，則歸入<其他>，這些詞彙通常表示比較不重要的詞彙，或是比較罕見的專有名詞。<事件>是表示具備動作性質的詞彙，通常也是腳本的觸發關鍵詞。<忽略>則表示該詞屬於可以忽略不計，只是發語詞或是贅詞。除了出

現在詞庫中的十六個詞類外，系統運作過程中尚有一類，<未知>，此類別用來標示未出現在詞庫中的字彙，亦即未知詞；此類詞彙在系統運作過程中，若被確認類別，該詞將會加入詞庫之中，日後再度出現時，就不再是未知詞。

而詞庫裡的字詞不該單從字詞去觀看，應該將之視為「具有詞類的字詞」，例如：我們應該將中央大學<組織>與中央大學<地點>視為兩個不同的「具詞類的字詞」。在這個觀念下，當新增了一個類別<學校>時，應該將中央大學<學校>視為是一個新字詞，而不應認為中央大學<組織/地點>是個錯誤的字詞。

表 2：詞庫類別與範例

問候：唷，早	肯定：是，OK	事件：吃飯，查	數值：19，21.3
情緒：^^，呵呵	否定：不是，否	物品：車，電腦	時間：日，：
結束：Bye，再見	疑問：嗎，？	地點：中央大學	組織：中央大學
忽略：和	驚嘆：吧，喔	人物：趙建銘	其他：反托拉斯

- 腳本

腳本是對話進行的重要依據，當中紀錄所代表的對話情境，以及所需資訊等，換言之，要進行怎樣的對話，僅需製造出相對應的腳本即可。本文以『』代表腳本，例如：『邀約』代表邀約此一腳本。腳本格式如圖 2。

Script	欄位名稱	說明	值域
Identification	腳本識別資料		
NO.	唯一的編號		01 \ 02 \ 03 ...
Name	腳本主題名稱		邀約 \ 訂票 \ ...
Event-Trigger	觸發腳本之條件		
Tone	語氣		肯定 \ 疑問 \ 不拘
Trigger Words	觸發關鍵詞		參加 \ 訂購 \ ...
Response	填充完成後之指示		
Mode	執行模式		立即回應 \ 不予回應
Chain	連鎖觸發		01 \ 02 \ 03 ...
Slot #1	腳本所需之資訊		
Name	插槽名稱		出發地 \ 日期 \ ...
TermType	值型態		地點 \ 日期 \ ...
Necessary	必須與否		是 \ 否
Candidates	候選值		華航 \ 長榮 ...
Slot #2			
Name			
TermType			
Necessary			
Candidates			
:			

圖 2：腳本格式（左）與腳本內容說明（右）

3.2 對話管理員

對話管理員負責邏輯處理，由七個元件所組成，包括 N.L.U、Director、Response Generator 以及四個狀態的處理程式：Script selection、Script Filling、Knowledge Transformation、Completion 等，其運作流程如圖 3 所示。當對話管理員在收到訊息後，首先進行自然語言了解，而後根據 Status 裡頭所描述的狀態來決定走哪個路徑來處理當下的對話狀態，不同的路徑由不同的元件負責，最後由對話產生器依據對話狀態產生回應句。針對自然語言了解部份，分為中文斷詞，字詞類別標示，字詞群組化，未知詞標示等四項工作。

1. 斷詞：根據[5]所簡化之斷詞系統。
2. 字詞類別標示：將斷詞結果，依「詞庫」，內容標示對應詞類；例如，「中央大學」<組織/地點>，表示此字詞可為組織或地點兩種系統詞類，對於可能造成的模糊則以大膽假設及回問使用者來處理歧義。當有詞彙沒有出現在詞庫中時，此一詞彙稱之為未知詞，此時標示成<未知>。

