

行政院國家科學委員會專題研究計畫 期中進度報告

子計畫五：潛在語意分析及概念構圖在文章摘要和理解評量的應用(2/3)

計畫類別：整合型計畫

計畫編號：NSC92-2520-S-003-006-

執行期間：92年08月01日至93年07月31日

執行單位：國立臺灣師範大學資訊教育系(所)

計畫主持人：張國恩

共同主持人：宋曜廷

報告類型：精簡報告

報告附件：出席國際會議研究心得報告及發表論文

處理方式：本計畫可公開查詢

中 華 民 國 93 年 6 月 9 日

行政院國家科學委員會專題研究計畫期中報告

潛在語意分析與概念構圖在文章摘要與閱讀理解評量的應用

計劃編號：NSC 92-2520-S-003-006

執行期限：92年8月1日至93年7月31日

主持人：張國恩

執行機構及單位名稱：國立台灣師範大學資訊教育學系

一、中文摘要

由於概念圖在知識結構測量上的有效應用與其評量上的缺失，本計畫將找出概念圖的LSA評量模式，並驗證其效果。本計畫所建立的LSA評量指標以國小自然科為領域，並從事教學實驗以驗證成效。

Abstract

Because concept map is a good representation of knowledge structure, it is used to evaluate students' knowledge. However, the evaluation of concept map is not efficient to be implemented in computer; it is possible better to integrate concept map and LSA approach. In the proposal, we will develop an LSA-based evaluation of concept maps. After finishing the system, we will make an educational experiment to verify the system for educational impacts. The system will use the elementary science as subject and be used in elementary school.

Keywords: LSA, Concept map.

二、緣起及目的

當學生開始著手心得及摘要寫作時，教師應即著手進行有關於閱讀理解或摘要寫作的評量。根據林寶貴(民89)的研究指出，國內目前用以評量

學生閱讀能力的測驗，絕大部分都集中在閱讀歷程中由下而上的基本閱讀技巧，如字義理解、語法理解等方面，較少將評量的層面放在整體的文本理解(text comprehension)、摘要文本大意，或是推論、分析上面。因此，建置一個有效率的寫作評量系統，就顯得非常地迫切與需要的。參考目前常見的摘要評量方法，潛在性語意分析(Latent Semantic Analysis, LSA)的做法是以兩篇文章中關鍵詞向量的餘弦夾角做為衡量的標準(Zhang, 2001)，有效地表現出評量效果。

Foltz, Gilliam and Kendall(2000)認為在寫作的過程中，學生通常把他們的焦點放在如何改進寫作的品質上，而不是把焦點放在如何提高他們概念中所表達的資訊總量，並且認為在寫作過程中對學生寫作的內容深度及所包含的資訊量上提供回饋是非常重要的。Kintsch(2000)相信除了寫作的技巧外，在獨立寫作活動中從所收到的即時回饋將使學生們從自我的後設認知中受益。

摘要和一個人的學習成果之間是具有高度相關的，然而要評量這方面的資訊卻相當不容易。首先必須要針對所學習的教材做全文掃描，並且找出重要的資訊。但是如果受試者的摘

要並沒有使用教材中的字，那麼就不可能以相同字做精確的比對，故需在語意上做學生摘要與教材間的比對，以找出兩者間語意的相似性(Flotz 1996)。對於摘要寫作的看法，Kintsch(2000)認為它比簡單的閱讀測驗具有更多的優點，包括：

1. 摘要寫作並不僅僅是提供寫作的練習，它更包含了重要的學習技巧，例如從細瑣的內容中找出主要的概念。
2. 對特定的目標而言，它需要做更深入的思考與分析，以找出相關的訊息。
3. 摘要從複雜的學習材料中，發展全方位認知(solid understanding)及清晰地表達自知的想法並與他人分享的有效方法。
4. 從適度簡潔地表達中，使學生瞭解他們需要學習摘要的策略，而這已遠遠地超過了單純的字句修改。
5. 摘要需要主動的意義建構，在程度上遠比複選題及簡答題的測驗深奧。
6. 摘要寫作是建構及整合新知識的有效方法。

