

자동 역점역을 가능하게 하는 한글점자 부호체계의 개선

길 태 영[†] · 김 석 일^{††} · 김 홍 기^{†††}

요 약

기존의 점자 부호 체계에서 한글문장을 점자문장으로 번역하는 작업은 1:1의 대응 관계가 있으므로 아무런 문제를 야기시키지 않는다. 그러나 점자문장을 한글문장으로 번역하는 작업은 점자부호와 한글부호간의 1:N의 대응 관계로 인하여 피할 수 없는 오류를 파생시킨다. 분석 결과, 점자 부호 체계에서는 한글 초성과 숫자와의 충돌, 한글 중성과 문장부호와의 충돌, 한글 소괄호 열기와 소괄호 닫기 부호를 동일한 부호로 처리하는 것, 영문 열기 및 닫기 부호와 한글 중성과의 충돌 등이 번역 오류의 중요한 원인임을 알 수 있다. 본 논문에서는 이러한 충돌을 해소하기 위한 역점역 규칙을 제정하여, 일차적으로 역점역시에 발생하는 애매성을 앞, 뒤 문맥에 따라 해결할 수 있는 경우에는 원래의 점자 부호를 사용하도록 하고, 만일 이 방법으로 해결되지 않을 경우에는 역점역시에 충돌을 근본적으로 방지하는 확장 점자 부호를 사용하도록 하였다. 여러 가지 파일에 대한 실험 결과, 기존의 점자 부호체계에서의 오류발생율이 평균 25%이었음에 비하여 확장된 점자 부호체계에서는 번역에 따른 오류가 발견되지 않았다.

Improvement of Korean Braille-Code System for Automatic Reverse Braille Translation

Tae-Yeong Kihl[†] · Sukil Kim^{††} · Hong-Gi Kim^{†††}

ABSTRACT

Translation of a Korean text into a braille text causes no problem under the existing Korean braille-code System that maintains 1:1 correspondence between Korean characters and braille codes. However, reverse translation of a braille text into a Korean text would cause unavoidable mis-translation due to 1:N correspondences between braille codes and Korean characters. The analysis shows that in Korean braille-code System, the major reasons for mis-translation are the conflicts between Korean Chosung (initial consonant) letters and numbers, between Korean Jongsung (final consonant) letters and punctuations, between Korean Jongsung (final consonant) letters and English mode delimiter codes, and overlapping use of the same braille code such as opening and closing parentheses. In this paper, we firstly established a reverse translation rule, that is to use an ordinary braille code unless the braille code causes any conflict in braille code sequences. If any ordinary braille code may cause some conflict during the reverse translation, an extended braille code that we propose in this paper must be used. In the experimentation, we cannot get any translation errors with the newly designed braille-code System compared with 25% of errors with the existing braille-code System.

† 준 회 원: 청주대학교

†† 중 심 회 원: 충북대학교 컴퓨터과학과

††† 정 회 원: 충북대학교 컴퓨터과학과

논문접수: 1997년 5월 10일, 심사완료: 1998년 1월 16일

1. 서 론

한글점자 체계[1, 4, 8]는 한글 초성, 중성 및 종성을 각각 6비트로 표현하는 부호체계이므로 하나의 자소를 5비트로 표현하는 한글 자모와 1:1 대응되나, 특수 기호 및 약자 처리, 또는 외국어 처리를 위한 별도의 부호가 마련되어 있지 않으므로 하나의 점자 부호가 서로 상이한 의미를 나타내는 여러 개의 한글 부호로 대응된다. 예를 들어 점자 문자열

(---4-6)(1-3--6)(--3-56)(12---6)(-2--56)

을 역점역(reverse translation)하여 한글 부호로 변환하면 점자 부호 (-2--56)은 마침표 (.) 또는 종성 'ㅍ'으로 인식이 가능하므로 번역시 '좋아.' 또는 '좋았'으로 번역된다. 따라서 아래의 문장

'좋았'은 '좋아.'의 잘못된 표기이다.

을 점자 문장으로 점역한 후 다시 한글로 번역하면

'좋아.'은 '좋아.'의 잘못된 표기이다.

로 번역되어 점역 및 역점역 이전의 의미가 상실되는 것을 피할 수 없다.

참고문헌[2, 7]에 따르면 기존의 점자 부호체계 하에서 한 영문 혼용 파일을 점역한 다음, 점자 파일을 다시 역점역하였을 때, 원래의 파일과 새로 생성된 파일을 비교하면 평균 25%의 번역 오류가 발생하며, 기술 논문류의 경우에는 번역 오류율이 증가하여 최대 38.5%에 이르고 있음을 알 수 있다. 기술 논문류의 역점역 오류율이 높은 이유는 기술 논문 파일에 한글과 영문 및 여러 가지 기호들이 많이 섞여 있기 때문이며, 특히 한글과 영문을 구분하는 점자 부호를 구분하지 못하면 매우 심각한 오류를 유발하고 있다. 즉, 영문 열기에 해당하는 점자 부호가 한글 따옴표 단기 및 종성 'ㅎ'과 동일한 점자 부호를 사용하며, 영문 단기 점자 부호도 한글 종성 'ㅍ' 및 마침표와 동일한 점자 부호를 사용하고 있기 때문이다. 즉, 점자 문장에서 마침표에 해당하는 점자 부호가 존재할 때, 이것을 영문 단기 부호로 해석할 것인지 아니면 마침

표로 해석할 것인지를 결정하기가 매우 모호하다.