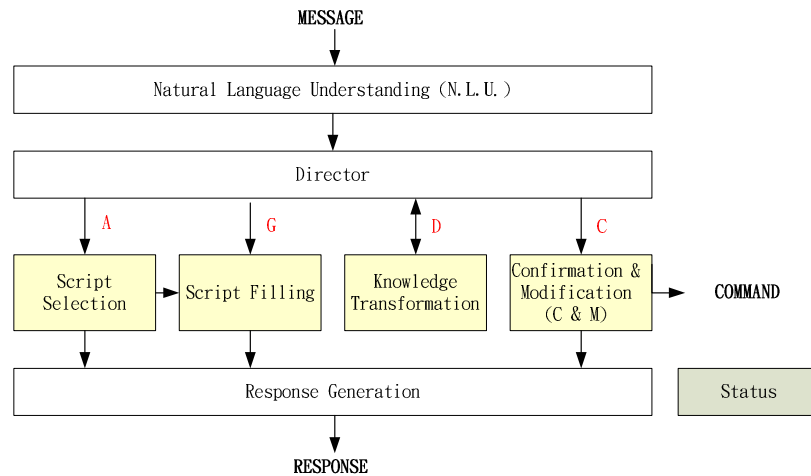


圖 3：對話管理員運作示意圖

3. 字詞群組化：經過斷詞後，句子成了一堆字詞的堆砌，但別忘了，這些字詞原本存在句子中的順序是有其意涵的，相鄰的字詞存在有相當的關聯性，因

此若將這些關係群組化起來，將可提升系統了解所接到之資訊。群組化分成同質群組與異質群組兩部份先後且循環執行，直到沒有任何相鄰字詞可以進行群組化為止。同質群組是將具有同樣詞類的相鄰字詞結合，異質群組則是透過特殊規則將相鄰字詞合併；在尋找可結合之相鄰字詞時，<忽略>一類之詞會被忽視，但仍會保留在群組化的結果之中。舉例來說，「讓我取消訂票行不行？」，經過斷詞與詞類標示之後為：讓<事件/其他>我<人物>取消<事件>訂票<事件>行<肯定>不行<否定>？<疑問>；「取消」與「訂票」這兩個詞首先會在首輪的同質群組時被組合，隨後，「能」與「不能」也在同輪的異質群組被合併，形成：讓<事件/其他>我<人物>取消訂票<事件>行不行<疑問>？<疑問>。因為還有可以群組化的字詞，所以進行第二輪群組化，形成：讓<事件/其他>我<人物>取消訂票<事件>行不行？<疑問>。結束後，已無任何相鄰字詞可以群組化，因此，群組化在第二輪後結束，並輸出結果。

4. 未知詞標示：當有詞彙未出現在詞庫中時，該詞所標示之詞類為<未知>；在群組化的過程中，所有相鄰的<未知>詞亦會被群組化成單詞，這些單詞在每輪群組化結束時，會反覆查詢這些組合後的詞彙是否出現在詞庫之中，若有，則變更詞類。

3.3 對話狀態 (Status)

對話管理員以對話狀態記錄目前情況，對各元件而言，狀態就如同是任務指示，標示著目前該執行的任務。對話狀態總共有 19 種(如表 3 所示)，可依其類型分成四類：《Assignment》、《Definition》、《General》、《Complete》。圖 3 中，Director 元件周圍的英文字母，A、D、G、C，即表示各 State 之縮寫。運作時，Director 會根據目前 Status 最上層狀態 (Top State) 決定將控制權移交給哪個元件，例如：A 就交給 Script Selection，G 就交給 Script Filling。

N.L.U.與狀態並無任何關聯，Director 以及 Response Generation 這兩個元件對於 States 是以偷窺 (Peek) 的方式，而非使用彈出 (Pop)，亦不會對狀態產生任何變化，其餘四個元件，則視情況異動狀態。

各種狀態轉換條件如圖 4；此圖為簡化資訊，僅針對四大狀態，不計其參數變化。雖系統運作可用有限狀態圖表示，但不代表 MAGEN 應歸類為有限狀態基礎之對話系統；相關研究中所提有限狀態基礎系統，其有限狀態圖乃應用於腳本，對話僅依該圖進行，不同之對話有不同之狀態圖，MAGEN 將有限狀態圖做為運作基礎，此一狀態圖乃固定不變之存在，是假定使用者所給予之訊息應代表何種涵義，用以引導系統運作，雖同屬有限狀態圖，用途、目標卻大不相同。

