

TR-88-005

評 估 標 準 鍵 盤 中 文
輸 入 法 之 模 擬

參 考 書
不 外 借

中研院資訊所圖書室



3 0330 03 000080 1

0080

TR-88-005

評 估 標 準 鍵 盤 中 文
輸 入 法 之 模 擬

中央研究院資訊研究所

鄭國揚

參與人員：楊順評

中華民國七十七年六月

目 錄

誌 謝	i
摘 要	i i
第 一 章 緒 論	1
第 二 章 系 統 架 構	9
第 三 章 視 覺 語 言 的 工 作 環 境	14
第 一 節 何 謂 "視 覺 語 言"	14
第 二 節 本 系 統 之 視 覺 語 言 環 境	15
第 四 章 評 估 模 式	26
第 一 節 服 務 項 目	26
第 二 節 S - P 圖 表 分 析 技 術	30
第 三 節 輔 助 評 估 模 式	39
第 五 章 分 析 結 果 輸 出 單 元	41
第 一 節 單 次 測 驗 分 析 結 果 輸 出	41
第 二 節 以 規 則 資 料 庫 為 基 礎 的 自 然 語 言 輸 出 子 系 統	44
第 三 節 整 體 分 析 結 果 輸 出	54
第 六 章 結 論	55
參 考 文 獻	57

摘 要

在中文電腦的推廣與中文資料的輸入過程中，中文輸入法是一個非常重要的要素，而在眾多的輸入法中，如何去選擇一套適當的輸入法，則必須有賴一套客觀的評估方法來提供有關輸入法基本特性的資料，以做為選擇之參考。

本論文的目的乃在於應用電腦提出一套評估標準鍵盤中文輸入法的模擬機。本系統將以視覺語言 (Visual Language) 的觀念提供一個視覺的人 - 機介面 (Human-Machine Interface)，由使用者定義其中文輸入法，並且以隨機的方式自動產生測驗的題目，分階段進行測驗。而利用 S - P 圖表的分析技術，配合吉尼係數，學習函數等進行評估工作，以發覺輸入法問題的癥結之所在，以期對輸入法之選擇以及設計過程能有所助益。

最後除將輸入法基本特性，如平均混淆率，學習曲線等輸出以外，並以規則資料庫的概念，將解釋評估結果的原則建立成知識庫，用一推理機 (Inference Engine) 以回溯式鏈結 (Backward Chaining) 的方式來推理解釋，並將答案與理由，以自然語言的方式輸出，使得系統更具有親和力 (User Friendly)。

第一章：結 論

在資訊發達的今天，電腦已由資訊中心，學校漸漸的走向工商業界，並且深入家庭。電腦的應用已經普及到社會上的各行各業，電腦不再是電腦專業人員的專利品。在推廣電腦的過程當中，電腦中文化是最重要的一項工作，而中文輸入法則又是電腦中文化的最根本要素。根據資訊工業策進委員會在1982年對中文輸入法的一項市場研究報告中 [1]，說明了中文輸入法在最近幾年來的成長情況。從圖 1-1 中，我們可以明顯的看出，雖然中文電腦的發展僅是最近十餘年的事，但是根據研究報告顯示，至1981年為止，中文輸入法中已經被市場化的方法，就有 24 種之多，而到 1982 年時更提高到 31 種，足足成長了 28 個百分點，可見得其成長之速。而另外未被市場化的輸入法，更是不計其數。在這麼多的輸入法中，每一種方法的好壞優劣，若是沒有一套客觀的評估模式來加以評估，則各家各自吹擂，莫衷一是。

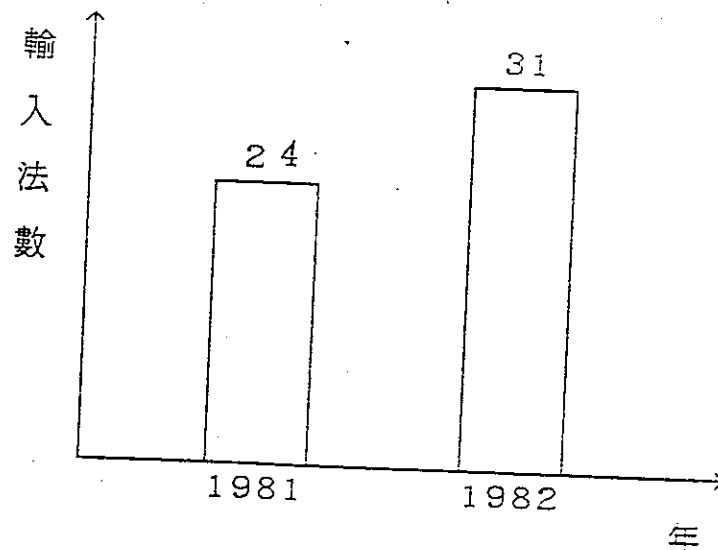


圖 1-1.

此外，對一套中文系統的發展者而言，選擇一套適當的中文輸入法是非常重要的工作，因此他們對各種輸入法的基本特性，諸如輸入速度，輸入錯誤率，學習困難度．．．．等等，都必須要有一全盤的瞭解，以便選擇一套適合自己所需的輸入法，以提高整個系統的效益。從以上的描述，我們可以想像得到，一套客觀的評估系統對中文電腦推廣的重要性。

在談中文輸入法評估模式以前，我們有必要對目前所發展出來的中文輸入法，做一個歸類與整理，以便從當中發覺中文輸入法的特性。就目前一般常見的輸入法來看，我們可以將中文輸入法分成四大類。

< 一 > . 形碼輸入法

所謂形碼輸入法，就是將中文字依照其平面結構，經過分析後歸納出一組的基本字碼 (Elementary Radical)，然後再根據一套拆解與取碼規則，將每個中文字拆解成若干個字碼，以這組字碼來代表此中文字。形碼輸入法中，字碼集的大小，對於整個輸入法佔有一相當重要的地位。字碼集大到一字一碼時，則無需拆解規則，也沒有重複字的現象發生，但是相對的，則必需應用較大的輸入器來當做輸入媒介，使用者必需花很大的工夫去記憶字碼在輸入器上的位置，而且輸入資料時亦必需花較長的時間去找尋字碼所在的位置，然後敲鍵。不能做到 "觸摸輸入" (Touch Typing)。若字碼集小的話，則拆解規則勢必會相當複雜，不容易學習。不過對形碼輸入法而言，無論字碼集大小，最大的優點就

是，它是以中文字的结构來拆解字碼，只要熟悉拆解規則，即使不認識的字，也能夠見字識碼，進行輸入資料的工作。

< 二 >. 拼音輸入法

拼音輸入法以我們所熟悉的注音符號法為主，由於國人從小學即開始學習注音符號，因此它最大的優點就是一般人對注音輸入法已經相當熟悉，不用花太多時間去學習，但是也有它的缺點，那就是同音字太多。輸入注音符號後常須要目視螢光幕，再從顯示在螢光幕上的同音字中，選擇到真正所希望的字，浪費時間。此外它亦不能像形碼輸入法一樣，可以見字識碼，因此不會注音符號的字就無法輸入。

< 三 >. 筆劃輸入法

它是根據中文字的基本筆劃來拆解中文字，優點是能夠見字識碼，而且拆解規則簡單，容易學習。但是平均來講，一個中文字所需的筆劃數不少，因此輸入速度有限。

< 四 >. 內碼輸入法

內碼輸入法即是每一個中文字都有它特定的內碼，能夠唯一來代表它，所以沒有混淆的現象(Ambiguous)，而且每個中文字敲鍵數一定。但這種方法的缺點是，必需花費很大的工夫去記憶每一個字的內碼，學習困難。

綜合以上所討論的各類中文輸入法，我們可以很清楚看出，選擇中文輸入法所考慮的因素莫過於，<1>. 輸入速度 (Input Speed)。<2>. 學習時間 (Learning Time)。從圖 1-2，一般輸入法學習時間與輸入速度的關係大略可分為底下四種：

方法 <一>. 輸入速度最快，但是相對的，學習時間也最長，因此在訓練的過程中，所需的時間與成本也最高，比較適合專業人員來使用。

方法 <二>. 輸入的速度最慢，但卻最容易學習，所花的訓練時間與成本都相對大大降低，因此比較適合於不重視輸入速度的一般使用者來使用，也比較能夠大眾化。

方法 <三>. 這是一個優秀的輸入法，也是目前中文輸入法的設計者所追求，理想的輸入法。不但輸入速度快，而且又容易學習，是一套令人振奮的輸入法，但是以目前而言，這種系統仍是一個理想中的方法，有待專家們去研究開發。

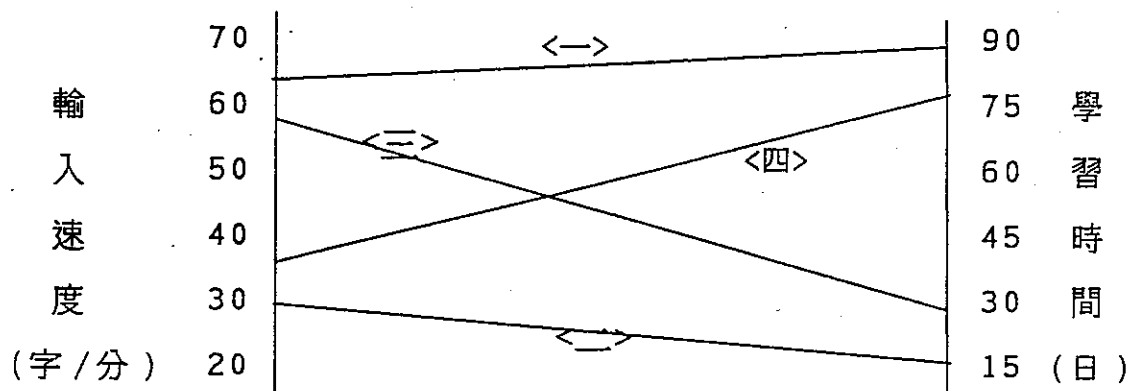


圖 1 - 2.

方法 <四>. 從各種方法的比較中，我們可以容易的想像得到，這是一個極差的方法，不但學習時間長，而且輸入速度又慢。

在兼俱輸入速度快，而又容易學習的中文輸入法尚未開發成功以前，使用者在選擇一套輸入法時，必需知道這套輸入法的各種特性，才能選擇一套適合於自己所需的方法，而這些特性即是本篇文章所要提供的服務項目 (Service Item)。

在還沒介紹我們所用的評估模式以前，我們先來看看以前學者所做過比較具有代表性的文章。在我們參考相關中文輸入法評估方面的雜誌以及各類參考文獻的過程中，就我們所發現的，到目前為止，有關這類的文章還是相當有限。但是在一些比較具有親合力 (User Friendly) 的中文輸入法 (如自然語言的語音輸入法，手寫字型輸入法以及光學印刷字型掃描輸入法) 還沒發展出來以前，鍵盤式的中文輸入法仍為中文資料輸入的主要方式。因此對於鍵盤式中文輸入法的評估，也就仍然是我們討論的重點。以下我們將簡單地概述兩篇比較具有代表性的相關文章。

首先我們來看看 Cheng-Kuang Chen，以及 Reng-Weng Gong 在 1984 年所提出的一篇 " 中文輸入法評估 " 的文章 [2]，他們在這篇論文中，就當時所盛行的輸入法，加以歸類整理，並且比較各類方法的優缺點，最後提出七項評估中文輸入法的評估因素，作為他們評估的根據。並且根據 Dr. Hirsch 以及 Dr. Yu 所提出的測量模式 (Measurement Model)，給予每項因素一個加權 (Weight)，然後利用各輸入法在各項因素所得的分數，乘以相對的加權，加總起來而得到一個總分，以得分的高低來判斷輸入法的優劣。

誠然在這篇文章裡，他們的確對於各種方法，做了一個非常詳細的分析，而且對各種方法的特性做了詳盡的介紹。但是在其評估模式的“加權”部分，全憑專家的經驗得來，缺乏客觀因素的依據，而且在測試的過程當中並沒有考慮到當時，受測試者的學習態度與狀況，因此測試所得資料的可靠度也值得商榷。另外他們的方法，也只限於已推出的系統才能去進行測試，收集評估的資料。

另外 1985 年 S. Y. Lo，在同一個雜誌裡也提出一篇有關中文輸入法評估方面，比較屬於科學模式(Scientific Model)的研究 [3]。在這篇文章裡，S. Y. Lo 為了改進一般評估方法花大量人力以及物力去做訓練與測驗，以獲得評估數據的缺失，因此他提出一套以資訊理論(Information Theory)為基礎的純粹數學理論分析模式，來評估各種中文輸入法。在他的方法中，就輸入法符號(Symbol)多寡，鍵盤大小對輸入步驟的輸入速度與學習時間之影響來評估。所謂輸入步驟是指，使用者從看到中文字，進行字碼的拆解，然後到找尋字碼在鍵盤的位置，最後敲打鍵盤。例如，倉頡輸入法的輸入步驟可分成，中文字——→拆解字碼——→找尋字碼在鍵盤上的位置——→敲打鍵盤。這篇文章最大的目的就是希望在理論分析的過程中，找出影響輸入速度以及學習速度的癥結所在。它的優點是利用理論分析，最適合於電腦的應用，而將人力的介入程度降至最低的地步。但是由於他的方法太過於理論，全由機率的觀點出發，卻忽略了人因因素(Human Factor)，因此降低了它的實用價值。

看過上述兩篇文章以後，我們可以看得出來，這兩篇文章所走的是全然不同的路線。Chen 以及 Gong 是以實際走向 (Physical Approach)，利用現有的系統，找一些人員來進行測驗，獲得實際的數據，而達到評估的目的。而 Lo 則是屬於理論走向 (Theoretical Approach)，純粹以數學理論模式去分析，但是各有缺點，前者是忽略了中文輸入法評估的一項重要因素——學習困難度，並且只適合於已開發完成的系統。而後者的缺點乃是太過理論化，僅考慮機率問題，而忽視了人類的行為因素 (Human Behavior)。

