

圖書分類的 自動化에 관한 研究 [2]

—圖書館學 및 情報學分野 書誌分類를 中心으로—

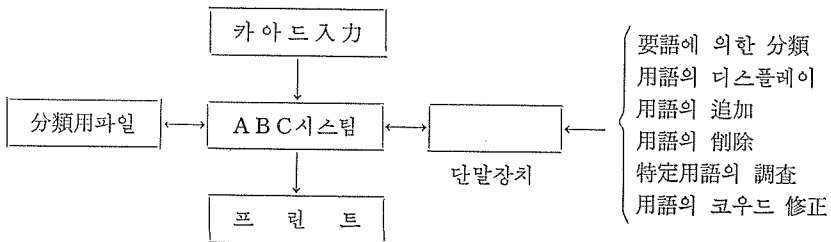
李 慶 浩
韓社實業專門大·專講

三. 分類的 自動化시스템 模型

1. 시스템의 概要

分類的 自動化시스템 (Automatic Book Classification System. 이하 ABCS라 稱한다) 은 우선 分析한 用語를 基礎로 하여 分類用파일을 構成한다. 따라서 分類는 文獻의 標題上에 나타난 要語를 分類用파일에서 探索함으로써 이루어 진다. 또한 이 ABCS 는 단말장치를 통하여 要語의 入力에 의한 圖書分類 뿐만 아니라, 蓄積된 用語의 디스플레이(display), 새로운 用語의 追加(addition), 不必要한 用語의 削除(deletion), 잘못 入力한 코우드에 대한 修正(correction) 및 特定要語의 調査(inquiry) 등 多様な 機能을 遂行한다.

이러한 機能을 가진 分類의 自動化시 스템인 ABCS의 概要를 그림으로 그려보면 <圖 1>과 같다.



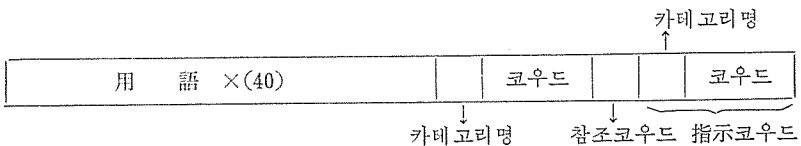
<圖 1> ABC시스템의 概要

2. 파일의 設計

分類用파일을 만들기 위하여 入力되는 入力要素의 構成은 다음과 같다.

- ① 用語.
- ② 카테고리名.
- ③ 分析된 用語의 코우드.
- ④ 參照코우드(A, X, R, S, Y, H)
- ⑤ 指示코우드

이러한 入力要素들은 하나의 레코오드를 構成하며 이들 레코오드가 모여 하나의 分類用파일을 이루게 된다. 入力되는 레코오드의 構成은 다음과 같다.



이러한 레코오드들이 컴퓨터의 디스크상에 蓄積되어 分類時에 要語를 단말장치로 入力하면 온라인 方式으로 分類가 된다. 이 파일은 컴퓨터의 蓄積媒體上에 알파벳 順으로 配列되어 있으며, 새로운 用語의 追加나 不必要한 用語의 削除 및 不正確한 데이터의 修正時마다 그때 그때 파일의 更新(updating)이 이루어진다.

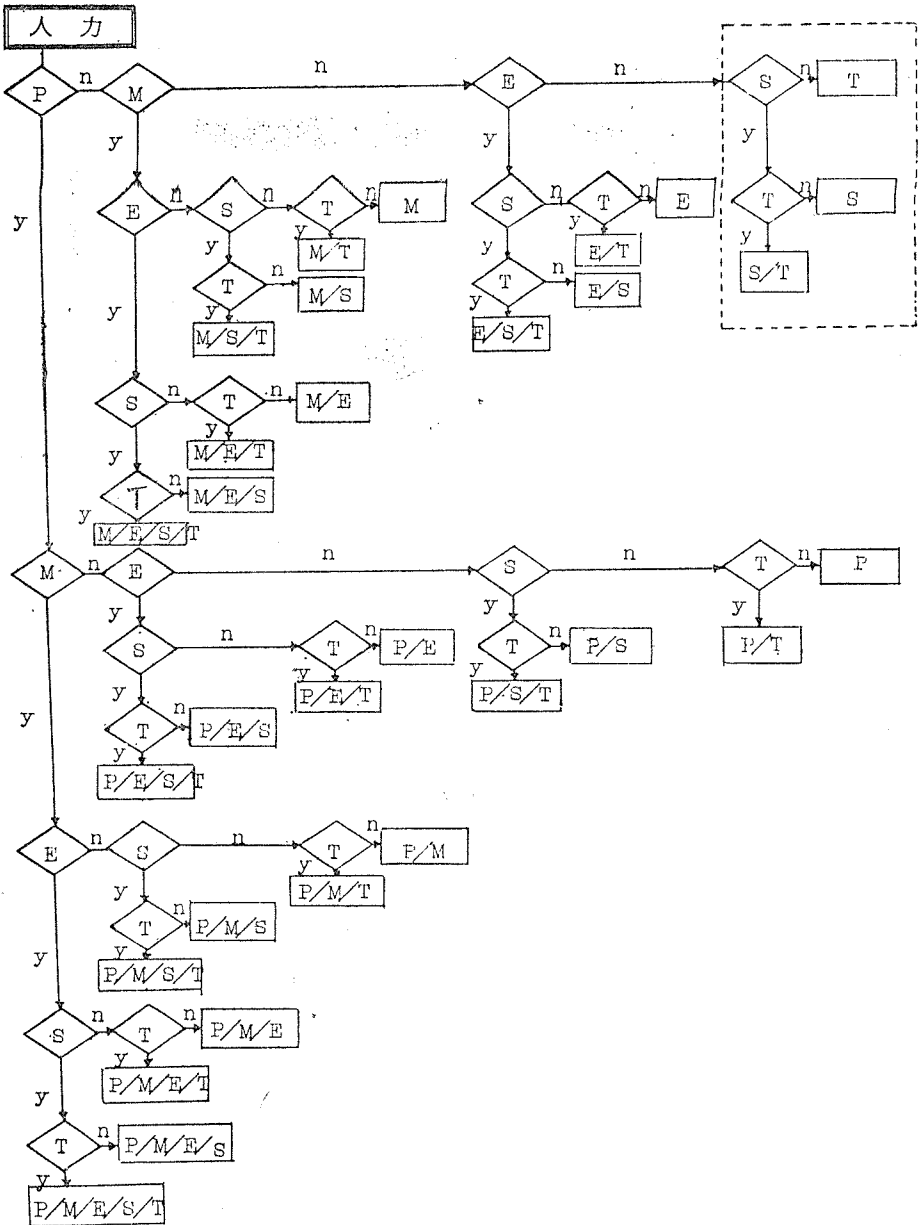
3. 分類의 自動化 原理

한 圖書에 대한 分類는 그 文獻에 대한 主題分析의 過程을 거쳐 그 文獻을 代表하는 用語를 抽出하여 이 用語에 適切한 記號를 부여함으로써 이루어지며, 이 때 主題分析은 주로 文獻의 標題나 內容 및 抄錄 등을 分析對象으로 삼고 있다.

그러나 本 研究에서는 文獻을 代表하는 要語를 標題에서 抽出하여 分類하도록 하였다. 이것은 무엇보다도 文獻의 內容이 標題上에 가장 最小한 用語로 가장 잘 묘사되어지리라고 믿기 때문이었다. 특히, KWIC(Keyword-in-Context)가 紹介된 이래로 著者들은 그 전의 어느 때 보다도 더욱 記述的이고 意味 있는 標題를 쓰려고 한다는 점¹⁶⁾은 이 研究의 信憑性을 뒷바침하여 주고 있다 하겠다.

