

E.D.P.S 한글화에 관한 연구

연구책임자 : 성기수 (전자계산실 실장)
연구원 : 이덕순 (콘트롤 데이터 코리아 S.E.)
유완영 (전자계산실 프로그래머)
백인섭 (전자계산실 S.E.)
변희성 (전자계산실 //)

차 례

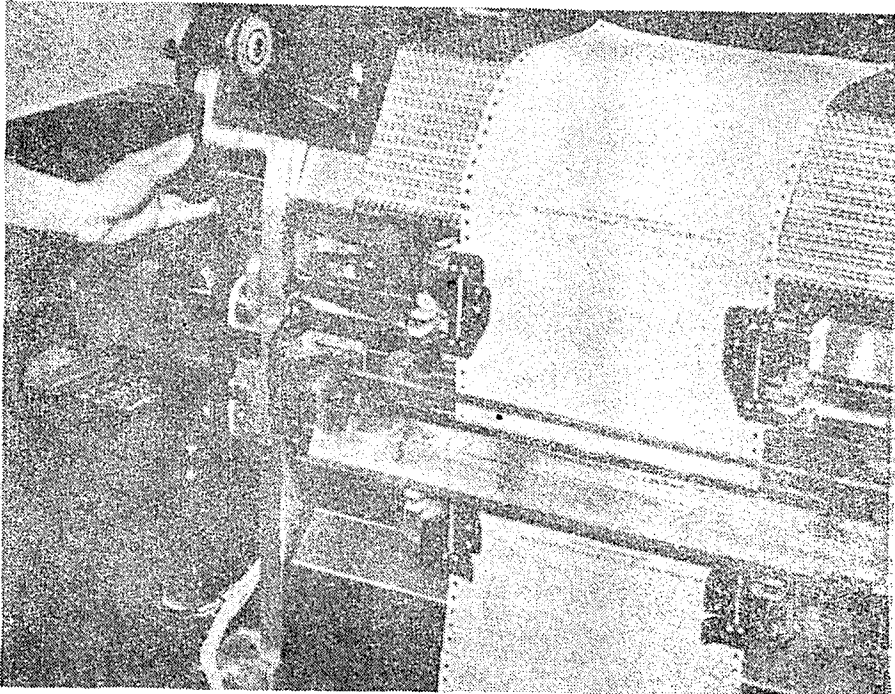
I. 서 론

1. 연구의 목적
2. 연구의 경위
3. 연구의 범위

II. 본 론

1. EDPS 한글화의 필요성
2. 한글의 구조

3. 한글 입력 시스템
4. 한글 출력 시스템
5. 한글-영문 혼용 2 줄모아쓰기의 구성
6. 한글 전용 3 줄모아쓰기의 구성
7. EDPS 한글화의 나아갈 길
8. 연구결론



한글 LINE PRINTER

미 리 마

최근 국가 경제의 급속한 발전과 더불어 지방기관, 금융기관
 및 각종 산업계에서 전자계산장비(EDPS)의 도입으로 표출하는
 업무의 기계화하여 국내의 시업, 경제, 학술 등 모든 분야의
 발전에 크게 공헌할 수 있는 중요한 일이다. 그러나 국내에서
 개발된 전자계산장비 업무의 위주로 개발되어 있어 하급
 언어로 사용하는 우리나라 사람에게는 현대문명의 총아인
 전자계산장비의 생활화하여 커다란 장애가 되어 왔다. 본
 전자계산장비 이러한 국내의 필요성에 따라 하급언어기
 국내 자체 연구사업의 하나로써 EDPS 한글화에 관한 1차 연구
 과제로 CDC 512 LIVE-PRINTER에 새로운 한글화자 제어
 시스템을 CDC 3300 MASTER 시스템에 한글자동모아보기 프로그램
 개발시켜 컴퓨터 한글 데이터를 읽어 기억할 수 있는 프로그램
 수 있는 한글 입력 시스템을 개발하여 EDPS 한글화사업의
 기초를 마련하였다. 이 연구사업이 전자계산장비의 생활화
 에 크게 공헌할 수 있도록 하며 본 연구사업의 중요성을 인식하여
 연구 수행을 후원하여 주신 연구소 간계자 어리부, 하급
 프로그래머 개발에 공로로 참여하여 주신 본 연구소 데이터 처리
 장위원회(CDK) 어리부, 그리고 많은 조언 및 자료를 제공하여
 주신 하급언어기 어리부께 감사 드리며 보다 완전한 EDPS
 한글화가 실현될 때까지 계속적인 연구를 바라는 바이다.

I. 서 론

1. 연구의 목적

본 연구의 목적은 EDPS (Electronic Data Processing System) 한글화 사업의 하나로 컴퓨터의 한글 입·출력 시스템 (Input-Output System) 을 연구하고 한글 Line Printer 를 개발하여 실용화함에 그 목적이 있다.

2. 연구의 경위

본 연구는 한국과학기술연구소 자체연구사업의 하나로 1969년 7월부터 1970년 6월까지 한글 입출력 시스템을 개발하고자 하였다. 그러나 미국 CDC 사에 발주한 한글 활자 체인의 도착이 지연되어 7개월간 연구기간을 연장하여 한글 입출력시스템에 관한 연구를 끝내게 된 것이다.

3. 연구의 범위

본 연구의 범위는 EDPS 한글화 사업의 가장 중요한 부분인 한글 입·출력 시스템을 최단시일 내에 실용 가능한 단계까지 개발하는 것이다. 따라서 현행 입·출력 시스템의 변경을 최소로 하도록 노력한 결과 한글의 입력 (Input) 방법으로 한글전용 입력장치를 신규 개발 또는 제조를 피하고 기존의 입력수단인 IBM-29 KEY PUNCH MACHINE 의 KEY BOARD 를 정부 제정 표준 자판 2벌식을 기준으로 재배열하여 한글을 입력 시키도록 하였다.

한글의 출력 (Output) 방법도 새로운 한글전용 Line Printer 를 제작하지 않고 기존의 체인 (Chain) 형 프린터에 한글활자 체인을 주조하여 삼입함으로써 풀어쓰기로 입력된 한글을 자동적으로 모아쓰기로 프린트할 수 있는 한글 프로그램과 함께 한글 입·출력 시스템을 형성하도록 하였다.

연구기간중 자동 모아쓰기 프로그램과 한글 자판의 설계가 끝나 한글체인의 주조를 CDC 에 발주하였고 70년 11월에 한글-영문 혼용프린터가 가동하기 시작하였다. 71년 8월에 한글 전용 체인이 도착하면 Line Printer 에 프린트된 수

자(0~9)를 OCR (Optical Character Recognition)을 거쳐 다시 컴퓨터에 입력시킬 수 있는 한글입출력의 Turnaround 시스템이 완성될 것이다.

짧은 연구기관중 자료의 불충분·인원교체, 한글 체인 제작의 지연 등으로 부족한 점이 많았고 특히 활자체의 개량 및 Line Printer 의 속도를 높이기 위한 활자 배열 등은 계속하여 보완되어야 할 문제들이다.