為了解決這些問題，在我們的系統中除了保存上述兩種方法所提出的一些重要評估因素 (如輸入速度，輸入錯誤率，混淆率，．．．等等)。除此之外在我們的系統中，我們將針對以標準鍵盤為輸入媒介的中文輸入法，提供一個視覺的工作環境，讓使用者藉著這個工作環境，來定義他們所構想的中文輸入法。並且著重輸入法的學習困難度的探討，把中文輸入法的學習過程分成幾個階段 (拆解中文字，敲鍵)。在每個階段中，系統根據使用者所定義的資料，以隨機的方式自動產生測驗題目，讓受測驗者直接在電腦上作答。系統記錄這些測驗後的資料後，然後再以 S - P 圖表 (Student Problem Chart) 的評鑑技術，以靜態與動態的方式來分析受測驗者，在每個階段的學習情況，以找出影響中文輸入法，輸入速度與學習時間的癥結所在。而且在我們的系統中，主要的輸出都以圖形的方式來表達，使得評估者對於評估結果能夠一目了然，而不是面對一堆煩人的數據。

另外在系統中，我們亦提供一個以規則資料庫 (Rule Base) 為基礎的解釋子系統，讓專家們將他們對於評估方面的經驗，建立成一個知識庫 (knowledge Base)，系統的推理機 (Inference Engine) 將根據知識庫的資料來解釋評估的各種現象。

現在我們把本篇文章的綱要簡述如下：

- 第二章 主要是介紹模擬評估系統的系統架構。
- 第三章 將介紹本系統的人 - 機介面 ~ 視覺工作環境。
- 第四章 將說明本系統的主要評估工具，S - P 圖表的特性、處理方式與應用，以及本系統所提供的服務項目。
- 第五章 將介紹系統所提供以規則資料庫 (Rule Base) 為基礎的解釋單元以及分析結果的自然語言輸出。
- 第六章 做一個簡單的結論與討論。

第二章：系統架構 (System Organization)

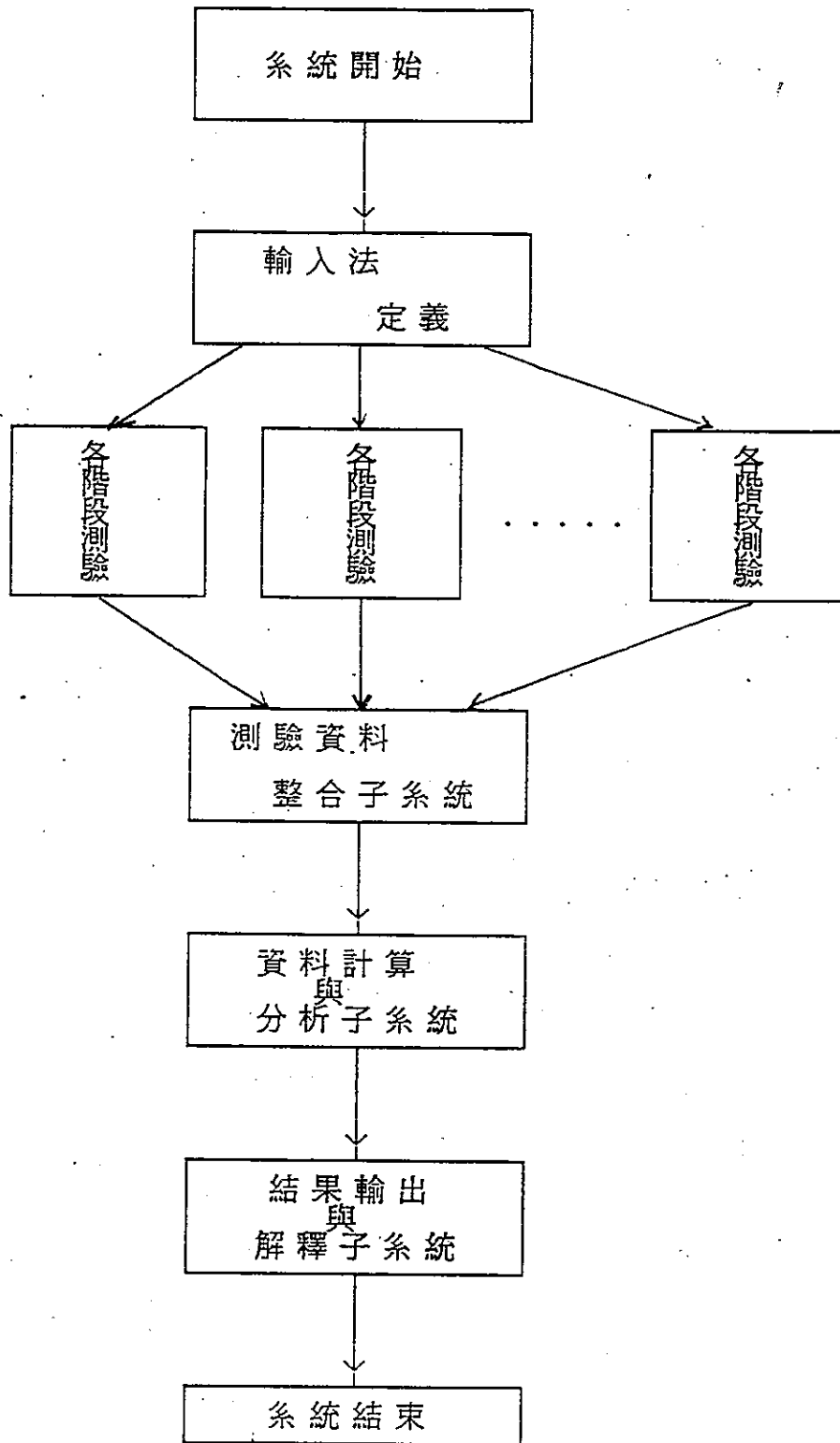


圖 2 - 1 . 系統架構

從圖 2 - 1，系統架構圖中，可以看出我們將整個模擬評估系統，分成五個子系統。

- < 一 >. 輸入法定義子系統
- < 二 >. 測驗子系統
- < 三 >. 測驗資料整合子系統
- < 四 >. 資料計算與分析子系統
- < 五 >. 結果輸出與解釋子系統

在此我們僅對這些子系統作一整體性的概略描述，致於其詳細的作業情形，將於以後的章節作詳盡的說明。

< 一 >. 輸入法定義子系統

這一個子系統部分是屬於人 - 機 (Human - Machine)，溝通的部分，也就是模擬評估系統獲得 " 知識 " 的部分。在這個子系統裡，我們將應用視覺語言 (Visual Language) 的觀念來達到人 - 機溝通的目的。透過定義子系統，使用者能夠將自己輸入法的字碼集，中文字拆解，鍵盤如何安排等資料告訴模擬評估系統。其架構如下圖 2 - 2。

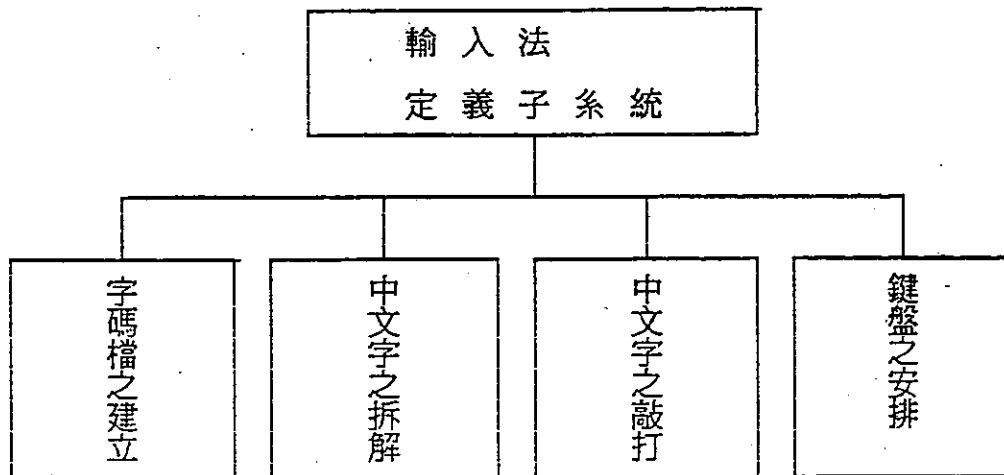


圖 2 - 2. 輸入法定義子系統

< 二 >. 測 驗 子 系 統

設計者透過定義子系統，將自己的構想一一描述給模擬評估機以後，系統將根據這些資料以隨機 (Random) 的方式自動產生測驗的題目，而評估者則可找一批人員，按照學習階段 (如中文字之拆解，中文資料之輸入等)，來作一連串的測驗，而系統並能自動的記錄每個測試人員答題的得分情況，所花費的時間...等，而成為一個原始的測試資料檔以作為將來計算與分析的根據。其架構如圖 2 - 3 所示。

< 三 >. 測 驗 資 料 整 合 子 系 統

因為本模擬系統測驗時，為了能夠讓所有的測驗者同時進行起見，可以以分散式的方式，根據設備的情況，將測試人員分成幾組或是一人一機，同時進行測驗。然後透過資料整合子系統，將這些資料整合成為一個完整的資料檔。其架構如圖 2 - 4. 所示。

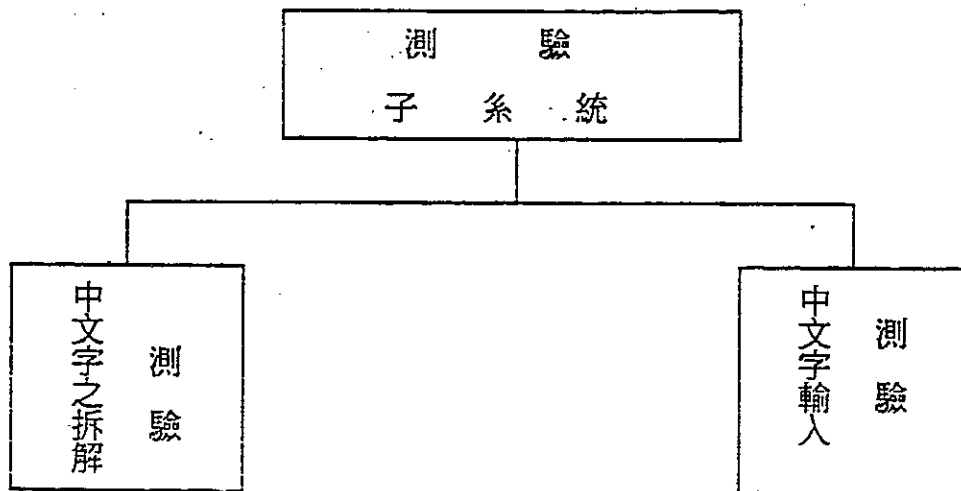


圖 2 - 3. 測 驗 子 系 統

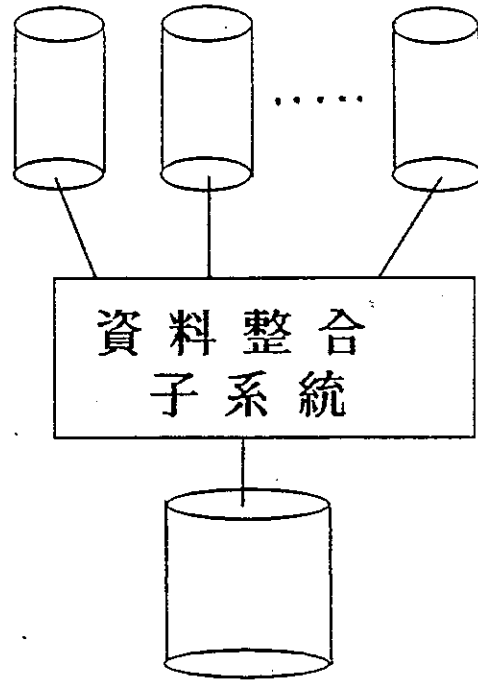


圖 2 - 4. 資料整合子系統

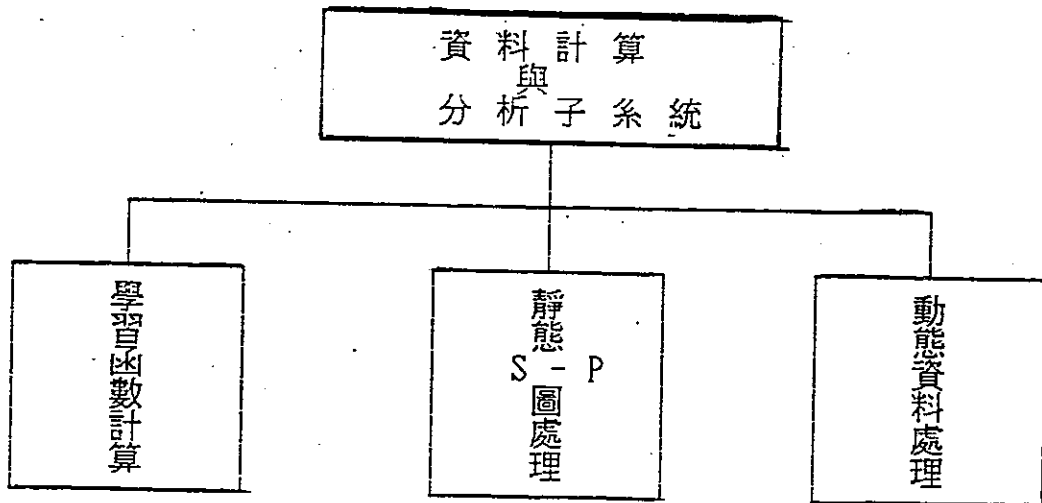


圖 2 - 5. 資料計算與分析子系統

< 四 >. 資料計算與分析子系統

經過分段的測驗，並且將資料整合後，我們可以得到學習過程中，每一個步驟的測驗資料，我們便可逐一分項地對這些資料，根據 S - P 圖表的評鑑技術以及學習函數，做一些計算以及分析的工作，並將結果儲存起來，然後在結果輸出子系統，將這些結果輸出。其架構如圖 2 - 5 所示。

< 五 >. 結果輸出子系統

經過資料計算子系統的處理以後，我們便可將分析與計算的結果利用螢幕的輸出，將結果輸出。包括單張 S - P 圖表的分析結果與圖形，一個以規則資料庫 (Rule Base) 為基礎的自然語言解釋輸出系統，以及以表格形式的評估總結果輸出。其架構如圖 2 - 6 所示。