그러나 分類를 하는데 있어 文獻의 內容 가운데에서 要語를 抽出하여 分類한다고 하여 물론 分類가 되지 않는 것은 아니다. 이것은 어디까지나 入力の 要語를 각 圖書館마다. 統一함으로써 同一한 文獻 내에서는 同一한 分類番號를 維持시켜 보자는 데 있을 뿐이며, 文獻의 標題가 그 文獻의 內容을 確實히 要語가 없을 경우에는 內容分析에 의한 分類도 考慮할 수 있다 하겠다.

16) Kowk, op.cit., p.202.



〈圖 2〉 分類의 自動化 原理

本 研究에서 採擇한 標題에 의한 分類方法은 文獻의 標題 가운데에 나타난 要語를 入力함으로써 이루어지며, 이들 要語들은 [P][M][E][S][T] 가운데의 어느 한 카테고리상에는 반드시 속하게 되어 있다. 여기서 入力の 要語가 각 카테고리마다 하나씩 있고, 또 各種 參照코우드의 影響을 받지 않는 狀態의 一般的 分類의 自動化原理는 <圖2>와 같다.

圖2에서 보는 바와같은 組合의 경우의 數는 P/M/E/S/T, P/M/E/S, P/M/E/T, P/M/E……S/T, S,T에 이르기까지 모두 31가지¹⁷⁾ 方法이 있으나, 각 카테고리마다 한 개 이상 入力되는 경우가 있으므로 그 組合方法의 數가 더욱 늘어나게 된다.

예를들어 위의 P/M/E의 하나의 경우에 있어서만도 組合에 대한 경우의 數는 [P] 112個, [M]74個, [E]146個이므로 $112 \times 74 \times 146 = 1,210,048$ 個의 組合方法이 可能하나, 카테고리의 重複이 있을 때에는 P/P/M/E, P/M/M/E, P/P/E/E, P/P/M…… 등과 같은 組合의 方法도 나올 수 있어 그 數는 무수히 늘어나게 된다.

그러나 여기에서 [S]와 [T]는 사실상 그 自體만으로는 圖書館學 및 情報學 分野 내에서의 基本카테고리가 아니기 때문에 入力の 要語가 [S]나 [T] 혹은 [S][T]만으로 主題가 構成되는 경우는 거의 없는 것으로 생각된다. 그러나 만약 [S]나 [T]의 組合만으로 되어 있는 圖書가 圖書館學의 主題를 가진 圖書라면 이러한 경우에도 分類가 可能하여야 한다. 따라서 이러한 경우에도 分類가 되도록 例外規定(圖2의 接線안 部分)을 두었다.

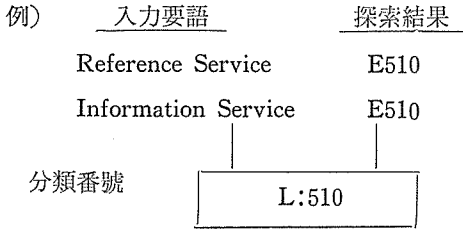
이상과 같은 分類의 自動化原理에 의하여 組合된 結果는, 각 要語의 코우드 앞에 連結符號를 부여하여 分類番號가 決定되어진다. 즉 [T]에는 ' (apostrophe), [S]에는 . (dot), [E]에는 : (colon), [M]에는 ; (semicolon), 그리고 [P]에는 , (comma)를 個個의 코우드 앞에 부여하되 [P]가 제일 처음 나타날 때는 [,]를 省略하였다. 그리고 또 앞으로 모든 主題分野를 分類할 경우에 主題分野의 區分을 위하여 그 圖書의 分類番號 앞에 해당 主題分野코우드(本 研究에서는 Library & Information Science의 첫자 "L")를 添加시켜 出力하도록 하였다.

여기서 分類에 관한 諸 規則을 들면 다음과 같다.

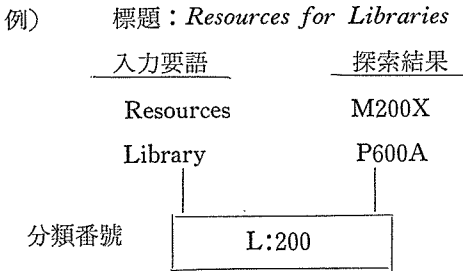
- ① 入力要語에 의한 各 用語의 探索 後에 分類番號는 [P][M][E][S][T]順으로 組合되며, 같은 카테고리 내에서는 코우드값이 큰 것에서 작은 것 順으로 配列되어진다.
- ② 入力要語 가운데에서 카테고리名과 코우드값이 모두 같은 경우에는 이들 중 하

17) N개를 가지고 特定的 順序로 組合possible한 경우의 數는 $5C_1 + 5C_2 + 5C_3 + 5C_4 + 5C_5 = 5 + 10 + 10 + 5 + 1 = 31$ 個가 된다.

나만을 使用하여 分類한다.

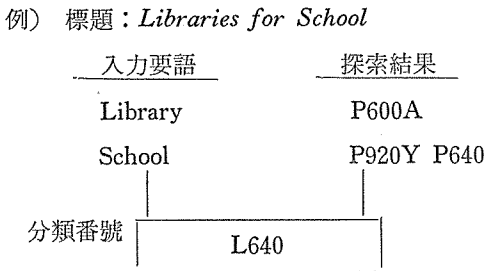


③ 入力要語 가운데에서 參照코우드 A와 X가 동시에 나타날 때에는 A코우드의 값이 무시되어 分類되어 진다.



따라서 이 경우에는 標題가 Library Resources가 되어 入力を Library Resources로 하여도 分類番號가 同一하게 되어 主題의 分散이 되지 않는다. 이러한 Resources와 같은 類型의 모든 用語에는 X參照코우드가 부여되어 있다.

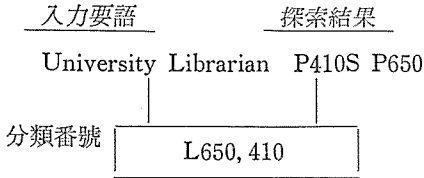
④ 入力要語 가운데 參照코우드 A와 Y를 가진 두 用語의 結合時는 A코우드의 코우드값은 무시되고 Y코우드의 뒤에 있는 指示코우드값에 의하여 分類되어 진다.



이것은 또한 標題가 School Library로 나타날 경우와 同一한 分類番號가 부여되게 되어 역시 主題의 分散을 防止할 수 있다.

⑤ 入力要語 가운데 參照코우드 S가 나타나는 경우에는 코우드상에 있는 것과 指示코우드를 함께 使用하여 分類하도록 하였다.

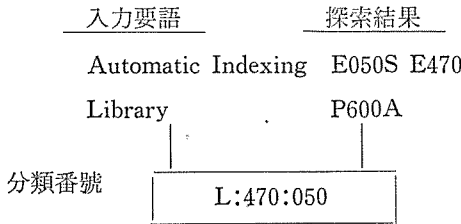
例) *University Librarian*



또한 이 때는 標題가 Librarians in University와 같이 나타날 경우 入力을 각각 Librarian, University Library로 入力하여도 同一한 分類番號를 이끌어 낼 수 있게 된다.

⑥ 入力要語 가운데 參照코우드 S를 가진 用語와 參照코우드 A를 가진 用語가 組合할 때는 A코우드의 값을 무시하고 S코우드의 값만으로 分類되도록 하였다.

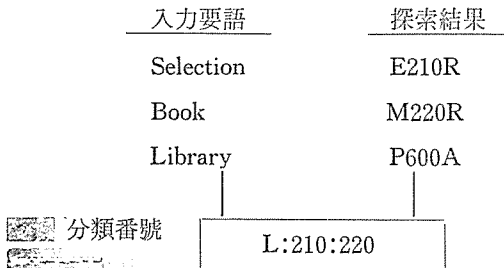
例) 標題 *Automatic Indexing in Libraries*



이 때는 Automatic Indexing 및 Automatic Indexing System 등과도 分類番號가 類似하게 되어 主題의 分散이 되지 않는다.