II. 본 론

1. EDPS 한글화의 필요성

1967년 5월에 컴퓨터가 한국에 처음 도입된 이래 국내 정보산업은 새로운 차원에 도달하였다. 즉, 산업의 발전과 더불어 기업의 대형화, 기술의 고도화 등은 각 기업체로 하여금 경영합리화라는 문제를 대두시켰고 또한 날이 갈수록 증대일로에 있는 업무량을 처리하기 위하여 컴퓨터라는 새로운 도구가 없이는 감당하기 어려운 단계에 이르러 컴퓨터를 어떻게 효율적으로 운영하느냐가 가장 중요한 연구문제로 등장하였다.

한국의 컴퓨터 이용도 이제 여명기를 벗어나 본격적인 단계에 이르렀다. 즉, 70년도말 국내에 도입된 컴퓨터는 18대로서 11대 (대형 1, 중형 3, 소형 7)가 이미 가동중이고 각기관 및 기업체에서 도입 예정중인 것까지 고려하면 71년 말에는 30대에 이를 것이며 계속하여 기하급수적으로 증가할 것으로 예상된다.

컴퓨터의 지속적인 도입과 더불어 지속적인 기술의 발전도 이루어져 69년 9월 한국과학기술연구소 전자계산실의 CDC 3300이 가동하여 처음으로 고급언어인 COBOL 및 FORTRAN Compiler 를 사용하기 시작하였고 71년 6월에는 국내 유일한 Multiprogramming 시스템인 CDC 3300 MASTER 시스템이 가동하였다. 70년 10월에는 경제기획원 예산국에 전용전화선으로 연결된 CDC 200 Users Terminal 이 설치되어 성공적으로 가동됨으로써 국내 최초로 데이터 통신이 시작되었다.

이와같이 국내에서 컴퓨터의 이용도가 증가함에 따라 컴퓨터로 인쇄된 각종 보고서가 생활화하게 되자 커다란 문제가 발생하였다. 즉, 컴퓨터는 미국을 중심으로 개발되어 모든 프로그램 및 데이터에 영문을 사용하여야 하므로 한글을 국어로 사용하는 한국사람이 컴퓨터를 100% 활용하기에는 먼 거리감이 생겼다. 즉, 컴퓨터의 입력방법은 프로그램 및 데이터를 Punch Card, Paper Tape 에 수록하여 Card Reader 또는 Paper Tape Reader 를 통하여 입력시키게 되어있고 최근에는 OCR (광학문자 해독기 : Optical Character Recognition)을 통하여 약속된 형태의 글자체로 된 서류를 직접 입력시킬 수 있으나 이러한 모든 방법은 컴퓨터 내부에서 처리된 프로그램 및 데이터를 최종적으로 Line Printer^F에 영문으로 프린트하게 되어 있어 과학계산에는 크게 문제될 것이 없다 하더라도 사무용 자료 처리에 응용시에는 누구나 한글로 된 보고서를 볼 수 있는 것이 가장 바람직한 일이다. 예를 들면 과세업무를 EDPS화 할 경우 최종문서인 납세고지서를 한글로 프린트하여 납세자가 누구나 손쉽게 식별하여야 함은 당연한 일이다. 이것은 하나의 예에 불과하고 우리들의 일상 통신의 매개체가 한글인 이상 전자계산기가 읽고 쓰는 모든내용이 한글이어야 함은 당연하고도 시급한 일이라 하겠다.

이와같이 전자계산조작을 한글화하여 컴퓨터 언어의 제한을 피할 수 있다면 가장 바람직한 일 이겠으나 한국 컴퓨터의 역사를 살펴볼 때 힘겨운 일이라 아니할 수 없다. 즉, 6만대 이상의 컴퓨터를 운영하는 미국과 같이 기술의 축적 및 경제적 지원이 이루어져 컴퓨터가 완전 실용화 될 때 가능하다 하겠다. 본 연구에서는 전 시스템을 한글화 하는 것은 어렵다 하더라도 가장 중요한 부분인 입출력 시스템을 한글화하여 우선적으로 한글로 입력된 데이터를 읽어 한글로 프린트할 수 있는 한글 입출력 시스템의 개발문제를 검토하였다.

2. 한글의 구조

우리말을 표시하는 데 사용되는 한글은 기본

적으로 14 개의 자음과 10 개의 모음의 조합으로 표시된다. 10 개의 모음 이외에 “ㄱ” 및 “ㄴ”은 복합모음이지만 사실상 단모음으로 발음되므로 단모음으로 취급하여 모음의 수를 12 개로 구분할 수도 있다(표 1 참조).

표 1. 한글자모의 구성

자	음	모	음	자	음	모	음
ㄱ		ㅏ		ㅈ		ㅡ	
ㄴ		ㅑ		ㅊ		ㅣ	
ㄷ		ㅓ		ㅋ		(개)	
ㄹ		ㅕ		ㅌ		(해)	
ㅁ		ㅗ		ㅍ			
ㅂ		ㅛ		ㅎ			
ㅅ		ㅜ					
ㅇ		ㅠ					
				14개		10개	(12개)

한글의 구성은 표 1의 24 개(26 개)의 자음·모음을 엘코도적(Ergodic)인 구성에 따라 무수하게 많은 것이 아니고 몇가지 원칙에 따라 자음 4 개, 모음 3 개, 합하여 7 개까지의 자모의 조합으로 이루어진다.

한글을 구성하는 자모의 구성원칙은 아래와 같다.

- (1) 한글의 첫자는 반드시 자음으로 구성되고

표 2. 한글 자모의 구성방법

CV	CV	CV	CVV	CVV	CVV
	C	CC		C	CC
CCV	CCV	CCV	CCVV	CCVV	CCVV
	C	CC		C	CC
CCV	CCV	CCV	CCVV	CCVV	CCVV
	V	V	V	V	V
		C	CC	C	CC
CV	CV	CV	CVV	CVV	CVV
V	V	V	V	V	V
	C	CC		C	CC
C	C	C			
V	V	V			
	C	CC			
CC	CC	CC			
V	V	V			
	C	CC			