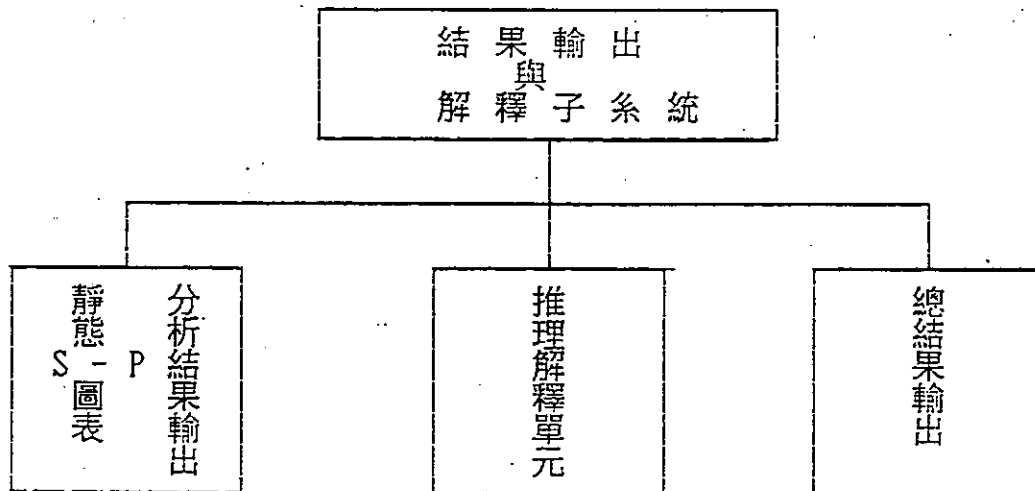


圖 2 - 6. 結果輸出子系統

第三章：視覺語言的工作環境

(Visual Language Environment)

第一節：何謂 " 視覺語言 "

所謂 " 視覺語言 "，就是以人眼睛所看見的影像，如圖形、表格等來表達意思，當作人類與電腦之間的溝通語言。它最簡單的描述就是『What you see is what you get.』以及『What you sketch is what you get.』。從廣義的眼光來看，它是使用者（包括系統人員，程式設計師，以及一般使用者。），與軟體或是程式之間的一種溝通介面 (Interface)，它可利用圖形的技巧，以靜態或動態的表現方式，將軟體或是程式的表現由傳統一維的循序式程式指令，拓展到二維，甚至三維的資訊。而以狹義的眼光來看，它則是一種新的程式設計技巧，它能透過填表格 (Form) 的方式，或是利用圖像 (Icon) 的處理技術，經過一些定義以後，由系統本身將這些使用者定義的資料，經過一個特殊的翻譯器 (Interpreter)，產生一個可執行的程式。或是利用轉換器 (Translator)，將它轉換成使用者所需的程式語言 (Programming Language) 程式 [4][5][6][7][8]。由於由傳統的程式語言來設計程式是一項非常枯燥，費時，而且又是一項非常須要具備專業知識的工作，因此在應用程式 (Application Program)，以及資訊需求量，日益大量成長的今天，傳統設計程式技巧，以及使用者與軟體系統之間的溝通方式，已經不敷需求，因此 " 視覺語言 " 就更顯得其重要性。





第二節：本系統的“視覺語言”環境


介紹過“視覺語言”的意義以及優點後，在本節中我們將利用“視覺語言”的概念來設計一個簡單的輸出／輸入環境。來作為評估系統與中文輸入法設計者，以及受測驗者之間的人-機介面 (Man - Machine Interface)。首先從中文輸入法的特性來看，中文字可以看成是一個個的圖像 (Icons)，而這些圖像則是由一組最基本圖像 (在此我們稱為“字碼”)，根據一些組合規則組合而成。這些基本的字碼，依不同的輸入法而有不同的字碼集合，比方說：三角號碼輸入法有一百個基本的字碼，包括 {言、亦、文、...、九}。而倉頡輸入法則由二十四個所構成，包括 {一、土、日、...、十}。由不同的基本字碼集合，根據不同的組合規則，便產生不同的中文輸入法。例如：



中文字	三角號碼輸入法	倉頡輸入法
麗	一一广	一一月月心
矮	人禾大	人大竹木女


因此在本系統簡單的人-機介面裡，我們將提供一個處理這些基本字碼的工作環境，讓輸入法的設計者從設計、測驗，乃至於評估工作，都能透過電腦來完成。底下我們將分成四點來詳加介紹。

< 1 >. 基本字碼的建立

在前面我們曾經介紹過，中文輸入法中，每個中文字是由基本的字碼所組合而成，因此基本字碼的建立是設計一套中文輸入法最基本的工作。中文輸入法字碼的設計過程中，必須考慮到鍵盤大小對資料輸入速度所造成的影響，因此基本字碼的多寡會受到鍵盤大小的限制。另外在中文字的拆解過程中，如果要把所有的中文字，直接對應到基本字碼，則其拆解的規則一定相當複雜，因此基於輸入速度與中文字拆解的困難度兩項因素來考慮，一般的中文輸入法又將字碼分成兩類。一是基本字碼，另一類是輔助字碼。所謂的基本字碼就是輸入法在取碼時代表一個字的那些字碼，例如，倉頡輸入法的基本字碼有：{ 一、土、月、. . .、十 } 等二十四個，而輔助字碼則是一些與某個基本字碼，意義相似或是字形類似，或由基本字碼變形所衍伸出來的字碼，而將這些輔助字碼分別歸屬於這個基本字碼。例如，倉頡輸入法中，基本字碼『月』，它的輔助字碼有：、、、，其中

『』：乃是取用『月』的外形。

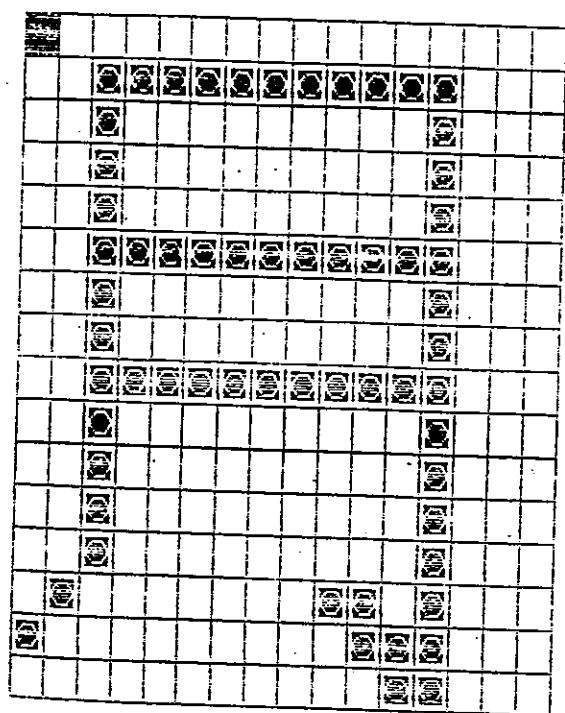
『』：則是由『』變形而來。

『』：也是取『月』之外形。

『』：是由『』變形而來。

輔助字碼最重要的角色，就是在於能使學習者便於“見字識碼”，而使得中文字的拆解規則能夠更有系統，更簡單，以增加中文字拆解的速度與準確性。

另外，有的中文輸入法為了解決受限於鍵盤大小的問題，輸入法設計者便提出，在基本字碼與鍵盤之間又加了一層對應。例如三角號碼輸入法，將一百個基本字碼對應到 00 ~ 99，因此雖然它有一百個基本字碼，但卻只需 0 ~ 9 十個鍵即可。因此在我們的系統中也提供定義敲鍵字碼的功能。圖 3-1 所示便是字碼定義子系統。在這個子系統中，我們將利用圖形的技巧，由設計者來建立這些字碼集合。每一個字碼的圖像用 16 X 16 像元 (Pixel) 的矩陣來表示，提供使用者設計並修改其輸入法的字碼集合。



移動

X Y

***** 功能鍵說明 *****

* 請按鍵盤右方的八個
游標控制鍵移動游標

* INS==>加點

* DEL==>刪點

* + ==>移動游標

* F1==>儲存

* F2==>修改

* F3==>離開

* F4==>清除




月


圖 3-1.

< 2 >. 中文字之拆解


前面提過對中文輸入法而言，中文字是由基本的字碼所組成，因此建立完字碼集以後，接下來的工作就是告訴系統，中文字的拆解過程與中文字的輸入字碼，而建立一個中文字拆解與敲鍵的資料庫，以便將來進行拆解中文字以及中文輸入的測驗。並且求得一些定量的數據（如混淆率、平均輸入字碼數等）。因為中文字有其平面的視覺架構，所以在中文輸入法中，大都會運用到這個特性來訂定一套拆解的規則，以便減少中文字拆解的困難度，加速學習的速度。


應用 Nagahashi 與 Nakatsugama [9] 所引述有關 Rankin[10]，提出的切割（Segment）觀念，一個中文字可以利用三個規則來對中文字進行切割的工作。

1. 東-西法則 (East-West), 。例如，明 → 日，明。
2. 南-北法則 (North-South), 。例如，盲 → 亡，目。
3. 內-外法則 (Border-Interior), 。

例如，國 → ，或。

運用這三個簡單的規則，即可將中文字初步的拆解成一組一組的次字元（Sub-Character）。底下我們將以『個』字來作一個完整的說明，它的拆解步驟如下：

步驟一：利用法則 1. 個 → ，固。

步驟二：利用法則 3. 固 → ，古。

步驟三：利用法則 2. 古 → 十，口。

根據拆解的結果，並且在次字元之間附於空間的關係（Spatial Relation），包括“左邊”（Left-of），“上面”（Above-of），以及“包含”（Include-of）等關係。另外還有階層關係（Hierarch Relation），如此就可建立一個中文字的拆解圖（Decomposition Graph），如圖 3-2。

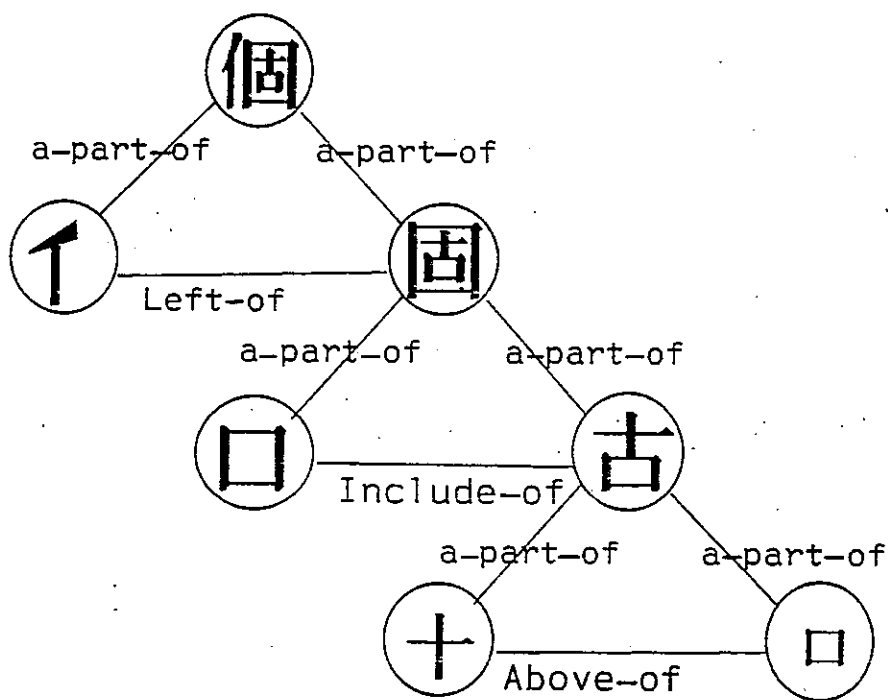


圖 3-2.

有了中文字的拆解圖以後，如果我們再根據不同輸入法的取碼規則以及輔助字碼與基本字碼之間的對應 (Map)，即可建立中文字的字碼拆解圖。以倉頡輸入法為例，圖 3-2 加上倉頡法取碼規則以及輔助字碼與基本字碼之間的對應關係，則可變成圖 3-3。如此我們便可由中文字的字碼拆解圖中，獲知『個』字在倉頡輸入法中，可分解成『人』、『田』、『十』、『口』。

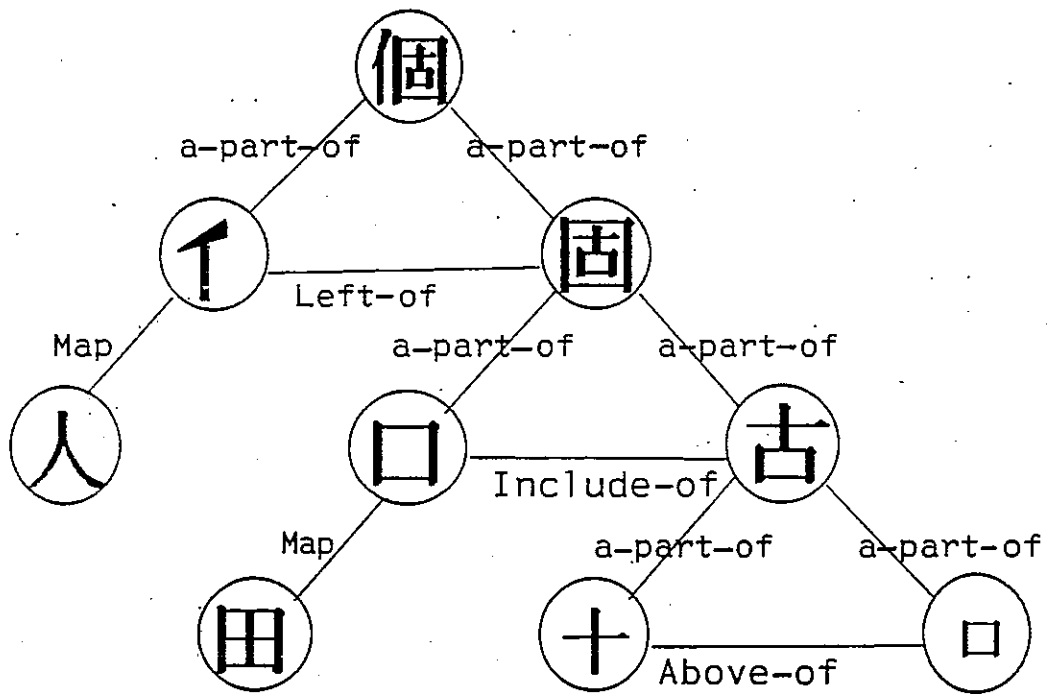


圖 3-3. 中文字的字碼拆解圖

為了容易實施 (Implement) 起見，我們可利用資料結構中，樹 (Tree) 的前序追循法則 (Preorder Traversal) 將字碼拆解圖以 Predicate Logic 的方是來表達，如