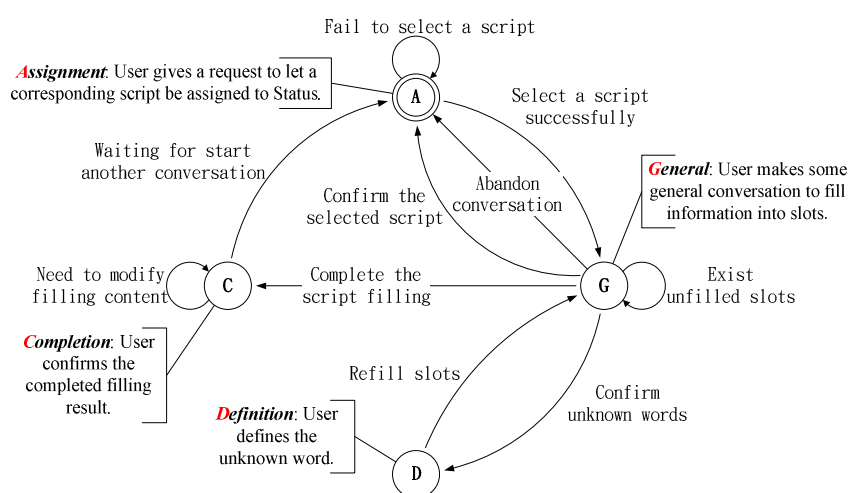


圖 4：狀態變化移轉圖

回應產生(Response Generation)會根據當時的 State 尋找對應之句法，結合狀態資訊，產生回應句。不同之 State 定義有不同的句法格式，如表 3 所示；各變數意義如下：{Script}表目前選定腳本名稱，{ScriptList}表多個腳本的名稱，{Slot}表狀態格式指定之插槽，{Slot.TermType}與{Slot.Name}各表插槽中的詞類限制與識別名稱，{Content}則是整個填充內容所改寫之確認句，而{Trigger word}表示狀態中被保留的字眼。另有一個函數：ASK()，此一函數是透過{Slot}的識別名稱、限制詞類以及候選值等資訊改寫成詢問句，如一個型態值為地點的插槽，產生的回應為：集合地點在哪裡？若是此插槽具有候選值，則以插槽名稱配合候選值產生回應句如：火車種類是自強，莒光還是復興？

《5》與《6》僅對用以內部狀態之識別，對使用者而言並無分別，故合用同一回應句法。《12》，《14》與《15》亦屬內部狀態識別，與使用者實際互動之對話狀態為《18》，《19》以及《9》。

表 3：狀態格式暨回應句法明細

編號	類型	情境描述	回應句法
1	A	初始對話	您好，我是 MAGEN。請問您有什麼需求？
2		詢問是否還要繼續對話	還有事嗎？沒事就先這樣囉！
3		結束對話	感謝您的使用，祝您今天愉快。
4		放棄對話	您的問題超出我的理解範圍，或許您該嘗試其他服務。
5		選取失敗	我的回答很有限，你必須問正確的問題才行。
6		選取失敗但保有關鍵詞	
7		多重腳本中選	我不太清楚，您是要進行 {ScriptList} 哪個話題？
8		詢問增添腳本關鍵詞	確認一下，「{Trigger-word}」與「{Script}」有關係嗎？
9		詢問對話目標是否正確	等一下！您確定我們進行的對話目的是：{Script} 嗎？
10		確認是否連鎖觸發主題	嗯，那您接下來要進行 {Script} 嗎？
11	G	提示性填充	ASK ({Slot})
12		針對未知詞填充	
13		二次提示性填充	請您認真一點！ASK ({Slot})
14		確認性填充	
15		指示對話目標確認	
16	C	首次確認填充內容	OK！跟您確認一下剛剛的資訊{Content}
17		確認修正後之填充內容	修正後的資訊是：{Content} 請確認。
18	D	確認未知詞之類別	請問「{Term}」是指「{Slot.TermType}」嗎？
19		確認填充內容之類別	確定 {Slot.Name} 是 {Term} 嗎？