C ; 자음
V ; 모음
계 30가지

표 3. 문교부 조사에 의한 한글 자모 사용빈도표

순 위	자 모		중 합		모 음		자 음		총 합		전 자 음 (초성)		후 자 음 (종성)			
	자모 빈도	1000개 당빈도	%	모음 빈도	1000개 당빈도	%	자음 빈도	1000개 당빈도	%	결 자음	빈도	1000개 당빈도	후 자음	빈도	1000개 당빈도	%
1	58,486	119.1	11.91	가	49,429	243.5	24.35	ㅇ	58,486	203	20.3	48,581	ㄴ	26,251	294.2	29.42
2	49,429	100.7	10.56	나	36,567	180.2	18.75	ㄴ	40,925	142.2	16.37	28,752	ㄷ	18,755	210.2	20.82
3	40,925	83.35	9.45	다	28,092	138.3	14.4	ㄷ	37,129	129	13.22	20,521	ㄹ	11,632	130.4	14.75
4	37,129	75.60	8.29	라	24,686	121.5	10.9	ㄹ	33,225	115.4	10.68	16,874	ㅁ	9,905	111.0	10.48
5	36,567	74.45	7.35	ㄴ	21,534	106.1	8.4	ㄴ	28,506	98.9	8.62	15,062	ㄷ	8,377	93.9	7.40
6	33,225	67.65	6.515	ㄷ	14,508	71.4	6.4	ㄷ	20,991	72.8	6.96	14,674	ㄹ	6,303	70.6	5.24
7	28,506	58.02	5.751	ㄹ	8,676	42.7	4.9	ㄹ	16,411	57	5.62	14,470	ㅂ	3,448	38.6	3.71
8	28,092	57.20	5.092	ㄴ	8,176	40.3	3.8	ㄴ	15,493	53.7	4.53	12,729	ㅅ	1,904	21.3	2.63
9	24,686	50.25	4.525	ㄷ	7,215	35.5	2.9	ㄷ	14,633	50.8	3.65	10,108	ㅇ	1,036	11.6	1.86
10	21,534	43.88	3.995	ㄹ	1,850	9.1	2.2	ㄹ	1,351	44.6	2.95	9,403	ㅈ	470	5.3	1.32
11	20,991	42.75	3.549	ㅂ	1,643	8.1	1.7	ㅂ	3,716	12.9	2.38	3,433	ㅊ	431	4.9	0.93
12	16,411	33.40	3.136	ㅅ	706	3.3	1.3	ㅅ	2,640	9.2	1.92	1,940	ㅌ	414	4.6	0.65
13	15,493	31.55	2.777	ㅈ			1.3	ㅈ	2,354	8.2	1.55	1,604	ㅍ	283	3.4	0.46
14	14,633	29.82	2.458	ㅊ				ㅊ	684	2.3	1.25	684	ㅑ	0	0	0.33
15	14,508	29.55	2.175	ㅌ				ㅌ								
16	12,851	26.10	1.928	ㅍ				ㅍ								
17	8,676	17.60	1.710	ㅑ				ㅑ								
18	8,176	16.60	1.518	ㅓ				ㅓ								
19	7,215	14.90	1.346	ㅕ				ㅕ								
20	3,716	7.60	1.187	ㅗ				ㅗ								
21	2,640	5.40	1.051	ㅛ				ㅛ								
22	2,354	4.80	0.931	ㅜ				ㅜ								
23	1,850	3.80	0.825	ㅠ				ㅠ								
24	1,643	3.30	0.731	ㅡ				ㅡ								
25	706	1.44	0.648	ㅣ				ㅣ								
26	684	1.39	0.574	ㅍ				ㅍ								
합계	491,126	1,000	100%		203,082	1,000	100%		288,044	1,000	100%	198,835		89,209	1,000	100%

자음+모음의-구성이 뒤따른다.

(2) 한글 1 자를 구성하기 위한 자모의 수는 최소로 자음 1 개, 모음 1 개이고, 최대로 자음 4 개 모음 3 개이다.

위의 두가지 원칙에 따라 한글의 자모로 구성할 수 있는 방법은 모두 30 가지이다(표 2참조).

표 2의 구성에 따른 한글을 컴퓨터에 입력 또는 출력시키는 여러가지 문제가 뒤따른다. 입력 문제는 풀어쓰기 또는 모아쓰기로 입력된 한글을 어떻게 Code 화 하느냐가 문제되고 출력문제에 있어서는 한글 자모의 배열, Line Printer Character Set 의 선정 및 Slug Position 의 배열, 한글자모의 design 등이 문제된다. 한글 Line Printer Character Set 선정에 참고로 한 문교부가 조사한 한글자모의 사용빈도 표는 표 3 과 같다.

II 장의 나머지 부분은 한글의 입출력 문제를 차례로 상세히 기술하고 한글 Line Printer 의 사용법 및 EDPS 한글화의 나아갈 길을 제시하였다.

3. 한글 입력 시스템

최근 컴퓨터의 입력매체(Input Media)는 놀라운 발전을 이루어 PCS (Punched Card System) 으로부터 paper Tape, OCR (Optical Character Recognition), MICR (MagneticInk Character Recognition) 나아가서는 사람의 음성을 컴퓨터에 입력시킬 수 있는 방안까지 연구되고 있다.

한글의 입력 방법으로는 여러가지 방법중 우선 2 가지 방법, 즉, Punch Card 에 의한 입력과 OCR 에 의한 입력을 채택할 수 있다.

Punch Card 에 의한 입력은 표준화된 방법으로 Hollerith Code 를 80 행(80 col) 카아드에 천공하여 6 단위 BCD (Binary Coded Decimal) 로 전환하여 컴퓨터에 입력시킬 수 있는 가장 안정된 방법이다. 즉, Punch Card 1 행(1Column) 의 1~12 Position 에 수자는 1 개, 한글의 자모는 2 개, 특수문자는 3 개까지의 구멍을 천공하여 Card Reader 의 광전지(Photocell) 를 통한 후 Internal BCD Code 로 전환 시킬 수 있다(표 4 참조).

Punch Card 에 의한 입력 문제는 현재 국내에

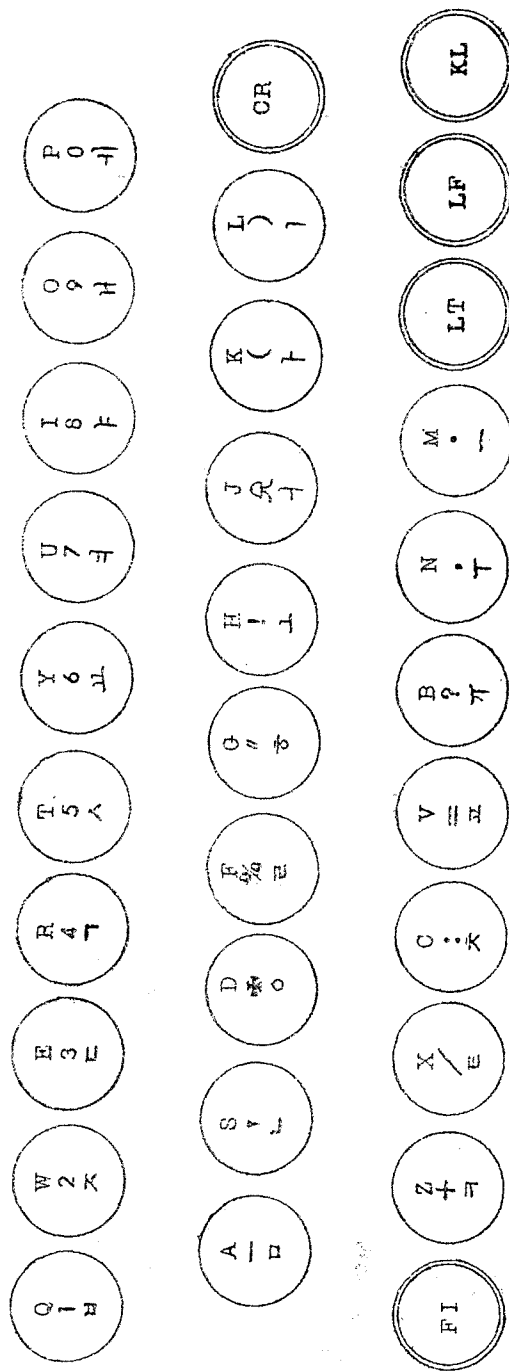


그림 1. 정부제정 표준자판 (인쇄전신기용 2벌)

