

情報處理時의 CDS/ISIS 利用

Using of CDS/ISIS on Information Processing

徐 眞 二*
(Seo, Jinny)

抄 錄

情報의 定義와 特性을 高찰해보고, 이 情報를 利用하기 위한 托탈시스템의 하나인 CDS/ISIS에 關係 소개한다. Micro ISIS의 主要기능과 特性을 서술하고 있으며, 또한 시스템 利用시의 機能향상을 促하는 方法등을 기술한다. 영문자는 물론 한글문자도 가능한 한글 버전 프로그램의 개발이 시급히 要求된다.

키 워 드

情報, 데이터, 데이터베이스, 정보검색, 托탈시스템, CDS/ISIS, 인버티드파일, 검색, 순서 배열, 워드프로세서, WordStar, Fastsort, 연속간행물, SLS

ABSTRACT

Look for definition and characters of information and data, CDS/ISIS which is one of the total system in information processing are introduced. Explain the main function, characters and advantages of CDS/ISIS and furthermore try to encourage of the function of current ISIS. We need program for Korean language are available.

KEYWORDS

Information, Data, Database, Information retrieval, Definition, Total system, CDS/ISIS, Inverted File, Search, Sort, Fastsort, WordStar, Serials, SLS

* 産業技術情報院 電算室 研究員.
Researcher, Computer System Division, KINITI.

I. 序 論

“情報의 홍수”라든가 “情報의 폭발(Information Explosion)”이라는 표현이 현재를 대표하듯이, 오늘날에는 ‘Information(情報)’이 중요한 단어중의 하나이다. 어떠한 인간 관계나 인간 활동에도 커뮤니케이션을 통한 정보의 전달, 생성 등의 과정이 상당히 중요시 된다. 나아가 모든 지식은 現在 발생되고 있는 정보나 연구중인 정보에서 출발한다. 이러하듯이 인간사회의 주변에는 情報가 될 수 있는 데이터가 산재해 있다고 볼 수 있다. 따라서 궁극적인 목적은 이러한 情報를 이용하여 원하는 결과를 얻는 데 있다.

II. 情報의 定義와 특징

데이터란 인간이 생활하거나 연구 개발한다던가 등의 과정을 통하여 生成되는 부산물이라고 할 수 있으며 이는 수치자료이거나, 신문, 잡지등의 도서 자료나 실험등을 통한 결과 등의 여러가지가 될 수 있다.

나아가 이러한 데이터를 이용 가능하게 변환한것을 情報라고 定義할 수 있다. 情報에 관한 학술적인 定義는 여러가지이다. ① Szaniawski는 ‘實用主義的인 情報는 의사결정을 해야할 문제에 있어서 선택 가능한 상황에 관한것으로 이때의 가치는 의사결정 문제의 해결정도에 의하여 결정된다’고 실용주의적 입장에서 定義했고, 田中靖政은 ‘情報란 不確實性を 감소 또는 제거시켜 주는 것’이라고 했다. ② 情報에 관한 질적인 定義를 보면 ‘정보는 데이터를 표시함에 있어서 既知의 약속이나 결정에 의하여 인간이 데이터에 부여하는 의미(meaning)이며, 또한 우리가 과거에는 알지 못했던 것에 대하여 타인이나 機械가 알려주는 경우에 얻어지는 것을 말한다’고 정의한다. 정보의 양은 불확실성이 감소되는 양에 의하여 결정되며, 情報量은 일어날지도 모르는데 실제로는 일어나지 않았던 것의 함수이다라고도 定義되어 지기도 한다. ③ 情報란, 사람이 어떠한 메시지를 선택(select)할 때의 선택(choice)의 자유도를 나타내는 척도이다. 情報의 개념은 개개의 메시지가 아니고 전체로서의 상황에 관하여 적용된다. 또한 情報는 우리가 선택(choice)을 추구할 때 필요로 하는 그 무엇이

다. 그 내용여부에 관계없이 필요로 하는 情報의 양은 선택대상의 多小(複雜性)에 의하여 결정된다고 Miller가 定義한다. ④ 情報은 어떠한 의미를 가지고 생각과 사상을 표현하는 것이다. 情報의 소재는 데이터(data)이며 데이터를 情報利用者의 요구에 따라서 일정한 형태로 처리하는 것이라고 定義한다.

이렇게 여러가지로 定義되는 情報라는 것은 수신자가 존재함으로써 비로서 情報로 될 수 있으며 수신자의 상황적 요인에 따라 그 가치(의미)가 달라진다. 이것이 물질, 에너지와 다른 점이다. 또한 情報에는 보편성이 없다. 그리하여 어떤 사람에게는 중요한 의미를 가지는 것이 다른 사람에게는 전혀 쓸모 없는 것일 수도 있다. 情報은 그 자체가 단독으로 존재하는 것이 아니고 수신자가 있고 또한 어떤 목적에 사용됨으로써 비로서 情報가 된다. 귀중한 논문 같은 것도 그 자체는 단지 인쇄물이란 물리적 존재에 불과하게 된다.