오프라인 번역의 경우에는 마침표 또는 영문 단기로 번역될 수 있는 점자 부호 다음에 나오는 일련의 문장 중에서 분명히 한글 음절을 구성하는 것이 확인될 때에는 이 점자 부호를 마침표로 처리할 수 있다. 마찬가지로 방법으로 만일 따옴표 또는 영문 열기로 번역될 수 있는 점자 부호 다음에 이어지는 문장 중에서 영문 단어를 구성하는 것이 분명할 경우에는 이를 영문 열기 부호로 처리할 수 있다. 그러나 이 방법은 점자 단어가 영문 단어나 한글 단어 중에서 어느 것으로 번역되어도 무방한 단어인 경우에는 완벽한 방법이 될 수 없다. 더군다나 이 방법은 입력 중인 점자를 한영문 음절로 실시간 번역을 하는 경우에는 사용될 수 없는 단점이 있다.

영문이 포함되지 않은 문장의 경우에도 앞에서 예로든 마침표와 종성'ㅍ'간의 경우와 같이 한글 종성과 문장 부호들 간의 충돌로 인하여 번역 상의 오류율이 10%에 달한다[2, 7]. 따라서 점자 문장을 정확한 한영문 문장으로 자동 번역하기 위해서는 역점역시에 충돌을 유발하는 점자 부호의 수정이 불가피하다. 본 논문에서는 역점역시에 충돌을 일으키는 점자 부호를 확장하여 점자 부호를 재정의하고 이를 점역 시에 사용하도록 하여 역점역시의 애매성을 줄이는 방안을 연구하였다.

본 논문의 제 2 절에서는 기존의 점자 부호체계[8]에서 점자 부호가 경우마다 대응되는 한영문 부호를 조사하였으며, 제 3 절에서는 역점역시에 발생하는 애매성을 최소화하기 위하여 점자 부호의 수정 방안을 제시하였다. 제 4 절에서는 본 논문에서 제안한 점자 부호를 이용하여 하이텔 등에 수록된 임의의 한영문 파일을 점역한 후 역점역하여 원래의 파일과 비교하여 오류 발생 여부를 조사하였다. 마지막으로 제 5 절에서는 본 논문의 결론을 도출하고 향후 계획을 수록하였다.

2. 애매성을 유발하는 점자와 한글의 대응 관계

한글점자 통일안[8]을 토대로 각각의 점자 부호에 대응하는 한영문 및 특수 기호로 사용될 때의 의미를 정리하면 <표 1>과 같다. <표 1>에서 특수 기호 항목의 설명에서 첫머리에 붙은 'M'은 해당 점자 부호를

수학 기호로 사용할 때의 의미를 나타내며, 'E'는 영어 문장 내에서 사용될 때의 의미를 나타낸다.

〈표 1〉에서 보는 바와 같이 한글의 경우에는 한글 음절 다음에 구두점 등의 특수 기호가 이어서 나올 경우에 애매성을 유발하는 점자 부호는 (-23-6), (-2--56) 및 (-3-56)의 세 가지이다. 즉, 점자 부호 (-23-6)는 중성 'ㅌ', 물음표 및 따옴표 열기 중의 하나로 번역될 수 있다. 또한, 점자 부호 (-2--56)은 중성 'ㄹ', 마침표 및 영문 닫기 기호로 번역될 수 있다. 마찬가지로 점자 부호 (-3-56)은 중성 'ㅎ', 따옴표 닫기 및 영문 열기 기호 중의 하나로 번역될 수 있어 정확한 번역이 어렵다.

요즈음 문서편집기에서는 문장 내에서 컴퓨터 자판상의 모든 특수 기호를 포함시킬 수 있다. 따라서 이러한 특수 기호들을 문장 내에서 사용할 수 있도록 허용한다면 더욱 많은 충돌을 발생시키게 될 것이다. 한글 중성과 충돌을 발생시키는 특수 기호로는 분수표(-34-), 뉘셈표(-3-5-), 덧셈표(-2--6) 등이 있으며, 1종 약자와 충돌을 일으킬 수 있는 특수 기호로는 곱셈표(1--6), 느낌표(--456)가 있다. 또한 말소표(123456)도 음절 'ㅇ'과 항상 충돌을 일으킨다. 한글 점자와는 달리 영문의 경우에는 특수 기호와 영문 알파벳간의 충돌이 발생하지 않도록 점자 부호가 할당되어 있다.

한국점자통일안[8]에 따르면 점자 문장을 한영문 문장으로 번역하는 과정에서 발생하는 오류의 많은 부분이 한영문 혼용 점자 문장에서 마침표와 영문 닫기 및 따옴표 닫기와 영문 열기 부호의 충돌에 기인한다. 이러한 부호의 충돌은 점자 문서를 역점역하여 얻은 문서의 완성도를 저하시키는 중요한 원인이 되어 왔다. 그러나 이제까지는 일반인이 작성한 문서를 시각 장애인들이 점역해서 이용하는 경우가 대부분이었으므로 역점역의 필요성이 크지 않았으나 점차 시각 장애인에 작성한 점자 문서를 일반인이 사용하는 경우가 늘어나고 있으므로 역점역의 중요성이 점차 중요하게 인식되고 있다. 따라서 본 절에서는 한국점자통일안을 준용하면서도 역점역시에 발생하는 충돌을 가능한 한 해소할 수 있는 방법을 우선 강구하여 보았다.