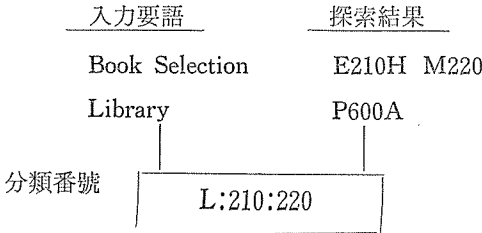
⑦ 入力要語 가운데 [E]카테고리와 [M]카테고리의 用語에서 각각 參照코우드 R이 나타날 경우에는 PEMST順으로 組合되도록 하였다.

例) 標題 : Selection for Books in Library



따라서 이 때는 標題가 Book Selection, 또는 Book Selection in Libraries 등과도 分類番號가 同一하게 되어 主題의 分散이 되지 않는다.

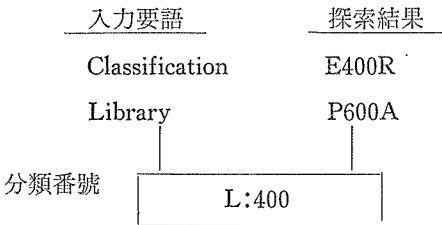
例) 標題 : *Book Selection in Library*



⑧ 入力要語 가운데 参照코우드 A와 H의 結合 때는 A코우드의 값은 無視되고 H코우드의 값으로 分類되어지며, 이들의 順序는 PEMST順가 된다(⑦의 例 參照).

⑨ 入力要語 가운데 参照코우드 R을 가진 用語와 A를 가진 用語의 組合 때는 A코 : 드의 값은 無視하고 R코우드값만으로 分類되도록 하였다.

例) 標題 : *Introduction to Classification in Library*



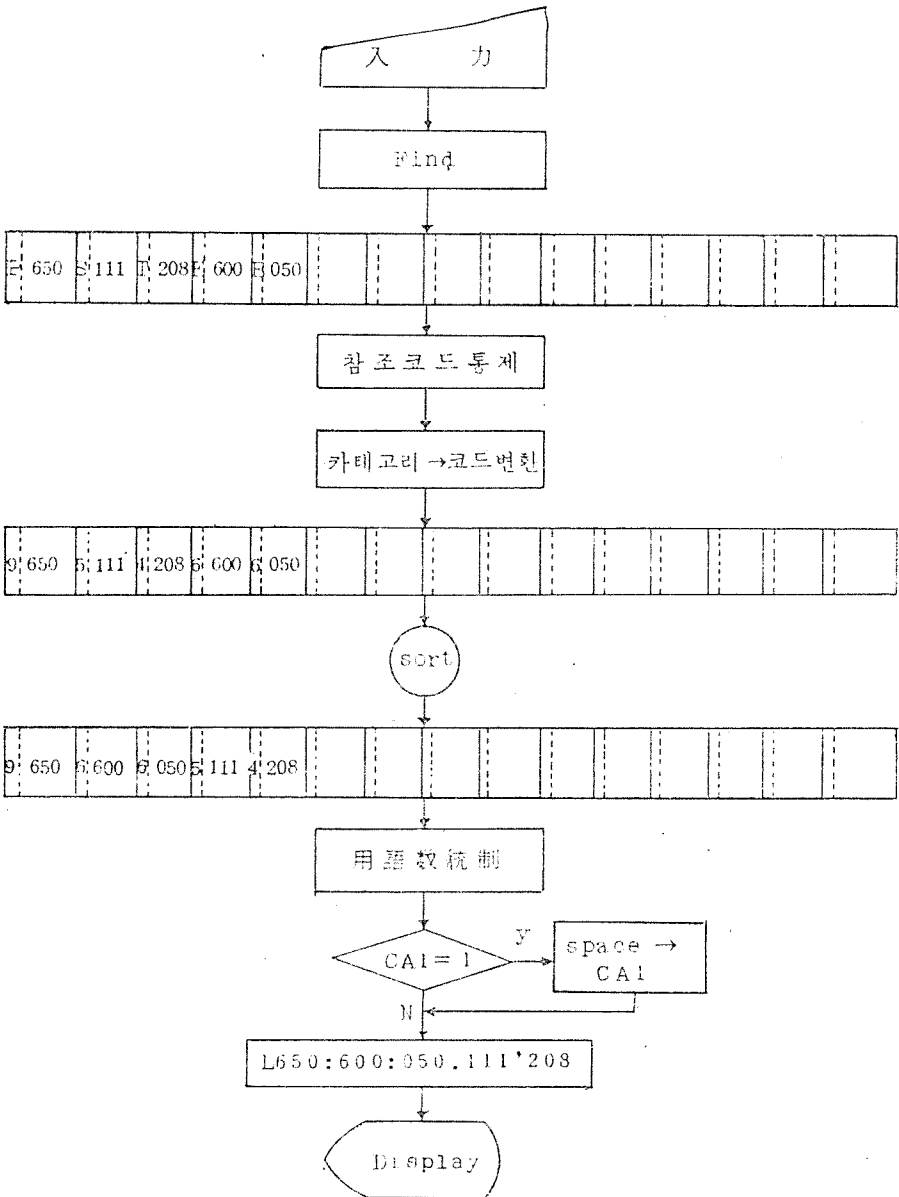
이 경우에는 圖書의 標題가 Library Classification 이 되어도 同一한 分類番號가 되며, 또 Book Classification, Classification for Books 등과도 主題分散이 안된다.

4. 프로그램의 흐름

分類番號가 決定되는 過程인 프로그램의 흐름은, 우선 分類作業을 행하는 司書가 文獻의 標題上에서 要語를 入力함으로써 始作된다. 이렇게 入力하는 用語는 컴퓨터에 이미 蓄積되어 있는 分類用파일에서 이들 用語를 探索하고, 그 結果를 컴퓨터의 단말장치에 디스플레이 함과 아울러 最終적으로는 이들 用語의 코우드를 分類原理에 따라 組合하여 分類番號를 出力하도록 하였다.

本 研究에서는 入力이 可能한 用語數는 最大 15個까지로 制限하고, 分類가 可能한 用語數는 8個까지로 하였다. 그것은 本 研究者가 調査한 바에 의하면 事實上 한 標題上에서 抽出할 수 있는 用語數는 平均 5個 이내가 거의 大部分인 것으로 나타나고 있지만¹⁸⁾ 그 이내이 되는 경우도 있기 때문에 좀 더 그 폭을 넓혀 入力은 15個,

18) 本 研究者가 Books in Print 1979年度에 收錄된 圖書館學分野 圖書들을 調査한 결과 한 標題上에서 抽出할 수 있는 用語의 數는 平均 5個 이내가 거의 大部分인 것으로 나타나고 있다.



〈圖 3〉 分類의 自動化 過程

分類는 8個까지 可能하도록 하였다.