個 $\implies L(\uparrow, \text{固})$
 $\implies L(M(\text{人}), \text{固})$
 $\implies L(M(\text{人}), I(\square, \text{古}))$
 $\implies L(M(\text{人}), I(M(\text{田}), \text{古}))$
 $\implies L(M(\text{人}), I(M(\text{田}), A(+, \square)))$

其中 L 代表：Left-of， M 代表：Map

A 代表：Above-of， I 代表：Include-of。

由於本系統在中文字視覺結構的應用，最主要的目的乃是在於探討，在一個輸入法中，中文字碼拆解圖，結構的複雜度與學習困難度之間的關係，因此在本系統的資料庫中，只需保留中文字碼拆解中，每一個基本字碼在拆解圖中的深度資料以及階層關係。

定義中文字碼拆解圖的過程，如圖 3-4(a) 以及圖 3-4(b)所示。圖 3-4(a) 是進入中文字碼拆解圖定義子系統的起始狀態，在這個子系統中，我們可以利用 " 老鼠 " (Mouse)，當作輸入工具，來選擇系統提供的一些基本功能。

- 一. ○ : 定義虛字元 (這裡所指的虛字元是指，在中文字碼拆解圖中，非基本字碼的節點。)
- 二. ● : 定義基本字碼所在的節點。
- 三. T : 定義中文字之敲鍵。
- 四. / : 定義節點之間的鏈結。
- 五. 定義結束。

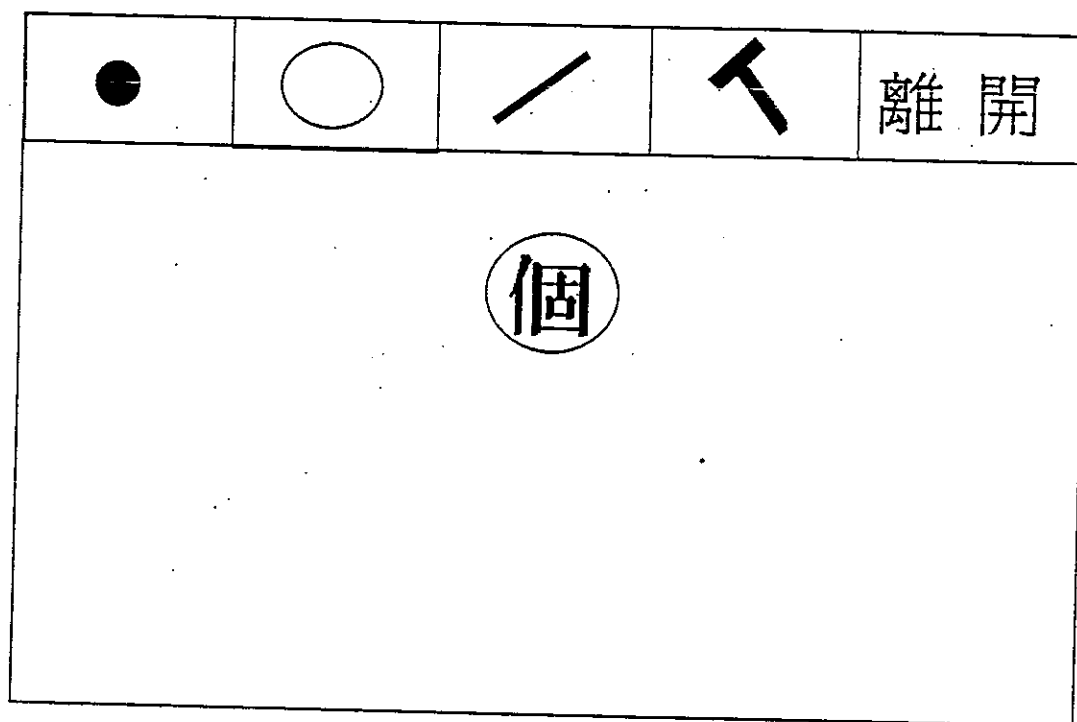


圖 3-4(a)

圖 3-4(b)是選擇了『●』定義基本字碼所在的節點後所出現的畫面，最上方是輸入法所有的基本字碼，使用者選擇基本字碼後，再指定應該將該字碼放到那個位置。建立完字碼拆解後，系統自動將其轉換成 Predicate Logic 的形式。

例如，圖 3-3 『個』字可轉換成：((人)((田)(十口)))。

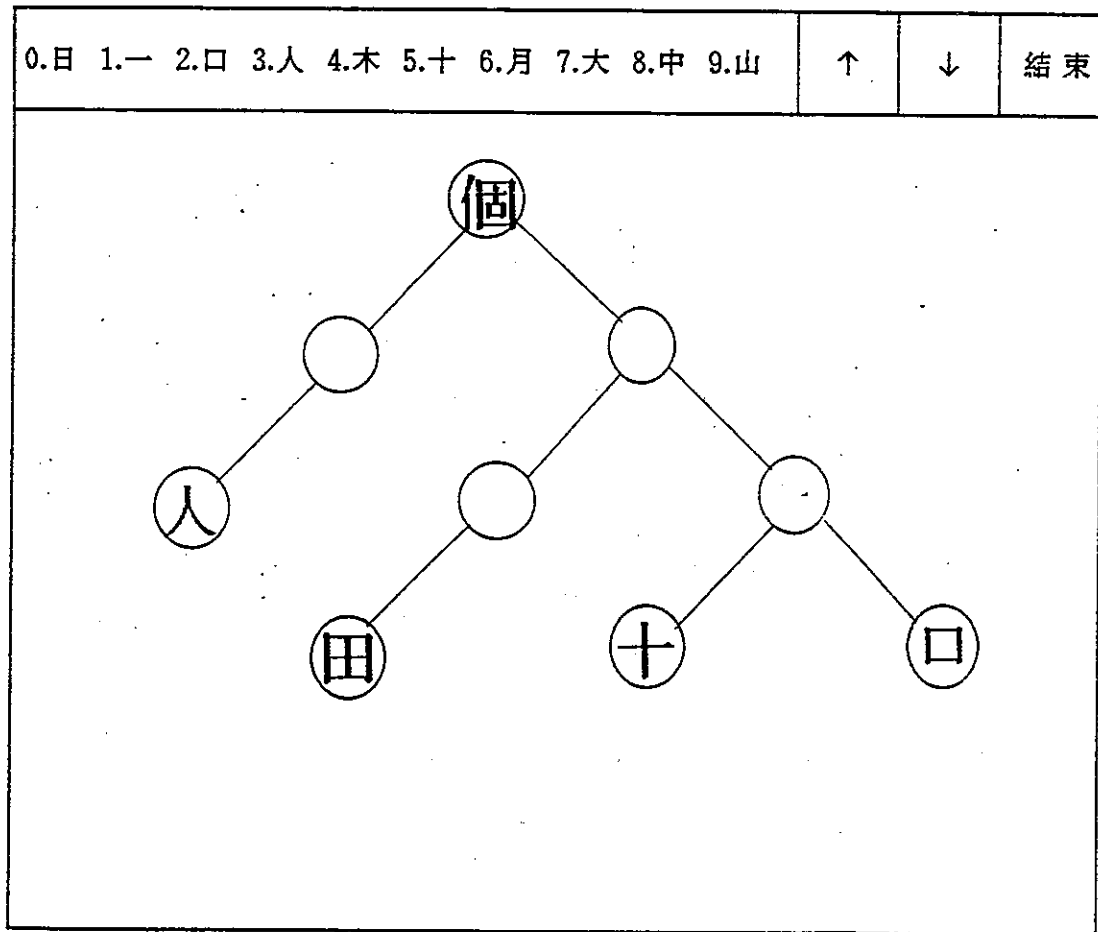
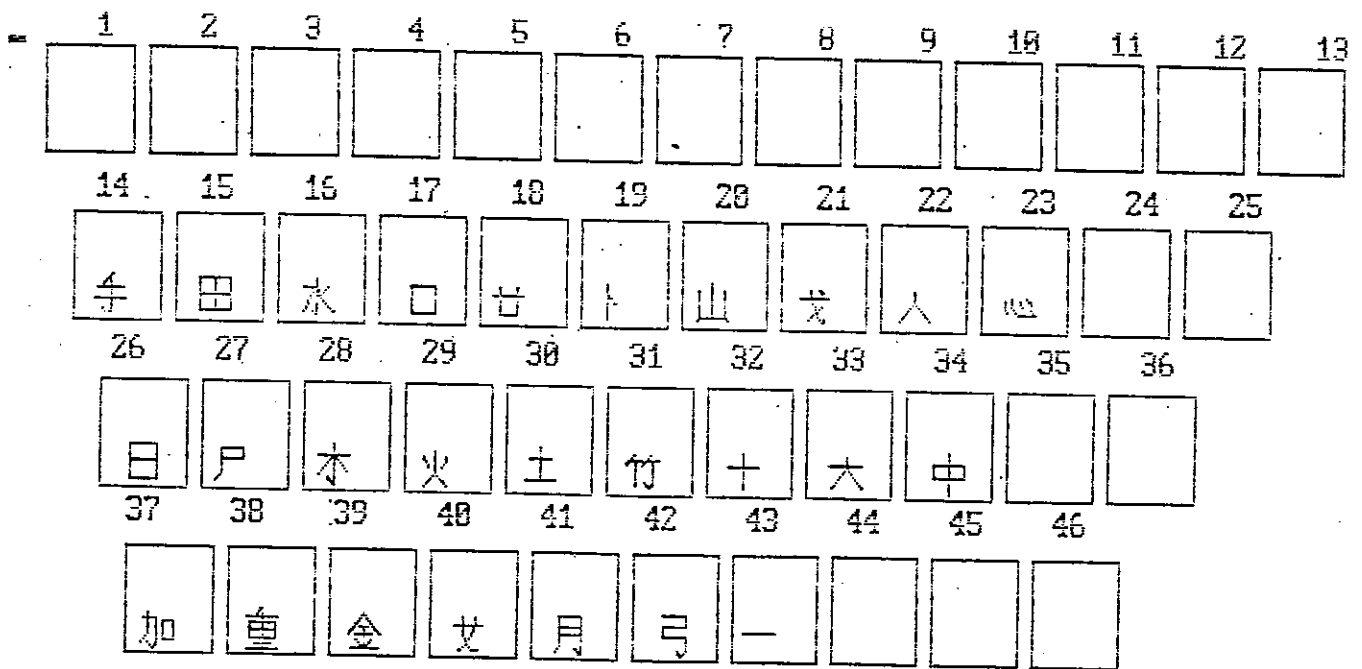


圖 3-4(b)

< 3 >. 鍵盤之安排

目前一般的輸入法都是經由鍵盤輸入，而且不同的鍵盤安排法，無論在輸入速度或是輸入時的錯誤率，都有很大的影響，因此鍵盤的安排也佔有一個重要的地位。在本系統中，所謂的鍵盤安排是指，以標準鍵盤為基礎，將敲鍵字碼安排在鍵盤上的方式。如圖 3-5，使用者可以將字碼，根據自己認為比較好的安排方式，安排在鍵盤上。以便於敲鍵的測驗。



< 0. — 1. 戈 2. 中 3. 土 4. 口 5. 日 6. 卜 7. 手 8. 田 9. 山 >

字根：
鍵盤代號：

F1: 儲存 F2: 修改 F3: 離開 PgUp: 上十個字根 PgDn: 下十個字根

圖 3-5.

< 4 >. 模擬敲鍵過程

在鍵盤速度的測驗中，對於一個還未發展完成的輸入法的評估，不能以實地的鍵盤操作來進形測驗，同時也無法將測驗者所敲打的结果顯示在螢幕上，因此不能達到真正鍵盤測驗的目的。所以我們透過鍵盤的安排，以現有的鍵盤來模擬設計者所安排的鍵盤，以便於鍵盤的實地操作，提高測驗數據的可靠度。如圖 3-6 所示。

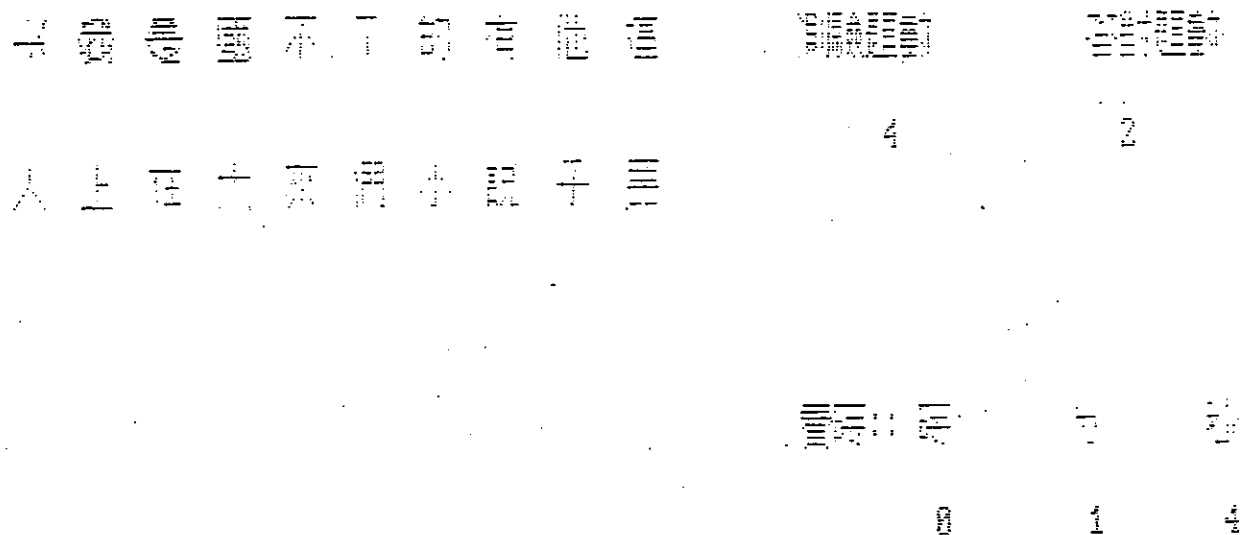


圖 3-6. 0.不 1.有 2.人

圖 3-6.