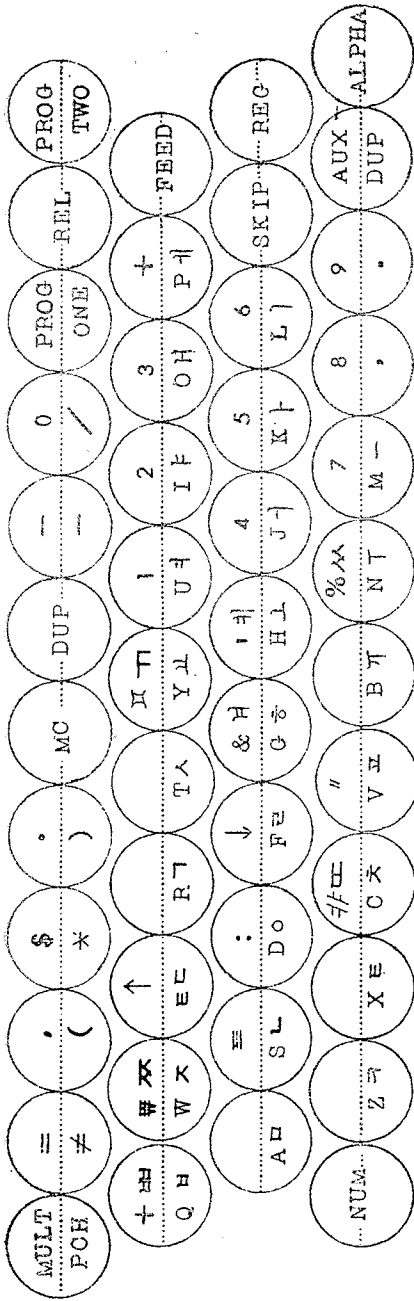


그림 2. 한글 영문 혼용 PUNCH KEY BOARD (정부제정 표준자판 2벌식 기준) IBM 029

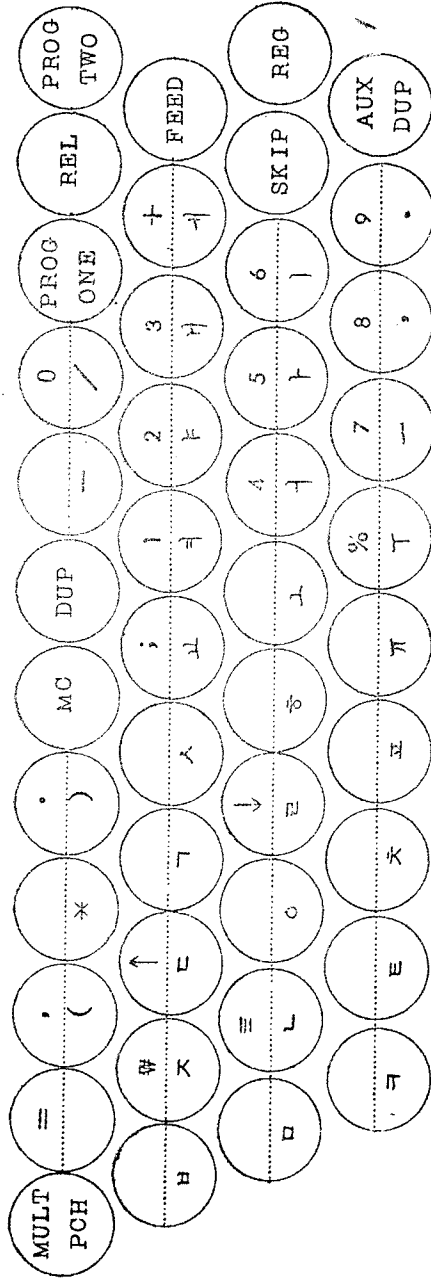


그림 3. 한글 영문 혼용 PUNCH KEY BOARD (정부제정 표준자판 2벌식 기준)

보급되어 있는 IBM-29 Punch Machine 을 효과적으로 이용한다면 입력매체 (Input Media) 를 새로이 제조 또는 개조하지않고 해결할 수 있다. 즉, 한글과 영문을 구분하는 특수문자(예를 들면 ₩≡C≠)를 한글의 전후에 배열하여 영문의 R,E,S.....등을 한글의 ㄱ, ㄴ, ㄷ.....등으로 손쉽게 전환시킬 수 있으므로 한글을 식별할 수 있는 특수문자와 함께 BCD Code System (2⁸=64 개 까지가능)을 그대로 사용한다면 한글의 자모를 각각 1 행에 천공하여 풀어쓰기로 한글을 입력시킬 수 있다.

OCR 를 입력매체로 선정할 때에는 여러가지 기술적인 문제가 뒤따른다. 한글의 복잡성에 따라 Hand Printing 또는 Font 로 된 한글을 입력시킬 수 있는 새로운 Optical Reader 의 설계가 가장 문제가 될 것이다. 본 연구에서는 우선 한글 Line Printer 활갈 활자 체인에 CDC SC 1700 컴퓨터와 연결된 CDC OCR-936 을 통해 재입력시킬 수 있는 USASOCR Size A Font (수자)를 삽입하여 컴퓨터 Output 를 회수하여 재입력시킬 수 있는 Turnaround 시스템을 채택하였다.

Punch Card 를 매체로 한글을 풀어쓰기로 입력시킬 경우 Punch Machine 의 자판의 구성 및 배열문제가 뒤따른다. 한글자판의 구성은 한글 영문 혼합 2 줄 모아쓰기 및 한글전용 3 줄 모아쓰기 등 각기 자모의 선전방법에 따라 93 자, 52 자(한글표시 특수문자 포함)로 각각 구성된다(그림 6, 7 참조). 한글자 판은 정부제정 표준자판 2 벌식을 기준으로 복자음 및 특수문자를 첨가하여 배열하였다(그림 1, 2, 3 참조).

4. 한글 출력 시스템

컴퓨터의 출력매체 (Output Media)는 Line Printer, Card Punch, Plotter, Microfilm Plotter, Tape, Disk 등 여러가지가 있으나 최종적으로는 Line Printer 의 이용이 대부분을 차지한다 하겠다. Line Printer 에 한글을 프린트하기 이전에 1969 년 5 월 미국 Battele Memorial Institute 에서 Computer Handling of Korean Language Texts(부록 참조)란 연구제목으로 Battele 연구소의 CDC 6400 컴퓨터 시스템과 Stromberg

Datagraphix 4060 Microfilm Plotter 에 한글자동모아쓰기 프로그램의 가능성이 증명되어(그림 4) Line Printer 에 한글을 프린트 할 수 있는 시스템 개발에 본격적으로 착수하게 된 것이다.