從以上的訊息可以發現 LSA 在摘要評量方面具有三項特色：第一是強調立即地回饋，第二是強調語意上的評量，第三是使學習者成長。

本計畫以一年的時間參考潛在性語意分析的方法，以關鍵詞為比對摘要的依據設計一套簡易的摘要寫作評量離型系統，使能夠對學生的摘要與心得寫作達到更有效的評量效果。

由於概念圖是知識表徵的有效工具，可用來做閱讀後之摘要練習與重點整理，因此結合概念圖與 LSA 兩者功能發展出新的知識結構評量是值得研究的主題。

三、研究方法

3-1 LSA 評量系統

潛在性語意分析 (LSA) 就是使用統計的方法，從一篇文章中將具有某些概念的詞，擷取出來並重新呈現的一種理論和方法 (Landauer & Dumais, 1997)。這種技術最初的發展是應用於資訊擷取方面 (information retrieval) (Deerwester, Dumais, Furnas, Landauer, & Harshman, 1990)，後來被 Landauer 和他們同事採用在心理語言學的分析上 (Landauer & Dumais, 1997; Landauer, Foltz, & Laham, 1998; Landauer, 1999)。

LSA 所計算出來的相似度，並不僅僅是單純地計算關鍵字出現的頻率、共同出現的次數，和使用上的關連性。此外它還藉助一項功能強大的數學分析工具 SVD (Singular Value Decomposition, 奇異值分解)，所以能夠正確地推理更深層次的關係 (因此稱為 "Latent semantic")。因此在以人類意識為主的判斷上，它比那些長久以來就被人們所唾棄的著重表像的方法，更能夠提供較佳的預測及效能 (Burgess and Lund, 1996)。

Onathan(2000) 將 LSA 的 MODEL 分成下列四個步驟：

1. 建立 occurrence matrix
2. 將 occurrence matrix 中的值做轉換運算
3. 利用 SVD 方法將矩陣分解
4. 利用 singular value 將矩陣重建。

現在分別來探討其內部的詳細運作情形：

3-1-1 建立 occurrence matrix

首先它會從文件中找出關鍵字做為比對及處理的依據。當關鍵字決定好之後，便開始根據關鍵字在文件中出現的位置，建立起一個矩陣。矩陣中的列代表每一個不同的關鍵字，矩陣中的行代表著一篇文件中不同的句子、不同段落、或是不同的文件。矩陣中的每個元素(element)值就代表著某一個關鍵字在文件中不同位置所出現的次數，記為 f_{ij} 。因此從矩陣的

列可以看出某一個關鍵字在不同的句子（段落或文件）中所出現的情形。另外從矩陣的行則可以看出每一個句子（段落或文件）中有那些重要的關鍵字。

3-1-2 occurrence matrix 中的值做轉換運算

在本步驟的主要目的是將 occurrence matrix 中每個元素的值 (f_{ij}) 做以下的轉換如下：

$$- \sum \left[\left[\frac{\log (f_{ij} + 1)}{\sum_j f_{ij}} \right] * \log \left[\frac{f_{ij}}{\sum_j f_{ij}} \right] \right]$$

這種轉換方法是以每一列為轉換單位。從原來的 occurrence matrix 中可以看出每個關鍵字在各句子（段落、文件）間的出現情況，轉換後可以得到每一個關鍵字在各句子（段落、文件）間的重要性。Harman(1996)認為類似這種轉換方式，長久以來就被認為在資訊擷取的效益上提供了顯著的改善。

3-1-3 利用 SVD 方法將矩陣分解

利用 SVD 的方法將矩陣分解成三個矩陣，若原來的矩陣為 $[IJ]$ ，則分解後的三個矩陣分別為 $[IK][KK][JK]'$ ，其中 $K \leq \max(I, J)$ 。矩陣 $[IK]$ 保留原來關鍵字的資訊，矩陣 $[JK]'$ 保留原來各文件的資訊。矩陣 $[KK]$ 為一個對角矩陣。

3-1-4 利用 singular value 將矩陣重建

重建的方法為取矩陣 $\{S\}$ 中最大 d 個 singular values，並將矩陣中未引用的值全設為 0 (這個步驟稱做 reduced-dimension)。然後再將轉換後的三個矩陣相乘，如此就可以得到原來矩陣的最小平方的最佳估計值 (least squares best approximation)。

3-1-5 LSA 的評分方式

Foltz (1996) 認為閱讀的文件和寫作的小論文之間，其語意上的相似程度可以用來評量學習的效果。因此寫作品質較好的人就表示他從文件中瞭解了更多語意上之訊息。

第一種評量的方法是找出學生的小論文和原文間語意相似的句子。其方法是將小論文中的每一個句子和原文件中的所有句子相互比較，找出內容最相似的句子以 cosine 的方法來評分，再以每個句子的平均成績作為小論文的成績。這種評量方法被視為一種機械性背誦的評量，如果學生的句子和原文非常相似的話，評定的成績將非常高。