第四章：評估模式 (Evaluation Model)

第一節：服務項目 (Service Item)

根據以往學者的研究，我們整理出一些輸入法的基本特性，當作本系統的服務項目。在還沒有正式談到本系統所提供的服務項目之前，因為在中文字的使用頻率上，大抵上可以分為『常用字』、『次常用字』、『罕用字』以及『死字』等四類。在一般的使用上，這四類字的使用頻率相差甚巨，因此在許多項服務項目的計算中，為了使這些數據更具有代表性，我們將根據交通大學所研究整理的資料，來當作中文字使用頻率的根據 [11]。

< 一 >. 平均敲鍵數 (Average Strokes per Character)

計算每個中文字輸入時，平均需要敲打幾個鍵，因為這項因素對輸入速度之影響非常大。

F_i ：表中文字 i 的使用頻率。

S_i ：表中文字 i 的敲鍵數目。

N ：表所有計算根據之中文字數。

其計算公式如下：

$$\text{則平均敲鍵數 (AS)} = \frac{\sum_{i=1}^N (F_i * S_i)}{\sum_{i=1}^N F_i}$$

< 二 >. 碰撞率或混淆率 (Collision/Ambiguous Rate)

所謂碰撞或混淆是指在中文字輸入時，同樣的一組字碼組合形式亦相同，但卻代表不相同的中文字。

定義：碰撞

X：字碼組合串， a_i ：中文字 i 。

f：字碼組合串對應中文字的函數。

若 $f(X) = \{ a_1, a_2, \dots, a_n \}$ 且
 $n > 1$ 則稱發生碰撞。

例如：注音輸入中，“一尤 / ” 的同音字（亦即碰撞數）

有 13 個，包括 { 羊、洋、陽、楊、.、颺 }。

當發生碰撞的現象時，輸入系統一般的處理方式是將這些字，同時顯示在螢幕上，資料輸入人員必須檢視螢光幕，再從這些字當中選擇真正想要的字，非常浪費時間。而且碰撞率越高者，相對的輸入速度就會降低。碰撞率公式如下：

F_i ：表中文字 i 的使用頻率。

C_i ：表中文字 i 所發生的碰撞數。

N：表所有計算根據之中文字數。

$$\sum_{i=1}^N (F_i * C_i)$$

$$\text{則平均碰撞率 (CR)} = \frac{\sum_{i=1}^N (F_i * C_i)}{\sum_{i=1}^N F_i}$$

< 三 >. 平均輸入速度 (Average Input Speed)

此服務項目乃說明，在測驗當中，受測驗者輸入資料的平均速度。其計算公式如下：

T_i : 表受測驗者 i 作答所花的時間。

PI_i : 表受測驗者 i 在 T_i 時間內個人所輸入的字數。

M : 表受測驗者 的總人數。

$$\text{則平均輸入速度 (AIS)} = \frac{\sum_{i=1}^M \frac{PI_i}{T_i}}{M}$$

< 四 >. 學習時間 (Average Learning Time)

此項目乃在說明，學習一套輸入法所需要的學習速度。其計算公式如下：

T_i : 表第 $i - 1$ 次測驗與第 i 次測驗之學習時間。

m : 表學習次數。

$$\text{學習時間 (LT)} = \sum_{i=1}^m T_i$$

< 五 >. 平均錯誤率 (Average Error Rate)

在要求輸入法輸入速度快時，也必需同時考慮到輸入資料時所造成的錯誤率，進而發覺造成錯誤的原因，以便改進，促進輸入速度的加速。其計算公式如下：

$E I_i$: 表受測驗者 i 答錯的題目數。

$P I_i$: 表受測驗者 i 答 的題目數。

M : 表受測驗者的總人數。

$$\sum_{i=1}^M \frac{E I_i}{P I_i}$$

則平均錯誤率 (AER) = $\frac{\sum_{i=1}^M \frac{E I_i}{P I_i}}{M}$

第二節：S - P 圖表分析技術

在評估的過程中，除了上節所介紹的基本參數 (Parameters)，另外我們將利用 S - P 的評鑑技術 [12][13] 來分析受測驗者的學習狀況以及輸入法的特異情形。由圖 4-1 中我們可以明顯的看出，在日本現行的主要評鑑方法 (S - P 圖表分析、一般統計分析、問卷調查分析、以及標準測驗評分報告)，而 S - P 圖表的分析方法乃是應用最為廣泛的評鑑方法。由單張的 S - P 圖表的分析中，我們可以由 S - P 曲線的面積，看出在某階段，受測驗者學習的結果，並可利用學生注意係數以及試題注意係數，分析受測驗者以及試題的特異現象。而由多張 S - P 圖的分析則可看出整個學習過程，受測驗者的學習狀況與輸入法的特異現象。

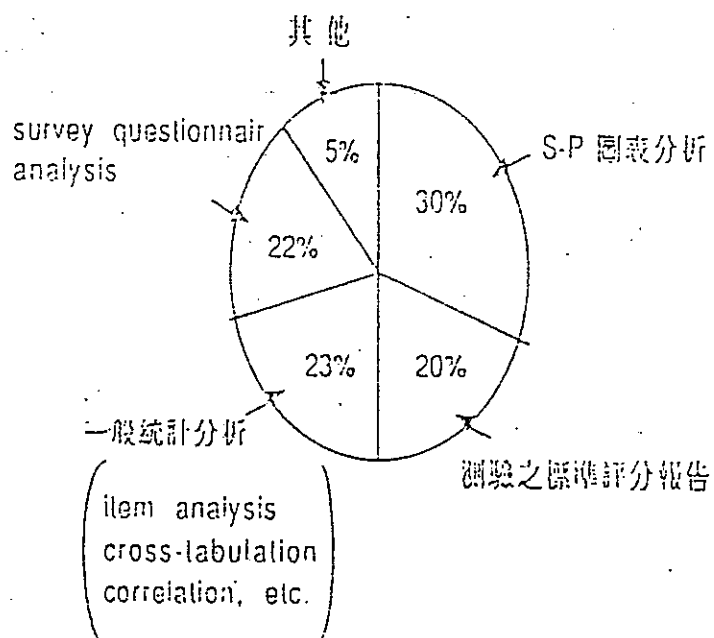


圖 4-1.

< 一 >. S - P 圖表的意義與處理過程

所謂 S - P 圖表乃是英文 Student 的『S』與 Problem 的『P』之縮寫，而其實質上所代表的意義是『學生』與『試題』之間答對與答錯的關係表格。我們可以將 S - P 表格看成一個 $M \times N$ 的矩陣 A。

$$A = [a_{ij}] \quad 1 \leq i \leq M, 1 \leq j \leq N。$$

$a_{ij} = 0$ ，如果編號為 i 的學生答對第 j 號試題。

$a_{ij} = 1$ ，如果編號為 i 的學生答錯第 j 號試題。

其中 M 為學生人數， N 代表試題數。

S - P 圖表的目的，乃在於希望從這簡單的『0』與『1』之矩陣，透過一些簡單的處理，而能從處理後的 S-P 圖表中得到有關教材與試題之適宜性，以及學生學習之有效性的資料。

S - P 圖表的處理可分成兩個步驟：

步驟一：將 S - P 表格作列方向與行方向的重排，重排原則乃是在列方面，根據學生成績的高低，將成績高者放在上端，而成績差者放在下端。在行的方面是根據試題的困難度，將被答對較高的試題放在矩陣的最左端，被答對較低者放到右端。

步驟二：根據重排過的 S - P 表格，繪出 S 曲線與 P 曲線，其畫法如下：

< 1 >. S 曲線：在每列中相等於該列得分數（即此學生答對之試題數）的行之右端畫一線段，然後將這些線段一一連接起來，便是 S 曲線。

< 2 >. P 曲線：在每一行中相等於該行答對數（即該試題答對的學生個數）的列之下端畫一橫線段，然後將這些線段一一連接起來。如圖

		試題編號												
		1	2	3	4	5	6	7	8	9	10			
學生編號	S	1	0	1	1	0	0	0	1	0	1	1	5	合計得分
	2	0	1	1	1	1	1	0	0	1	0	6		
	3	0	1	0	0	0	0	1	0	1	0	3		
	4	1	1	1	1	0	1	1	0	1	0	7		
	5	1	1	1	1	1	1	1	0	1	1	9		
	6	1	1	0	1	1	0	0	0	0	1	5		
	7	1	1	1	1	1	1	1	1	1	1	10		
	8	1	0	0	0	0	0	1	0	0	0	2		
	9	0	1	1	1	1	1	1	1	1	0	8		
	10	0	1	0	0	1	1	0	1	0	1	6		
	11	1	1	0	0	1	0	0	1	0	0	4		
	12	0	0	1	0	0	0	0	0	0	0	1		
	13	0	1	1	1	0	1	0	1	0	0	5		
	14	1	1	1	1	0	0	1	0	0	0	5		
	15	1	0	1	1	0	0	1	0	0	0	4		
		8 12 11 9 6 7 10 4 8 5												

答對人數

圖 4-2(a).

		試題編號												
		1	2	3	4	5	6	7	8	9	10			
學生編號	S	7	1	1	1	1	1	1	1	1	1	1	10	合計得分
	5	1	1	1	1	1	1	1	0	1	1	9		
	9	0	1	1	1	1	1	1	1	1	0	8		
	4	1	1	1	1	0	1	1	0	1	0	7		
	10	0	1	1	0	0	1	1	0	1	1	6		
	2	0	1	1	1	1	1	0	0	1	0	6		
	14	1	1	1	1	0	0	1	0	0	0	5		
	1	0	1	1	0	0	0	1	0	1	1	5		
	13	0	1	1	0	1	1	0	1	0	0	5		
	6	1	1	0	1	1	0	0	0	0	1	5		
	15	1	0	1	1	0	0	1	0	0	0	4		
	11	1	1	0	1	0	0	1	0	0	0	4		
	3	0	1	0	0	0	0	1	0	1	0	3		
	8	1	0	0	0	0	1	0	0	0	0	2		
	12	0	0	1	0	0	0	0	0	0	0	1		
			8 12 11 9 6 7 10 4 8 5											

答對人數

圖 4-2(b).

		試題編號												
		1	2	3	7	4	9	1	6	5	10	8		
學生編號	S	7	1	1	1	1	1	1	1	1	1	1	10	合計得分
	5	1	1	1	1	1	1	1	1	1	1	0	9	
	9	1	1	1	1	1	0	1	1	1	0	1	8	
	4	1	1	1	1	1	1	0	0	1	0	0	7	
	10	1	1	1	0	1	0	1	0	1	0	0	6	
	2	1	1	0	1	1	0	1	1	0	0	0	6	
	14	1	1	1	1	0	1	0	0	0	0	0	5	
	1	1	1	0	1	0	1	0	0	1	0	0	5	
	13	1	1	0	1	0	1	0	0	1	0	1	5	
	6	1	0	0	1	1	0	1	0	1	0	0	5	
	15	0	1	1	1	0	1	0	0	0	0	0	4	
	11	1	1	0	0	0	0	1	0	1	0	1	4	
	3	0	0	1	0	1	0	0	0	0	0	0	3	
	8	0	0	1	0	0	1	0	0	0	0	0	2	
	12	0	1	0	0	0	0	0	0	0	0	0	1	
		12 11 10 9 8 8 7 6 5 4												

答對人數

圖 4-2(c).

		試題編號 (左邊排答對數多的順序排出)												
		2	3	7	4	9	1	6	5	10	8			
學生編號	S	7	1	1	1	1	1	1	1	1	1	10	合計得分	
	5	1	1	1	1	1	1	1	1	1	0	9		
	9	1	1	1	1	1	0	1	1	0	1	8		
	4	1	1	1	1	1	1	0	0	0	0	7		
	10	1	1	1	0	1	0	1	0	1	0	6		
	2	1	1	0	1	1	0	1	1	0	0	6		
	14	1	1	1	1	0	1	0	0	0	0	5		
	1	1	1	1	0	1	0	0	0	1	0	5		
	13	1	1	0	1	0	0	1	0	0	1	5		
	6	1	0	0	1	0	1	0	1	1	0	5		
	15	0	1	1	1	0	1	0	0	0	0	4		
	11	1	0	0	0	0	1	0	1	0	1	4		
	3	1	0	1	0	1	0	0	0	0	0	3		
	8	0	0	1	0	0	1	0	0	0	0	2		
	12	0	1	0	0	0	0	0	0	0	0	1		
			12 11 10 9 8 8 7 6 5 4											

答對人數

圖 4-2(d).