현재 각 Maker 에서 생산하고 있는 Line Printer 는 드럼 (Drum)형과 체인 (Chain)형의 2 가지로 구분할 수 있다. 드럼형 프린터는 드럼에 Character 가 일렬로 배열되어 있어 드럼이 일정한 속도로 회전하면 (분당 300~1000 회전) Cabor 뒤에 있는 Hammer 가 프린트하려는 글자가 정 위치에 왔을 때 충격을 주어 Print Sheet 에 글자가 적히도록 하는 것이고 체인형 프린터는 드럼형 프린터와는 달리 Character Set 가 고정되어 있지 않아 필요에 따라 새로운 Character Set 를 삽입할 수 있는 것이 특징이다. 한글 전용 프린터를 개발한다면 이에 따른 막대한 기술적·경제적 문제가 뒤따를 것이지만 체인형 프린터의 특성을 고려한다면 한글 프린터를 별도로 개발하지 않고 기존의 영문체인을 새로 구조한 한글 체인으로 대체시키고 이에 따른 소프트웨어 (Software)를 개발하여 한글 출력문제를 해결할 수 있는 것이다.

한글 모아쓰기 프로그램은 Subroutine 형식으로 하여 User 프로그램에서 Call 하거나 Line Printer Driver 프로그램을 일부 수정하여 한글 모아쓰기 프로그램과 함께 사용할 수 있으나 Subroutine 형식은 Compiler 마다 Calling Sequence 가 다른 등 사용법이 복잡해질 것이므로 Driver 프로그램을 일부 수정하여 MASTER 시스템과 한글 COMPASS 프로그램을 Interface 시켜 최종적으로 Driver 프로그램에서 한글을 Check 하도록 하였다.

한글용 체인에 한글 Character Set 를 구조하여 삽입하는 데는 자모의 수 및 상관도 이에 따른 프린트 속도 등을 고려하여야 한다. 한글 전용 512 Line Printer 의 특성 및 Character Set 수에 따른 프린트 속도는 아래와 같다(표 4, 5 참조).

한글 Chain 의 Character Set 는 한글 영문 혼용 2 줄 모아쓰기에서는 89 개, 한글전용 3 줄 모아쓰기에서는 48 개 (한글표시 특수문자(₩≡) 및

표 4. 한글 Line printer 의 물리적 특성

Cabinet name	512 Printer
Height	50 inches (1.27 meters)
Depth	28 inches (.710 meters)
Weight (not including controller or controller power supply)	1200 lbs. (544kg)
Maximum depth, including all access doors extended	70 inches (1.802meters)
Maximum width, including all access doors extended	88 1/4 inches (2.243 meters)
Heat (with integral controller) (without integral controller)	9370BTU/hr. (2360kg-cal/hr) 7670BTU/hr. (1930kg-cal-hr)
Line Current ※50/60 cycle, 208 volts, three phase	10 amperes per phase (maximum)
Breaker size, 50/60 cycle, 208 volts, three phase	20 amperes
Air required at inlet	400 CFM at 1 inch H ₂ O static pressure
Source of cooling air	Room
Power Cable Length	5 feet (1.53meters)

표 5. Character Set 에 따른 Print Speed

Character Set	Lines Per Minute (Nominal)
144	470
96	670
72	855
48	1150
36	1430
32	1500
24	1800
16	2400
Maximum	2500

blank 제외)로 구성되어 한글 1 자를 프린트 하려면 최대 3 번까지 프린트하여야 하므로 89 개 일 때는 250 줄, 48 개일 때는 400 줄이 프린트 가능하다. 한글 자모의 상관도에 따른 Character Set 내에서의 자모의 배열 및 글자형태를 보기 좋게 하기 위한 자체의 보완 등은 계속하여 연구되어야 할 문제들이다.

영문프린터 Character Set 의 Slug Position 및 한글 프린터의 6 단위 BCD Code 시스템은 아래와 같다((표 6, 7 참조)

표 6. 영문 프린터 Character Set 의 Slug Position

SET NO.	SLUG POSITION IN SET											
	1	2	3	4	5	6	7	8	9	10	11	12
1	0123	4567	BZ/	89AC	(GH)	RIFL	MNOP	[(%)]	DUST	VWXY	+ - .E	JKQ+
2			\$↗ ≠ =		BZ,/			(GH)		↕ ↘ ^ ~		VWXY
3			JKQ*	< > < >				BZ,/		(GH)		: ; ≡ =
4			VWXY	JKQ*				[%]		BZ,/		(GH)
5			\$↗ ≠ =		VWXY			JKQ*		↕ ↘ ^ ~		BZ,/
6			(GH)	< > < >				VWXY		IKQ*		: ; ≡ =

표 7. 한글 Code 시스템

Internal BCD	External BCD	Printer Character	HANGUL CHARACTER	8-Bit ASCH Character	Binary ASCH Code	Punched Card Codes
00	12	0	0	0	00110000	0
01	01	1	1	1	00110001	1
02	02	2	2	2	00110010	2
03	03	3	3	3	00110011	3
04	04	4	4	4	00110100	4
05	05	5	5	5	00110101	5
06	06	6	6	6	00110110	6
07	07	7	7	7	00110111	7
10	10	8	8	8	00111000	8

Internal BCD	External BCD	Printer Character	HANGUL CHARACTER	8-Bit ASCH Character	Binary ASCH Code	Punched Card Codes
11	11	9	9	9	00111001	9
12	00	:(colon)	:	:	00111010	2.8
13	13	=(equals)	=	=	00111101	3.8
14	14	≠(not equal)	≠	,	00100111	4.8
15	15	≤(less than & equal)	≤	&	00100110	5.8
16	16	% (Percent)	%	%	00100101	6.8
17	17	[(open bracket)	[[01011011	7.8
20	60	+(plus)	+	+	00101011	12
21	61	A	A	A	01000001	12.1
22	62	B	B	B	01000010	12.2
23	63	C	C	C	01000011	12.3
24	64	D	D	D	01000100	12.4
25	65	E	E	E	01000101	12.5
26	66	F	F	F	01000110	12.6
27	67	G	G	G	01000111	12.7
30	70	H	H	H	01001000	12.8
31	71	I	I	I	01001001	12.9
32	72	<(less than)	<	<	00111100	12.0
33	73	.(Period)	.	.	00101110	12.3.8
34	74) (closed parenthesis)))	00101001	12.4.8
35	75	≥(greater than or equal)	≥	^	01011110	12.5.8
36	76	¬(NOT)	+	"	00100010	12.6.8
37	77	;(semicolon)	;	;	00111011	12.7.8
40	40	-(minus)	-	-	00101101	11
41	41	J	J	J	01001010	11.1
42	42	K	K	K	01001011	11.2
43	43	L	L	L	01001100	11.3
44	44	M	M	M	01001101	11.4
45	45	N	N	N	01001110	11.5
46	46	O	O	O	01001111	11.6
47	47	P	P	P	01010000	11.7
50	50	Q	Q	Q	01010001	11.8
51	51	R	R	R	01010010	11.9
52	52	V(OR)	V	!	00100001	11.0
53	53	\$ (dollar sign)	\$	\$	00100100	11.3.8
54	54	*(asterisk)	*	*	00101010	11.4.8
55	55	↑(arrow up)	↑	#	00100011	11.5.8
56	56	↓(arrow down)	↓	\	01000000	11.6.8
57	57	>(greater than)	>	>	00111110	11.7.8
60	20	blank		Blank	00100000	(blank)
61	21	/ (slash)	/	/	00101111	0 "
62	22	S	S	S	01010011	0 "
63	23	T	T	T	01010100	0 "
64	24	U	U	U	01010101	0 "
65	25	V	V	V	01010110	0 "
66	26	W	W	W	01010111	0 "
67	27	X	X	X	01011000	0 "