또한 情報은 소모되지 않고 이동하지도 않으며 복제가 가능한 특징이 있다. 情報은 물질이나 에너지와는 달리 사용하여도 마모되지 않고 타인에게 제공하여도 소모되지 않는다. 따라서 情報을 복제하여 많은 사람들에게 전달하기도 하고 장기적인 축적이 가능하다. 또 가격을 매겨서 판매할 수도 있다.

그러므로 情報전달의 과정은 情報源이 제공하는 情報(data)로 부터 필요한 정보(information)를 색출해서, 이것을 분석 가공하여 사업활동(의사결정)에 도움이 되는 情報(intelligence)를 얻으며 이러한 情報가 보편화되어 지식(knowledge)으로 축적되는 것이라고 말할수 있다.

결론적으로 창조적 지적활동의 소산인 情報은 독창성(신규성, originality), 우선성(priority)을 가지는 것으로 그것이 사회적으로 인식되고 객관적 가치를 보증받기 위해서는 그것이 가능하도록, 어떠한 물리적 형태를 기록하여 공개되어져야 한다. 이렇게 여러가지의 특징을 가진 情報를 수집, 축적하여 결국은 원하고자 하는 情報를 바로 서비스 받고자 함이다. 즉, 많은 양의 데이터, 情報를 축적매체에 저장하여 컴퓨터를 통하여 쓰는 것을 의미한다. 컴퓨터를 통한 情報의 데이터베이스화가 필요하다.

Ⅲ. 데이터베이스

데이터베이스(database)란 각종 데이터(data)를 컴퓨터 처리가 가능하도록 (컴퓨터 가독형) 수집, 정리, 가공, 편집한것을 말한다. 즉 데이터를 정리, 종합하여 컴퓨터에 의한 검색과 갱신처리가 가능하도록 구조화한 형식의 정보파일 또는 그 집합체를 말하며, 이것을 제작하는 활동과 이것을 이용함을 의미한다.

또한 데이터베이스는 일정 기간마다 신규 情報의 추가 또는 데이터의 교환이라는 데이터베이스의 갱신(update)이 필요하다. 데이터베이스는 각 情報의 분야나 목적등에 따라 각각 다르게 만들어지며 특정되어 질 수 있는데 각각 데이터베이스안에서의 실제적인 내용이 얼마나 원하는 내용을 다 망라하느냐, 얼마나 최신의, 최적의 情報를 제공하느냐 등에 따라 데이터베이스를 평가한다.

이러하듯이 주변에 산재해 있는 모든 데이터, 情報를 수집, 분류, 분석하여 원하는 데이터베이스의 특성대로 定義하며 실제적으로 데이터를 축적하여 그 축적된 데이터를 이용한다.

과거에는 데이터를 처리하는데 있어 情報의 분석, 축적, 색인어 파일의 작성은 情報관리자가, 검색은 실제 이용자가 하는등의 분리된 체계였으나 요즘은 실제적인 이용자 입장에서라도 데이터를 분석, 定義, 입력, 축적하는 등의 토털 시스템(total system)의 필요성이 대두되고 있다. 情報처리의 최근 동향은 ① 데이터베이스의 관리 시스템(DBMS)를 운용하고, ② 온라인 처리를 주체로 하며, ③ Total System을 지향하는 추세이다. 이런 추세에 힘입어 DBMS와 Total System이 가능한 시스템은 여러가지가 있다. 그중 하나가 유네스코에서 개발된 CDS/ISIS(Computerized Documentation System / Integrated Sets of Information Systems)이다.

Ⅳ. CDS/ISIS

CDS/ISIS란 구조화된(structured) 비수치형 데이터베이스를 컴퓨터로 관리하기 위해 1985년 UNESCO에서 특별히 고안한 메뉴 방식의, 일반화된 情報축

적 및 검색 시스템이다. 비수치형 데이터, 즉 문장으로 구성된 데이터를 데이터베이스화하고 관리한다. 일반화된 시스템 설계기능을 제공하는데 CDS/ISIS는 데이터베이스 수에는 제한없이 취급할 수 있으며, 각 데이터베이스는 완전히 다른 데이터 요소(data elements)로 구성되는 장점이 있다. 정보관련의 전문가가 오랜 기간 동안 분석하고 실제의 이용자의 요구 등을 다 고려하여 개발한 소프트웨어이다.

또한 유네스코 본부의 다국어 지원체제의 일환으로 영어, 불어, 스페인어 버전에 이어 아랍어와 중국어 버전이 개발되어 있다. 모든 프로그램은 전부 대화식이므로 이용자가 이용하기 용이하게 되어 있으며 화면에 보여지는 시스템 메뉴와 메시지 언어도 다국어지원에 힘입어 여러 언어가 가능하게 했다.

1. 주요 기능

CDS/ISIS는 문자나 단어들을 취급하고 워드프로세서 패키지에서 일반적으로 나타나는 특징들을 많이 제공하므로 text processing보다 더 많은 기능을 제공한다고도 볼 수 있다. 또한 CDS/ISIS의 특징은 가변길이의 필드(결과적으로 레코드)를 취급하기 위하여 특별히 고안 설계되어, 한편으로는 디스크의 용량을 최적화했다.