3.4 系統運作

MAGEN 的運作與狀態移轉息息相關，因此本文將由圖 4 進行說明。MAGEN 是目標導向之對話系統，因此首先必須偵測使用者之目的，狀態為《Assignment》。目標偵測時，有多種例外狀態，例如：多腳本同時吻合、具有新觸發字須確認添加，這些都會在此狀態中完成；在目標確定之後，狀態移轉至《General》進行腳本填充。

依據所擇定的腳本插槽進行詢問，直到完成後，交由使用者確認；同樣，填充過程中可能會有多种例外狀態，例如：辨識失敗、未知詞干擾等，這些都需做

進一步確認。當然，系統在第一時間中，並不會立即發出確認句，而是採用假設與推論的方式猜測未知之資訊，例如：當系統詢問地點時，使用者回應一未知詞，此時，系統將大膽假設該未知詞意指地點；當然，在無法猜測之時，還是需要透過《Definition》進行未知詞的確認。填充時，未免因腳本設計有誤或其他因素導致對話卡死，單一插槽最多僅嘗試填充五次，超過，將會中止該段對話，當然，使用者也可以隨時提出中止對話之要求，在確認之後即會中止。終止之後，狀態回到《Assignment》，並等候下一個對話主題的開始。

填充過程中，包含有系統的判讀與猜測，因此不可能全然無誤，因此在完成填充之後，狀態移轉為《Completion》，將填充內容改寫成確認句，交由使用者進行確認或是修正。由於填充過程採用詞類為填充之依據，若採用相同策略進行修正，恐難有良好效能，於是我們改採以內容為修正策略。以〈肯定〉{內容}〈否定〉{內容}的句法為判定，例如：是今天晚上，不是明天晚上，此時使用者表示已經填充的明天晚上有誤，應以今天晚上取代。

待使用者完成修正，確定填充內容之後，檢視該腳本是否設定有連鎖觸發腳本資訊，若有進行確認，是否要直接進行該話題，若無則回到初始狀態；不論何者，系統將是把狀態移轉至《Assignment》等候下一段對話的開始。

3.5 對話範例

MAGEN 是以 Microsoft® .net framework© 平台實做，可移植至多種裝置之上；表 4 為一完整對話範例，並附加重點說明，加深運作流程與實際對話之聯結。除了將對話核心實作完成之外，更實際完成應用形態。藉由開放式元件 DotMSN 的協助，成功與 MSN Messenger 進行溝通。搭配 IBM ViaVoice 與 Microsoft Speech 讓 MAGEN 擁有語音輸入輸出的效果。當然，以 MAGEN 的特性，移植到 Pocket PC 之上更是絕對沒有問題。為了更清楚展示 MAGEN，我們實錄兩段展示影片，可於 <http://140.115.155.87/MAGEN/DemoVedio.htm> 下載。

表 4：對話範例

User	唷！要不要去吃個飯？
MAGEN	地點在哪裡？ 【說明】「去吃個飯」一詞觸發『邀約』，緊接著首次嘗試填充腳本；成功填充事件插槽為「去吃個飯」。
User	魔法咖哩，西門店
MAGEN	參予的有誰？ 【說明】成功填充地點插槽，接著詢問參予人士。
User	我
MAGEN	時間是什麼時候？ 【說明】代換指稱詞，「我」將被換成「你」再進行腳本填充。
User	週末晚上，如何？
MAGEN	OK！跟您確認一下剛剛的資訊：週末晚上，在魔法咖哩西門店，和你去吃飯。這樣對嗎？ 【說明】『邀約』填寫完成，改寫填充資訊為確認句，並由使用者進行確認與修正的步驟。
User	對！
MAGEN	我會盡快給您答覆。還有事嗎？沒事就先這樣囉！ 【說明】使用者肯定填充內容，加上無連鎖觸發之腳本，故以《2》結束對話。