- 1) 한글 중성과 문장부호간의 충돌 - 이 경우에는 번역하려는 점자 부호의 바로 앞에 있는 점자 부호

가 한글 중성이면 중성으로 인식하도록 하며, 앞의 점자 부호가 중성이나 특수 부호일 경우에는 문장부호로 간주한다. 또한, 다음에 나오는 점자 부호가 공백(space)이면 문장부호로, 한글 초성이거나 중성이면 한글 중성으로 번역하도록 하여 문제를 해결할 수 있으나 중성이 없는 음절의 경우에는 적용될 수 없다.

- 2) 큰따옴표가 다른 문장부호 또는 외국어 표시 부호와의 충돌 - 따옴표 열기 또는 중성 'ㅌ'용 점자 부호(-23--6)의 구분은 중성 'ㅌ'앞에는 항상 한글 초성('ㅌ'생략 점자 부호의 경우)이나 중성이 위치하나 따옴표 열기의 경우에는 공백이 있는 경우가 빈번하므로 이 점자 부호를 번역할 때 앞에 공백이 있을 경우에는 따옴표 열기로 이외의 경우에는 중성 'ㅌ'으로 번역하도록 하여 문제를 해결할 수 있다. 마찬가지로 따옴표 닫기와 영문 열기를 의미하는 점자 부호(-3-56)의 경우에도 번역할 점자 부호에 이어서 공백이 오는 경우에는 따옴표 닫기로 공백이외의 점자 부호가 올 경우에는 영문열기 점자 부호로 번역할 수 있다. 그러나 [7]의 실험에서 밝혀진 바와 같이 점자 문서를 작성하면서 따옴표에 이어서 공백을 삽입하는 것은 문서를 작성하는 개인의 습관에 달린 문제이며 따라서 점자 문서를 작성하는 개인의 습관이 역점역의 신뢰도에 미치는 영향이 매우 큰 단점이 있다.

외국어 표시 부호 또는 따옴표 닫기 부호를 어떻게 번역할 것인가를 판단하는 또 다른 방법은 이어지는 점자 부호 한 단어를 검사하여 이 단어가 영문인가 한글인가를 확인한 후 영문일 경우에만 외국어 표시 부호로 인식하고 나머지 경우에는 따옴표 닫기로 처리하는 것이다. 이 방법은 영문 닫기 또는 마침표(.)를 의미하는 점자 부호(-2--56)의 경우에도 적용할 수 있는 장점이 있다. 그러나 이 방법 역시 경험적인 방법에 근거를 두고 있기 때문에 완벽하지 못하다. 특히 〈표 1〉에서 볼 수 있듯이 한글 자모와 영문자간의 중복성이 크기 때문에 충돌의 애매성을 근본적으로 제거할 수 없다. 실험[2]에 따르면 한영문 혼용 점자 문서를 역점역할 때 발생하는 대부분의 오류가 외국어 표시 부호와 문장부호의 충돌에 의하여 발생하

〈표 1〉 한글점자체계에서의 충돌 발생
 (Table 1) Conflict in korean braille code system

점자 부호	초성	중성	종성	1종약자	2종약자	영문 숫자	복수기호	점자 부호	초성	중성	종성	1종약자	2종약자	영문 숫자	복수기호
1----			ㄱ			A 1	그러나(+14) 그리고(+136) 그래서(+234) 그런데(+1345) 그러면(+25) 그러므로(+26)	1---6				연			M곱셈표
								-2--6			ㅁ			M덧셈표 M부등호 '>' (+26)	
								12--6	ㅏ				M계급표시		
-2----			ㄴ				M자릿점 E염표	-3--6						붙임표 중표-괄표(+36) 소괄호열기 소괄호닫기	
12----			ㄷ			B 2									
--3---			ㄹ				M소수점 E어퍼스트로피 E말줄임(+3+3)	1-3--6	ㅑ					U	
1-3---			ㅍ			K		-23--6			ㅓ				물음표 따옴표열기 중괄호열기(+23) 대괄호열기(+3)
-23---			ㅑ			:									
123---			ㅓ			L		123--6	ㅕ	ㅕ				V	
---4--	ㄱ														
1--4--	ㄴ			나		C 3		---4-6	ㅗ		자				E이텔릭기호
-2-4--	ㄷ			다		I 9		1--4-6	ㅛ						M세계급표시
12-4--	ㅋ			카		F 6		-2-4-6	ㅜ						
--34--	ㅋ		ㅍ				M나눗셈표(+34) M분수표 (분모, 분자)	12-4-6			가				
								--34-6	ㅝ						
1-34--	ㅌ 기(+1235)					M		1-34-6			옥			X	
-234--	기					S		-234-6			울				
1234--	기 계(+1235)					P		1234-6			울				
---5-	ㄴ			라			염표 콜론(+2) 가운데점(+23) 말줄임표(+5, +5)	1--56	ㅋ						세미콜론(+23) 중괄호열기(+356) E철자기호
								-2--56			ㅍ				마침표 영문닫기
1--5-	ㅁ			마		E 5		12--56			열				
-2--5-			ㄴ			:	M등호(+25)								
12--5-	ㅓ			타		H 8		--3-56			ㅎ				영문열기 따옴표닫기 작은따옴표닫기 (+3)
--3-5-			ㄷ				M뺄셈표 M부등호 '<'(+35) 널표(+35)	1-3-56			은			Z	
1-3-5-	ㅣ					O									E소괄호열기 E소괄호닫기 E대괄호열기(+3)
-23-5-			ㅋ			!		-23-56			ㅇ				
123-5-	ㅎ					R	판 ' '								
---45-	ㅍ			바				123-56			은				
1--45-	ㅑ			과		D 4		---456			것 (+234)				느낌표 M빗금 ' ' (+34)
-2-45-	ㅎ			하		J 0		1--456			역				
12-45-				훈		G 7		-2-456	ㄱ					W	
--345-	ㅌ 계(+1235)							12-456			영				
1-345-	ㅋ					N		--3456							수표
-2345-				얼		T		1-3456	ㅓ					Y	
12345-				인		Q		-23456			언				
-----6	ㅏ						원소리기호 긴소리표(+3) 작은따옴표열기 (+236) 대괄호열기(+356) E영문대문자 E대괄호닫기 (+2356)	123456			용				말소표 E대문자기호 E이중대문자기호 (+6)

