

模糊詮釋資料及其在圖書館分類編目之應用

林信成

副教授

蕭勝文

研究生

淡江大學資訊與圖書館學系

摘要

本研究提出模糊詮釋資料(fuzzy metadata)之概念，並將其應用於圖書館的分類與編目上，以加強書目資料的關聯性，提升檢索結果的廣度。本文首先對於分類編目、詮釋資料、模糊理論等不同領域的研究進行探討；接著，闡述模糊詮釋資料之定義，並以模糊理論模型解釋其意涵；再者，提出圖書模糊分類與編目之實務作法，實際將模糊詮釋資料以中國機讀編目格式註錄至書目資料庫中；最後，發展一套圖書模糊分類與編目實驗系統，利用模糊詮釋資料在相關圖書與類別之間建立不同程度的聯繫，使用者透過此系統，可以從檢索出的資料中，掌握書籍間的相關性，以便進一步查找相關書籍，達到更加滿足其實際資訊需求的目標。

關鍵詞：詮釋資料，模糊理論，模糊詮釋資料，模糊分類，模糊編目

前言

圖書分類與編目乃圖書館的重要工作之一，是個複雜且繁瑣的專業領域，圖書館編目員通常須要經過特殊專業訓練方能勝任。一般而言，編目員進行圖書分類與編目的過程，可歸納為如下程序(註1)：(一)編目員瀏覽書名、目次表、篇章名、前言、序跋、參考書目、附註、正文及書評等之書籍內容做初步主題分析；(二)決定書籍內容的學科重點與層面關係；(三)主客觀性給予適當大類的類號(類表)；(四)再經由更詳細的書籍內容大綱、性質、用途及重點，而給予更明確的複分號與子目。所以，一本書籍要完成分類，必須依次查看「類表」、「簡表」、「綱要表」以決定適當的類目及類號，由類號再進而查其「詳表」找到最能代表圖書資料的類號。因此，在分類與編目的過程中，編目員除了須要具備圖書館學

專業素養外，對於各領域學科知識也要有一定程度瞭解，以利於掌握更明確的學科領域，而給予適當的類號及書目資料。

圖書館自動化系統遵循既定的各種標準化分類法，達到將館藏資料分門別類之目的。不過現今的資訊社會中，各類出版品內容廣泛，主題多變，跨領域書籍屢見不鮮，天生具備極大的「模糊性」。對於牽涉較廣以致涵蓋二種以上主題範圍的書籍，編目員在進行分類時，通常必須依賴個人主觀判斷或既定規則，硬將該冊圖書歸類至某個特定類別中。此種「硬性分類法」(hard classification)抹煞了書籍內容因涉及多重主題而與生俱來的模糊性質，也增加使用者在查找資料時的困難。茲以《圖書館學的哲學》(註2)一書為例，此書至少可歸屬於「總類」、「哲學類」兩類別。若將其歸屬於「總類」，則在「哲學類」便無法找到此冊圖書；反之，若將其歸屬在「哲學」類，則在「圖書館學」類又找不到，導致分類查找時的回現率降低。由此可知，若要使讀者在分類查找時能找到更多相關的圖書資料，可捨硬性分類法而採用「軟性分類法」(soft classification)。而模糊分類是軟性分類法的一種，可增強不同類別的關聯性，使分類結果更平順自然，也更符合一般人的思維模式。

本研究首先對於分類編目、詮釋資料、模糊理論等不同領域的研究進行探討；接著，提出模糊詮釋資料(fuzzy metadata)之概念，並將其應用於圖書館的分類與編目上，以便在相關圖書與類別之間建立輕重程度不同的關聯性；再者，我們發展一套圖書模糊分類與編目實驗系統，實際將模糊詮釋資料著錄至書目資料庫中，使用者透過此系統，可以從檢索出的模糊詮釋資料，掌握書籍間的相關程度，以更滿足實際的資訊需求。

二、相關研究

(一)分類編目

圖書分類編目是依主題或性質將圖書分門別類，編定書目資料，以使讀者可以迅速找到所需要的圖書。圖書館的圖書分類規則通常遵循既定的共通規範，國外以「美國國會分類法」(Library of Congress Classification : LCC)(註3)或「杜威十進分類法」(Dewey Decimal Classification : DDC)(註4)居多，國內則以「中國圖書分類法」(New Classification Scheme for Chinese Libraries : CCL)(註5)較普遍。而在編目方面，則有機讀編目格式 (Machine-Readable Cataloging : MARC)(註6)、國際標準書目著錄規則(International Standard Bibliographic Description : ISBD)(註7)、英美編目規則第二版(Anglo-American Cataloging Rules : AACR2)(註8)等相關標準。

圖書分類編目是一項思考性工作，必須經過一番審慎的過程與時間，所以為了簡化此編目過程，及節省更多的人力、時間與經費，發展出合作編目，並分享

館藏目錄以便進行自動化的抄錄編目之工作，也就是圖書館自動化過程中俗稱的「書目共用中心」(Bibliographic Utility：BU)，如國內的NBINET，國外的OCLC、RLIN等。這都是仰賴機讀編目格式(MARC)的標準化之賜。而圖書館自動化作業的成功關鍵，標準(standards)更佔有極重要的地位。早期，圖書館自動化的發展是由各國圖書館自行設計機器可讀的格式，一旦各館之間需要資料交換或分享時就會造成困難。因此，統一的機讀編目格式發展，是很重要的旅程，而MARC也成為國際的標準(ISO-2709)，如美國的USMARC、英國的UKMARC、台灣的CMARC等。其輸出格式都必須符合ISO-2709格式才能做轉換，且書目資料要以機讀編目格式處理，系統才能執行轉出或轉入，才能達到書目資料庫的交換動作(註9)。