Internal BCD	External BCD	Printer Character	HANGŬL CHARACTER	8-Bit ASCH Character	Binary ASCH Code	Punched Card Codes
70	30	Y	ㅇ	Y	01011001	0 ,
71	31	Z	ㅋ	Z	01011010	0 ,
72	32] (closed bracket)]	01011101	0.2.8
73	33	, (comma)	,	Comma	00101100	0.3.8
74	34	((open parenthesis)	((00101000	0.4.8
75	35	→(right arrow)	↔	~	01011100	0.5.8
76	36	≡(identity)	≡	—	01011111	0.6.8
77	37	∧(AND)	∧	?	00111111	0.7.8

5. 한글 영문 혼용 2줄모아쓰기의 구성

한글의 자음 14자 모음 10자로 글자를 구성할 수 있는 방법은 30가지임을 이미 기술하였다(표 2 참조). 여기에 영문·수자·특수문자 등을 합하면 훨씬 구성이 복잡하여져서 Driver 프로그램에서 한글을 Checking 하는 데 많은 시간을 소모할 것이다. 한글 영문 혼용 2줄모아쓰기에서는 표 8 과 같이 한글 자음 19자, 모음 14자, 영문 26자, 특수문자 20자 외에 한글 표시 특수문자(≡)와 공란(blank)을 합하여 93자로 Character Set 를 구성하였다. 이 중 한글 33자를 사용하면 한글을 9가지 조합의 하나로 구성할 수 있다(표 9 참조).

이와같이 한글을 30개의 복잡한 조합에 복자음 5자(ㄱ ㅋ ㆁ ㅅ ㅆ), 복모음 4자(ㅘ ㅙ ㅚ ㅛ)를 추가하여 한글 표시 특수문자 속에 포함시켜 풀어쓰기로 Coding 하여 입력시키면 Driver Program 인 DRV512H 에서 9가지 형태의 하나로 구분하여 Print Buffer Area 에 기억시키게 된다.

한글 영문 혼용 2줄 모아쓰기에 입력될 한글의 Coding 방법은 아래와 같다.

- ① 모든 한글은 풀어쓰기로 한칸에 자모 1개씩 현재 쓰고 있는 순서대로 Coding 한다.
- ② 한글표시는 2개의 특수문자 ≡를 사용한다. 즉, 모든 한글자모는 ↑로 시작해서 ↓로 끝나야 한다. 연산부호(+-*/)를 포함한 대부분의 특수문자 및 수자는 ≡에 관계없이 그대로 사용할 수 있다.
- ③ 한글의 경우 풀어쓰기 입력(Input)와 모아쓰기 출력(Output)의 차이에서 발생하는 Print

표 8. 한글 영문 혼용 2줄 모아쓰기 자판의 구성

종류	구분	자	모	계
		자 음	ㄱ ㅋ ㆁ ㅅ ㅆ ㅈ ㅊ ㅌ ㄴ ㄷ ㄹ ㅍ ㅎ	
한글	모 음	ㅘ ㅙ ㅚ ㅛ ㅜ ㅠ ㅡ ㅣ		10자
	복자음	ㄱ ㅋ ㆁ ㅅ ㅆ		5자
	복모음	ㅘ ㅙ ㅚ ㅛ		4자
	계			33자
영 문		A B C D E F G H I J K L M N O P Q R S T U V W X Y Z		26자
숫 자		0 1 2 3 4 5 6 7 8 9		10자
특수문자		+ - (= , & / # @ % * \$ % ≡ * * * * *)		20자
계				89자 (≡ 포함하지 않음)

Position 의 손실을 방지하기 위하여 특수문자 를 사용한다. 특수문자 의 다음에는 반드시 3 자리 수자(001~136)가 따라야 하며 그 다음에 오는 데이터의 Print Position 을 지적해 준다.

④ 복자음은 초성인 경우에는 한칸에, 종성인 경우에는 두칸에 나누어 Coding 한다.

⑤ 한글 Coding 의 예

표 9. 자모 33자에 의한 한글의 구성

순위	구 성 방 법	예	비 고
1	C + V ₂	고, 르, 드	V ₁ 윗모음
2	C + V ₂ + C	공, 목, 똥	V ₂ 아랫모음
3	C + V ₂ + C + C	흙, 흙	C 자음
4	C + V ₂ + V ₁	와, 과	
5	C + V ₂ + V ₁ + C	광, 권	
6	C + V ₂ + V ₁ + C + C	왔, 왔	
7	C + V ₁	가, 예, 에, 빠	
8	C + V ₁ + C	갈, 빵	
9	C + V ₁ + C + C	얹, 겹	

인쇄되지 않는다<특수문자 다음에 사용된 Position Indication(001~136) 포함>

⑤ 프린트의 예

국·영문 혼용 한글체인의 특성은 아래와 같다.

① 속도 250 줄/분……한글(Hangul)

150 줄/분……영문(Alphabet)

② Line Size……최대 136 자이며 136 자 이상은 Truncate 되고 Error 메시지 “Positioning Error”가 프린트된다.

③ Carriage Control……첫 Character(Standard MASTER Control)에 의하여 Control 된다.

④ Activation……512 프린터의 Image Buffer를 Image TR09으로 Load 함으로써 자동적으로 시동된다(Operator의 임무).

한글 영문 혼용 2줄 모아쓰기에 사용되는 한글 체인의 Character Set의 배열·한글 자모의 Design, 모아쓰기 프로그램의 Flow Chart는 아래와 같다.

보 한글자모인 한글과 하기수어인소와 CDK가 2도개바하
 한글 CHAIN으로 인쇄한 기입니다. 한글 모아쓰기 PROGRAM은
 한글 33자를 비롯하여 ALPHABET 26자, NUMERIC 10자, 특수문자
 20자 등 총 89자를 혼합하여 인쇄할 수 있습니다.
 71년 1월에 한글 모아쓰기 PROGRAM이 개발되고 앞으로는
 8자를 COMPUTER에 입력 시킬 수 있는 OCR이 도입되면 각
 기관 및 산업계에 크게 이바지할 것입니다.