므로 실용적인 역점역을 개발하기 위해서는 이 부분의 충돌을 근본적으로 방지하는 방안을 강구하여야 한다.

- 3) 점자 부호(--3--6)의 지나친 중복 사용으로 인한 충돌-점자 부호(--3--6)은 괄호 열기와 괄호 닫기 부호로 사용되므로 괄호 열기인지 닫기인지를 구분할 수 없다. 플래그(flag)를 이용하여 괄호 열기와 괄호 닫기를 구분할 수 있으나 중첩 괄호인 경우나 괄호 열기만 있고 괄호 닫기가 없는 문장의 경우에도 괄호 열기와 닫기가 반복되는 것으로 해석할 수 있는 단점이 있다. 이 문제와 관련하여 개정을 둘러싼 많은 토의가 있었으나 사용자에게 혼동을 준다는 이유로 그 동안 개정되지 못하였으나 역점역시에 치명적인 오류를 발생하는 점자 부호이므로 이의 개정이 매우 시급하다[10].

이상의 관찰로부터, 한영문의 시작과 종료를 나타내는 점자 부호는 다른 점자 부호와 충돌이 일어나지 않는 특별한 부호를 사용하는 것이 바람직하며, 이외의 점자 부호도 역점역시에 충돌이 발생하는 경우에는 특별히 정의된 점자 부호를 사용하도록 하여 점자 문서를 한영문 문서로 번역함에 따른 정확성을 향상시킬 필요가 있음을 알 수 있다.

3. 점자 부호의 개선

본 절에서는 기존의 점자를 확장하거나 번역기법을 개선하여 점자를 한글로 번역하는 과정에서 애매성으로 인한 번역오류를 줄일 수 있는 점역방법 및 점자 부호 확장 방안을 설계하였다. 기존의 한글점자 부호 체계를 전면적으로 개편하는 경우에는 점자를 한글로 번역할 때의 오류를 원천적으로 해소할 수 있는 방안이 강구될 수 있으나 새로운 점자 부호체계가 도입됨으로 인하여 시각장애인들이 새로운 점자 부호 체계에 익숙해질 때까지 학습이 필요하므로 충분한 학습이 이루어질 때까지 많은 혼란을 일으킬 수 있다. 더군다나 새로운 점자 부호체계를 도입하는 전략은 컴퓨터를 이용하여 보다 편리하게 점자 문장을 한영문 문장으로 번역하자는 취지를 무색하게 하므로 본 논문에서는 기존의 점자 부호체계를 사용하되, 역점역을 하는 과정에서 분명한 애매성이 발생하는 경우

만 새로 정의된 점자 부호를 사용하도록 하였다. 이를 위하여 우선 몇 개의 문서를 이용하여 역점역시의 애매성을 유발하는 부호의 충돌 사례를 조사하였다.

〈표 2〉는 하이텔에서 임의로 선택한 3개의 기술은 문류의 문서와 3개의 신문 사설류를 조사하여 이들 문서를 점역한 후에 다시 역점역할 경우 애매성을 유발하는 횟수를 측정한 것이다. 조사 결과 마침표 부호의 사용횟수가 6가지 문서의 경우에 가장 빈번하였으며, 기술 관련 문서에서는 영문 닫기 부호가 두 번째로 많이 사용되고 있음이 밝혀졌다. 그런데 마침표와 영문닫기 부호가 충돌을 일으키므로 이를 구분하는 것이 필수적임을 알 수 있다. 그러나 신문 사설의 경우에는 같은 점자 부호를 사용하는 불임표(hyphen), 소괄호열기 및 소괄호 닫기가 많이 사용되지 않음을 알 수 있다.

기술 논문에서는 따옴표 닫기 또는 영문 열기들의 미하는 점자 부호(--3-56)의 대부분이 영문 열기부호로 사용되었음을 알 수 있으며, 신문 사설 등의 문서에서는 영문열기부호의 사용횟수가 기술관련 문서에서의 사용횟수에 비하여 크게 작은 것을 알 수 있다. 따라서 점자 부호(--3-56)를 무조건 영문 열기부호로 사용하여도 98%이상의 정확성을 얻을 수 있음을 알 수 있다. 따라서 본 논문에서는 영문 닫기 점자 부호와 마침표 점자 부호를 분명히 구분하기 위하여 영문 닫기 점자 부호를 명확히 구분하고 나머지 점자 부호에서 발생하는 충돌은 기존의 점자 부호를 그대로 사용하면서 충돌을 해소할 수 있는 방법을 우선 찾고자 하고 불가능할 경우에만 새로운 점자 부호를 정의하여 사용하도록 하였다.