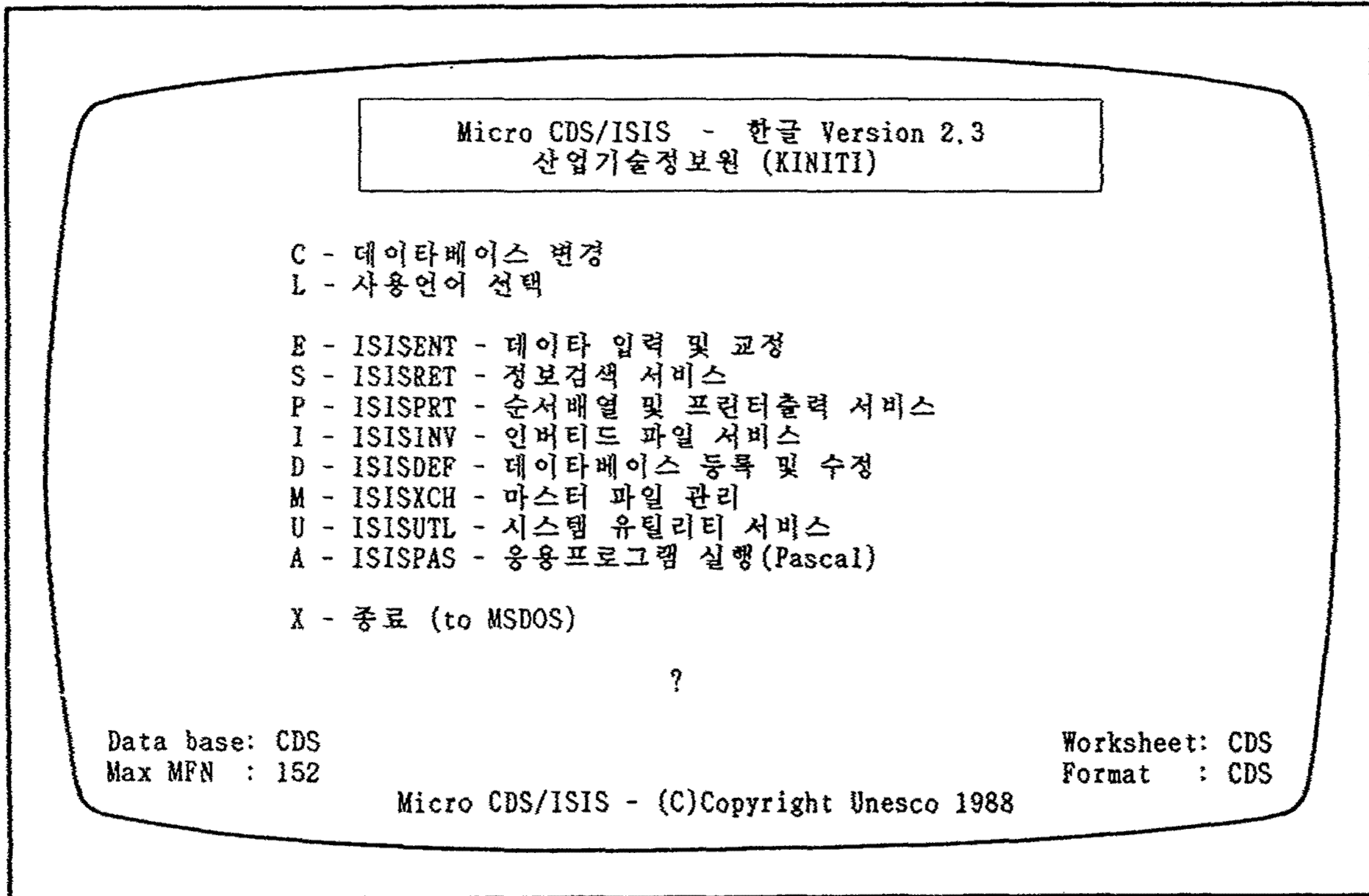
CDS/ISIS는

- ① 요구되는 데이터요소를 포함하고 있는 데이터베이스를 定義
- ② 주어진 데이터베이스에 새로운 레코드를 입력
- ③ 기존의 레코드를 수정 혹은 삭제가 가능
- ④ 검색 속도를 최대화 하기위해 각 데이터베이스의 빠른 접근 파일(fast access file)을 자동으로 생성, 유지
- ⑤ 탁월한 검색언어에 의해 내용에 따라 레코드를 검색
- ⑥ 요구에 따른 일부분 혹은 전체 레코드를 화면에 디스플레이
- ⑦ 원하는 순서로 레코드를 소트
- ⑧ 목록과 색인의 일부분 혹은 전체를 printer로 출력
- ⑨ CDS/ISIS자체의 집약된 프로그래밍 기능을 이용해서 특별한 응용프로그램 개발이 가능 등 여러가지 주요기능이 있다.