1970년 11월

한국과학기술연구소 전자계산실

표 10. 한글 영문 혼용 Character Set의 Slug Position

Set No.	SLUG POSITION IN SET											
	1	2	3	4	5	6	7	8	9	10	11	12
1	0123	4567	89GH	+-.E	ABCD	RIFL	MNOP	서타비 (=나스키	포스트고	₩ &, ㅋ	ㅁ非@%	
2	0123	4567	89ㅎㅇ	\$-. ㄷ	ㅁㅂㅅㅇ	가랴리	—가해	서타비*/나스키	포스트고	ㅈㅊㅋ	기비비씨	
3	0123	4567	89GH	KIK	ㅁㅂㅅㅇ	가랴리	—가해	서타비 J F나스키	포스트고	MNL ㅋ	ABCD	
4	0123	4567	89ㅎㅇ	ㄷ, -+	ㅁㅂㅅㅇ	가랴리	—가해	J K Q*/STU	VWXY	\$ &, Z	(=·)	
5	0123	4567	89ㅎㅇ	+₩: ㄷ	ㅁㅂㅅㅇ	가랴리	—가해	서타비) #나스키	포스트고	ㅈㅊㅋ	기비비씨	
6	0123	4567	89ㅎㅇ	OPQ	ㅁㅂㅅㅇ	가리	—가해	서타비 R S나스키	포스트고	TU Z ㅋ	VWXY	

6. 한글전용 3줄 모아쓰기의 구성

컴퓨터의 편리함은 “Speed” “Accuracy” “Cost” 등이 다른 정보처리 수단보다 우월한 데 있다. 그중에도 10-6 sec의 Cycle Time으로 연산을 수행하는 CPU(Central Processor Unit)의 “Speed” 및 “Accuracy”는 1분당 1000매까지의 카아드를 읽어 1000줄의 속도로 프린트시키는 주변기기(Peripheral Equipment)의 입출력(Input

Output)속도보다 훨씬 우수하여 주변기기의 능력을 향상이 컴퓨터 전체 시스템에 주는 영향은 크다. 특히 한글Line Printer의 경우 활자의 배열 방법에 따라 한글 한자를 2줄 또는 3줄에 프린트 하여야 하므로 한글 프린터의 속도는 한글의 Character Set가 영문보다 많고 한글 Checking 시간까지 고려한다면 영문 프린터의 1/3~1/4 정도의 속도를 갖는다.

한글전용 3줄모아쓰기 구성은 위와 같은 한글

Line Printer의 성능을 향상시키기 위하여 영문을 완전 제거하고 특수문자의 수도 제한하여 한글의 빈도를 증가시켜 주변기기의 능률을 향상토록 한 것이다. 즉, 아래와 같이 한글자판을 자음 14자, 모음 12자, 수자 10자, 특수문자 12자, 합계 48자로 구성하여 한글 체인(Chain) Slug Position 288자에 6 Set의 한글이 배열됨으로써 한글의 빈도를 영문 Line Printer와 같은 6개 Set로 증가시켜 분당 수자 1150줄, 한

줄 400줄의 프린트 속도를 갖는다(표 11, 12참조)

표 11. 한글 3줄 모아쓰기 Character Set의 구성요소

구분	자 음, 모 음	계
종류		
자 음	ㄱ ㄴ ㄷ ㄹ ㅁ ㅂ ㅅ ㅇ ㅈ ㅊ ㅋ ㅌ ㅍ ㅎ	14개
모 음	ㅏ ㅑ ㅓ ㅕ ㅗ ㅛ ㅜ ㅠ ㅡ ㅣ ㅞ ㅟ	12개
수 자	0 1 2 3 4 5 6 7 8 9	10개
특수문자	* : . , / W () + - = %	12개
총 계		48개

표 12 한글 3줄 모아쓰기 자판의 구성

계 288자

SET NO.	SLUG POSITION IN SET											
	1	2	3	4	5	6	7	8	9	10	11	12
1	0123	4567	89ㅎㄱ	+-. ㄷ	ㅁㅂㅅㅇ	ㄱㅌㄹㅇ	-ㅌㅑㅑ	ㅑㅑㅑ*	/ㄴㅌㅑ	ㅑㅑㅑㅑ	W%, ㅋ	{= ;}
2	0123	4567	89ㅎㄱ	+-. ㄷ	ㅁㅂㅅㅇ	ㄱㅌㄹㅇ	-ㅌㅑㅑ	ㅑㅑㅑ*	/ㄴㅌㅑ	ㅑㅑㅑㅑ	W%, ㅋ	{= ;}
3	0123	4567	89ㅎㄱ	+-. ㄷ	ㅁㅂㅅㅇ	ㄱㅌㄹㅇ	-ㅌㅑㅑ	ㅑㅑㅑ*	/ㄴㅌㅑ	ㅑㅑㅑㅑ	W%, ㅋ	{= ;}
4	0123	4567	89ㅎㄱ	+-. ㄷ	ㅁㅂㅅㅇ	ㄱㅌㄹㅇ	-ㅌㅑㅑ	ㅑㅑㅑ*	/ㄴㅌㅑ	ㅑㅑㅑㅑ	W%, ㅋ	{= ;}
5	0123	4567	89ㅎㄱ	+-. ㄷ	ㅁㅂㅅㅇ	ㄱㅌㄹㅇ	-ㅌㅑㅑ	ㅑㅑㅑ*	/ㄴㅌㅑ	ㅑㅑㅑㅑ	W% 'ㅋ	{= ;}
6	0123	4567	89ㅎㄱ	+-. ㄷ	ㅁㅂㅅㅇ	ㄱㅌㄹㅇ	-ㅌㅑㅑ	ㅑㅑㅑ*	/ㄴㅌㅑ	ㅑㅑㅑㅑ	W%, ㅋ	{= ;}

특히 한글전용 3줄 모아쓰기에 사용되는 한글 Chain에 배열된 수자(0~9)는 USSASCOCR size A font로 되어 있어 Print Sheet에 찍힌 수자를 OCR(Optical Character Recognition; 광학문자해독기)를 거쳐 컴퓨터에 다시 입력시킬 수 있다.

한글전용 3줄모아쓰기 구성방식은 풀어쓰기로 입력된 한글을 Line Printer에 프린트되기 직전 Line Printer Driver 프로그램인 LPEX512 LPD-R512 Routine에서 Check하여 한글표식(☺≡)이 있을 때는 KISTCHK Routine에서 한글을 30가지형태(표 2)의 하나로 구분하여 자음, 모음을 3개의 Print Buffer에 배열하였다. 즉, 아랫모음이 뒤따르는 초성은 BUFF1에 기타 초성 및 모음은 BUFF2에 종성은 BUFF3에 기억시켜 프린트시 1, 2, 3 단계에서 각각 BUFF1, BUFF2, BUFF3를 프린트시켜 3줄로 한글의 1줄을 형성토록 구성한 것이다.

- 한글전용 Line Printer의 성능은 아래와 같다
- ① 속도 400줄/분.....한글 (HANGUL)
1150줄/분.....수자(NUMERIC)
 - ② Line Size.....136자까지 프린트 가능하며 이

상은 truncate되고 Error 메시지 "RECORD OVERFLOW"가 프린트된다.

- ③ Carriage Control...첫 Character (Standard Master Control) Control된다.
 - ④ Activation.....CDC 512 HANGUL Printer에 한글 Image TR08을 Load하면 자동적으로 시동된다(Operator의 임무).
- 한글전용 3줄모아쓰기 구성을 위한 코오딩은 한글-영문 혼합의 경우와 같으며 Output이 3줄인 점을 제외하면 사용법은 동일하다.