3.1 한영문 전환 점자 부호

한영문 점자 문장을 한영문 일반문장으로 번역하는 경우에 가장 많이 발생하는 오류는 〈표 2〉에서 보는 바와 같이 영문 닫기 점자 부호와 마침표, 중성 ‘ㅍ’과의 충돌에 기인한다. 그런데 마침표는 영문 점자 문장에서도 빈번하게 사용되므로 영문 점자문장을 번역할 때 점자 부호(-2--56)를 마침표로 번역하고 계속해서 영문 점자가 계속될 것인지 아니면 마침표가 아닌 영문 닫기에 해당하는가를 결정할 수 없다. 따라서 이 문제를 분명하게 하기 위해서는 영문 열기와 영문 닫기 점자 부호가 다른 점자 부호와 확실히

〈표 2〉 여러 가지 문서에서 충돌 가능성 조사
 〈Table 2〉 Survey of possible conflicts on various documents

부호구분	문서 종류 문서 크기 문서명	기술논문문서			신문사설			계
		11,508	8,852	34,504	2,223	5,943	10,397	
		test1	test2	test4	test15	test16	test17	
(-2--56)	마침표(.)	380	147	391	29	64	8	1,019
	중성(ㅍ)	0	1	9	1	2	0	13
	영문 닫기	202	127	321	7	3	3	663
(-23--6)	따옴표열기	3	6	0	2	0	0	11
	붙임표	2	0	18	0	3	0	23
	중성(ㅌ)	0	5	29	0	0	0	34
(--3-56)	영문열기	202	127	321	7	3	3	663
	따옴표닫기	3	6	0	2	0	0	11
	중성(ㅎ)	0	2	6	0	1	0	9
(--3--6)	소괄호열기	35	36	107	4	3	7	192
	소괄호닫기	35	36	107	4	3	7	192
	붙임표	24	29	71	6	0	8	138

계 구분될 수 있도록 점자 부호를 확장하여 사용할 필요가 있다. 따라서 본 논문에서는 기존의 특수 기호에 해당하는 점자 부호를 확장하는 경우에는 느낌표(---456)를 점자 부호에 선행시킨 확장 부호를 사용하도록 하고, 영문 닫기(또는 한글 열기) 점자 부호로 영문자 'K'에 해당하는 점자 부호(1-3--)를 느낌표에 후치시켜 표시하도록 하였다. 동일한 방법으로 영문자 열기부호와 중성 'ㅎ'과의 충돌로 인한 오류도 피하기 위하여 영문 열기 점자 부호로 점자 부호(1--5-)를 사용하도록 하였다. 즉,

영문 열기:(---456)(1---5-)

한글 열기:(---456)(1-3--)

이 방법은 한 문서 내에 여러 외국어를 혼용하는 경우에도 유리하다. 예를 들어 일본어 열기 부호를 영문자 'J'에 해당하는 점자 부호(-2-45)를 느낌표에 후치시켜서 사용할 수 있다.

3.2 역점역 순서에 의한 충돌 해소

기존 점자 부호체계에서의 중요한 오류 중의 하나인 한글 중성과 구두점이 같은 점자 부호를 사용하므로써 발생하는 충돌은 역점역시에 동일한 점자 부호가 문장부호나 한글중성으로 해석될 수 있는 경우에는 그 중의 한가지를 우선적으로 적용하여 오류를 줄일 수 있다. 예를 들면 〈표 2〉에서의 조사와 같이 마침표의 사용빈도가 중성 'ㅍ'에 비하여 많으므로 점자 부호(-2--56)을 우선 마침표로 사용하도록 하면 중성 'ㅍ'으로 번역하는 경우에 비하여 역점역시의 오류를 줄일 수 있다. 그러나 점자 부호(-23--6)이나 점자 부호(--3-56)의 경우에는 각각 중성 'ㅌ'과 'ㅎ'으로 사용되는 경우가 구두점으로 사용되는 경우에 비하여 빈번하므로 본 논문에서는 이들 점자 부호를 우선 한글 중성으로 번역하도록 우선 순위를 부여하였다. 즉, 한글 음절은 초성, 중성, 종성의 조합으로 구성되므로 한글 중성과 구두점으로 인식될 수 있는 점자 부호가 있다면 중성 다음에 나오는 점자 부호인 경우에는 구두점으로 한글 중성이나 초성 다음에 나오는 점자 부호는 우선 한글 중성으로 처리하는 것이다. 이렇게 하므로써 사용하는 경우의 수가 많은 한글

중성의 경우에는 점자 부호를 확장하지 않고 구두점을 나타내는 점자 부호를 확장하도록 하여 문장에서 사용되는 확장 점자 부호의 수를 줄일 수 있도록 하였다.

본 논문에서는 한글점자 문장을 한글 문장으로 번역할 때,

- 1) 한글 초성
- 2) 중성
- 3) 종성
- 4) 문장부호(구두점)

의 순으로 점자 부호가 사용되는 의미를 정하였다. 따라서 쌍받침 중성 점자 부호 다음에 오는 점자 부호는 항상 구두점으로 번역할 수 있으며, 초성이나 중성 다음에 오는 점자 부호는 중성으로 처리할 수 있다.