4. 實驗與分析

評估一個人機互動的對話系統，最直接的方法即是以人來進行評估，然而僅由數人評論，恐有失客觀。未免流於主觀，過程中除了人為觀感外，另有「成功率」、「主題擇定回合數」、「對話回合數」的統計輔助，一來驗證這些實驗對象所言，也提供不同角度之觀點。

「成功率」計算方式為：正確內容之插槽數 / 對話腳本之插槽數，其值愈高愈好。以有四個插槽的『邀約』為例，在修正到最後送出後，仍只有三個插槽填充內容正確，則該對話的成功率為 0.75。「主題擇定回合數」與「對話回合數」皆以「回合數」為計量單位；所謂「回合」，使用者送出一訊息且系統對之產生回應合為一回合。前者表示系統辨識使用者目標所需花費之成本，當然愈低愈好；後者表示對話之經歷，從問候語開始到送出外部命令為止，此數值會與插槽數有高度相關性，但基本而言，數值愈低表示愈快達成目的。

實驗採用透過 MSN Messenger 的方式進行，並事後透過網路問卷回收其滿意度。我們從 MSN 聯絡人中挑選出八位從未接觸過 MAGEN 之對象進行「邀約吃飯」、「邀約看電影」、「留言」、「問題詢問」以及「查詢網站資料」等五個主題之對話。

實驗過程中，八位對象各自獨立進行。知識庫中存有四張腳本以及 1327 個具詞類的字詞（同一字詞，不同類別視為不同之具類別字詞）；這些詞彙是透過人工定義的方式與先建立各種詞類常見之字詞，但刻意忽略與電影相關之字詞。

五個對話主題中，「邀約吃飯」與「邀約看電影」共用『邀約』，進行不同之邀約行為，是故腳本僅有四張；旨在探究能否讓多個近似之主題共享單一腳本，且由於詞庫缺乏電影知識，此情況下，對話是否能夠進行。由表 5 可發現，「邀約看電影」確實是五個主題中，數值表現最差之主題，然整體成功率卻仍有 0.76，顯見此等模糊之主題，系統依舊能夠與使用者進行對話。

「問題詢問」的主旨在查詢系統內部所擁有之知識，而「查詢網站資料」則是針對外部如搜尋網站進行遠端資訊查詢，目的近似，唯所需資訊不同，後者須知搜尋網址，插槽數比前者多一。此二者不論在目標或是所需資訊皆極為相似。由於 MAGEN 棄分類器而以觸發關鍵字與語氣詞偵測使用者之對話主旨，因此在辨識這類主題時，的確會產生困擾並且付出較多之回合數為代價(表 5)。

人機互動，人之情感亦不可少，因此，增加「滿意分數」之評量標準。其值由一至十，以十為佳；表 5 中，此數值與成功率有正向之關係，可見使用者之滿意程度多半與對話之完成性相關。

MAGEN 是個強調高度適應性之系統，因此，實驗的另一個重點即在於學習能力。為此，各對象再次對五個進行另一個循環之對話，此為循環二，其值如表 6。由平均值來看，第二循環之數值皆比第一循環來得好，表示系統與使用者間確實達到某種程度之契合。系統之所以能夠適應使用者之對話，全然是因為字詞的成長所致，而 MAGEN 的字詞成長有兩部份，一者，詞庫，另者，觸發字。在循環一中，總共成功添加 16 個新字詞，以及 6 個觸發字；循環二中，依舊有

16 個新字詞成功添加，而僅有 1 個新的觸發字。

表 5：循環一評估結果

評量 主題	滿意分數	成功率	主題擇定 回合數	對話回合數
邀約吃飯	8.50	0.95	1.38	10.5
邀約看電影	5.63	0.76	1.63	9.50
留言	9.25	1.00	1.88	4.13
問題詢問	9.13	0.88	2.25	5.13
查詢網站資料	7.25	0.94	4.50	9.25
平均值	7.95	0.91	2.33	7.70