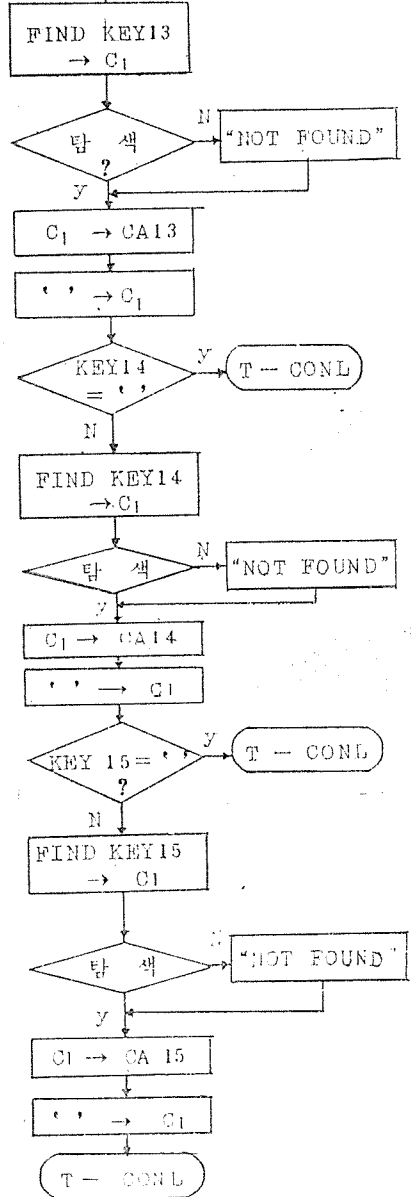
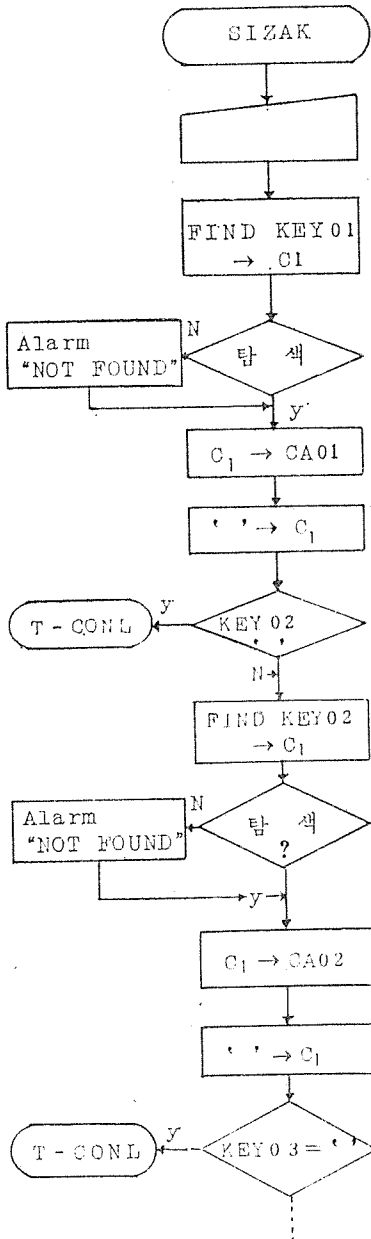
이 때 用語統制의 方法은 다음과 같다. 入力의 用語數가 모두 8個를 초과하는 경우에 있어서는 一次的으로 어떤 카테고리든지 3個 이상을 使用할 수 없도록 하고, 여기서 또 8個가 넘는 경우에는 2個 이상을 使用할 수 없도록 한다. 이러한 統制를 하여도 역시 8個를 초과할 경우에는 마지막으로 [T]와 [S]의 카테고리에 대한 統制를 한다. 統制를 받는 用語는 같은 카테고리 내에서 코우드값이 큰 것으로 하였다.

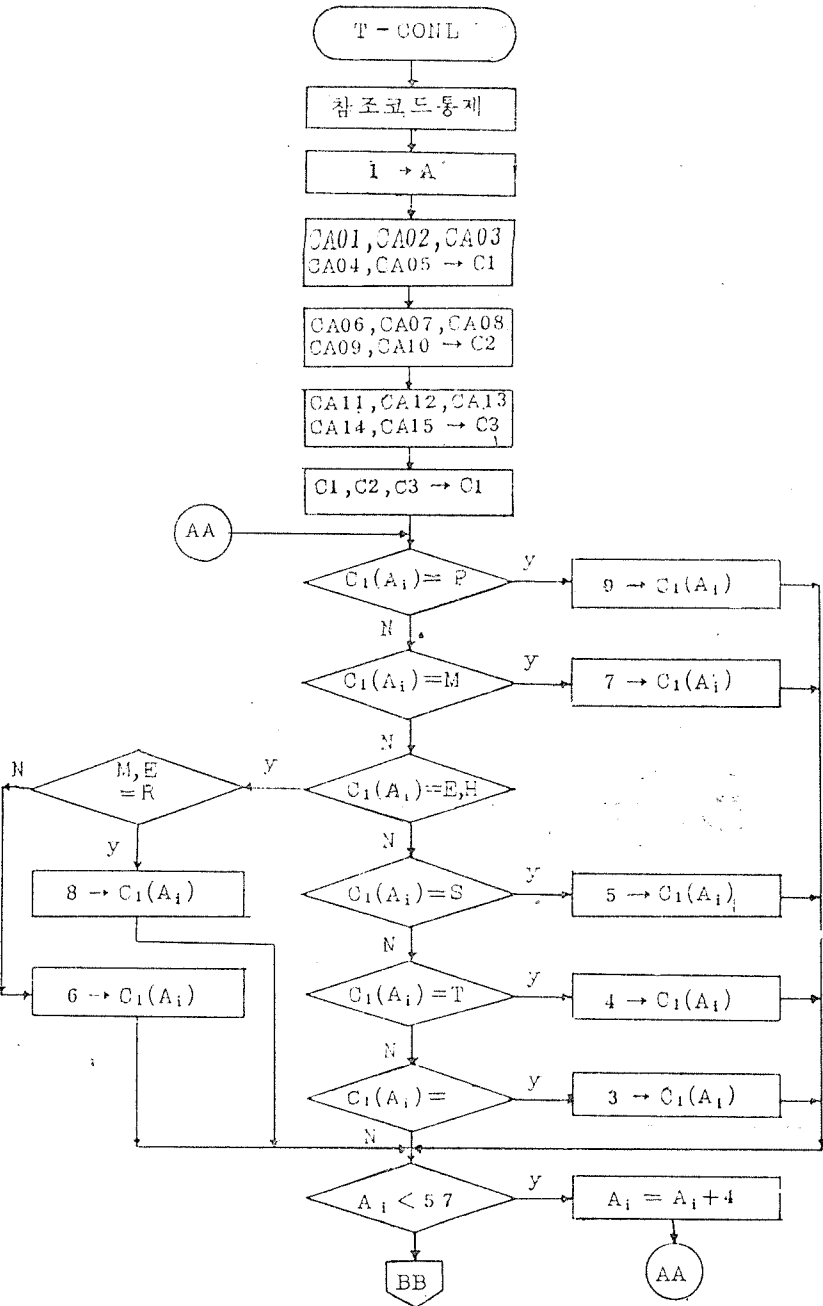
分類의 自動化흐름을 위해 간단한 예를 하나 살펴보기로 한다. 만약 入力의 用語가 각각 University Library, Korea, 1980, Library Cooperation, 그리고 Computer-Aided와 같은 경우라 한다면, 컴퓨터는 이들 用語를 探索하여 단말장치에 出力함과 동시에 記憶裝置에서의 過程은 <圖3>과 같이 進行되게 된다. 이 때 入力과 동시에 단말장치에는 다음과 같이 出力된다.