(一) 詮釋資料

在網路當道、資訊超載的今日，人們需要面對的一個嚴肅課題是如何精確快速的找到所要的資料？於是，知識組織(Knowledge Organization，簡稱KO)逐漸發展成一個極重要的研究領域，著重詮釋資料(metadata)(註10)的著錄，藉由對原始資料進行詮釋、標示等加值處理，而使資料更具語意層次的自我描述性，進而能更精確的被檢索出來。為了讓詮釋資料發揮更大的功效，人們乃陸續制訂各種詮釋資料標準以供遵循。目前，已發展成熟或正發展中的詮釋資料標準眾多，適用於不同的領域及用途，例如：廣為圖書館界使用的MARC家族(USMARC、UKMARC、CMARC)；專為描述科技文獻的BibTex、EELS、EVL、RFC 1807等；專為描述人文社會科學資源的TEI Headers、ICPSR SGML Codebook Initiative等；專為描述政府資訊的GILS Core Elements；專為描述地理空間資源的CSDGM(Content Standard for Digital Geospatial Metadata)；專為描述博物館藏品與檔案特藏的CDWA(Categories for the Description of Works of Art)、CIMI(Computer Interchange of Museum Information)、EAD(Encoding Archival Description)等；以及專為描述大量網路資源的Dublin Core(Dublin Metadata Core Element Set)、IAFA/WHOIS++Templates、LDIF(Lightweight Directory Access Protocol)、SOIF(Summary Object Interchange Format)、URC(Uniform Resource Characteristics)等(註11)。

(二) 模糊理論

模糊理論(Fuzzy Theory)基本上是為了處理模糊概念所發展出來的。它實際上是模糊集合(Fuzzy Set)、模糊關係(Fuzzy Relation)、模糊邏輯(Fuzzy Logic)、模糊控制(Fuzzy Control)、模糊量測(Fuzzy Measure)…等理論之泛稱(註12)，是一門用以將模糊概念量化的學問，起源於1965年扎德(L.A. Zadeh)教授所發表的著名論文〈模糊集合〉(註13)。模糊理論以模糊集合為基礎，以研究不確定事物為目標，接受模糊現象存在的事實，根據不清晰訊息，透過近似推理

(Approximation Reasoning)(註14)過程而得到正確結果。這與人腦「過程模糊，結論清晰」的思維方式極其類似，因此已被廣泛的應用於各種不同領域的智慧型系統中(註15)。

傳統明確集合(Crisp Set)的特徵函數(Characteristic Function)採用非0即1的二分法，而模糊集合的基本精神則是將其擴展成由0至1的任何值，稱為歸屬函數(Membership Function)。當一個元素屬於某集合的程度越大時，其歸屬程度就越接近於1，否則越接近於0。若以X代表論域(Universe of Discourse)，A代表X中的任一離散型模糊集合，則可將A表示成如下型式：

$$A = \sum_{i=1}^n \mu_i / x_i = \mu_1 / x_1 + \mu_2 / x_2 + \dots + \mu_n / x_n \quad (1)$$

其中， $x_i \in X$ 為X中任一點， $\mu_i \in [0, 1]$ 為歸屬函數值，代表 x_i 歸屬於模糊集合A的程度。藉由歸屬函數對模糊概念進行量化之後，便可以利用精確的數學方法進行模糊資訊的分析和處理了，而電腦也因此而稍具人腦思維的雛形。

三、模糊詮釋資料

在解釋模糊詮釋資料的意涵之前，我們先對模糊詮釋資料下一個簡單的定義：「所謂模糊詮釋資料，乃是指描述事物之模糊性質或彼此間之模糊關係的資料而言」。

綜觀目前大多數詮釋資料的發展而言，其目的大都在於「明確的」描述事物，對於各項資料之著錄皆有一定的規範以供遵循，極少具有模糊地帶，或涉及「模糊化」資訊者。然而，在真實世界中，模糊概念到處存在，許多事物的性質是模糊的，關係是模糊的，甚至人類有許多語言和思維也都是模糊的。例如：這本書和那本書「很類似」；他開車「有點快」；那所圖書館「很雄偉」；這張照片拍得「相當漂亮」……等；其中，很類似、有點快、很雄偉、相當漂亮……等，都是模糊語意詞(Fuzzy Linguistic Terms)，人腦都可以很容易的處理並理解，此種模糊現象顯然與傳統詮釋資料有著極大的差異。因此，若要讓電腦可以處理這些模糊概念，便必須藉助模糊理論的原理，將這些模糊語意轉換成電腦能理解的模糊詮釋資料，以便使得所描述之事物，能更廣泛的涵蓋無法明確描述的模糊資訊。

茲以圖書館編目員進行分類編目為例，現有《圖書館學的哲學》一書，其內容至少涵蓋了圖書館學(總類)與哲學(哲學類)兩大領域，若以中國圖書分類法進行分類，此書至少可歸屬於「總類」與「哲學類」兩類別。若將其歸屬於「總類」，則強制切斷此書與「哲學類」之關聯；反之，若將其歸屬於「哲學類」，卻又硬生生的使其脫離「總類」的範疇。但圖書館編目員基於分類規則，魚與熊掌無法兼得，只能在兩者之間取其主要者而捨其次要者，給予一個分類號；即使利