表 6：循環二評估結果

評量 主題	滿意分數	成功率	主題擇定 回合數	對話回合數
邀約吃飯	9.13	0.98	1.00	6.63
邀約看電影	7.83	0.93	1.75	9.75
留言	9.38	1.00	1.13	3.88
問題詢問	9.38	1.00	1.75	3.75
查詢網站資料	9.00	1.00	1.38	4.63
平均值	8.85	0.98	1.40	5.73

近似的「問題詢問」與「查詢網站資料」兩主題在循環二中的表現明顯優於循環一，尤其是「查詢網站資料」在主題擇定回合數與對話回合數上大幅減少。這是系統成長後的結果嗎？「問題詢問」在循環一中，並沒有增添任何新的觸發字，然而在循環二中，主題擇定回合數卻以 1.75，低於循環一的 2.25，表示系統花費更少的回合數正確辨識出使用者之目的，顯然是使用者的適應性，是使用者有經驗之後，能夠以更快速的方式與系統對話，相同現象也發生於「留言」。「查詢網站資料」的成長性在觸發詞的成長、未知詞數量減少以及使用者適應等三個現象同時發生情形下，是為五個主題中最顯著之主題。

「邀約看電影」不論是「主題擇定回合數」或是「對話回合數」都出現異於其他四個主題之「負成長」現象。究其緣由，一般邀約看電影時，所用之詞通常不會有「電影」一詞，較常直接講述電影片名；此現象讓選擇腳本時產生困擾，

自是少不了需多耗幾個回合。再者，電影相關的字眼多變，詞庫中又相當缺乏，循環二的未知詞數量不僅沒有低於循環一的 6 個字，更增加到 7 個字；額外的未知詞確認，自是讓對話回合數向上攀升。反觀「成功率」與「滿意分數」，在循環二比循環一高的情況而言，字詞確認並不影響使用者對於對話的感覺，同時成功率攀升至 0.93，顯見在循環一的成長後，對話變得更加容易。

5. 結論與未來展望

本篇論文所建構之系統，MAGEN，是一個有目標之對話系統，跟以往的系統不同，以屬於淺層分析器的斷詞系統為出發點，配合上群組化技術、詞類標示等技術，在對話是不合文法、中英字句夾雜的情況下，自然語言辨識能力略勝語意分析器一籌。由於將對話資訊由文法句型回歸到文字內容本身，因此在知識編輯上更加自由、簡單，這使得對話領域的移轉、線上知識成長變得更加簡單，在應用情境與對話風格的適應性上都有傑出表現。

線上知識成長通常需要使用的一些確認來配合，這很容易讓系統變得很囉唆，可是不足の確認又難以解決模糊不解之處；MAGEN 將所有確認機制完整建立，但搭配推論與猜測的策略，讓某些確認機制作而不用，既可保證不會有模糊地帶且不會有過於囉唆之對話發生。

MAGEN，強調高度適應性，在三個層次中充分展現。首先，以斷詞系統為出發點之下，移轉對話領域變得容易，使得系統得以適應在各個對話領域之中；其次，線上成長搭配詞類推論，系統得以在對話過程中，漸漸適應使用者的語言風格，進而成長；最後，以 Microsoft® .net Framework©實做，搭配三種版本的對話核心，不論在手持式環境，在語音環境，亦或是網路通訊環境，MAGEN 都可以適應其中，達到最大經濟效益。