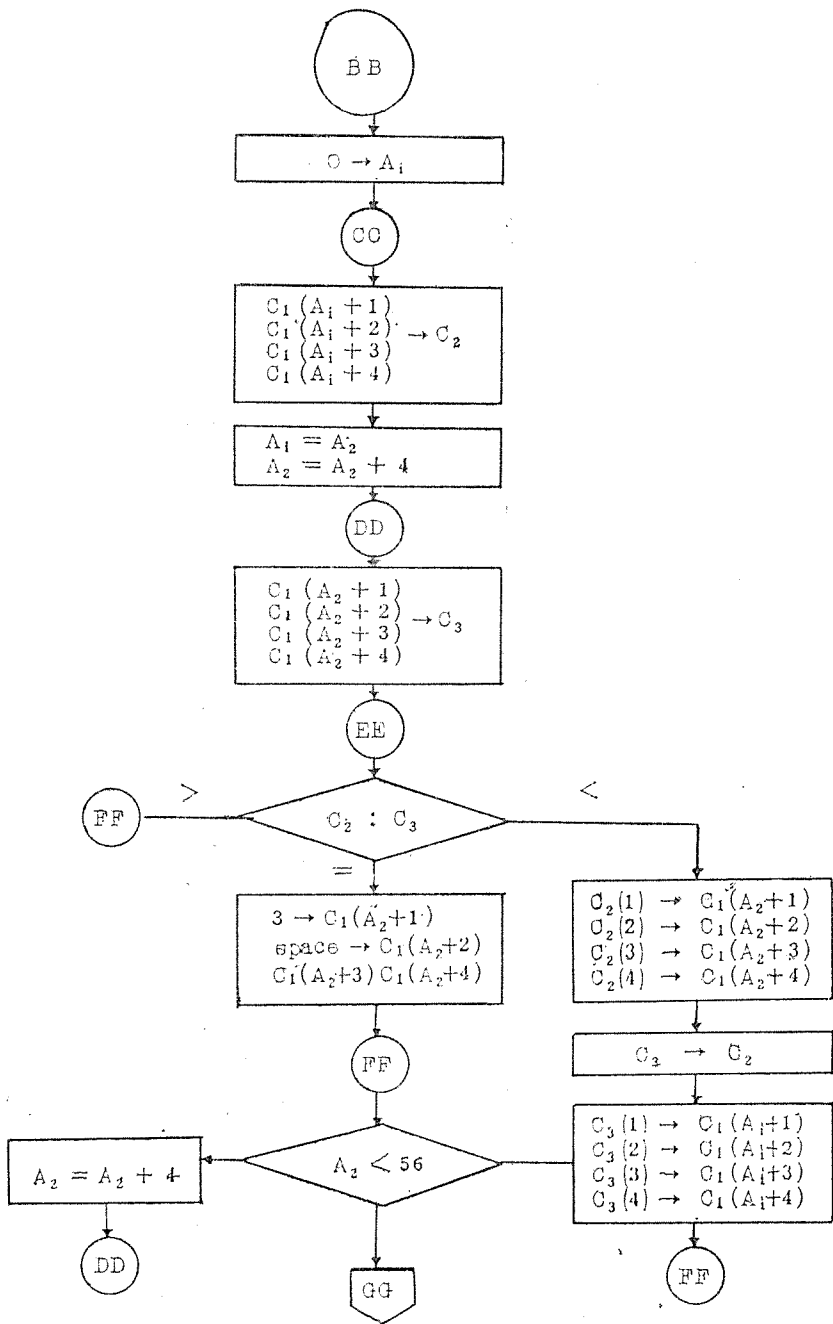
UNIVERSITY LIBRARY	P 650
KOREA	S 111
LIBRARY COOPERATION	E 600
COMPUTER-AIDED	E 050
1980	T 208
L 650 : 600 : 050. 111, 208	

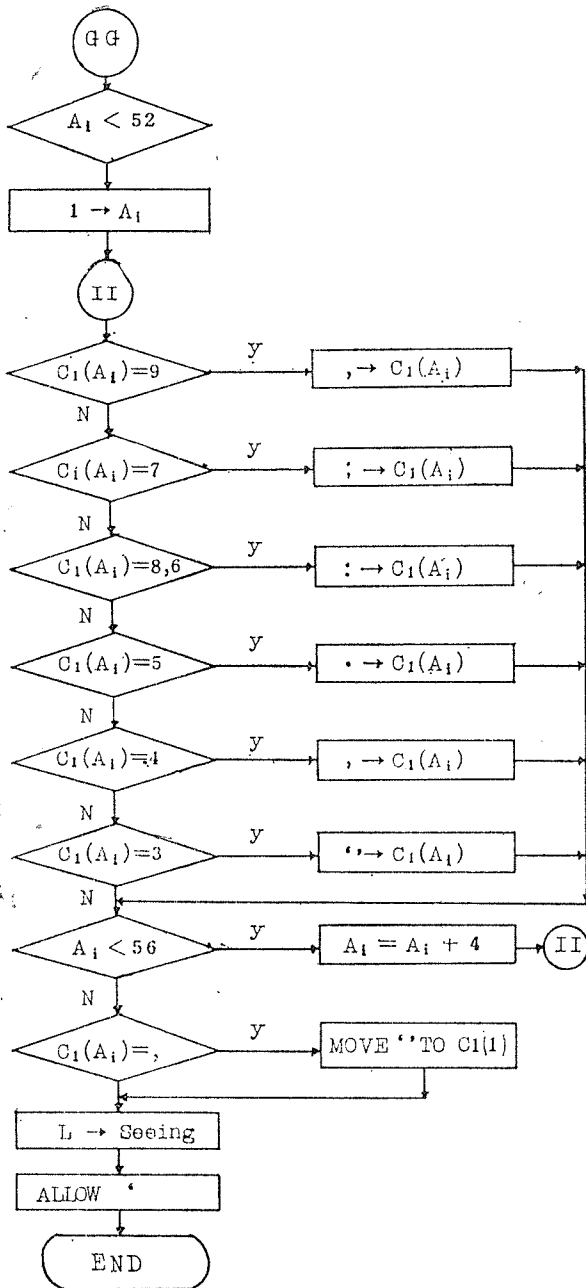
이와 같은 分類番號가 出力되는 過程을 보면 다음과 같다. 단말장치를 통하여 用語를 入力함과 동시에 프로그램에 의하여 이들 用語의 각각에 대한 코우드값을 探索하고, 이어 여러 參照코우드에 의한 組合關係를 決定한다. 그리고 나서 각 카테고리에 대하여 P는 9, M은 7, E는 6, S는 5로, 그리고 T는 4로 變換하되, 入力要語 가운데 M카테고리의 用語와 카테고리의 用語에서 다같이 參照코우드 R을 지니고 있을 경우나 參照코우드 H가 있는 경우에는 P를 9, M을 7, E를 8, S를 5 및 T를 4로 代置시키고, 이러한 狀態를 내림차 순으로 配列(sort)하면 이들의 論理的 順序는 各各 PMEST 및 PEMST가 된다.

이와 같은 配列 후 프로그램은 다시 P가 9로 바뀐 곳은 ,(콤마), M이 7로 바뀐 곳은 :(세미콜론), E가 8 및 6으로 바뀐 곳은 :(콜론), S가 5로 된 곳은.(점) 그리고 T가 4로 바뀐곳은 '(아프스트로피)를 부여하되, 만약 제일 앞에 P카테고리가 나타날 때에는 [,]를 省略하도록 하였다. 이어 主題番號 "L"을 제일 앞부분에









〈圖 4〉 프로그램의 흐름

부여되도록 하였다.

이상에서 言及한 프로그램의 흐름을 그림으로 나타내어 보면 다음 圖4와 같다.

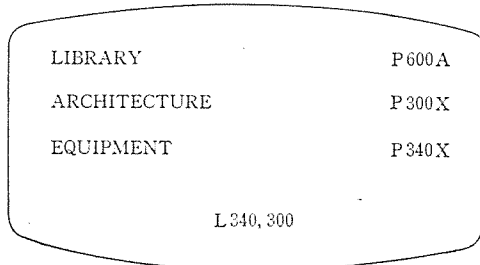
5. ABCS의 運用

이 分類의 自動化시스템(ABCS)은 다음과 같은 여섯가지 機能을 遂行한다.

- ① 要語에 의한 分類
- ② 蓄積된 用語의 디스플레이(display)
- ③ 不必要한 用語의 削除(deletion)
- ④ 새로운 用語의 追加(addition)
- ⑤ 잘못 入力된 코우드의 修正(correction)
- ⑥ 特定要語의 調査(inquiry)

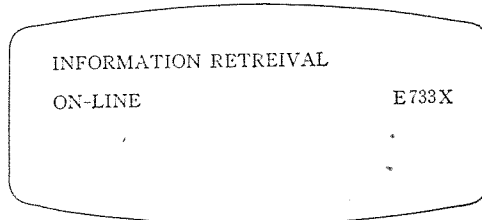
가. 要語에 의한 分類

要語에 의한 分類는 단말장치로 要語를 入力함으로써 이루어진다. 이 때 入力は 文獻의 標題上에 나타나 있는 用語가 운데에서 要語를 入力하면 入力한 數만큼 단말장치에 <圖5>와 같이 일단 디스플레이됨과 동시에 參照코우드에 의한 組合과 用語의 統制 및 分類의 原理 등에 의하여 分類番號가 단말장치의 下段部에 역시 出力되어 나온다.