한글 문장에서는 중성이 없는 음절이 매우 자주 사용되므로 '초성+중성' 또는 중성 'ㄷ' 생략 형태의 점자 문장에서 이어서 나오는 점자 부호를 우선 순위에 따라 항상 중성으로 번역하면 중성이 없는 음절을 표현할 수 없다. 따라서 이 경우에는 구두점을 나타내는 기존의 점자 부호를 사용하면 충돌을 피할 수 없으므로 이 경우에는 새로이 정의한 확장 점자 부호를 사용해야 한다.

3.3 중성 'ㄷ'의 생략 규칙 변경

개정한국점자통일안[8]에서는 '초성+ㄷ+중성'으로 구성되는 음절을 점자로 번역할 때 중성 'ㄷ'를 생략하도록 한다. 그러나 중성 'ㄷ'를 생략한 점자 문장은 한글 문장으로 번역할 때 오류를 유발할 수 있다. 예를 들어 중성 'ㅅ'과 중성 'ㅈ'가 점자 부호(--34--)를 같이 사용하므로 '났습니다'를 점역한 점자 문장

(1-4-)(-34-)(---6)(2-4-6)(12---)(1-4-)(1-3-5)(-2-4--)

를 역점역하면 (-34-)점이 3.2절의 우선 순위에 의해 중성 'ㅈ'로 인식되어 '네습니다'로 번역되어야 한다. 그러나 이것은 분명히 오류이므로 중성 'ㄷ'에 해당하는 점자 부호 (12--6)를 생략하지 않는 것이 역점역시의 오류를 줄일 수 있는 방법이 된다. 예를 들어 같은 문장을

(1-4-)(12--6)(-34-)(---6)(-2-4-6)(12---)(1-4--)
(1-3-5)(-2-4--)

로 표기하도록 하면 애매성이 해소될 수 있다. 여기서 '네습니다'는 실제로 사용되지 않는 단어이므로 맞춤법 검사기를 이용하면 이러한 단어로 번역되는 것을 방지할 수 있으나 한글의 특성상 비록 의미는 없으나 이 단어를 표현할 수는 있어야 할 것이므로 중성 'ㄷ'를 생략하지 않도록 점자 문법을 수정하는 것이 좋다고 본다.

3.4 한글 중성과 충돌하는 특수 기호

〈표 1〉에서 한글 중성과 구두점이 같은 점자 부호를 사용하고 있는 점자 부호는 초성 또는 중성 다음에 나올 때 3.2의 규칙에 의거하여 역점역을 하더라도 모호한 번역이 될 수밖에 없다. 이러한 충돌을 일으키는 점자 부호는 〈표 2〉에서 조사한 바와 같이 (-2--56), (-23--6) 및 (--3-56)의 세 가지이다.

점자 부호 (-2--56)은 중성 'ㄹ'과 마침표로 사용되며, 3.2의 규칙에서 이 점자 부호는 중성으로 우선 사용되므로 중성이 있는 한글 음절에서는 마침표를 의미하는 점자 부호로 (-2--56)을 사용하되, 중성이 없는 한글 음절에 이어서 쓰인 마침표를 표현할 때에는 느낌표(---456)에 마침표(-2--56)을 후치시킨 확장 점자 부호를 사용하도록 하였다. 마찬가지로 따옴표단기가 중성이 없는 음절에 이어서 사용될 경우에는 따옴표단기를 나타내는 점자 부호를 (---456)(-3-56)으로 결정하였다.

점자 부호 (-23--6)의 경우에는 따옴표 열기와 물음표의 두 가지 기호로 사용되므로 앞서의 경우와 같이 확장할 수 없다. 따라서 물음표의 경우에는 느낌표를 전치시켜 확장 부호를 만들고 따옴표 열기는 따옴표단기의 경우와 유사한 점자 부호 (-23-56)을 느낌표에 후치시켜 확장 부호를 구성하였다.

점자 부호(-3--6)은 한글 중성과 충돌하는 점자 부호는 아니지만 소괄호 열기, 소괄호단기 및 붙임표(hyphen) 등과 같이 여러 가지 의미로 사용되어 역점역시에 치명적인 오류를 발생시키는 부호이다. 따라서 이들 부호를 한영문 문장에서 중복해서 사용하는 경우에도 이를 점자로 번역한 점자 문장에서도 여러 가지로 해석될 수 있으므로 큰 혼란을 야기할 수 있

다. 따라서 점자 부호 (-3-6)이 나타내는 의미별로 점자 부호를 부여하자는 요구가 많으나 아직까지 개 정되지 못하고 논란이 계속되고 있다[10].

역점역의 경우에는 이들 각각의 의미를 분리하여 점자 부호를 지정하지 않으면 매우 잦은 혼동을 일으킬 수 있다. 예를 들어 소괄호 열기와 닫기를 플래그를 이용하여 구분하더라도 중첩 괄호일 경우나 열기가 반복되거나 닫기가 반복되는 경우에는 플래그에 의한 처리가 근본적으로 해결되지 못한다. 따라서 이 문제를 근본적으로 해결하기 위해서는 소괄호 열기와 소괄호닫기 및 붙임표를 각각 별개의 점자 부호로 구분하여야 한다. <표 3>은 이상의 방법으로 충돌이 발생할 때 사용할 네 가지 확장 점자 부호이다.