이상의 여러 기능을 8가지의 주요 서비스로 제공하는데, 2가지의 큰 범주로 분류할 수 있다. 기존의 데이터베이스를 작동시키는 이용자 서비스 기능과 새

〈圖 1〉

CDS/ISIS의 메인 메뉴표시



로운 데이터베이스를 생성시키기 위한 데이터베이스 설계 및 여러가지 시스템 일을 수행하는 시스템 서비스 기능으로 분류한다. 이용자 서비스 측면은 CDS/ISIS의 기본작동 지식만 요구되는 반면 시스템 서비스는 모든 시스템 구성요소의 심도있는 기술적인 지식과 더 나아가 컴퓨터 프로그래밍의 익숙함을 요구한다.

이용자 서비스 측면으로 ISISENT (데이터의 입력과 레코드의 교정), ISISRET(정보검색), ISISPRT(목록이나 색인과 같은 출력), ISISINV(인버티드 파일 유지 및 유틸리티 기능)이 제공된다. 반면에 시스템 서비스의 측면으로는 ISISDEF(새로운 데이터베이스 定義 및 기존 데이터베이스의 수정), ISISUTL (여러가지 시스템의 유틸리티 기능), ISISXCH(다른 시스템과의 데이터 교환기능과 마스터 파일 유틸리티 기능), ISISPAS (CDS/ISIS 응용 프로그램 개발)로 구분된다. 〈圖 1〉은 CDS/ISIS의 메인 메뉴를 디스플레이 한 것이다.

2. ISISDEF & ISISENT

ISISDEF를 이용해서 각각의 필드를 定義(서브 필드나 반복필드도 가능)하

며, 기존의 종이 형태의 워크시트를 실제 화면(screen)에 display해서 데이터를 입력하게 한다. 필드 에디터로 모든 시스템의 메뉴나 워크시트를 이용자 임의대로, 이용자 위주로 만들수 있다. ISISENT를 써서 이미 定義한 데이터베이스에 실제로 데이터를 추가, 교정, 삭제하고 이 모든 데이터는 마스터 파일에 저장된다.

3. ISISINV & ISISRET

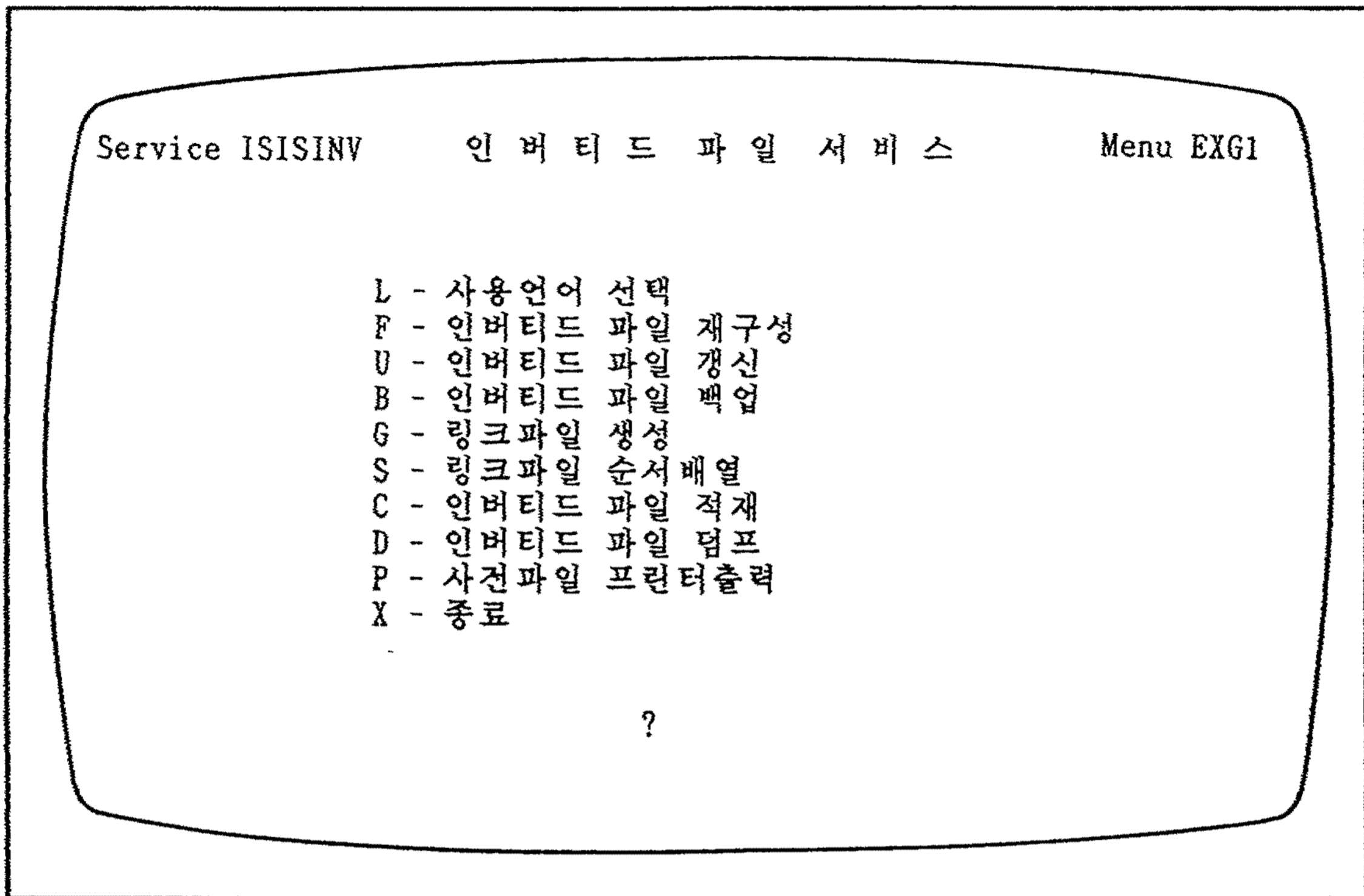
모든 情報는 결국은 情報檢索을 하기 위함이다. 즉, 축적된 情報 가운데서 어떤 특정한 속성을 가진 情報를 선택하는 것이다. 속성이라는 것은 情報의 집합에서 그 부분적 집합을 도출할 수 있는 것으로, 일반적으로 흔히 사용되는 속성은 저자명, 표제(서명), 정보내용을 표현하는 어그, 기호(키워드, 디스크립터, 주제명표목, 분류기호 등)들로서 색인어, 검색어, 액세스 키 등으로 불리어 진다. 이 ISISINV 서비스는 색인어 파일을 만드는 것이다.