用805段可以彈性給予多重類號，也無法標示此書究竟屬於總類多些或屬於哲學類多些？多幾分？反之，如果採用模糊分類方式，便可將此書歸屬於「總類」和「哲學類」的程度以不同的歸屬函數值加以標示，例如，前者給予0.6的歸屬度，而後者給予0.3的歸屬度。若以模糊集合表示式則可寫成：

$$\mu(\text{圖書館學的哲學}) = 0.6/\text{總類} + 0.3/\text{哲學類}$$

接著，只要能將這份模糊詮釋資料著錄到書目資料庫中，便可以此書與總類和哲學類建立起輕重程度不一的關聯度，同時也提升了詮釋資料的詮釋能力與範圍。

總而言之，模糊詮釋資料使得自然直覺與電腦世界之間搭起了一座交流的橋樑，更符合人類軟性思維的本質。

四、圖書模糊分類與編目

(一) 明確分類模式

傳統圖書館的分類模式為明確分類(Crisp Classification)，屬於硬性分類法。主要作法是將某一冊圖書歸屬於特定類別，並給予分類號。書籍與類別的關係為明確關係(Crisp Relation)，只有「屬於」和「不屬於」兩種可能。

假設某一圖書館館藏中共有 n 冊圖書，其中第 i 冊記為 d_i ，並以 D 表示所有圖書的集合，即 $d_i \in D, i=1, 2, \dots, n$ ；現有 m 個圖書類別，第 j 類記為 c_j ，並以 C 表示所有類別的集合，亦即 $c_j \in C, j=1, 2, \dots, m$ 。對於某冊圖書 d_i 與某個類別 c_j 的歸屬關係，我們可以用明確集合表示法表示如下：

$$\sigma_{ij} = \begin{cases} 0, & d_i \notin c_j \\ 1, & d_i \in c_j \end{cases} \quad (2)$$

其中， σ_{ij} 為特徵函數，若 d_i 屬於 c_j ，則 $\sigma_{ij}=1$ ，否則 $\sigma_{ij}=0$ 。由於 σ_{ij} 的值非0即1，所以又稱為二值分類法，簡稱二分法。不論圖書館採用何種分類法(杜威十進分類法，美國國會分類法，或中國圖書分類法)，皆屬此種分類法，適用此一數學模式。

(二) 模糊分類模式

如果圖書內容之主題涵蓋了好幾個學科領域，且有輕重程度之別，則明確分類模式便不適用。此時，屬於軟性分類模式的模糊分類法便可派上用場。若將 d_i 視為一個定義在論域 C 上的離散型模糊集合，且將 d_i 歸屬於某個類別 c_j 的程度以 μ_{ij} (介於0~1)表示，則此一模糊集合可記為：

$$d_i = \sum_{j=1}^m \mu_{ij} / c_j \quad (3)$$

亦即：

$$d_i = \mu_{i1}/c_1 + \mu_{i2}/c_2 + \dots + \mu_{im}/c_m$$



$$d_2 = \mu_{21}/c_1 + \mu_{22}/c_2 + \dots + \mu_{2m}/c_m$$

:

$$d_n = \mu_{n1}/c_1 + \mu_{n2}/c_2 + \dots + \mu_{nm}/c_m$$

舉例而言， $d_1 = \mu_{11}/c_1 + \mu_{12}/c_2 + \dots + \mu_{1m}/c_m$ 代表 d_1 此本圖書在 c_1 這一類別裡的歸屬值是 μ_{11} ，在 c_2 這一類別裡的歸屬值是 μ_{12} ，在 c_m 這一類別裡的歸屬值是 μ_{1m} 。若將其歸屬函數畫成離散式模糊集合，則如圖1所示。圖中顯示出 d_1 圖書分別在 $c_1、c_2 \dots c_m$ 類別中所佔的歸屬函數值高低，愈接近於1，表示屬於某類別的程度愈高，反之，愈接近於0，表示屬於某類別的程度愈低。

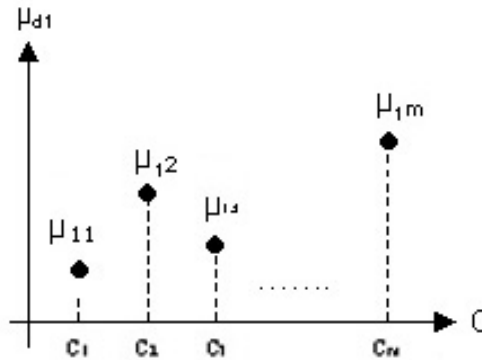


圖1 將圖書以離散型模糊集合表示

再者，我們可將各圖書($d_1 \sim d_n$)在各個類別($c_1 \sim c_m$)中分屬的歸屬函數值($\mu_{11} \sim \mu_{nm}$)，以一個 $n \times m$ 的模糊關係矩陣表達如下：

	c_1	c_2	\dots	c_m
d_1	μ_{11}	μ_{12}	\dots	μ_{1m}
:	:			
:	:			
:	:			
d_n	μ_{n1}	μ_{n2}	\dots	μ_{nm}

圖2 圖書與類別之模糊關係矩陣

由此可知， d_i 與 C_j 之間具備一種「模糊歸屬關係」，從語意的觀點而言可描述成「 d_i 歸屬於 C_j 的程度為 μ_{ij} 」。因此，上述的模糊關係矩陣實際上等同於下圖所示的模糊語意網路。我們可以很清楚的看出，模糊歸屬函數值在此網路中，扮演了一個將相關概念以不同的強度加以連結的重要角色。