<표 3> 중성과 구두점이 충돌하는 점자 부호
<Table 3> Conflicts between korean jongsung and punctuation codes

점자부호	원래의미	확장방안	
(-2--56)	중성(ㅂ)	(-2--56)	
	마침표(.)	(--456)(-2--56)	
	영문닫기	(--456)(1-3--)	
(-23--6)	중성(ㅌ)	(-23--6)	
	따옴표열기(")	(--456)(-23--6)	
	물음표(?)	(--456)(-23--6)	
(-3-56)	중성(ㅎ)	(-3-56)	
	따옴표닫기(")	(--456)(-3-56)	
	영문열기	(--456)(1--5-)	
(-3--6)	붙임표	(--456)(-3--6)	
	소괄호열기	(--456)(123-56)	
	소괄호닫기	(--456)(-23456)	

3.5 특수 기호와 한글 초성의 충돌

특수 기호와 한글 초성 또는 초성이 없는 중성과의 충돌이 있는 경우에도 역점역이 애매해질 수 있다. 예를 들어 (---5-)는 쉼표 ‘,’와 초성 ‘ㄱ’로 해석될 수 있다. 따라서 이것을 방지하기 위해서 (-3-5-)에 이어서 중성이나 중성 점자 부호가 이어지는 경우에는 (-3-5-)를 초성 ‘ㄱ’로 간주하고 공백이나 다른 특수 기호들이 있을 때에는 쉼표 ‘,’로 간주하는 방법을 이

용하여 해결할 수 있다. 그러나 아래 문장 ‘아버지,어머니,오빠’는 ‘아버지, 어머니, 오빠’와 같이 띄어써야 한다.

을 통일 점자안에 따라 점역한 후 역점역하면,

‘아버지러머니로빠’는 ‘아버지, 어머니, 오빠’와 같이 띄어써야 한다.

로 번역되어 점역 및 역점역 이전의 문장이 지닌 의미를 잃어버리게 된다. 따라서 이러한 현상을 방지하기 위해서는 쉼표에 이어서 한글 중성이나 중성이 오는 경우에는 쉼표를 (---5-)를 사용하지 않고 확장 점자 부호를 사용하도록 하였으며, 이외의 경우에는 원래의 점자 부호 (-3-5-)를 사용하도록 하였다.

느낌표의 경우에도 비슷한 경우가 발생한다. 예를 들어 문장

‘좋다!어때?’는 ‘좋다! 어때?’의 잘못이다.

을 점역한 후 역점역하면

‘좋다것때?’는 ‘좋다! 어때?’의 잘못이다.

으로 번역된다. 따라서 이러한 현상을 방지하기 위해서 느낌표의 경우에도 충돌이 일어날 경우에만 확장 점자 부호를 정의하여 사용하도록 하였다.

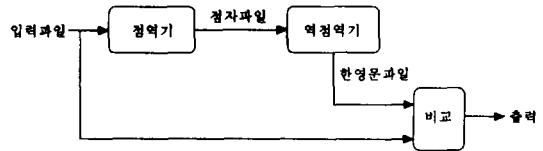
<표 4> 여러 가지 특수 기호로 사용되는 점자 부호
<Table 4> Special braille codes

기호	명칭	기존 점자 부호		확장 코드
		한글	영문	
,	쉼표	---5-	-2---	(--456)(-2---
!	느낌표	--456	-23-5-	(--456)(-23-5-
'	apostrophe		--3---	(--456)(--3---
<	less than	35+35	12--6	(--456)(12--6)
>	greater than	26+26	--345-	(--456)(--345-
+	plus	-2--6	--34-6	(--456)(-34-6)
*	star	35+35	1--6	(--456)(1--6)

점표와 느낌표의 경우에는 한글 문장과 영문 문장에서 각각 정의된 점자 부호가 다르다. 따라서 이들을 하나로 통일할 필요가 있다. 또한 한영문 문장에서 표현하는 점자 부호가 다른 점자부호는 일관성을 부여하기 위하여 <표 4>와 같이 확장하도록 하였다. <표 4>에서 어퍼스트로피(apostrophe)는 한글점자로 정의되지 않았으나 영문에서 자주 사용되며, 따라서 한글에서도 사용될 가능성이 있으므로 이를 정의하였다. 또한 한글 문장에서 부동호 ‘<’와 별표 ‘*’가 동일한 점자 부호를 사용하므로 확장 부호에서는 느낌표에 영문 기호에 해당하는 점자 부호를 후치시켜 구분하도록 하였다.

3.6 한글 문장에서 사용 가능한 아스키 (ASCII) 부호
 컴퓨터를 이용한 문서 편집이 보편화되면서 그 동안 한글 문장에서는 사용되지 않던 아스키(ASCII)기호들을 문장 중에 사용하는 경우가 생긴다. 예를 들면, 대괄호 열기 ‘[’, 대괄호 닫기 ‘]’ 등의 특수 부호가 그것이다. 이들 특수 기호는 컴퓨터용 점자 기호나 영문 점자 부호로는 정의되어 있으나 한글 문장 내에서 사용하는 경우에 대해서는 정의되어 있지 않다. 또한 이들을 한글 문장에서 사용하였을 경우 아스키 점자 부호로 번역한 후 역점역하면 여러 가지 충

돌을 야기할 수 있다. 따라서 이러한 오류를 방지하기 위하여 충돌을 피할 수 없는 경우에는 <표 5>와 같이 정의된 확장 부호를 사용하도록 하였다. <표 5>에서 볼 수 있듯이 이들도 의미 없는 부호인 느낌표를 사용하여 점자 부호를 확장하였다.