정보검색에 있어서 검색된 문헌정보가 항상 정보요구를 충족시키느냐는 재현율과 정도율로 평가할 수 있다. 불필요한 문헌이 검색되는 검색 잡음과, 반대로 검색되어야 할 情報가 검색되지 않는 검색누락으로 검색의 효율을 평가함을 의미한다.

재현율(再現率 : recall ratio)은 검색 누락에 관한 척도로서, 검색된 적합정도가 시스템 전체의 적합정보 가운데서 차지하는 비율이고, 정도율(精度率 : precision ratio)은 검색잡음에 관련되는 것으로서 검색된 情報 가운데서 차지하는 적합정보의 비율을 나타내는 척도이다. 재현율과 정도율은 대립적 관계를 가지고 있어서, 단순히 한쪽을 높이고자 하면 다른 한쪽의 수치는 내려가게 된다. 재현율과 정도율은 카드목록, 색인지, 오프라인 검색에 있어서는 매우 중요한 평가 척도가 되기도 하지만, 온라인 검색에 있어서는 그 중요성이 오프라인 검색에 비해 낮으며, 시스템의 응답시간 속도가 더욱 중요한 평가 기준으로 되어 있다. CDS/ISIS에선 색인어에 해당하는 마스타 파일의 위치 포스팅 값을 가진 색인 파일(inverted file)을 만드는데, 검색속도의 향상 때문에 B-Tree 파일의 구성을 가진다. 마스터 파일의 레코드 수가 많더라도 파일을 만드는 속도에는 별다른 영향이 없다. <圖 2>는 ISISINV 서비스의 메뉴를 디스플레이한 것이다.

〈圖 2〉

ISISINV 서비스의 메뉴



ISISINV 서비스를 이용하면 우선, 각 마스타 파일의 레코드에서 FST(Field Selection Table)에서 定義한대로 키워드를 추출하며, 그 키워드에 실제적인 그 단어의 위치, occurrence 등의 情報를 함께 가지고 파일을 구성한다. 추출된 키워드는 단어의 길이가 10자보다 길고 짧음을 기준으로 file(Lφ1, Nφ1 : 길이가 짧은 용어 / Lφ2, Nφ2 : 긴 용어)을 두가지로 분류하여 순서배열(sort)한다. 다시 말하면 인버티드 파일이라고 불리는 특별한 파일을 만들어서 각 레코드에 대해 가상적으로 무한대의 접근점(access point)을 제공한다. 인버티드 파일은 주어진 데이터베이스를 검색하는 동안 접근점으로 사용되는 모든 용어들과 그들 용어들이 추출된 마스타 파일 레코드를 참조한 리스트를 포함하고 있다. 주어진 데이터베이스의 모든 접근점의 집합을 사전(dictionary)이라고 부른다. 즉 인버티드 파일은 마스타 파일의 내용에 대한 색인을 의미한다.

각각의 용어는 일반적으로 각 용어로 색인된 다른 레코드 번호(MFN : Master File Number)를 가지고 있기 때문에 인버티드 파일의 논리적인 길이는 가변적이다. 여기에서 다시 각 접근점의 빠른 검색을 위하여 실제적으로 인버티드 파일은 여러개의 물리적인 파일로 구성된다. 실제적으로는 확장자가 .CNT, .Lφ1, .Lφ2, .Nφ1, .Nφ2, .IFP 등의 여러 파일로 검색을 가능케 해 준다.