< 二 > . 典型的 S - P 圖表及其意義

< 1 > . 標準化測驗的 S - P 圖表

如圖 4 - 4 (a)，從 S 曲線和 P 曲線所佔的面積，我們可以看出其平均答對率為 50 %，而從 P 曲線其試題答對率大致分佈在 20 % ~ 80 % 之間，且得分率在 50 % 的學生最多，愈靠近滿分及零分的學生人數也愈少，並且下降之弧度不快。因此從整個 S - P 圖表看來，無論在試題之困難度或學生學習的分佈情形，都顯得相當調和，此乃標準化測驗之 S - P 圖表。

< 2 > . 達到度測驗的 S - P 圖表

如圖 4 - 4 (b)，從 S、P 曲線所佔的面積來看，其平均答對率亦是在 50 % 左右，且由 S 曲線可以看出學生答對率，由高而低形成向左下成直線的傾斜，對於鑑別到達度的個別差異，可一目瞭然，極為方便。

< 3 > . 部分學生到達度急降的 S - P 圖表

如圖 4 - 4 (c)，從 S、P 曲線所佔的面積來看，其平均答對率在 70 % ~ 75 % 之間，但從 S 曲線的下半部可以看出少數學生得分有急速下降現象，是值得注意的地方。

< 4 > . 高低得分率顯然分離的 S - P 圖表

如圖 4 - 4 (d)，得知其平均答對率在 60 % 左右，但從 P 曲線來看，我們可以明顯的看出一奇怪的現象——在試題中，分成高答對率的試題群和低答對率的試題群，而且是明顯的分離，這對試題的編排來說是一個不良的現象。

< 5 > . 編序學習後的 S - P 圖表

如圖 4 - 4 (e)，平均答對率大約在 80 % 左右，除少數學生的答對率稍低以外，其餘的學生都有相當不錯的答對率，此類型是屬於編序學習 (Programmed Learning) 所常見的 S - P 圖表。

< 6 > . 學前測驗的 S - P 圖表

如圖 4 - 4 (f)，平均答對率在 25 % 左右，試題的答對率以及學生的達到度都非常的低，這是屬於標準的學前測驗的 S - P 圖表。

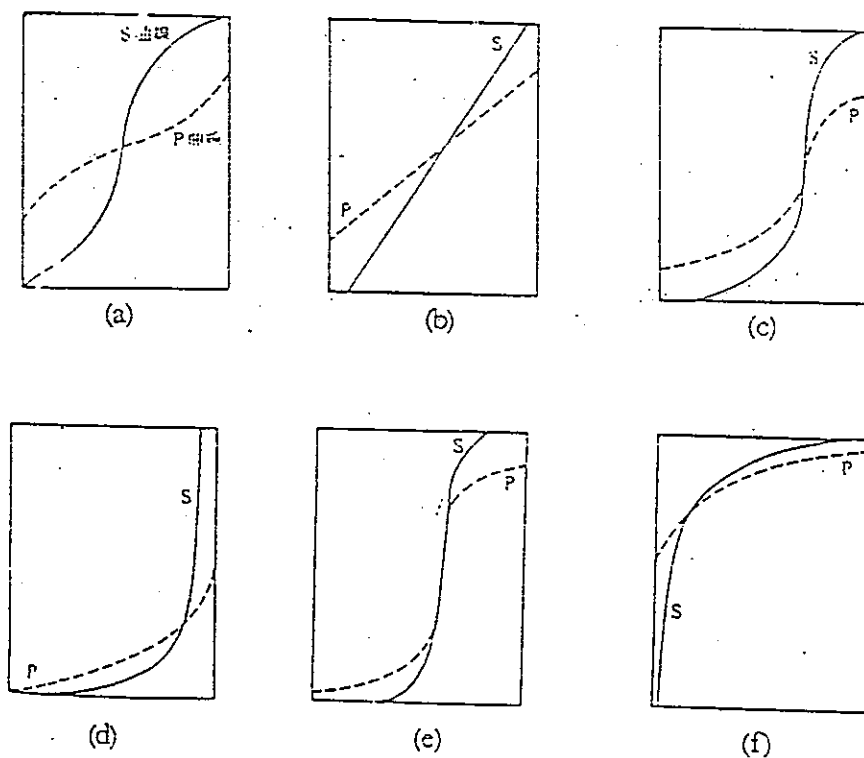


圖 4 - 4 .

< 二 >. 差異係數與注意係數

在 S - P 圖表的分析過程中，我們除了可以根據 S 曲線與 P 曲線的形狀很容易看出學生整體的學習狀況，此外更可將 S - P 圖表的資料作更進一步的處理，而得到更多的資訊。

< 1 >. 差異係數

所謂差異係數是『實際所得 S - P 圖表的 S, P 曲線所夾面積與隨機 S - P 圖表 S, P 曲線所夾面積的比值』。根據 [12] 引述佐藤隆博博士提出的算法，對於差異係數可以得到一個簡便且容易實施 (Implement) 之計算公式：

$$\text{差異係數 } (D^*) = \frac{\text{S, P 曲線之間 "0" 與 "1" 個數}}{4 N n P (1 - P) * D(M)}$$

P : 代表答對题目的機率，N : 代表學生人數。

n : 代表試題數，M = (N * n) + 0.5 取整數值。

D(M) 的值由表 4-1 可查得。

差異係數之應用在於考查學生的整體學習力，以 $D^* = 0.5$ 時為標準，差異係數若大於 0.6 或小於 0.4 皆表示『要注意』，代表著試題中有異質因素的存在，必需加以檢討。

表 4 - 1.

M	$D_s(M)$	N	$D_s(N)$
11	0.278	29	0.353
12	0.285	30	0.358
13	0.291	31	0.360
14	0.296	32	0.362
15	0.302	33	0.362
16	0.307	34	0.366
17	0.312	35	0.367
18	0.317	36	0.369
19	0.321	37	0.370
20	0.326	38	0.372
21	0.330	39	0.373
22	0.334	40	0.375
23	0.337	41	0.377
24	0.341	42	0.378
25	0.344	43	0.388
26	0.347	44	0.381
27	0.350	45	0.382
28	0.353		

< 2 > . 注意係數

在 S - P 圖表中，我們可以從某個學生在 S 曲線左右兩邊的答題情況（"0" 與 "1" 的分佈情形），或是從某個試題在 P 曲線上下兩端的答對學生人數，分別看出某學生或試題是否特異性或值得注意的地方。根據所分析的結果，進一步的針對值得注意的學生或是有特異現象的試題，加以輔導或改進，對於學生的學習進度以及輸入法修正有很大的幫助。其公式如下：

CS_i (S-P 圖表，由上而下第 i 位學生的注意係數)

$$\frac{\left(\begin{array}{l} i \text{ 生的 S 曲線左邊與 "0"} \\ \text{對應的答對學生數之和} \end{array} \right) - \left(\begin{array}{l} i \text{ 生的 S 曲線左邊與 "1"} \\ \text{對應的答對學生數之和} \end{array} \right)}{\left(\begin{array}{l} i \text{ 生的 S 曲線左邊} \\ \text{試題答對者數之和} \end{array} \right) - \left(\begin{array}{l} i \text{ 生的得分} \\ \text{合計} \end{array} \right) * \text{學生平均}} \\ \text{答對數}$$

CP_j (S-P 圖表，由左而右第 j 題試題的注意係數)

$$\frac{\left(\begin{array}{l} \text{試題 j 的 P 曲線上端} \\ \text{與 "0" 對應的得分和} \end{array} \right) - \left(\begin{array}{l} \text{試題 j 的 P 曲線下端} \\ \text{與 "1" 對應的得分和} \end{array} \right)}{\left(\begin{array}{l} \text{試題 j 的 P 曲線上端} \\ \text{學生得分之和} \end{array} \right) - \left(\begin{array}{l} \text{試題 j 的} \\ \text{答對者數} \end{array} \right) * \text{平均得分}}$$

例如 如圖 4-5, $N = 16$, $n = 10$

$$\begin{aligned} \text{學生平均答對數} &= \frac{1}{10} (15+14+13+12+11+11+11+9+5+5) \\ &= 10.6 \end{aligned}$$

$$(15+14+13) - (11+11+9)$$

$$CS_{13} = \frac{(15+14+13+12) - (4 \times 10.6)}{(15+14+13+12) - (4 \times 10.6)} = 0.948$$

同理 $CS_{11} = 0.156$.

所以可以看出，編號 13 的學生注意係數很大，值得注意。我們再回到 S - P 圖表來看，編號 11 的學生，由其 S 曲線左右兩端 "0" 與 "1" 的分佈情況得知，此生雖然得分不高，但一些簡單的題目他亦能答對，而答錯者都是屬於較難的問題，這是一合理之現象。反觀編號為 13 的學生其 S 曲線左右兩端 "0" 與 "1" 的分佈情形，非常奇怪，答對了一些困難的題目，一些簡單的題目卻反而答錯了，顯然這是一個反常的現象值得注意。

注意係數的判別原則如圖 4-6(a), 4-6(b)。

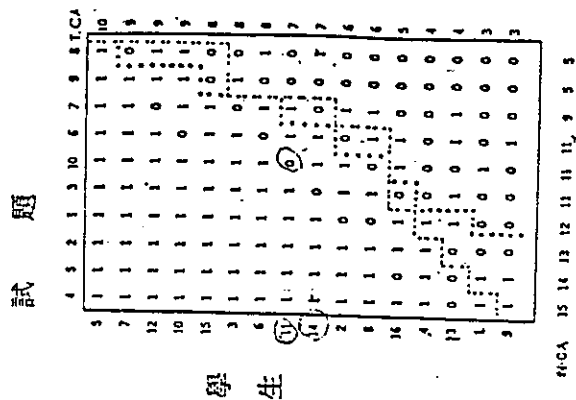


圖 4-5

滿分

學生的答對得分

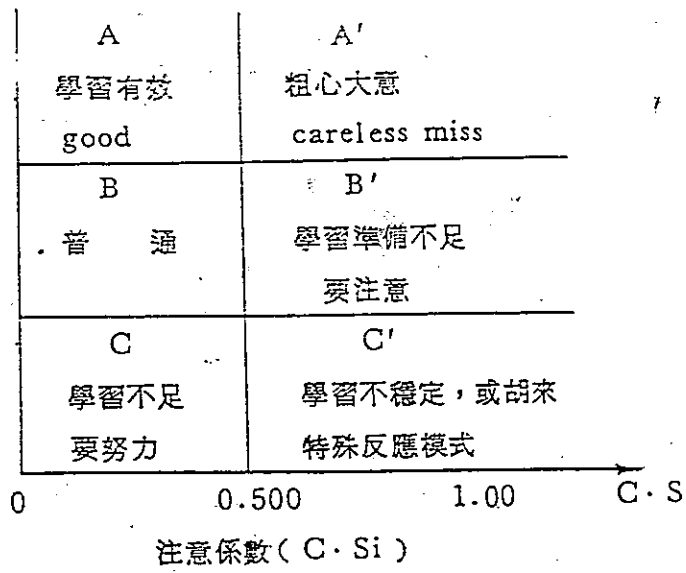


圖 4-6(a). 學生注意係數之判別

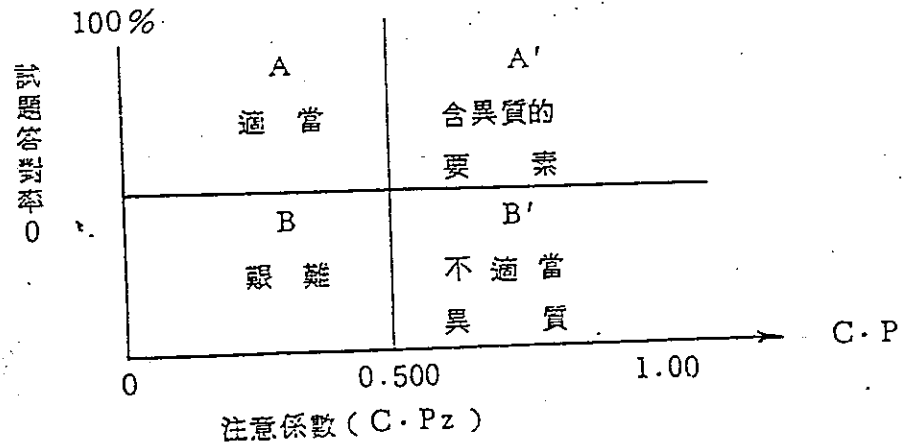


圖 4-6(b). 試題注意係數之判別

第三節：輔助評估模式

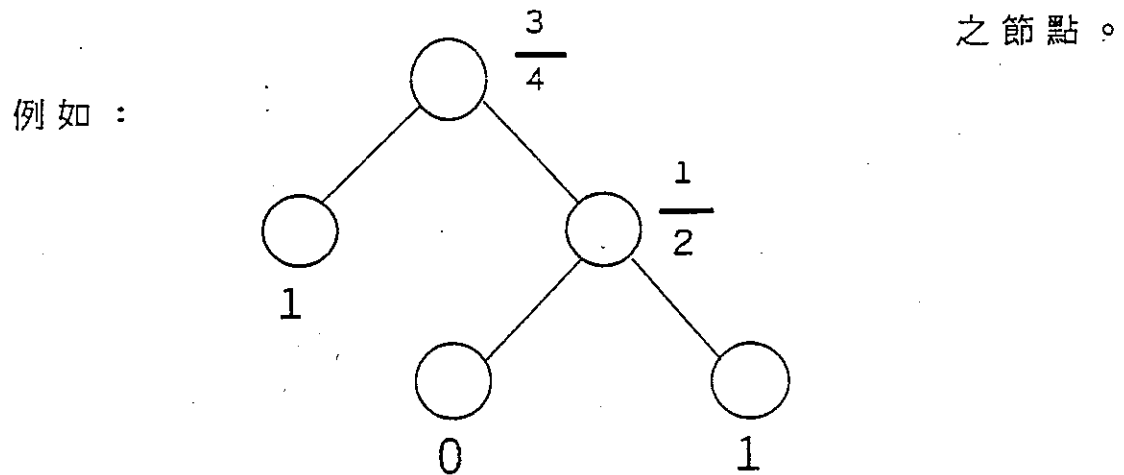
由於 S - P 圖表所能提供的資訊有限，因此我們擴展 S - P 圖表分析，提出幾項輔助評估的項目來說明測驗的現象。

< 1 >. 學習函數 (F(u))