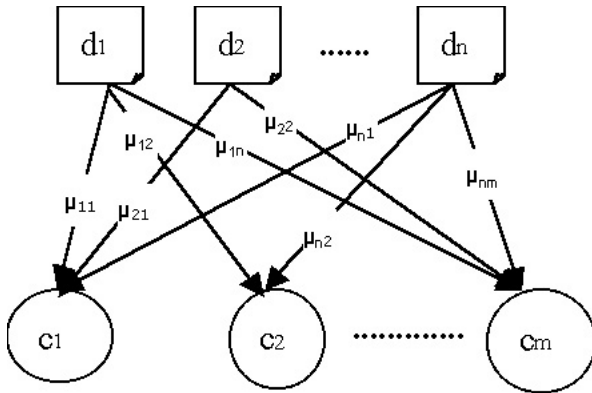


圖3 圖書與類別之模糊語意網路

㊦ 模糊編目

經由上述模糊分類後所得到的模糊詮釋資料，若要註錄至圖書館的書目資料庫中，必須先探討MARC格式中是否有適合註錄之處？其次，再研究註錄格式為何？

我們以中國機讀編目格式第四版(CMARC4)簡篇作為研究對象，先審視其註錄格式，最後發現在CMARC4簡篇「8XX各館使用段」中的「805館藏紀錄」部分，其分欄 \$d\$ 可重複註錄分類號，而分欄 \$n\$ 則作為註記之用，可彈性使用。此兩者皆可重複，詳如表1所示。因此，我們認為805段非常適合作為註錄本研究所述的模糊詮釋資料之用。註錄方式是以分欄 \$d\$ 標示分類號，而以分欄 \$n\$ 標示該類號所對應的歸屬函數值；若某一本書所歸屬的類別不只一個，則可重複著錄不同的類號與歸屬函數值。

表1 [805館藏紀錄]分欄說明

805 館藏紀錄			
分欄識別	分欄內容	說明	備註
\$d	分類號	若為複本書而總館與分館之分類不同時，本欄可重複填寫	可重複
\$n	註記	如出版品形式或其他等	可重複

再舉《圖書館學的哲學》一書為例，假設此書的館藏單位是淡江資圖系實驗圖書館(簡稱資圖)，使用的分類法是中國圖書分類法CCL，並且在分類時認定：此書屬於0類(總類)的程度為0.6，分類號為020.11；屬於1類(哲學類)的程度為0.3，分類號為170。據此，我們便可註錄如表2所示的機讀編目格式，其中，805段是為了凸顯書籍的模糊性質而加入的。

表2 《圖書館學的哲學》 機讀格式簡篇

欄位	指標		分欄 / 內容	說明
	1	2		
010	0		\$a 9578708033 \$b 平裝 \$d 新臺幣 400 元	ISBN
100			\$a 20030804 0chi09 e	一般性資料
101	0		\$a chi	作品語文
200	1		\$a 圖書館學的哲學 \$f 賴鼎銘著	題名及著者
205			\$a 初版	版本項
210			\$a 臺北市 \$c: 文華 \$d, 民 82	出版項
215	0		\$a [11], 295 面 \$c: 圖 \$d; 21 公分	稽核項
300			\$a 含索引	一般註
327			\$a 參考書目: 面 241-280	內容註
606			\$a 圖書館學 \$x -哲學, 原理 \$2 csh	主題標題
681			\$a 020.11 \$b 8366 \$v 增訂七版	中國圖書分類法
700	1		\$a 賴 \$b 鼎銘	人名_主要著者
801	0		\$a tw \$b 資圖 \$c 20030804 \$g CCR \$m4	出處欄
805			\$a 資圖 \$d 020.11 \$n 0.6 \$d 170 \$n 0.3 \$t CCL	館藏紀錄

從表2，我們可以很清楚看到，805段加入了「\$a 資圖 \$d 020.11 \$n 0.6 \$170 \$0.3 \$t CCL」資料，其中的模糊詮釋資料「\$d 020.11 \$n 0.6」和「\$d 170 \$n 0.3」所描述的模糊特性即是：

$$\mu(\text{圖書館學的哲學}) = 0.6/020.11(\text{總類}) + 0.3/170(\text{哲學類})$$

此一作法由於採用標準機讀編目格式來註錄模糊詮釋資料，因此，圖書館自動化系統並不需大幅改寫，只需在解讀805段的分欄 \$d和分欄 \$n後，適當的予以加值應用，便可以輕易的具有支援模糊詮釋資料之能力了。

五、實例探討

為了更進一步說明模糊詮釋資料在圖書分類與編目之應用，我們以系統實作方式建構實驗系統加以驗證。

(一) 實驗資料之界定

我們以中國圖書分類法0~9類為基準，從淡江大學圖書館中，每一類找出兩本圖書，共取得20本圖書及其書目資料作為實驗用途。這些書的內容大都不只涵蓋單一領域，如《圖書館學的哲學》、《新聞心理學概論》、《土地倫理》……等，但由於傳統上，圖書館仍採用明確的分類與編目規則，所以每本書都僅歸屬於一個單一類別，給予一個單一類號，我們將原始的分類在圖4左側A部分標示出來。此部分顯示了圖書與類別之間的明確歸屬關係。

接著，我們經由檢閱這些書籍的內容，重新歸納其涵蓋的主題範圍；再依據前述的模糊理論概念，將這些書籍進行模糊分類，然後彈性的給予每本書籍多重類號及適當的模糊歸屬函數，圖4中央B部分說明了模糊分類後圖書與類別之間的模糊歸屬關係。如此一來，模糊分類不但直接的在圖書與類別之間，建立了輕重