각각의 검색가능한 필드에 대하여 인버티드 파일이 분리되어 있는 검색시스템의 인버티드 파일과는 달리 ISIS는 어떤 주어진 데이터베이스에 대해서는 하나의 인버티드 파일을 사용한다. 그러나 이 파일의 특별한 구조 때문에 여러개의 인버티드 파일로 접근하는것과 기능적으로 동일하다. 사실 각 포스팅은 MFN(Master File Number)뿐 아니라 데이터가 어느 필드에서 추출되었는지 정확하게 구별하는 情報와 필드내 단어의 상대적 위치를 포함한다.

이와같이 ISISINV를 이용해서 만든 파일로 ISISRET 기능을 서비스 받는다. Micro ISIS 검색언어는 논리적 연산자들을 표현하는데 편리한 기능을 제공하는 불리언대수(Boolean Algebra)를 기초로 한다. 주어진 레코드와 관련된 각 탐색어는 사실상 용어와 관계있는 모든 레코드를 표현하는것 처럼 볼 수 있다. 따라서 탐색어 사이의 논리적인 연산자의 표현에 의해 요구의 응답이 검색되어지는 레코드를 명확하게 定義할 수 있다.

나아가 옵션 'T'를 이용해 화면에 실제적인 그 해당 데이터베이스의 용어사전을 직접 화면으로 보고 원하는 탐색어를 선택한다. 또한 ANY file(용어에 관한 집합을 定義한 파일 : 확장자 .ANY)와 불용어 파일(확장자 .STW)을 사용한 검색도 물론 가능하다. <圖 3>은 인버티드파일 생성 중 링크파일 및 인버티드파일 실행레이고, <圖 4>는 ISISRET 검색레이, <圖 5>는 Term Dichionang 검색레이이다.

<圖 3.a> 인버티드 파일 생성중 링크파일 생성 실행레이

```

Data Base : CDS
└ MFN limits?

      ID   IT   Format      ** Field Select Table **
      70   0   MHU, (V70/)
      24   4   MHU, V24
      69   2   V69
Unsorted Link file generation
Short terms -   1335
Long terms  -   1050

Unsorted Link file generation completed --->
    
```

4. 기 타

ISISPRT기능은 이용자가 제공된 출력형식 언어(print formatting language)를 사용하여 이용자 임의대로 원하는 형태를 정한다. 예를 들면, 1,2,3단 편집, 또는 필드별 순서배열 등의 여러 종류의 출력이 용이하다.

ISISXCH의 서비스는 ISO 2709형태로 되어있는 외부 파일을 ISIS내에서 이용할수 있도록 파일을 Import하고, 또 ISIS내의 유효한 데이터베이스를 외부로 옮기는 파일 Export 기능을 가능하게 한다. 이 서비스를 사용하여 개인용 PC를 이용하는데, 호스트나 타 기종 컴퓨터의 데이터를 데이터 베이스로 이용할 수 있어서 더욱 효율적이다.

이상에서 언급되지 않은 ISIS의 기능도 많이 있지만, 지면 관계상 간략히 주요기능과 특징만을 서술하였다.

V. 適用事例

여러 장점을 많이 가진 ISIS의 이용이 많아지면서 이용자 측면의 새로운 요구가 많이 도출되고 있다.

1. Wordstar와 CDS/ISIS와의 Interfacing

데이터의 입출력시 워드프로세서인 Wordstar와 Micro ISIS와의 Interface를 통해 ISIS의 기능을 향상시키고자 Nepal의 Kathmandu의 ICIMOD(International Centre for Integrated Mountain Development)에서 수행했다.

Interfacing을 하여 철자 체크(spelling check)와 출력강화(print enhancement)의 기능이 더해진다. 기존의 데이터베이스의 새로운 입력이나 수정시 철자의 체크를 매뉴얼로 해왔는데 꽤 많은 시간이 소요되었다. 그리하여 대부분의 이용자들의 요구가 있으므로 WordStar 2000에 interface하는 방법을 이용하게 된 것이다.

텍스트 에디터에 의해 만들어진 파일을 생성하기 위해서 CDS/ISIS의 report generator program인 ISISPRT를 사용한다. 그러나 텍스트를 교정하는 등의 manual handling 때문에 error는 항상 나타나기 마련이다. 이것은 기계

적인 interface에 의해 행하게 하였고, 이는 가능한한 텍스트 에디터의 기능을 최대화하고자 함이다.

왜냐하면 Micro ISIS에 의해 일반화된 리포트들은 워드프로세서에서 읽혀질 수 있고 이용자가 원하는 방법으로 파일을 처리할 수 있다. 그리고 모든 텍스트 에디터에서는 텍스트 기능을 변환하는것이 가능하기 때문이다. Micro ISIS 출력형태언어(print formatting language)는 어떠한 문자나 상수의 출력을 가능하게 하고 이용자의 요구에 따라서 상수는 필드의 앞이나 뒤에 출력할 수 있다. 이러한 각각의 프로그램을 interface한 것이다.