<圖 5> 要語에 의한 分類의 出力 例

그러나 入力한 要語가 分類用파일에 蓄積되어 있지 않거나 혹은 蓄積은 되어 있 을지라도 잘못 入力되어 있는 경우에는 그 해당 要語를 入力하여도 探索이 不可能 하기 때문에 코우드값이 없이 그대로 나타나게 되며, 分類番號의 決定에도 <圖6>와



<圖 6> 入力要語가 파일상에 蓄積되어 있지 아니 한 경우의 出力 例

같이 아무런 影響을 미치지 않는다. 그러나 이러한 경우에는 파일상의 用語를 디스플레이하여 보고 蓄積이 되어 있지 않은지, 그렇지 않으면 잘못 入力되어 있는지 確認하여 볼 必要가 있다.

나, 蓄積된 파일의 디스플레이(display)

이 디스플레이에는 파일상의 特定部分을 調査하여 볼 必要性이 있을 경우에 利用된다. 예를들어 <圖6>에서와 같이 information retrieval을 단말장치로 入力하였으나 探索이 되지 않는 경우가 있다. 따라서 이 때에는 이 用語가 제대로 入力되어 있지 않은지, 그렇지 않으면 入力は 되어 있으나 잘못 入力되어 있는지를 알아보기 위한 경우에 주로 使用한다.

이 蓄積된 파일은 入力要語의 알파벳順으로 되어 있기 때문에 出力 또한 알파벳 順이다. 여기서 이 information retrieval의 用語가 위치한 部分을 디스플레이하여 본 結果는 <圖7>과 같다.

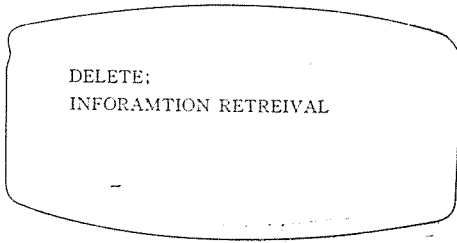
INDEX	M120
INDEXING	E470X
INDIA	S152
INDIVISUAL LIBRARY	P690
INDONESIA	S147
INFORAMTION RETRIEVAL	E750X
INFORMATICS	P200

<圖 7> 蓄積된 파일의 出力 例

이 디스플레이 상에서 본 바와 같이 information retrieval이 아닌 information retrieval로 蓄積되어 있음을 알 수 있다. 그러므로 이 用語는 파일상에서 削除하고 새로이 information retrieval로 入力하여야 한다는 것을 알 수 있다.

다. 不必要한 用語의 削除(deletion)

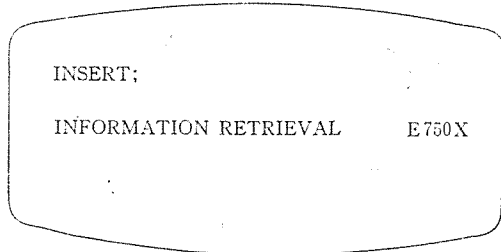
위의 <圖7>에서 디스플레이된 것을 보면 information retrieval이 information retrieval로 잘못 入力되어 있기 때문에 information retrieval을 削除하여야 한다. 削除하는 경우에는 “D”란 키(key)를 누르면 <圖8>과 같이 단말장치에 “DELETE”란 用語가 出力되어 나온다. 이 때 계속하여 削除하고자 하는 用語 즉, inforamtion retrieval을 入力하면 파일상에서 自動적으로 削除되어진다.



〈圖 8〉 削除의 例

라. 새로운 用語의 追加(addition)

分類用파일상에는 蓄積되어 있지 않지만 文獻의 標題上에서는 자주 要語로 나타나는 用語에 대하여는 파일상에 이들 用語를 入力하여 蓄積하여 두어야 한다. 이러한 경우 위의 information retrieval을 계속 適用하여 본다면, 지금 파일상에는 information retrieval이란 用語가 蓄積되어 있지 않기 때문에 이 用語를 入力하고자 할 때에는 追加키 “N”을 누르고 나서 “INSERT”란 메시지가 出力되어 나오면 이 때 入力하고자 하는 用語와 코우드를 함께 〈圖 9〉와 같이 入力하면 된다.



〈圖 9〉 새로운 用語의 追加 例

마. 파일상의 레코오드 修正(correction)

파일상에 蓄積되어 있는 用語 가운데 用語가 잘못 入力되어 있는 경우에는 修正이 不可能하기 때문에 削除 후 다시 入力追加를 하여야 한다. 그러나 用語가 아닌 코우드가 잘못 入力되어진 경우에는 修正만으로 可能하다. 만약 파일상에 “book classification” 이 E400X로 되어 있는데 이것을 E412X로 修正하고자 할 때에는 修正키 “F”를 누르고 “CORRECT”란 메시지가 나오면 계속하여 “book classification E412X”를 入力하면 바로 修正이 된다.

바. 特定要語의 調査(inquiry)

特定要語에 대한 디스플레이를 하고자 하는 경우에는 調査키 “i”를 누르고 나서 “INQUIRY”란 메시지가 나오면 調査하고자 하는 要語를 入力하면 단말장치에 出力되어 나온다.

6. 運用에 따른 分類의 實例

이상에서 다른 分類의 自動化시스템을 中心으로 實際 文獻의 標題上에서 나타난 要語를 入力하여 分類하여 본 結果는 表와 같다.

<表 5> 分類의 實例(_____은 入力要素의 表示임)

標 題	分 類 番 號
<u>Introduction to librarianship.</u>	L100
<u>Research Method on librarianship.</u>	L100 : 121
<u>The foundation of education for Library Science.</u>	L100 : 800
<u>Introduction to Information Science.</u>	L200
<u>Education in Information Science.</u>	L200 : 800
<u>Building in Libraries.</u>	L300
<u>Library Building.</u>	L300
<u>Library equipment.</u>	L340
<u>Libraries, architecture and equipment.</u>	L340, 300
<u>Library personnel.</u>	L400
<u>Personnel in Libraries.</u>	L400
<u>The teacher-librarian's handbook.</u>	L412
<u>The Serials Librarian</u>	L413
<u>Library law.</u>	L500
<u>American library law.</u>	L500, 420
<u>Manual of book classification and display for Public Libraries.</u>	L630 : 400 ; 220
<u>Libraries for School.</u>	L640
<u>School Library.</u>	L640
<u>School Library personnel task analysis.</u>	L640, 400 : 134
<u>School Librarian.</u>	L640, 410
<u>New education and School Library.</u>	L640 : 800
<u>The role of the beginning Librarian in University Libraries.</u>	L650, 410
<u>University Librarian.</u>	L650, 410
<u>Librarians in University Library.</u>	L650, 410
<u>The administration of the College Libraries.</u>	L650 : 100
<u>Interlibrary loan in Academic Libraries.</u>	L650 : 610
<u>Staff management in University and College Libraries.</u>	L651, 650:130:100
<u>Special Libraries & Information Service in India.</u>	L670 : 510, 152
<u>Picture file in School.</u>	L920 ; 132
<u>Computerized Library Catalog.</u>	L110 : 050
<u>Libry Resources.</u>	L200
<u>Resources for Libraries.</u>	L200
<u>Materials in Libraries.</u>	L200
<u>Periodicals and serials.</u>	L231 ; 230
<u>Administration of Archives.</u>	L246 ; 100