程度不一的模糊關係，更間接的使得圖書與圖書彼此之間也建構起了模糊關係，如圖4右側C部分所示。

這些由模糊關係所交織出來的模糊語意網路並非無中生有，而是原本就存在於這些書籍之間，只不過因為傳統上採用明確分類法，使得編目員在分類時，不得不取主要類別而捨次要類別，以致強制切斷了書籍之間天生的關聯性，模糊分類僅是重建這些原已存在的關聯性罷了。書與書之間原本就不應被視為獨立了個體，藉由加入模糊詮釋資料，使得圖書的主題就像地圖一樣能互通相連，讀者便能經由這個主題類號找到特定的圖書，也能在另一個類號裡找到相同之書籍，如《圖書館學的哲學》，雖其正規明確的類號屬於「0類：總類」，但模糊分類除可歸屬於「0類：總類」外，尚可歸屬於「1類：哲學類」，所以當讀者查找《圖書館學的哲學》這本書時，除了可在「0類：總類」找到之外，也可以在「1類：哲學類」中找到，而且還可以藉由書籍間的模糊語意網路查找到相關書籍，進行延伸閱讀。例如，由圖中可知，借閱《圖書館學的哲學》的讀者有可能對哲學有興趣

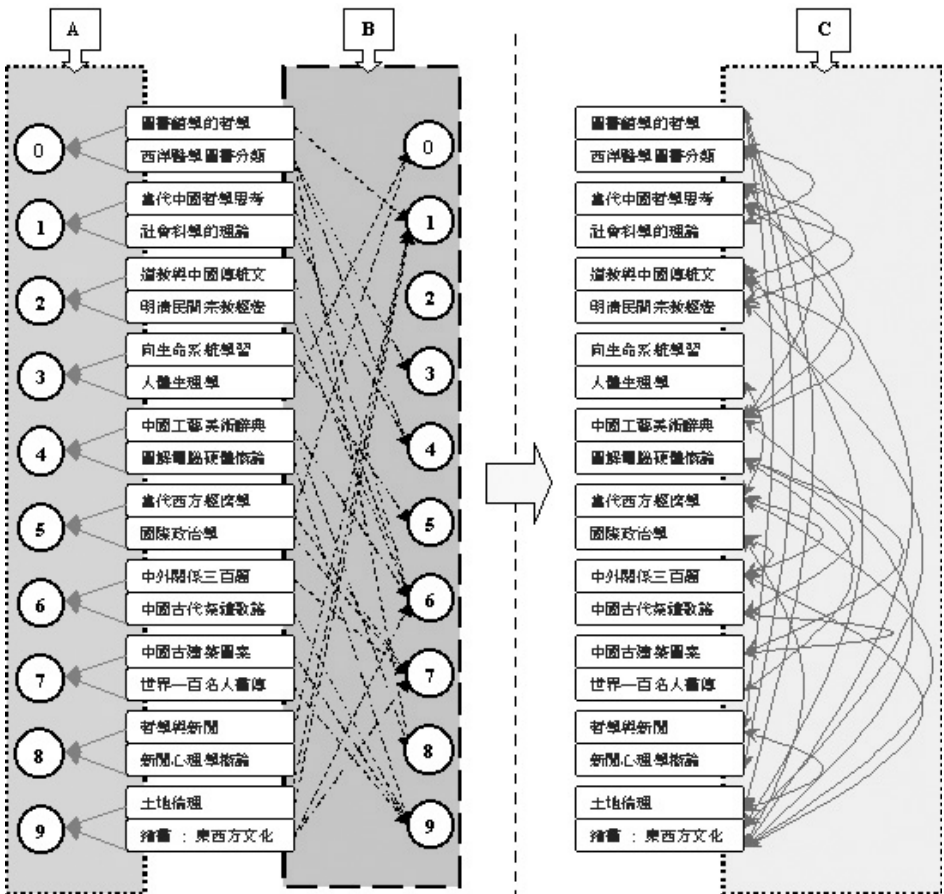


圖4 明確分類與模糊分類關係圖

，因此，有可能成為其延伸閱讀的相關書籍有《哲學與新聞》、《新聞心理學概論》等；也有可能對圖書館建築有興趣，則《土地倫理》等書則較相關。此外，各書籍之關係度更可從歸屬函數判斷其輕重，數值越接近於1越相似，反之，越接近於0則越不相似。因此，假使一本書籍其內容涵蓋多個領域，其所屬類別也將分屬多重類別，同類型的相關書籍也將越廣。

(二) 實驗系統之實作

接著，我們著手進行實驗型圖書模糊編目與檢索系統之設計。此系統藉由模糊分類與加註模糊詮釋資料，進而可進行模糊查詢，透過此機制實際將模糊理論概念加以應用於圖書分類與編目上，可以將某書籍「同類」與「模糊類」相關書籍查找出來，藉以提高檢索結果的廣度。

關聯式資料庫是目前最通用的資料庫系統，因此，在本研究中採用關聯式資料庫系統。依據前述所分析的模糊分類模式，可以很容易的將模糊關係矩陣轉換為關聯式資料表的模糊分類欄位，如表3所示，並給予歸屬函數值，使得模糊分類檢索模組可藉以判別圖書所屬之模糊類別。

表3 模糊關聯式資料表

	c_1	c_2	...	c_m
d_1	μ_{11}	μ_{12}	...	μ_{1m}
d_2	μ_{21}	μ_{22}	...	μ_{2m}
:	:	:	...	
d_n	μ_{n1}	μ_{n2}	...	μ_{nm}