誌謝

本研究由國科會編號 NSC 94-2524-S-008-002 贊助。

參考文獻

1. J. Allen, "Natural Language Understanding," *The Benjamin/Cummings Publishing Company*, 1995.
2. J. F. Allen, D. K. Byron, M. Dzikovska, G. Ferguson, L. Galescu, and A. Stent, "Towards Conversational Human-Computer Interaction," *AI Magazine*, 2001.
3. H. Aust, M. Oerder, F. Seide, and V. Steinbiss, "The Philips automatic train timetable information system," *Speech Communication*, vol. 17, pp. 249-262, 1995.
4. M. J. Chen, "Intention Extraction for Intelligent Medical Query System," *National Cheng Kung University, Master dissertation*, Jun. 2003.
5. K. J. Chen, S. H. Liu, "Word Identification for Mandarin Chinese Sentences," *COLING*, pp. 101-107, 1992.
6. A. E. Cheyer and D.E. Martin, "The Open Agent Architecture," *Autonomous Agents and Multi-Agent Systems*, 2001.
7. J. Chu-Carroll, "MIMIC: An Adaptive Mixed Initiative Spoken Dialogue System for Information Queries," *The Sixth Conference on Applied Natural Language*, pp. 97-104, 2000.
8. K. M. Colby, "Artificial Paranoia," *Artificial Intelligence*, vol. 2, 1971.
9. K. D. Gottschalk, S. Graham, H. Kreger, and J. Snell, "Introduction to Web services architecture," *New Developments in Web Services and E-commerce*, vol. 41, Nov. 2001.
10. R. Higashinaka, N. Miyazaki, M. Nakano, K. Aikawa, "Evaluating discourse understanding in spoken dialogue systems," *ACM Transactions on Speech and Language Processing*, vol. 1, pp. 1-24, 2004.

11. R. Kaplan, S. Riezler, T. King, J. Maxwell, A. Vasserman, and R. Crouch, "Speed and accuracy in shallow and deep stochastic parsing," *HLT-NAACL*, 2004.
12. C. J. Lee, E. F. Huang, and J. K. Chen, "A Multi-keyword Spotter for the Application of the TL Phone Directory Assistant Service," *Workshop on Distributed System Technologies & Applications*, pp. 197-202, 1997.
13. X. Li and D. Roth, "Exploring evidence for shallow parsing," *The Annual Conference on Computational Natural Language Learning*, 2001.
14. Y. C. Lin, T. H. Chiang, H. M. Wang, C. Peng, and C. Chang, "The Design of Mandarin Chinese Spoken Dialogue System," *International Conference on Spoken Language*, vol. 1, pp. 230-233, 1998.
15. M. Lundeberg, J. Gustafson, and N. Lindberg, "The august spoken dialogue system," *Eurospeech*, 1999.
16. D. Martin, A. Cheyer, and D. Moran, "The Open Agent Architecture: a framework for building distributed software systems," *Applied Artificial Intelligence*, vol. 13, pp. 91-128, 1999.
17. M. F. McTear, "Spoken Dialog Technology: Enabling the Conversational User Interface," *ACM Computing Surveys*, vol. 34, pp. 90-169, Mar. 2002.
18. H. Meng, P. C. Ching, S. F. Chan, Y. F. Wong, and C. C. Chan, "ISIS: An Adaptive, Trilingual Conversational System With Interleaving Interaction and Delegation Dialogs," *ACM Transactions on Computer-Human Interaction*, vol. 11, pp. 268-299, Sep. 2004.
19. A. Nguyen, and W. Wobcke, "An Agent-Based Approach to Dialogue Management in Personal Assistants," *International conference on intelligent user interfaces*, pp. 137-144, Jan. 2005.
20. A. Stolcke, E. Shriberg, R. Bates, N. Coccaro, D. Jurafsky, R. Martin, M.

- Meteer, K. Ries, P. Taylor, and C. V Ess-Dykema, "Dialog act modeling for conversational speech," *AAAI Spring Symposium on Applying Machine Learning*, pp. 98-105, 1998.
21. E. Voorhees, "The TREC-8 Question Answering Track Report," *Eighth Text Retrieval Conference*, pp. 77-82, 1999.
 22. J. Weizenbaum, "ELIZA—A Computer Program for the Study of Natural Language Communication between Man and machine," *CACM*, vol. 10, 1967.
 23. Y. Wilks, "Human-Computer Conversation," *International Workshop on Human-Computer Conversation*, vol. 1, pp. 1-14, Jun. 1999.
 24. M. Z. Yang, "Semantic Dependency Based Natural Language Understanding in a Medical Dialogue System," *National Cheng Kung University, Master dissertation*, Jun. 2004.