<u>Library computerization.</u>	L 050
<u>Computerization in Libraries.</u>	L 050
<u>Library automation.</u>	L 050
<u>Computer-based information system.</u>	L 050
<u>Computer-based Library & Information Systems.</u>	L 050
<u>Library Management.</u>	L 100
<u>Management in Libraries.</u>	L 100
<u>Management Techniques for Libraries.</u>	L 100
<u>Library Administration.</u>	L 100
<u>Organization & Administration in Libraries.</u>	L 110 : 100
<u>Scientific Management in Libraries.</u>	L 120
<u>Library system analysis.</u>	L 174
<u>System analysis in Libraries.</u>	L 174
<u>Library Acquisition.</u>	L 200
<u>Acquisition for Libraries.</u>	L 200
<u>Acquisition of Library Materials.</u>	L 200 ; 200
<u>Book selection.</u>	L 210 ; 220
<u>Selection of books.</u>	L 210 ; 220
<u>Selection of Atlas.</u>	L 210 ; 222
<u>Selection of Nonbook Materials.</u>	L 210 ; 240
<u>Introduction to Cataloging.</u>	L 300
<u>Cataloging U.S.A.</u>	L 300, 420
<u>Library classification.</u>	L 400
<u>Classification in Libraries.</u>	L 400
<u>Book classification.</u>	L 400 ; 220
<u>Classification of Nonbook Materials.</u>	L 400 ; 240
<u>Introduction to Cataloging & classification.</u>	L 400 ; 300
<u>Automatic Keyword classification.</u>	L 420 ; 050
<u>Computerization for Keyword classification.</u>	L 420 ; 050
<u>Computerization in indexing.</u>	L 470 ; 050
<u>Automatic indexing.</u>	L 470 ; 050
<u>Automatic indexing system.</u>	L 470 ; 050
<u>Automatic indexing in the Libraries.</u>	L 470 ; 050
<u>Classification & Indexing.</u>	L 470 ; 400
<u>Introduction to Reference works.</u>	L 510
<u>Reference & Information service.</u>	L 510
<u>Community service.</u>	L 520
<u>Dissemination of materials in the Library.</u>	L 520 ; 200
<u>Dissemination of Library materials.</u>	L 520 ; 200
<u>Information Dissemination.</u>	L 520 ; 201
<u>Dissemination of Information.</u>	L 520 ; 201
<u>Information processing.</u>	L 730 ; 201
<u>Data processing for Libraries.</u>	L 730 ; 210

<u>Data processing in 1900's.</u>	L 730 ; 210, 200
<u>Information retrieval systems.</u>	L 750
<u>Information retrieval & processing.</u>	L 750 : 730 ; 201
<u>Computer-aided information retrieval.</u>	L 750 : 050
<u>Information retrieval on-line.</u>	L 750 : 733

四. 分類의 自動化에 관한 提言

1. 概念分析上的 提言

가. 카테고리

分類의 自動化에 있어 概念分析에 관한 關心은 世界的 추세로 나아 가고 있다. 특히 概念分析에 의한 경우에는 카테고리의 分析이 시스템의 成敗를 좌우할만큼의 큰 比重을 차지하고 있다. 그러나 어떤 用語의 分析에 있어서는 카테고리의 明確性不在로, 카테고리의 意味附與에 다소의 混亂을 야기시키고 있다 하겠다.

따라서 分類의 自動化을 위하여는 무엇보다도 카테고리의 明確한 分析을 위하여 이에 대한 하나의 基準을 定立하여 두는 것이 바람직하다고 하겠다.

나. 意味論

概念分析時에 대두되는 意味論的 問題는 주로 類似性 및 意味의 階層關係가 있는 用語에 대한 分析이라 하겠다. 여기서 類似性의 用語는 가능한 한 모아서 같은 코우드를 附與하면 그만큼 分類番號가 짧아질 수 있는 長點이 있지만 이를 適用하지 않을 때는 分類番號가 길어지는 短點이 있다.

또한 意味의 階層關係에 있어서는 어디까지를 上位概念으로 하고 어디까지를 下位概念으로 할 것인지를 決定하기가 어렵다 하겠다.

따라서 意味論的 側面에서의 分析은 모든 用語를 全部 包括하여서 다룬 디소오러스(Thesaurus)의 使用이 무엇보다 큰 意味를 가져다 주리라 믿는다. 그러므로 分類의 自動化도 概念分析의 方法에 의한 경우라면 디소오러스와 같은 用語統制가 先行되고 난 후 自動化하는 것이 더욱 바람직하다고 본다.

2. 標題의 表現

文獻의 標題는 그 文獻이 무엇에 관하여 記錄되어 있는지를 가장 簡單 明瞭하게 表現하여야 하기 때문에 단지 指示의 情報만으로 나타낸다. 그러나 우리는 文獻의 標題가 물론 애매모호한 것도 있지만 大部分은 文獻의 內容을 充分히 나타내고 있

다고 본다. 이미 前述한 바와 같이 KWIC가 紹介되면서부터는 著者들은 過去 어느 때 보다는 더욱 記述的이고 意味 있는 標題를 쓰려고 한다는 것은 文獻의 標題가 文獻檢索의 自動化에 따라 그 重要性이 認識되어진 한 例라 하겠다.

특히 文獻의 標題上에 나타난 用語에 의한 分類는 檢索의 機能도 함께 지니고 있으므로 너무 包括的인 用語도 피하여야 하겠지만 너무 많은 用語에 의한 表現도 分類番號가 길어지는 短點이 있어 考慮하여야 할 問題라 하겠다.

따라서 이러한 諸 問題點을 考慮하여 文獻의 內容을 가장 잘 나타낼 수 있는 具體的인 用語로 用語가 記述되어질 수 있도록 著者나 出版業者에게 그 必要性을 널리 周知시키는 것이 바람직하다 하겠다.

3. 分類番號의 길이

分類番號의 길이는 同意語나 類似語 및 參照코우드에 의한 組合이 없는 한 入力하는 用語의 數에 比例하기 때문에 많은 用語를 入力하는 경우에는 그만큼 分類番號의 길이가 길어진다. 그러나 이렇게 分類番號의 길이가 길어지게 될지라도 書架 分類가 아닌 書誌分類이기 때문에 큰 問題가 誘發되는 것은 아니지만, 可能하면 짧은 分類番號를 가질 수 있도록 적절한 用語의 統制가 要望된다 하겠다.

4. 機械的 處理

分類를 하는데 있어서 컴퓨터가 스스로 文獻의 標題를 分析하여 分類番號를 이끌어 낼 수 있다면, 文獻의 標題에서 要語를 入力하여 分類하는 것 보다는 더욱 바람직하다고 하겠다.

그러나 本研究에서는 단말장치로 要語를 入力함과 동시에 온라인으로 分類番號를 디스플레이되도록 設計하였기 때문에, 司書가 文獻의 標題로부터 要語를 抽出하여 단말장치로 入力하여야 하는 問題點이 있다.

따라서 機械上의 處理에 있어서는 要語에 의한 入力보다는 文獻의 標題를 全部 入力함으로써 分類番號를 이끌어 낼 수 있는 方法의 開發이 더욱 바람직하다고 하겠다.

五. 結 論

오늘날 圖書館 및 情報센터 등은 文獻量의 急增에 따른 効率的인 情報利用을 위하여 情報의 組織 및 接近에 대한 새로운 方法論의 摸索이 活潑히 進行되고 있으며

分類도 또한 例外는 아니다. 이것은 무엇보다도 情報의 組織이 情報檢索의 成敗를 좌우할만큼 重要한 意味를 지니고 있기 때문이라고 할 수 있다.

따라서 本 研究에서는 그와 같은 當面問題를 다소나마 解決하여 보고자 하는 意圖에서 란가나단의 分類原理에 立却하여 檢索上的 分類는 물론, 書誌上的 分類도 可能할 수 있도록 分類의 自動化에 관한 實驗的인 模型(案)을 定立하여 보았다.

本 研究의 結果를 要約하여 보면 다음과 같다.