本系統架構於Microsoft Windows 2000 Advance Server(NT技術平臺)之上；

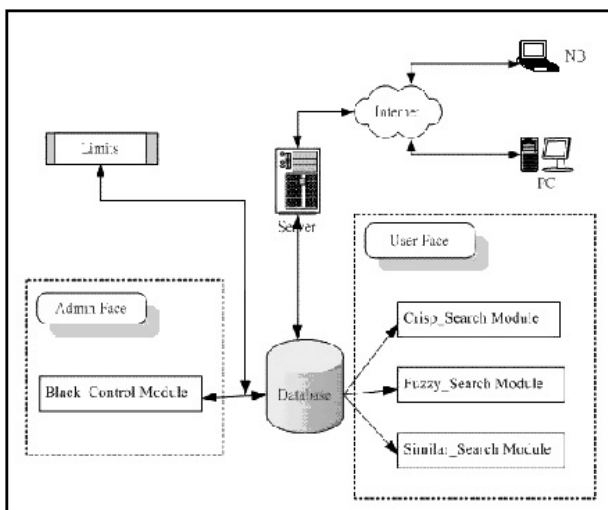


圖5 系統架構圖

伺服器管理介面採用內建Microsoft IIS5.0；程式部分採用ASP(Active Server Pages)程式語言開發；後端資料庫採用Microsoft Access 2000；網頁呈現採用Microsoft Front Page 2000；前端使用者可透過Microsoft Internet Explorer 5.0或Netscape 6.0以上版本等瀏覽器使用。本系統主要分為A.前端使用模組：提供讀者透過個人需求選擇適當檢索方式進行書目檢索，依功能分為：1.明確分類檢索模組(Crisp Search Module)，2.模糊分類檢索模組(Fuzzy Search Module)，3.相似書籍檢索模組(Similar Search Module)；B.後端管理模組：提供系統管理者維護書目資料之新增、刪除、修改以及模糊分類度數選擇之動作，進入該模組前必須鍵入特定帳號與密碼提供系統驗證，才能進入進行管理動作。系統架構如圖5所示。

各模組之功能及實驗結果進一步說明如下：

1. 明確分類檢索模組

本模組提供使用者鍵入關鍵字並選擇以書名、作者、出版社、出版地、出版年、館藏地、ISBN 進行比對查詢。檢索結果如圖6所示。

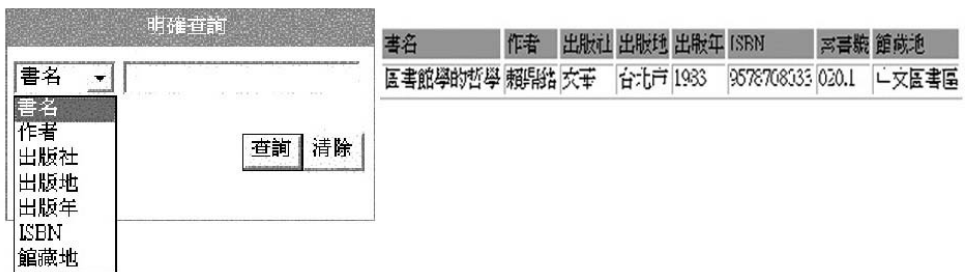


圖6 明確分類檢索模組與結果

2. 模糊分類檢索模組

本模組提供使用者選擇圖書模糊分類歸屬值，搭配圖書類別進行查詢，如圖7所示。類別部分提供使用者以下拉式方式選擇中國圖書分類法0~9類，模糊度部份也採用下拉表單方式供選擇其模糊歸屬函數值，以決定檢索資料時的關聯度，查詢結果之模糊度數越接近1之書籍，表示越接近所選擇之圖書類別。

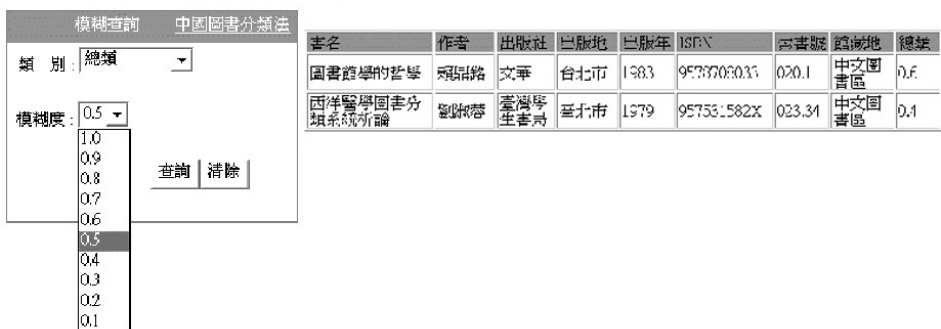


圖7 模糊分類檢索模組與結果

使用者選定類號與模糊度後，系統會依內部書目資料庫中的模糊詮釋資料進行比對，如核對成功系統會將書目資料依照所選擇類別之圖書模糊度數，依度數高至低排序出來。如圖8所示是使用者選擇類號為「總類」、模糊度為「0.1」之檢索結果，所有模糊度數高於0.1之總類圖書全部會呈列出來，越接近於1的某筆資料則會排序在最上面，以便使用者快速瀏覽。從圖7也可得知，《圖書館學的哲學》、《西洋醫學圖書分類系統析論》這二本書依內容來分類都含有「0類：總類」之類別，表示使用者可以透過此模糊檢索之功能檢索出同類型之書籍，由模糊度數來判斷《圖書館學的哲學》模糊度數0.6，比《西洋醫學圖書分類系統析論》0.4還接近於1，因此屬於總類的程度較高，但兩者之間因度數相近，也代表兩者屬於同一類之書籍，如此一來，使用者更能掌握所檢索圖書的關聯性，進而提升檢索效率。