一個中文字是由數個字碼所組成，測驗者答錯一個字，並不能代表所有的字碼都錯，然而在 S - P 表格中，卻只考慮 "對" 與 "錯" 兩個極端值 (即 "0" 與 "1")。它所代表的意義是全懂或是完全不懂，不能表達受測驗者對錯誤問題的瞭解程度，因此我們利用模糊集 (Fuzzy Set) 的概念提出一學習函數，以 0 ~ 1 的值來表示學生對問題的瞭解程度。其公式如下：

$$F(u_j) = \frac{\sum_{j=1}^M F(u_{i+1})_j}{M}$$

0 or i : 基本字碼所在之節點。
 : 非基本字碼所在



我們將利用受測驗者的平均 F(u) 與平均得分之差異來解釋拆解的一些現象。

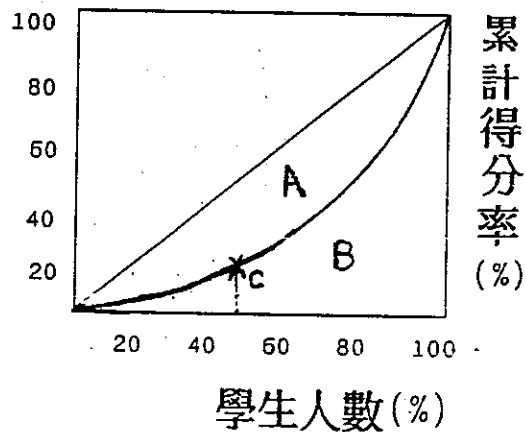
< 2 >. 吉尼 (Gini) 係數

如圖：X 軸表所佔的學生比例。

Y 軸表相對學生比例之得分比例數。

如 C 點表分數比較低之 50 % 的學生得分，佔總得分之 20%。

Gini = $\frac{A}{A + B}$: 值愈大者表學生得分愈呈高低分佈。



< 3 >. 鍵盤熟悉度

其公式如下：

$$\text{熟悉度}(X) = \frac{\sum_{i=1}^m \frac{(n_i - e_i)}{n_i}}{m}$$

n_i : 表字碼 i 在此次測驗出現之總次數。

e_i : 表字碼 i 在此次測驗錯誤之總次數。

m : 表此次測驗出現之字碼總數。

也就是說，我們以字碼為單位 (取代以字為單位)，並加字碼出現的頻率，來探討對鍵盤的熟悉度。

第五章：分析結果輸出單元

第一節：單次測驗的分析結果輸出

為了瞭解學習過程的學習狀況，因此在每次測驗之後，我們便把當次測驗的分析結果輸出，並配合相關的圖形輸出，以便讓使用者對每一次的測驗結果都能有所瞭解。

在這個部分的輸出項目包括：

< 1 >：原始的 S - P 表格

這是測驗後而未經過處理 S - P 表格。如圖 5-1。

< 2 >：處理後之 S - P 表格，並且包括學生的注意係數、試題注意係數等，如圖 5-2。

		試 題 代 號 (problem number)																																				
student number		1	2	3	4	5	6	7	8	9	0	1	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2	2	3	3	3							
學 生 編 號	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	0	1	1	1	0	1	0	0	1	1	0	0	0	0	0	0	1	1	0	0	0		
	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1		
	3	1	0	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
	4	0	1	1	1	1	1	1	0	1	0	0	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
	5	1	1	1	1	1	1	1	1	1	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
	6	1	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
	9	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
	10	1	1	1	1	1	1	1	1	1	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
	11	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
	12	1	1	1	1	1	1	1	1	1	1	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	13	1	1	1	1	1	1	1	1	1	1	0	0	0	1	0	0	0	0	0	0	0	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0
	15	1	1	1	1	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	16	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	18	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	19	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	20	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	21	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	22	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	24	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	26	1	1	1	1	1	1	0	1	1	0	0	0	1	1	1	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	27	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	28	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	29	1	1	1	1	0	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	30	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	32	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	34	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	35	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	36	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	37	1	0	1	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	38	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

圖 5-1.

student number	試題代號																				得分CA	得分率%	注意係數CS
	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1			
16	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	31	100	0.00
24	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	29	94	0.04
36	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	28	90	0.41
27	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	27	87	0.38
30	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	25	81	0.22
18	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	25	81	0.33
12	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	25	81	0.73
11	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	24	77	0.27
2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	24	77	0.44
32	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	23	74	0.02
19	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	22	71	0.12
28	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	22	71	0.25
3	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	22	71	0.30
10	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	22	71	1.13
34	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	21	68	0.15
38	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	21	68	0.17
29	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	21	68	0.63
35	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	21	65	0.01
22	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	20	65	0.04
20	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	20	65	0.05
5	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	20	65	0.30
21	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	20	65	0.43
15	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	18	58	0.28
6	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	18	58	0.56
26	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	18	58	0.73
9	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	17	55	0.11
4	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	17	55	0.18
1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	16	52	0.36
37	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	13	42	0.73
13	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	11	35	0.10
答對人數	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	1	1	1
	9	8	8	8	8	7	7	6	5	5	5	5	5	4	4	3	2	2	1	0	3	3	2
答對率	9	9	9	9	9	9	9	8	8	8	8	8	8	8	7	7	7	7	6	6	4	4	4
	7	3	3	3	3	3	0	7	3	3	3	3	3	0	7	3	3	0	7	3	3	0	3
	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
試題注意係數	4	4	0	2	0	0	3	8	1	2	3	5	5	5	8	6	2	6	6	2	3	4	
	6	8	3	3	5	1	8	7	3	5	1	4	5	0	0	9	6	1	4	6	5	2	

< 3 >. S - P 曲線

利用處理後之 S - P 表格畫出 S 與 P 曲線，提供使用者，來判別學生與試題的整體情況。如圖 5-3。

< 4 >. 平均輸入速度

< 5 >. 平均錯誤率

< 6 >. 吉尼係數，並繪出圖形供參考，如圖 5-4。

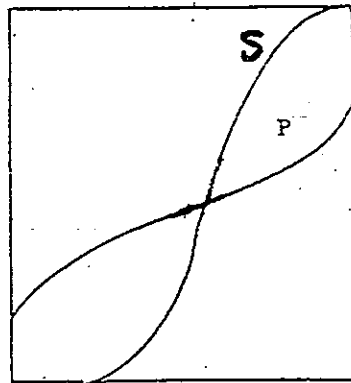


圖 5-3.

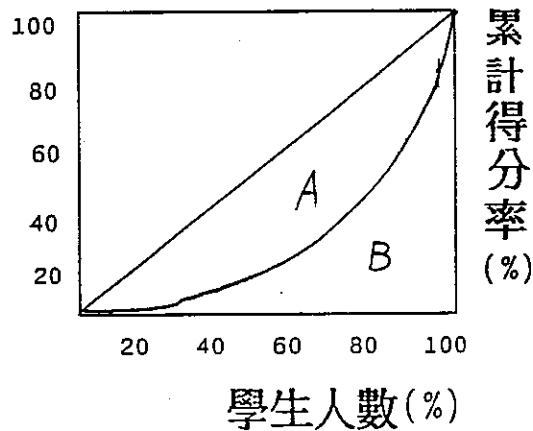


圖 5 - 4.

第二節：以規則資料庫為基礎的解釋子系統

(Rule Base Explanation System)

在第四章中，我們已經對 S - P 圖表所代表的意義以及一些參數，諸如差異係數、注意係數、學習函數、彈性係數等，作過詳盡的介紹。然而對一位評估人員來講，要根據這些圖表和數據來評鑑學生的學習狀況以及試題是否有特異現象，除對 S - P 圖表要有相當之認識以外，還必須要有豐富的經驗，才能對這些現象作一個客觀的評鑑與合理的解釋。

因此在本節當中我們將根據，專家系統的概念，將專家們在這方面豐富的經驗建立成一知識庫 (Knowledge Base)，再根據處理後所得的事實，利用一個推理機 (Inference Machine) 來推衍結果並以自然語言輸出，以達到諮詢的目的。幫助中文輸入法評估者，在測驗的過程中得知受測驗者的學習狀況以及輸入法的特性，找出輸入法的困難處。

< 一 > . 解釋子系統的基本架構

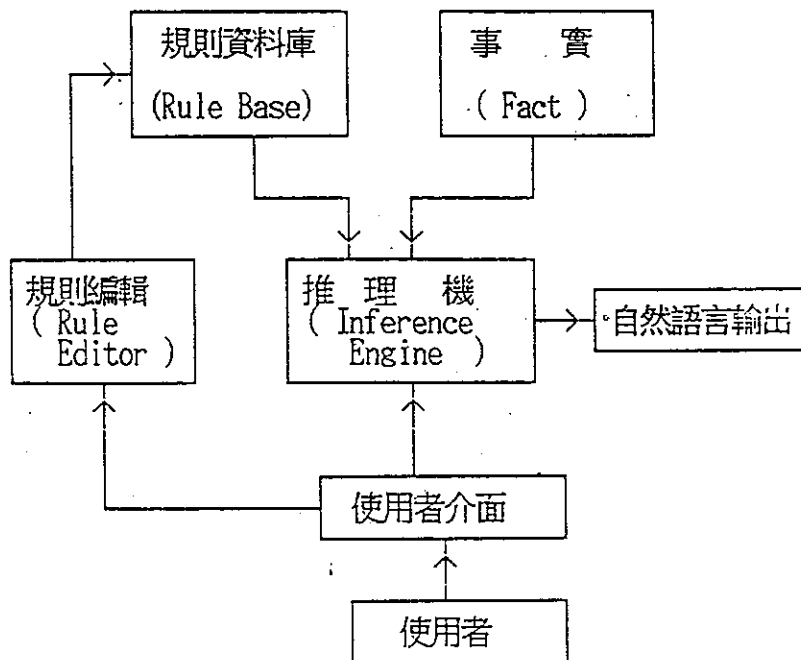


圖 5-5. 解釋子系統架構圖

如圖 5-5，即是解釋子系統的基本架構，一共包括了規則資料庫、事實、規則編輯單元、推理機、結果輸出單元以及使用者介面。

< 二 >. 知識庫 (Knowledge Base)

此部分包括經由 S - P 圖表分析處理後，所得的一些數據 (FACT)，諸如某受測驗者的得分數，注意係數，... 等，以及 S - P 圖表分析專家們，根據 S - P 圖表的特性，透過規則編輯常規，將以 HORN 子句 (HORN Clause) 的方式建立在知識庫中的規則，並可根據實際的需要而加以修改或擴充。包括：

- < 1 >. 根據 S-P 圖表之學生注意係數來解釋受測驗者的行為狀態。
- < 2 >. 根據 GINI 係數來解釋學生答題分佈情形。
- < 3 >. 根據 D 與 $F(u) - S$ 來解釋拆解的現象。
- < 4 >. 根據鍵盤熟悉度 X 與鍵盤測驗之 S 來解釋鍵盤的安排現象 [16]。

所謂 HORN 子句必須是以下的格式 [14]：

$$A_1 \wedge A_2 \wedge \dots \wedge A_n \rightarrow B_1 \vee B_2 \vee \dots \vee B_n$$

其限制為：<1>. 所有的 A_i, B_i (Atomic Formula) 都是第一階 Predicate Calculus，也就是說對於所有的 A_i, B_i 都是 $P(t_1, \dots, t_k)$ 的形式，其中 P 代表屬性 (Predicate) 而 t_1, \dots, t_k 為變數，函數表示式或是常數。

<2>. $n \geq 0, 0 \leq m \leq 1$

<3>. 必需是 Well Formula。

當 $m = 1, n = 0$ 時稱為 EDB 子句 (Extensional Data Base), 表可斷言的事實 (Assert Fact)。

當 $m = 1, n > 0$ 時稱為 IDB 子句 (Intensional Data Base), 表推衍的規則。

例如 : LF : 學習函數值 DT : 拆解現象 S : 成績

<1>. 事實部分

若第二次測驗學習函數值為 $LF=0.8$, 得分率 $=0.3$

則可表成 : $LF(2, 0.8), S(2, 0.3)$ 。

<2>. 規則部分

若規則中有 :

IF 學習函數值與得分率之差 = "大"

AND 得分率 < 0.35 THEN 拆解法等級 = "好"

IF 學習函數值 - 得分率 > 0.4 THEN

學習函數值與得分率之差 = "大"

若 $D(X, Y)$ 表第 X 次測驗學習函數值與得分率差之等級 "Y"。

則規則可表成 :

$D(X, "大"), S(X, Y), Y < 0.35 \rightarrow DT(X, "好")$

$LF(X, Z), S(X, Y), Z - Y > 0.4 \rightarrow D(X, "大")$

< 三 >. 推理機 (Inference Engine)

在專家系統中的推理機，是一個根據已建立在知識庫中的一些專家們所提供的規則，再配合資料庫資料或使用者所提示的事實，根據自己的推衍策略，來推衍結論的機器。推理機的運作方式大抵可分為兩類：一是前推式鏈結 (Forward Chaining)，另一種是回溯式鏈結 (Backward Chaining)。而這兩種方法的差異就是在於推理機如何達成其目標 (Goal) 的方式。

< 1 >. 前推式鏈結 (Forward Chaining)

前推式鏈結有時又稱為資料驅動法 (Data-Driven Method)，因為其推理原理是根據使用者所提供的事實，在邏輯的『且』以及『或』的網路中移動，直到找到目標為止。因此前推式鏈結最主要的精神就是由某些資訊開始，嘗試著去找尋符合這些資訊的標的物。

< 2 >. 回溯式鏈結 (Backward Chaining)

回溯式鏈結與前推式鏈結，推理方式恰巧相反，它是由一個假設的目標物為起點，要求使用者提供資訊或由資料庫讀取資料，來確認或否定這個目標物。如果這個目標物被否定，則再找另一個目標物來驗正。因此回溯式鏈結的基本精神就是以一個假設的目標物為起點，利用已有的或是使用者提供的資訊來驗證它，所以有時又稱目標導向法 (Object-Driven Method)。

在我們的推理部分，將採用回溯式鏈結的方式來推理，根據 Kowalski [15]指出，Horn Clauses 可分為四類：

< 1 >. $A_1 \wedge A_2 \wedge A_3 \cdots \wedge A_n \longrightarrow B$ ($n \neq 0$) [IDB 子句]
 其中 B 稱為 Procedure Name, $A_1 \wedge A_2 \wedge \cdots \wedge A_n$
 稱 Procedure Body, A_i 稱 Procedure Call。

< 2 >. $\longrightarrow B$ ($n = 0$) [EDB 子句]

表為可斷言的事實 (即由 S-P 圖表處理所得之資料)。

< 3 >. $A_1 \wedge A_2 \wedge \cdots \wedge A_n \longrightarrow$ ($m=0$)

稱為目標指示 (Goal Statement)。

< 4 >. \square ($m=0, n=0$)

稱為終止指示 (Halt Statement)。

而推理的過程就是根據 IDB 與 EDB 子句由目標指示推衍到終止指示。其推衍過程可分為二類：

< 1 >. 事實推衍 (EDB 子句)

令 $\alpha = G_1 G_2 \cdots G_i G_{i+1} \cdots G_n$ 為以 Horn 子句表

示的詢問句，而目前已處理到 G_i ，又令 G^1 為一 EDB 子句，且 G_i 與 G^1 "Match" (即有相同的 Predicate Name，以及有一個取代式 θ (Substitution) 使得此二 Atomic Formula 對應，則 α 可推衍成：

$$\alpha' = (G_1 G_2 \cdots G_i^* G_{i+1} \cdots G_n) \theta$$

G_i 之後的 * 表示已處理過。

例如：

$$\alpha = G_1(x, y)G_2(x, a)G_3(y, z) \text{ 且 } G^1 = G_2(b, a)$$

則 $\theta = \{ b / x \}$ ，所以

$$\begin{aligned}\alpha' &= (G_1(x, y)G_2(x, a)*G_3(y, z)) \{ b/x \} \\ &= G_1(b, y)G_2(x, a)*G_3(y, z)\end{aligned}$$

< 2 >. 規則推衍 (IDB 子句)

若 $A_1A_2A_3 \cdots A_k \longrightarrow G^1$ 為 IDB 中之一個子句，
且 G^1 與 G_i " Match "

$$\text{則 } \alpha' = (G_1 \cdots G_i*(A_1 \cdots A_k)G_{i+1} \cdots G_n)\theta$$

例如： $\alpha = G_1(x, y)G_2(x, a)G_3(y, z)$

$$\text{且 } G_4(u, v)G_5(v, w) \longrightarrow G_2(u, w)$$

$$\text{則 } G^1 = G_2(u, w), \theta = \{ x / u, a / w \}$$

所以 α'

$$\begin{aligned}&= (G_1(x, y)G_2(x, a)*(G_4(u, v)G_5(v, w))G_3(y, z)) \{ x/u, a/w \} \\ &= G_1(x, y)G_2(x, a)*(G_4(x, v)G_5(v, a))G_3(y, z)\end{aligned}$$

以推理解釋拆解法現象的例子來說明：

若 EDB：LF(2, 0.8)，S(2, 0.3)

IDB：D(X, "大"), S(X, Y), Y < 0.35 \longrightarrow DT(X, "好")

LF(X, Z), S(X, Y), Z - Y > 0.4 \longrightarrow D(X, "大")

且詢問句為 ? DT(2, P) [即第二次測驗拆解法等之級?]