철자체크의 기능을 보면 데이터베이스에서 데이터를 입력함에 있어 title index와 author index, keyword index를 프린트한다. 그리고 나서 이 파일들의 hard copies분은 WordStar 2000에서의 정상 파일로 import된다. title index는 WordStar내의 주요 또는 내부사전에 의해 철자가 체크되고, author index와 keyword index는 WordStar안에서 제공되는 개인사전 설치 기능을 사용시의 두 개의 사전을 참조하여 체크된다. 체크된 에러는 마크되어 프린트된다. 한글 데이터의 입력을 위한 완벽한 한글 ISIS의 개발뿐 아니라 완벽한 한글, 한자 철자 체크가 가능한 프로그램 개발이 요구되는 바다.

프린트 강화는 마치 볼드체의 문자출력이나 underline과 같은 특수한 프린트 효과를 가능케 함이다. 한 페이지의 많은 데이터 요소의 정확한 배치기능도 강화시켰다. 이 방법은 1987년 ICIMOD가 New Acquisitions in Library를 격월로 출판하면서 시작했다. 각 데이터 요소의 블록화 및 중앙위치(centerixed)가 되고, 볼드체 및 이탤릭체등이 출력된다.

ISIS의 이용이 더 보편화되면 이와같은 요구도 대두될 것임이 틀림없으므로 한글 워드프로세서를 이용하는 등의 연구가 필요하리라 생각된다.

〈圖 6〉 프린트 강화를 위한 출력형식 언어의 예

- [CENTER] 그 라인의 중앙에 위치
- [B] 볼드체
- [I] 이탤릭체
- [U] 언더라인(밑줄)