書名	作者	出版社	出版地	出版年	ISBN	索書號	館藏地	總類
圖書館學的哲學	賴鼎銘	文華	台北市	1983	9578708033	020.1	中文圖書區	0.6
西洋醫學圖書分類系統析論	劉淑蓉	臺灣學生書局	臺北市	1979	957531582X	023.34	中文圖書區	0.4

圖8 模糊分類檢索「0類：總類」之結果

圖9所示則為檢索類別選「哲學類」，模糊度數選「0.3」之結果，共有《人體生理學》、《社會科學的理路》、《當代中國的哲學思考》、《哲學與新聞》、《新聞心理學概論》、《圖書館學的哲學》以及《土地倫理》七本書，其中可以很清楚的看到《圖書館學的哲學》一書分屬總類與哲學類兩個領域的程度分別為0.6和0.3。

書名	作者	出版社	出版地	出版年	ISBN	索書號	館藏地	總類	哲學
人體生理學	黃基礎	藝軒圖書出版社	臺北市	2002	9576165156	397.8347	中文圖書區	0	0.9
社會科學的理路	黃光國	心理出版社公司	臺北市	2001	9577024203	143.8365	中文圖書區	0	0.8
當代中國的哲學思考	肖明	經濟科學出版社	北京	1996	7505809598	120.855	中文圖書區	0	0.7
哲學與新聞	徐光春	新華書店	北京	1991	7200015814	890.1.8564	中文圖書區	0	0.7
新聞心理學概論	劉京林	北京廣播學院出版社	北京	1993	7810040073	890.14.8727	中文圖書區	0	0.7
圖書館學的哲學	賴鼎銘	文華	台北市	1983	9578708033	020.1	中文圖書區	0.6	0.3
土地倫理	張元茜	富邦藝術基金會	臺北市	1999	9579775303	920	中文圖書區	0	0.3

圖9 模糊分類檢索「1類：哲學類」之結果

此外，為了使對圖書分類法架構不清楚的使用者，能順利掌握圖書分類號，使檢索時所下的檢索策略更精確，我們設計了一個中國圖書分類法明細項目，供使用者參考使用，如圖10所示。

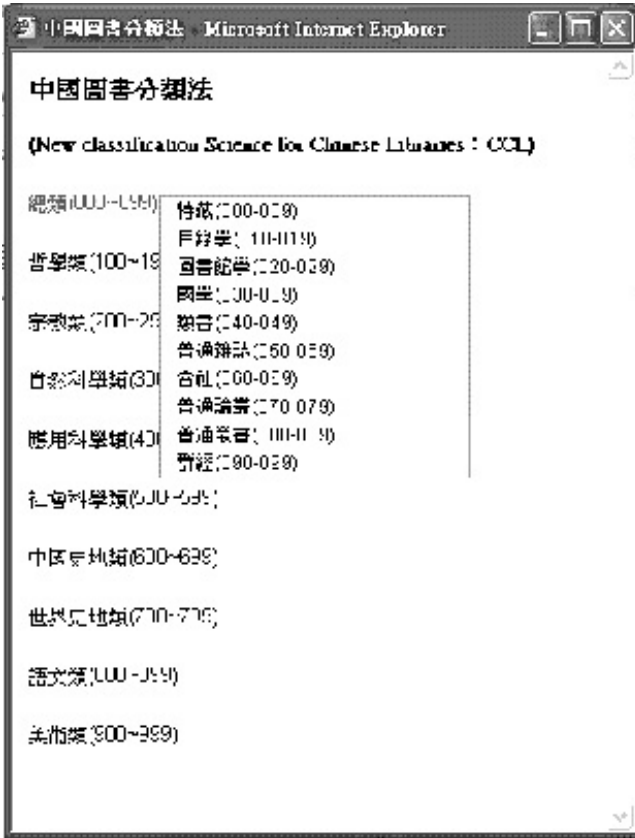


圖10 中國圖書分類法詳細類表

3.相似書籍檢索模組

本模組提供使用者選擇圖書類別，系統會將資料庫裡的相似書目資料，依照此類別的模糊度數高低全部呈列出來。



圖11 相似書籍檢索模組與結果

4.後端管理模組(Black—Control Module)

此模組進入前必須登入特定帳號與密碼，提供管理者維護書目資料新增、刪除、修改以及模糊分類度數選擇之異動等動作。

編目模組		模糊度數	
■ 登入號	<input type="text"/>	■ 總類類_C0	<input type="text" value="0.0"/>
■ 書名	<input type="text"/>	■ 哲學類_C1	<input type="text" value="0.0"/>
■ 作者	<input type="text"/>	■ 宗教類_C2	<input type="text" value="0.0"/>
■ 出版社	<input type="text"/>	■ 自然科學_C3	<input type="text" value="0.0"/>
■ 出版地	<input type="text" value="台北市"/>	■ 應用科學_C4	<input type="text" value="0.0"/>
■ 出版年	<input type="text"/>	■ 社會科學_C5	<input type="text" value="0.0"/>
■ ISBN	<input type="text"/>	■ 中國史地_C6	<input type="text" value="0.0"/>
■ 館藏地	<input type="text" value="中文圖書區"/>	■ 世界史地_C7	<input type="text" value="0.0"/>
■ 索書號	<input type="text"/>	■ 語文類_C8	<input type="text" value="0.0"/>
		■ 藝術類_C9	<input type="text" value="0.0"/>
		<input type="button" value="新增"/> <input type="button" value="清除"/>	