1. 本 研究의 實驗對象은 圖書館 및 情報學 分野로 局限하였다.
2. 分類用파일의 構成을 위한 用語의 蒐集은 既存 分類表와 *Library Literature*, 그리고 諸 圖書館學 및 情報學關係 關聯文獻으로 하였다.
3. 蒐集된 用語의 分析은 一次的으로 란가나단의 카테고리(P,M,E,S,T)에 의하여 分析한 후 二次的으로는 意味論的 側面에서 同意語 및 類似語를 中心로 分析하였다.
4. 概念을 組合할 때 惹起되기 쉬운 主題의 分散을 防止하기 위하여, 參照코우드(A,X,Y,S,R,H)를 使用하여 分析하였다.
5. 分析된 用語로 이루어진 分類用파일은 要語의 入力에 의한 分類 뿐만 아니라 蓄積된 用語의 디스플레이, 不必要한 用語의 削除, 새로운 用語의 追加, 잘못된 코우드의 修正 및 特定要語의 調査 등이 항상 可能하도록 하였다.
6. 分類는 단말장치로 要語를 入力하면 이들 用語를 分類用파일에서 探索한 후 [P][M][E][S][T] 順, 그리고 각 카테고리 내에서는 코우드값이 큰 것에서 작은 것의 順으로 分類되도록 하였다.
7. 同意語나 類似語의 경우에는 코우드값이 같기 때문에 하나의 番號만으로 分類되도록 하였다.
8. 主題의 分散을 防止하기 위하여 參照코우드를 使用하여 特定用語의 省略指示 結合이 可能하도록 하였다.

또한 本 研究의 實驗結果에 따라 期待할 수 있는 效果는 다음과 같다.

1. 分類의 迅速, 正確性を 維持할 수 있다.
2. 分類表(分類用파일)의 更新이 항상 可能하다.
3. 迅速한 分類로 利用者에게 最大限의 奉仕를 할 수 있다.
4. 利用者は 컴퓨터와의 對話로 자기가 원하는 資料의 檢索 및 資料에 대한 分類 番號도 探索할 수 있다.
5. 概念組合方法이므로 複雜한 主題일지라도 主題의 表現이 容易하다.
6. 모든 書誌分類 機關은 統一된 分類番號를 維持할 수 있어 凡國際的인 情報流

通의 手段으로도 使用되어질 수 있다.

이상과 같이 概念分析의 方法에 의한 圖書分類의 自動化시스템(ABCS)은 既存分類體系가 지니고 있는 많은 問題點을 解決할 수 있는 效果를 期待하여 볼 수 있기 때문에, 앞으로 分類의 自動化를 위한 研究는 檢索의 機能도 함께 지닐 수 있는 概念分析의 方法에 의한 코우드화된 시스템을 開發하는데 主眼點을 두어야 할 것으로 생각되며, 이러한 시스템開發을 통하여 凡國際的인 情報流通도 다소 圓滑化되어질 수 있도록 하여야 할 것이다.

〈參 考 文 獻〉

- 남기섬. 이정민. 이홍배. 언어학개론. 서울: 탑출판사, 1977.
- 嶺南心理學會 編, 敎養心理. 서울: 嶺南心理學會, 1974.
- 이재철 편. 주제명표목표. 서울: 연세대학교, 1961.
- 韓國圖書館協會 編. 圖書館用語集. 서울: 韓國圖書館協會, 1975.
- 韓國圖書館協會 編. 韓國十進分類法. 修正版四刷. 서울: 韓國圖書館協會, 1975.
- Artandi, Susan. *An Introduction to Computers in Information Science*, 2nd ed. Metuchen: The Scarecrow Press, c1972.
- Austin, Derek. "The Development of PRECIS: a Theoretical and Technical History." *Journal of Documentation*, Vol.30, No.1 (March 1974), pp.47~102.
- Bichter, Julice. and Parsons, Ronald G. "Document Retrieval by means of an Automatic Classification Algorithm for Citation." *Information Storage and Retrieval*, Vol. 10(1974), pp. 267 ~278.
- Black, Donald. V. "Automatic Classification and Indexing for Libraries?" *Library Resources & Technical Services*, Vol.9, No. 1(Winter 1975), pp.35~52.
- Borko, Harold, and Bernick, Myrna. "Automatic Document Classification." *Journal of the Association for Computing Machinery*, Vol.10, No.2 (April 1963), pp.151~161.
- Brown, A.G. *Introduction to Subject Indexing: a programmed text*. London: Clive Bingley, c1976. 2 Vols.
- Classification Research Group. "Bulletin no.9." *Journal of Documentation*, Vol.24, No.4 (Dec. 1968), pp.273~291.
- Coates, E.J. "Classification in Information Retrieval: The Twenty Years Following Dorking." *Journal of Documentation*, Vol.34, No.4 (December 1978), pp.289~299.
- Coates, E.J. "A Manual of Concept Analysis and Subject Indexing." *Journal of Documentation*, Vol. 32, No.1 (1976), pp.85~96.
- Dattola, R.T. "A Fast Algorithm for Automatic Classification." *Journal of Library Automation*, Vol 2, No.1 (March 1969), pp.31~48.
- Dewey, Melvil dev. *Dewey Decimal Classification and Relative Index*. 18th ed. New York: Forest Press, Inc., 1971.
- Farradane, J. "Concept Organization for Information Retrieval." *Information Storage and Retrieval*, Vol.3 (1967), pp.297~314.

- Farradance, J.E.L. "Analysis and Organization of Knowledge for Retrieval." *Aslib Processing*, Vol. 22, No.12 (December 1970), pp.606~616.
- Feinman, R.D., and Kwok, K.L. "Classification of Scientific Documents by means of Self-Generated Groups Employing Free Language." *Journal of the American Society for Information Science*, Vol.24 (Sept-Oct. 1973), pp.382~396.
- Field, B.J. "Towards Automatic Indexing: Automatic Assignment of Controlled-Language Indexing and Classification from Free Indexing." *Journal of Documentation*, Vol. 31, No.4 (December 1975), pp.246~265.
- Foskett, A.C. *The Subject Approach to Information*, 3rd ed. London: Clive Bingley, c1977.
- Gardin, J.C. "Document Analysis and Information Retrieval." *Unesco Bulletin for Libraries*, Vol.XIV, No.1 (January 1960), pp.1~5.
- Jones, Kevin P. "The Environment of Classification Part II: How we classify." *Journal of the American Society for Information Science*, Vol.25, No.1 (Jan.-Feb. 1974), pp.44~51.
- Kent, Allen. *Information Analysis and Retrieval*. New York: Backer and Hayes, Inc., c1971.
- Kessler, M.M. "Bibliographic Coupling between Scientific Papers." *American Documentation*, Vol.14, No.1 (January 1963), pp.10~25.
- Kumar, Krishan. *The Theory of Classification*, New Delhi: Vikas Publishing House PVT. LTD., c1979.
- Kumar, P.S.G. "International Study Conference on Classification Research (3)." *Herald of Library Science*, Vol.14, No.1 (January 1975), pp.8~29.
- Kwok, K.L. "The Use of Title and Cited Titles as Document Representation for Automatic Classification." *Information Processing & Management*, Vol.11 (1975), pp. 201~201.
- Lancaster, F.W. "The Third International Study Conference on Classification Research," *Bulletin of the American Society for Information Science*, Vol.2, No.4 (November 1975), pp.23~24.
- Library & Information Science Abstracts. London: The Library Association, c1969.
- Listin, David M., and Howder, Murry C. "Subject Analysis." *Annual Review of Information Science and Technology*, Vol.12(1971), pp.81~118.
- Meincke, Peter P.M., and Atherton, Pauline. "Knowledge Space: A Conceptual Basis for the Organization of Knowledge." *Journal of the American Society for Information Science*, Vol.27, No.1 (Jan.-Feb. 1976), pp.18~24.
- Metcalf, John. *Information Retrieval, British & American, 1876~1976*. Metuchen: The Scarecrow Press, Inc., 1976.
- Moberg, Zandra. "Automatic Classification: Direction of Recent Research." *Drexel Library Quarterly*, Vol.10, No.4 (October 1974), pp.90~104.
- Moss, R. "Categories and Relations: Orgins of two Classification theories." *American Documentation*, Vol.15, No.4 (October 1964), pp.296~301.
- Needham, R.M., and Sparck Jones, K. "Keyword's and Clumps." *Journal of Documentation*, Vol.20, No.1 (March 1964). pp.5~15.
- Neil, S.D. "Farradane's Relation as Perceptual Discrimination," *Journal of Documentation*, Vol.31, No.3 (September 1975), pp.144~157.