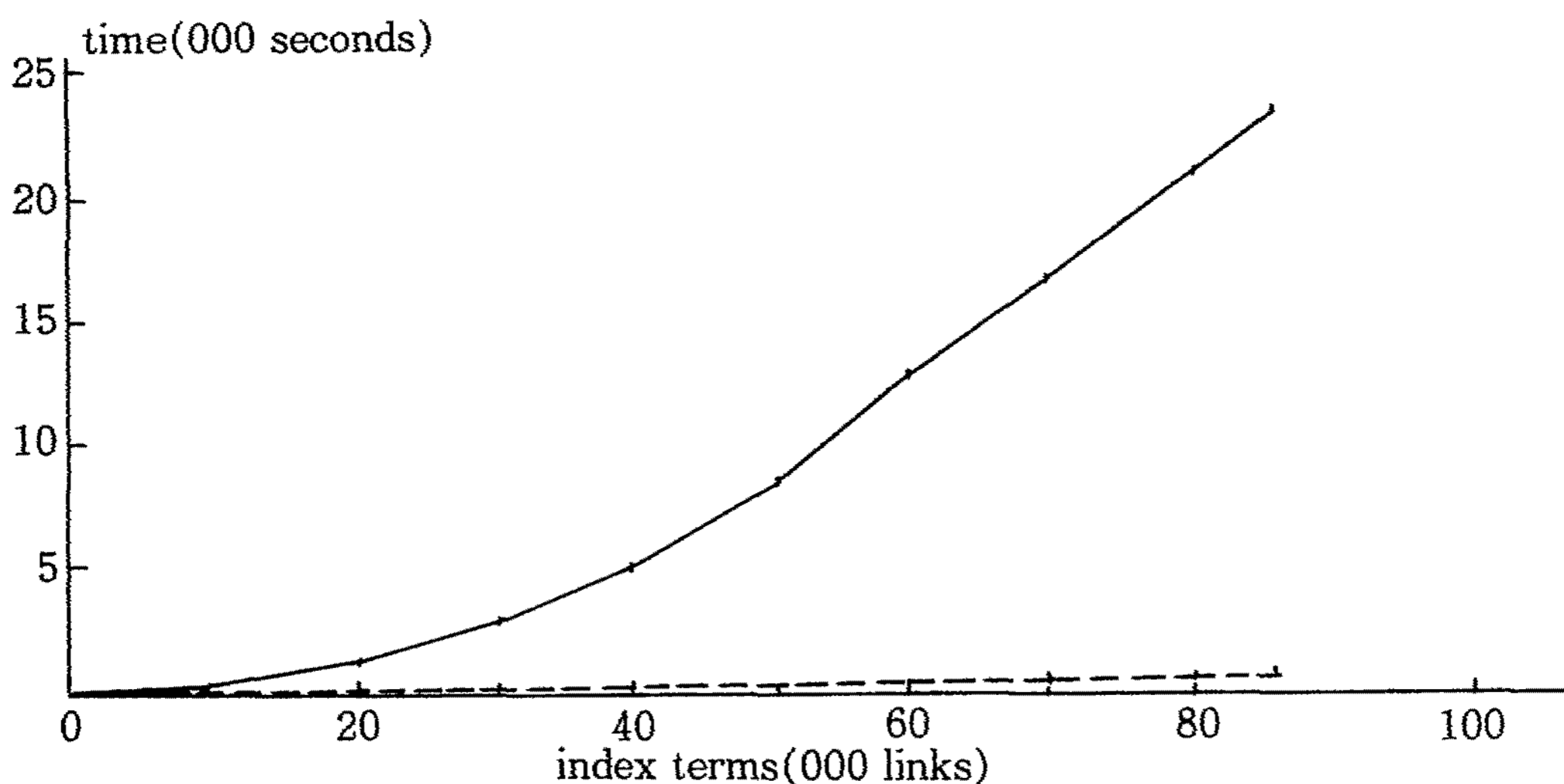
2. CDS/ISIS에서의 파일 生成 시의 Fastsort 이용

ISISINV의 서비스가 인버티드 파일을 生成시킨다. 인버티드 파일은 용어의 인덱스와 각각 데이터베이스 안에서의 주소 코드를 가진다. ISIS의 ISISINV 서비스는 ①순서배열 되지않은 링크파일(.LNU)을 生成하고 ②링크파일을 순서배열(sort)한 후 ③인버티드 파일을 로드하여 비로소 검색 가능하게 한다. 이 세가지 과정 중 순서배열(sort)에 가장 많은 시간이 소요된다. 많은 양의 데이터베이스일수록 더 많은 시간이 걸린다. 이에 Fastsort를 개발하여 보다 더 빨리 full inverted file을 生成한다. Fastsort를 이용한 결과를 다음의 <表 1>에서 비교했다.

이 데이터는 속도 16Mhz의 IBM PS/2 컴퓨터를 이용했다. Fastsort Program을 데이터베이스 파일과 같은 디렉토리에 복사하여 사용한다. 우선 ISIS에 의해 .LNU가 生成되면 Fastsort가 그 파일의 각 문자의 ASCII값을 순서배열한다. 여러 옵션으로 실행시킨다.

<表 1>은 CDS/ISIS sort algorithm과 Fastsort를 이용했을때의 시간 비교표이고, <圖 7>은 비교 그래프이다. 즉 fastsort를 이용하면 ①빠른 순서배열, ②디스크 스페이스의 오버헤드의 요구를 줄이고, ③요구되는 프리 디스크 스페이스를 확인할 수 있고, ④기존에 불용어 리스트 파일의 크기의 제한이 없어지며 또한, ⑤다중 인덱스 테크닉(MIT : Multiple Indexing Technique)이 가능해진다.

<圖 7> 속도의 비교 그래프



〈表 1〉 Inverted file 生成時 ISIS와 Fastsort 이용시의 시간 비교

레코드 수	Index의 수	시 간(초)		
		ISIS	Fastsort	차 이
50	1,220	11	10	1.10
100	2,400	26	15	1.73
200	4,727	96	30	3.20
300	6,970	177	39	4.54
400	9,434	346	56	6.18
500	11,971	465	69	6.74
750	18,288	1,140	111	10.27
1,000	24,623	1,860	154	12.08
1,500	38,399	4,440	250	17.76
2,000	50,520	8,100	390	20.77
2,500	63,328	14,520	506	28.69
2,961	75,083	18,240	608	30.00
3,411	84,622	23,100	735	31.43

VI. 期待效果 및 결론

지금까지 ISIS를 이용하여 情報를 처리하는 전체적인 흐름을 논했고, 나아가 앞으로 계속 활발한 이용이 기대된다.

여러방면으로 이용자의 요구를 분석, 연구하여 만들어진 CDS/ISIS시스템은 다국어 지원체제의 일환으로 여러 나라의 언어로 데이터의 입력, 수정, 검색을 가능하게 했다. 이에 동양권 내의 2바이트 체제라는 문자의 특이성과 문법상의 복잡성을 모두 고려한 완벽한 한글 버전의 개발이 시급히 요구되는 바이다. 나아가 기존의 ISIS를 이용하면서 타 프로그램이나, 소프트웨어와의 interfacing 측면도 고려해 볼만하다. 보다 나은 情報를 이용하고자, 情報를 처리함에 있어서 기존 시스템의 효율성을 향상시키기 위한 지속적인 노력이 경주되어야 할 것이다.

〈參 考 文 獻〉

- 김두홍, 우길호, 정영재, 채옥동, 「도서관정보학개론」, 서울:아세아출판사, 1988
- 사공철, 김태수 「정보관리론」, 서울:구미무역, 1989

- 「정보관리」, 서울:산업연구원, 1989
- 정영미, 「정보검색론」, 서울:정음사, 1986
- 사공철, 「정보검색론」, 서울:아세아출판사, 1977
- 이석호, 「데이타베이스론」, 서울:정익사, 1985
- *UNESCO Mini-micro CDS/ISIS Reference Manual*, 1989
- K.Shailendra, R.N.Neogi, K.L.Gogia, "Use of CDS/ISIS for creating India's database of serials", *Program*, vol. 23 no.3, July 1989, pp. 287~294
- Anwar Ali Chaudhry, 'Interfacing CDS/ISIS with WordStar:ICIMOD's experience' *Program* vol.23 no.1, January 1989, pp. 65~72
- Anwar Ali Chaudhry 'File inversion in CDS/ISIS:Fastsort helps to do it faster' *Program*, vol.23 no.3, July 1989, pp. 295~305