圖12 後端管理模組

六、結 論

知識浩瀚無涯，圖書館界無論使用哪一種分類系統，都無法百分之百的記載著文化資產與社會變遷的圖書資料，設計出完美的組織架構。再加上時代科技的進步，人類知識日新月異，圖書資料內容日漸增加、複雜，用來組織我們人類知識的分類系統，必須設法隨著知識創新快速的節奏跟上腳步。

本研究提出模糊詮釋資料的概念與作法，藉由圖書模糊分類與編目的過程，重建知識間的關聯性，使其更加環環相扣，解決傳統分類系統的二分法問題，進而使得一般使用者可以利用分類檢索系統搜尋相關資料，使查檢資料更直接、更有效率。而從編目館員的角度來看，也較不會因個人主觀因素造成取捨困難或認知誤差，使讀者查詢時找不到相關書籍，同時也可以降低參考館員回覆諮詢者問題的需求量。此外，依據本研究提出的方法，不但能為書籍作跨領域的分類，更可以將模糊詮釋資料以機讀編目格式註錄於書目資料庫中，解決模糊分類可能為圖書館自動化系統的書目共用帶來的困擾，更符合了ISO2709標準。我們期望本研究可作為未來進一步發展線上檢索系統的改進參考。

為了使圖書模糊分類與編目更具資訊加值應用之價值，我們未來研究方向將著手於「讀者模糊分類」之研究，以便合成讀者與圖書之模糊關係，提供圖書館界服務項目中最具特色的「專題選粹服務」(Selective Dissemination of Information, SDI)(註16)，使此系統更具彈性化、個人化、人性化及智慧化。

註 釋

註 1 陳和琴、吳瑤瑛、江綉瑛，圖書分類與編目(臺北縣，空中大學，民國85年)。

註 2 賴鼎銘，圖書館學的哲學(臺北市，文華，民國82年)。

註 3 Library of Congress. Cataloging Policy and Support Office, *Library of Congress Classification. H. Social Sciences* (Washington, D.C. : Library of Congress, Cataloging Distribution Service, 1994).

- 註 4 Mary C. Bauer, *Dewey Decimal Classification : 200 schedules expanded for use* (Catholic Library Association, 1988).
- 註 5 同註2。
- 註 6 MARC Standards, <<http://lcweb.loc.gov/marc/>>(29 July 2003)。
- 註 7 國際圖書館協會聯盟(International Federation of Library Associations and Institutions : IFLA), <<http://www.ifla.org/>>(29 July 2003)。
- 註 8 盧秀菊, 〈英美編目規則原則之探討〉, 圖書與資訊學刊, 32期(民國89年)。
- 註 9 黃世雄、黃鴻珠、鄭麗敏, 圖書館網路與資源運用(臺北:空大, 民國87年)。
- 註10 詮釋資料(metadata)是「用來描述資料的資料」(Data describes other data)或「關於資料的資料」(Data about data), 其譯名有「詮釋資料」、「後設資料」、「元資料」、「描述資料」、「超資料」...等, 並不統一。本文採用「詮釋資料」一詞。
- 註11 陳雪華, 〈網路資源與Metadata之發展〉, 圖書館學刊, 12期(民國86年):頁23-28。
- 註12 林信成、彭啟峰, *Oh! Fuzzy* 模糊理論剖析(臺北:第三波, 民國83年)。
- 註13 L. A. Zadeh, "Fuzzy sets," *Information and Control*, 8 (1965) : 338-353.
- 註14 L. A. Zadeh, "Outline of a new approach to the analysis of complex systems and decision processes," *IEEE Trans. on Syst., Man and Cybern. SMC*, 3 (1973) : 28-44.
- 註15 Jyh-Shing Roger Jang, Chuen-Tsai Sun, Eiji Mizutani, *Neuro-Fuzzy and Soft Computing : A computational approach to learning and machine intelligence* (Prentice Hall, 1996).
- 註16 SDI乃是針對讀者個別所需, 將最新的資訊選出, 以主動積極的方式, 定期提供給讀者, 以節省讀者檢索資料的時間, 並協助讀者獲取更多相關資料, 詳細內容可參考:王錫璋, 圖書館的參考服務: 理論與實務(台北, 文史哲, 民國86年)。

Fuzzy Metadata and Its Application to Library Classification and Catalogue

Sinn-Cheng Lin

Associate Professor

E-Mail : sclin@mail.tku.edu.tw

Sheng-Wen Hsiao

Graduate Student

Department of Information & Library Science, Tamkang University

Taipei, Taiwan, R.O.C.

Abstract

This study proposes the concept of Fuzzy Metadata, and then applies it to library classification and catalogue for the purpose of reinforcing the connection of bibliography, such that the broadness of search results could be improved. In the first, the paper discusses the classification, catalogue, metadata and fuzzy theory. And then, it expounds the definition of Fuzzy Metadata and uses the model of fuzzy theory to explain its meaning. Furthermore, we propose a practical approach of fuzzy classifying and cataloging for the library to record the Fuzzy Metadata to its bibliography databases in CMARC format. Finally, we develop a classifying and cataloging system; it uses Fuzzy Metadata to build up distinct levels of connection in related books and categories. The users who use the system could grasp the relationship among the books from the search results, so that they could further search related ones to more satisfy their information needs.

Keywords : Metadata; Fuzzy theory; Fuzzy metadata; Fuzzy classification; Fuzzy catalogue