第二種評量的方法是先由評分者

從同質性文件中選出幾個最重要的句子。然後將學生的每一個句子和被選出的句子相比較，找出最相似的句子並以 cosine 的方法算出兩者之間的成績。以每個句子的平均成績作為小論文的成績。

3-2 概念構圖在評量上的應用

概念圖是由命題所組成，命題包括兩個概念節點及概念間的連結語，而利用概念節點及連結語可以建構出特定知識領域的結構。因此，由學習者建構出來的概念圖，可以視為他對專業知識領域的知識結構，由該知識結構自然可以看出其理解程度以及是否存有錯誤概念。由於傳統選擇題測驗在評量學習成就上常被批評流於瑣碎，無法測得學習者完整的知識狀態 (Frederiksen & White, 1990)，而概念構圖法強調學習者將其所具備知識以具結構性、組織性、完整性的方式表達出來，可直接就其知識的組織與結構情形，了解其學習狀態。因此，以概念構圖法作為新的評量方式，愈來愈為研究者所重視 (Beyerbach & Smith, 1990; Anderson & Chiu, 1989; Novak & Gowin, 1984; Herl, Baker & Niem, 1996; 余民寧 et al., 1996; 邱上真, 1989)。

以概念構圖對學生知識進行評量，大致有下列方向：

(一)、將學生的概念圖與專家(教師)的概念圖相比較，找出異同，進行診斷，或根據學生對概念間關係描述的正確性、概念間的階層性及組織性加以給分。此一取向假定學科知識本身有一潛在結構，學生進行概念構圖時，即表達他們對此一結構的掌握情形，因而可藉概念圖進行評量與診

斷。

(二)、概念構圖應用在評量上的另一個取向，乃是利用概念構圖以了解學生學習時，知識結構的「改變」情形。此派的研究者多受建構主義的影響，認為學習是知識建構的歷程，而構圖的改變，正是反應知識結構改變與建構歷程的最佳寫照。因此，他們較強調學生個別構圖的歷程，希望藉由此歷程讓學生察覺到自身知識改變的特性，進而察覺與修正自身的潛在知識結構。採用此一取向的研究者較不著重於學生概念圖與專家概念圖的比較。

雖然不同的研究者對於應用概念圖於評量上的作法會因其對知識形成歷程的看法不同而互異，但基本上強調概念構圖法用以評量學生知識結構，則是一致持肯定的看法。然而此兩種評量計算皆需參考專家概念圖，用已知的概念名稱與連結語做依據，若學生概念圖用了不明顯或與專家概念圖不同的概念名稱與連結語，則無法用上述方法評量。由於 LSA 具有語意分析的功能，若在概念圖評量時能應用 LSA 技術，對知識結構測量將具重大貢獻，這也是本計畫的研究目的之一。

3-3 實驗設計

為瞭解以潛在性語意分析的方法所設計的評量系統與概念圖的 LSA 評分是否具有評量學生摘要與心得寫作的效果，本計畫將針對國小五六年級或國中一年級的學生來進行實証考驗研究。

3-3-1 實驗對象

本實驗的研究對象為國小五、六

年級與國中一年級學生

3-3-2 研究設計

本研究採相關研究法，以教師的摘要評分成績和學生的單元成就測驗成績作為效標，探討 LSA 與概念圖評分的關係。

3-2-3 實驗教材

本實驗之教材以國民小學五、六年級自然科審定版之教材為主，並參考其他出版社的教材。評量的內容以參與實驗學校所採用的教材版本為主，其餘之相關教材的主要目的在提供系統關鍵詞的擷取及比對。

3-2-4 實驗工具

1. 評量系統：由研究者自行開發的軟體，用以評量學生寫作的成績。
2. 評量試題：參照受測學校所採用之教材，另行委託實驗學校的自然科教師編製，為一份 20 題選擇題型的評量試題，做為外在效標的檢測依據。

四、結論

本年度完成下列工作：

1. 搜集國小自然科各出版社的教材，由專業老師做分析與自然科詞庫建置
2. 依據第一年的實驗結果修改系統功能
3. 擇一實驗小學與國中做自然科的全學期試用，驗證經由摘要與小論文訓練(採評量與回饋方式)後的學習成就
4. 分別探討不同年級、高低成就、男女學生對摘要與小論文訓練後的學習成就

5. 與總計畫的主系統做初步整合
6. 發展概念圖 LSA 評分模式，並建立程式
7. 探討概念圖 LSA 評分模式與詞庫關係，研究其語意分析的功能