7. EDPS 한글화의 나아갈 길

제 2장에서 한글의 입력, 출력, 한글·영문혼용 2줄 모아쓰기, 한글전용 3줄 모아쓰기에 대하여 기술하였다. 이상의 모든 것은 모두 한글이 영문의 Alphabet과 같이 풀어쓰기 방식이 아닌 모아쓰기에서 기인된 것이다. 사람이 있는 곳에 언어가 있고 또 그 언어를 보전하기 위하여 문자가 생겼다. 즉, 문자는 그것을 쓰는 사람들 자신의 것인 것이다. 지구상의 300이 넘는 언어를 살펴볼 때 선진국가의 문자는 모두 풀어쓰기 방식을 택하였고 우리말과 같은 형식은 찾기 어

럽다. 특히 20세기 후반과 같이 정보산업의 시대에 사는 우리들의 문자는 무엇보다도 신속하고 간단하게 다루어져야 할 것이다. 물론 수백년 동안 모아쓰기로 하던 우리의 관습을 하루아침에 풀어쓰기로 한다는 것은 힘든 일이다. 여기서 컴퓨터를 중심으로 한 한글의 과학화가 이루어져 한단계 한단계씩 사회의 저항없이 풀어쓰기로 옮길 것을 제안한다. 먼저 받침을 우측에 프린트하는 단계, 다음은 한글자모의 수직성을 인정하면서 수평으로 풀어쓰는 단계, 다음에 완전히 풀어쓰는 단계로 옮겨질 것을 제안한다 (그림 15 참조).

EDPS의 한글화 문제는 입출력 시스템에 그치는 것이 아니다. 프로그램 자체도 한글화 하여야 할 것이다. 우리가 흔히 쓰는 FORTRAN, COBOL과 마찬가지로 한글로 프로그램을 할 수 있어야 할 것이다. 나아가서는 하드웨어(Hardware) 자체, 즉, 우리손으로 만든 컴퓨터에 한글 프로그램 및 한글 데이터를 처리하여 보고서를 작성할 수 있을 때 EDPS 한글화 및 생활화가 이루어질 것이다.

한글 풀어쓰기 전환 단계(그림 15 참조)
 현행 : 2줄 모아쓰기

오	ㅣ	ㅈ	ㅊ	ㅊ	ㅣ	ㅈ	ㅣ													
ㄴ		ㄴ	ㄴ	ㄴ																

3줄 모아쓰기

ㅇ			ㅍ		ㅅ	ㅅ														
ㅈ	ㅣ	ㅈ	ㅊ	ㅊ	ㅣ	ㅈ	ㅣ													
ㄴ		ㄴ	ㄴ	ㄴ																

풀어쓰기 단계 : 제 1 단계

ㅇ	ㅣ	ㅈ	ㅊ	ㅊ	ㅣ	ㅈ	ㅣ													
ㄴ		ㄴ	ㄴ	ㄴ																

제 2 단계

ㅇ	ㅣ	ㅈ	ㅊ	ㅊ	ㅣ	ㅈ	ㅣ													
ㄴ		ㄴ	ㄴ	ㄴ		ㅈ	ㅣ													

제 3 단계

ㅇ	ㅈ	ㅊ	ㅊ	ㅈ	ㅣ	ㅈ	ㅣ	ㅈ	ㅣ	ㅈ	ㅣ									
ㄴ		ㄴ	ㄴ	ㄴ		ㅈ	ㅣ	ㅈ	ㅣ	ㅈ	ㅣ									

그림 15

III. 결 론

1. 연구결론

“EDPS 한글화에 관한 연구”는 풀어쓰기로 입력된 한글을 모아쓰기로 프린트할 수 있는 새로운 한글 입출력 시스템을 실용가능한 단계까지 완성시켰다. 즉, Key Punch Machine에서 한글표식 특수문자와 같이 풀어쓰기로 입력된 한글은 한글 Check Routine을 거쳐 실제로 프린트할 Buffer Area에 기억된 후 한줄씩 Line Printer에 프린트되어 3줄로서 한글 1줄을 형성할 수 있게 되었다. 이미 소개되어 사용되고 있는 한글 영문 혼용 체인이나 71년 3월에 도착예정인 한글전용 OCR Font수자 체인이 EDPS 한글화의 바탕을 이룬은 당연한 일이나 본 연구기간 중 EDPS 한글화에 있어서 해결되어야 할 중요한 사항이 지적되었다.

첫째, 한글은 궁극적으로 OCR을 통하여 입력되어 풀어쓰기로 사용되어야 할 것이다.

둘째, 컴퓨터 관계자뿐 아니라 전 한글관계자를 망라한 EDPS 한글화 사업이 추진되어야 할 것이다.

셋째, EDPS 요원 특히 시스템 담당자들의 자질 향상이 이루어져야 할 것이다.

네째, 활자체의 시각적 개량 및 한글 체인의 속도향상에 계속 노력하여야 할 것이다.

다섯째, 한글자모의 사용빈도 및 상호연관도를 세밀히 분석하여 한글 체인의 Character Set의 Slug Position을 합리적으로 선정하여야 할 것이다.

2. 건의사항

(1) EDPS 한글화에 관한 연구가 다음 단계로 갈 수 있도록 계속적인 지원이 필요하다.

(2) 현행 한글 활자체 및 구성방법을 계속하여 개량하여야 한다.

(3) 한글 관계 담당 S.E요원의 장기근무 및 자질향상이 필요하다.

컴퓨터에 연결된 한글인쇄기가 71년 11월에 가동한 이래 71년에는 서울시에서 재산세고지서

발행 및 자동차세고지서 발행업무에 이용한 것을 필두로 국세청, 체신부, 은행등에서 각종 행정 및 경영자료 처리에 활발히 이용되고 있다. 한글 기계화의 진척은 앞으로 행정능율화, 경영합리화에 더욱 박차를 가할 것이고, 이에 따라서 한글 기계화의 기술수준도 더욱 높아질 것이다. 한국의 현대화를 촉진하는 강력한 수단으로서 컴퓨터와 한글자동인쇄기가 더욱 더 활용되기를 바라는 마음 간절하다.

참 고 문 헌

David L. Smith and Ki-Soo Sung "Computer Handling of Korean Language Texts" April 15, 1969. Battelle Memorial Institute, Columbus, Ohio, USA.
Hayashi H., Duncan S., "Graphical Input/Output of

Nonstand Characters" Communication of ACM, Vol 11, No 9, pp 613-618, Sept 1968.
SD 4060 stored program Recording System, Description and Specifications, Stromberg Datagraphix, No 9500209, Revised April 28, 1967
송제범 "한글타자기의 전반배열에 관하여" 1953. 3.
CDC 3300 Computer System Reference Manual Pub No. 60157000
CDC 3300 3500 Computer Systems Master Version 3 Reference Manual Pub No. 60281000 Mar. 1970
CDC 3300 Computer Systems COMPASS Reference Manual Pub No. 60236800 July 1968.
CDC 3300 Master Version 3 Installation Handbook Pub No 60280500 Feb. 1970
CDC Internal Maintenance Specifications MIOCS