則其推理過程如下：

1. $DT(2, "好")$
2. $DT(2, "好") * [D(2, "大"), S(2, Y), Y < 0.35]$
3. $DT(2, "好") * [D(2, "大") * [LF(2, Z), S(2, Y), Z - Y > 0.4], S(2, Y), Y < 0.35]$
4. $DT(2, "好") * [D(2, "大") * [LF(2, 0.8) * , S(2, Y), 0.8 - Y > 0.4], S(2, Y), Y < 0.35]$
5. $DT(2, "好") * [D(2, "大") * [LF(2, 0.8) * , S(2, 0.3) * , 0.8 - 0.3 > 0.4 *], S(2, 0.3) * , 0.3 < 0.35 *]$

至此我們可得 $DT(2, "好")$ 成立，表示第二次測驗分析的結果顯示此拆解法的等級是 "好"。

< 四 >. 自然語言輸出單元

經過推理得到結果以後，我們希望從推理的過程中找出理由，並以自然語言的方式輸出。

< 1 >. 結論與理由之抽取

假設 $(G_1 * [A_1 * [A_2 * A_3 *] A_4 *] G_2 * [A_5 * A_6 *])$

為我們所推衍的結果，則我們可用一 PUSH-DOWN Automaton 來抽取結論與理由。其規則如下：

1. 預設 Level 為 0。
2. 由左而右掃瞄，碰到左刮號則 Level 加 1，而前面之 Atomic Formula 則為一規則之結論。

3. 遇到右括號則 Level 減 1，而在此左右括號之 Atomic Formula 則為 2. 當中規則之條件部分。

4. 否則為事實。

5. Level 為 0 則為答案。

例如：DT(2, "好") * [D(2, "大") * [LF(2, 0.8) * S(2, 0.3) *
0.8 - 0.3 > 0.4 *], S(2, 0.3) * , 0.3 < 0.35 *]

結論與理由之抽取的過程如下：

Level	Input
1. 0 :	DT(2, "好") 答案
2. 1 :	D(2, "大"), S(2, 0.3), 0.3 < 0.35 [規則 → DT(2, "好")
3. 2 :	D(2, "大") 事實
4. 2 :	LF(2, 0.8), S(2, 0.3), 0.8 - 0.3 > 0.4 [規則 → D(2, "大")
5. 2 :	LF(2, 0.8) 事實
6. 2 :	S(2, 0.3) 事實
7. 2 :	0.8 - 0.3 > 0.4 0.8 - 0.3 > 0.4 事實
8. 1 :]
9. 1 :	S(2, 0.3) 事實
10. 1 :	0.3 < 0.35 0.3 < 0.35 事實
11. 0 :]

< 2 >. 自然語言輸出

分析專家們建立知識庫的同時，並賦於每條規則與事實一句中文文句模型 (Template)，以便將規則與事實套到這些中文文句模型上，便可以自然語言輸出。

例如：

答案：DT(2,"好")

理由：D(2,"大"),S(2,0.3), $0.3 < 0.35 \longrightarrow$ DT(2,"好")

LF(2,0.8),S(2,0.3), $0.8-0.3 > 0.4 \longrightarrow$ D(2,"大")

事實：LF(2,0.8)

S(2,0.3)

$0.8 - 0.3 > 0.4$

$0.3 < 0.35$

英文文句模型如下：

DT：% 第 @1 次測驗分析結果，拆解法等級為 @2 %

D：% 第 @1 次測驗，學習函數值與 S 的差 @2 %

LF：% 第 @1 次測驗，學習函數值為 @2 %

S：% 第 @1 次測驗，S 值為 @2 %

>：% @1 大於 @2 %

<：% @1 小於 @2 %

自然語言輸出為：

答案：第 2 次測驗分析結果，拆解法等級為 "好"。

理由：1. 第 2 次測驗，學習函數值與 S 的差 "大"

『且』 第 2 次測驗，S 值為 0.3

——> 第 2 次測驗分析結果，拆解法等級為 "好"。

2. 第 2 次測驗，學習函數值為 0.8 『且』

第 2 次測驗，S 值為 0.3 『且』

$0.8 - 0.3$ 大於 0.4

——> 第 2 次測驗，學習函數值與 S 的差 "大"。

3. 第 2 次測驗，學習函數值為 0.8

4. 第 2 次測驗，S 值為 0.3

5. $0.8 - 0.3$ 大於 0.4

6. 0.3 小於 0.35

應用專家系統推理技巧來作解釋系統最大的好處，就是知識庫與推理機互相獨立，雖然規則不同的知識庫，但仍可用相同的推理機來推論，不像傳統演繹式的程式，當規則修改時則必須修改原來的原始程式。

第三節：整體分析結果輸出

為了讓使用者對於輸入法的整體特性有所瞭解，因此經過一連串的測驗以後，我們將整合多張 S - P 圖表的分析結果以及輸入法的基本特性，在此部分輸出，而給予使用者一張中文輸入法之成績單，其輸出項目包括：

- < 1 >. 平均碰撞率
- < 2 >. 平均敲鍵數
- < 3 >. 輸入速度曲線圖
- < 4 >. 拆解速度曲線圖
- < 5 >. 吉尼係數變化圖
- < 6 >. 差異係數以及學習函數變化圖

如下圖 5-6：

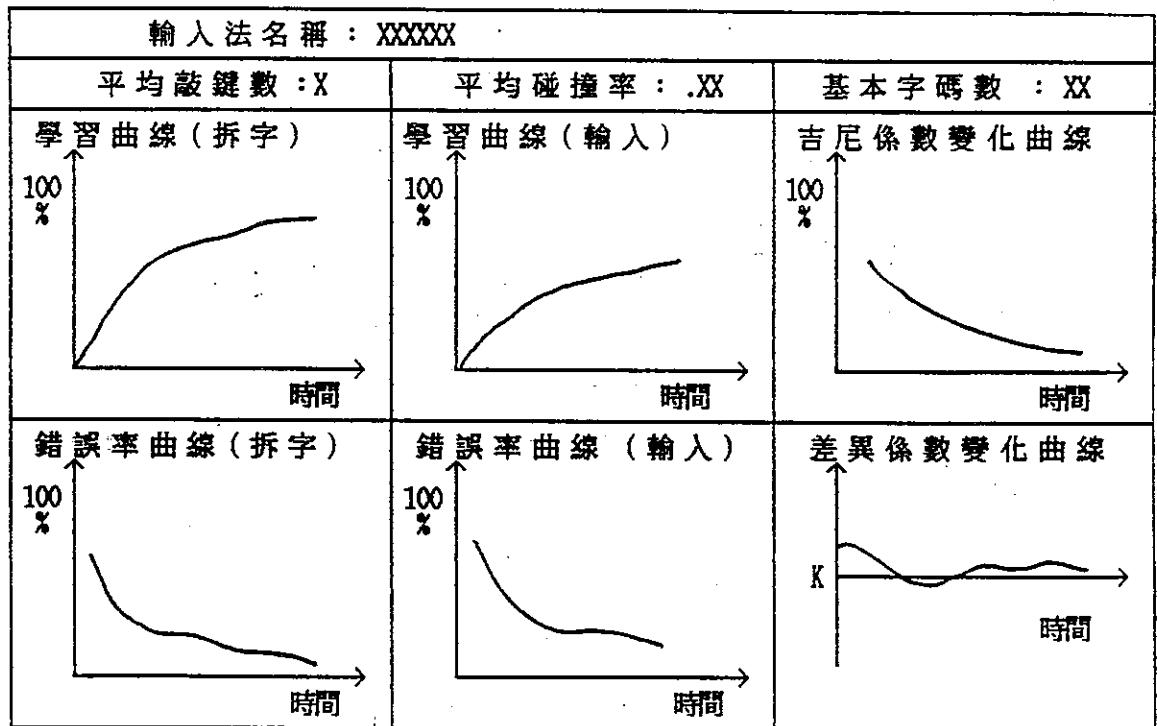


圖 5-6：

第六章：結 論

在這篇文章裡，我們應用電腦，對以標準鍵盤當作輸入媒介的中文輸入法，提出一套整體的評估模式。我們利用視覺語言 (Visual Language) 的觀念提供一個使用者與電腦之間的人-機介面 (Man - Machine Interface)，讓評估者或輸入法設計者能夠透過建立基本字碼，拆解中文字碼，安排鍵盤等來定義其輸入法，則系統自動產生測驗的題目，分階段進行測驗。然後透過 S - P 圖表的分析技術，學習函數，Gini 係數與鍵盤熟悉度等計算分析。最後並以規則知識庫的概念，將專家們的知識以 Horn Clause 的形式儲存在知識庫中，利用一個推理機來推理解釋，並將解釋的結果與原因，以自然語言的方式輸出。

本系統的優點有：

- < 1 >. 利用電腦大量的儲存空間，與快速的計算能力，來取代人工的計算與分析，大大地減少人力與物力的投資，降低成本。
- < 2 >. 對於尚未發展成型的中文輸入法，可透過本系統的模擬，來進行測驗與分析，以瞭解其基本特性與問題之所在。以便修正與改良，使得設計過程更有彈性，而以真實鍵盤來進行輸入測驗，更有真實性。
- < 3 >. 分析過程中，利用 S - P 圖表分析的學生注意係數，將影響輸入法速度與學習時間的人為因素也考慮進去。使分析的結果更具代表性。
- < 4 >. 利用規則知識庫的觀念，將專家們的經驗放到知識庫中，可以隨時更改或擴充，而不必修改原始程式。

< 5 >.. 以自然語言配合相關圖表的輸出，使得使用者對分析結果，更能一目了然。

然而，如何能夠將專家們的知識與經驗，透過一個簡單的方式（如，以自然語言輸入）建立到知識庫中，是一件非常困難的工作，有待我們作進一步的研究。而如何以更簡單的方式從輸入法設計者中，獲得更多的資訊也是將來可繼續探討的地方。

参 考 文 献

- [1]. L. C. Chen, " 1982 Chinese Computer Market Report ", III Technical Report C20, PP. 19-35, 94-95, Taipei, June, 1983.
- [2]. Cheng-Kuang Chen And Reng-Weng Gong, " Evaluation of Chinese Input Methods ", Computer Processing of Chinese & Oriental Languages Vol. 1, NO. 4, November, 1984.
- [3]. S. Y. Lo, " A Scientific Model For Comparing Various Methods of Inputting Chinese Characters Into Computer ", Computer Processing of Chinese & Oriental Languages Vol. 2, No. 1 May, 1985.
- [4]. Robert B. Grafton, Office of Naval Research, " Visual Programming ", IEEE Trans. on Computer, PP. 6-9, 1985.
- [5]. S. Feiner, S. Nagy, and A. VanDam, " An Experimental System For Creating and Presenting Interactive Graphical Documents ", ACM Trans. Graphics, Vol. 1, No. 1, Jan. 1982, PP. 59-77.
- [6]. Nan C. Shu, IBM Los Angeles Scientific Center, " FORMAL: A Forms-Oriented, Visual-Directed Application Development System ", IEEE Trans. on Computer, August, 1985, PP. 38-49.

- [7]. M. H. Brown and R. Sedgewick, " Techniques for Algorithm Animation ", IEEE Software, Vol. 2, No. 1, Jan. 1985, PP. 28-39。
- [8]. K. Y. Cheng, C. C. Hsu, I. P. Lin, M. C. Lu, and M. S. Hwu, "VIPS, A Visual Programming Synthesizer ", Second IEEE Computer Society Workshop on Visual Language, June. 1986。
- [9]. P. S. P. Wang and Y. Y. Zhang, " Decomposition and Recognition of Handprinted Chinese Characters ", College of Computer Science Northeastern University Boston, MA。
- [10]. Kiang, T. Y., " Construction of Chinese Characters In Terms of Basic Strokes, Proc. Int. Conf. of Chinese Language Computer Society, 1982.
- [11]. 科技研究報告, " 中文電腦基本用字研究 ", 國立交通大學電子計算機科學。
- [12]. 陳騰祥, " S - P 表分析 (Student-Problem Chart Analysis) 在學習診斷的應用法及其實作感受之探究 ", 國立台灣教育學院, 輔導學報, 民 75 年, 9 期, 275-311 頁。
- [13]. 謝文瑞, 胡明森, 盧世屏, 鄭國揚, " S - P 圖表分析與自然語言之輸出 (Natural Language Output for S - P Chart Analysis) ", Proceedings of National Computer Symposium 1987, Vol. 1 PP. 221-229。

- [14] . Jack Minker and Patricia B. Powell , " Answer and Reason Extraction , Natural Language and Voice Output for Deductive Relation Data Base " , Department of Computer Science University of Maryland College Park , MD 20742 USA .
- [15] Kowalski , R. A. , " Predicate Logic As Programming Language " , IFIP-74 , North Holland , 1974 , 569-574 .
- [16] 鄭國揚 , 曾志明 , " 標準鍵盤中文輸入法之評估 " , 中興大學應用數學研究所碩士論文 .