(그림 1) 점자 부호체계의 실험 장치
 (Fig. 1) Testbed for a braille code experimentation

4. 실험 및 고찰

제 3절에서 제안한 확장 점자 부호의 역점역 기능을 확인하기 위하여 제안한 점자 부호체계를 지원하는 역점역기를 구현하여 점자 파일을 역점역하여 얻은 한영문 파일의 정확성을 확인하도록 하였다. 실험의 일관성을 위하여 기존 점자 부호체계와 확장 점자 부호체계에 따라 한영문 파일을 점역한 후, 점역 결과를 본 연구진이 개발하여 사용중인 점자 문서편집기인 ‘소리문’[3]으로 읽어들여 전자점자로 출력하면서 점역의 정확성을 확인하였다. 실험에 사용된 파일은 하이텔(Hitel)의 자료실에 있는 완성형 파일 18개를 선정하여 각 파일별로 점역의 정확성을 확인하였다. 18개 파일중의 6개는 2절에서 역점역시에 발생하는 오류의 종류를 파악하는데 사용된 6개의 파일을 포함하고 있다.

점역의 정확성을 확인한 후에는 점역하여 얻은 점자 파일을 다시 각각의 역점역기를 이용하여 한영문 파일로 번역하고, 역점역을 통하여 얻은 한영문 파일을 원래의 한영문 파일과 비교하여 역점역시의 오류의 종류와 이유를 확인하였다. 이때 두 파일은 크기가 약간 다른 것을 알 수 있었는데 그 이유는 점역시에 하이픈이나 공백이 추가되며, 역점역시에는 이들 공백과 하이픈이 제거되지 않기 때문이다. 따라서 두 파일을 비교할 때에는 공백이나 하이픈의 경우는 역점역의 오류로 계산하지 않았다.

<표 5> ASCII 용 점자 부호의 확장
 <Table 5> Extended braille codes

ASCII	기호	명칭	확장 코드
(-2-4-6)	{	left brace	(--456)(-2-4-6)
	[left bracket	(--456)(123-6)
(12-456)	}	right brace	(--456)(12-456)
]	right bracket	(--456)(1--456)
(--4--)	@	at	(--456)(--4--)
(-3456)	#	number	(--456)(-3456)
(12-4-6)	\$	dollar	(--456)(12-4-6)
(1--4-6)	%	percent	(--456)(1--4-6)
(--45-)	^	caret	(--456)(--45-)
(1234-6)	&	ampersand	--456)(1234-6)
(--456)	_	underline	(--456)(1-3--6)
(12--56)	\	back slash	(--456)(12--56)

기존 점자 체계 및 제안한 점자 체계에 따른 두벌의 점역기와 역점역기의 정확성을 확인한 후에는 (그림 1)과 같은 실험 환경을 구축하여 여러 가지 파일을 입력시켜 출력 결과와 원래의 파일을 비교하여 오류의 수를 측정하였다. 실험에 사용된 파일들은 여러 가지 기술 논문 및 일반 문서들이었으며, 각 파일별로 실험한 결과는 <표 6>과 같다. 실험에서 점역기와 역점역기는 기존의 점자 체계에 따라 한영문 전환부호를 별도로 사용하지 않고 3.2절의 번역 우선 순위만을 적용한 것이다. 그 결과, <표 6>에서 볼 수 있듯이 기존의 점자 부호체계를 사용한 점역기 및 역점역기의 경우에 발생한 오류율은 실험에 사용된 파일의 종류에 따라 변화하는 것을 알 수 있다.

예를 들어 5번 파일의 경우에는 40%에 이르는 오

류율을 보이고 있다. 그 이유는 5번 파일 내에 한영문 혼용이 매우 많기 때문이었다. 영문의 사용횟수가 적은 17번 파일의 경우에도 8%이상의 오류율을 보이고 있음을 알 수 있다. 이렇듯 기존 점자 체계 하에서는 전체적으로 오류율이 25%에 달하고 있으므로 점자 체계의 애매성이 극심함을 알 수 있다.

영문열기 및 닫기의 오류가 심각한 오류를 발생할 수 있다는 점은 3.1절의 방법만을 적용하여 실험한 경우로부터 알 수 있다. <표 6>에서 한영문 전환부호만을 삽입하고 나머지 점자 부호는 3.2절의 충돌 해소 방안을 적용하여 실험한 결과는 전체적으로 오류율이 13%이하로 낮아진 것을 보여준다. 따라서 한영문 점환 점자 부호를 명확히 규정함으로써 역점역에 따른 오류율을 크게 낮출 수 있음을 알 수 있다. 또한

<표 6> 실험 결과
<Table 6> Test results for various text files

파일 번호		파일 크기 (Byte)	기존 점자부호체계		한영문전환부호만 사용결과		확장된 점자부호체계	
번호	종류		오류개수	오류율(%)	오류개수	오류율(%)	오류개수	오류율(%)
1	한영문 혼용 기술논 문류	11,508	3,870	33.63	1,420	12.3	0	0
2		8,852	2,490	28.13	1,237	13.9	0	0
3		23,551	6,381	26.83	2,648	11.2	0	0
4		34,504	8,590	24.90	5,105	14.7	0	0
5		2,583	1,026	39.72	456	17.6	0	0
6		37,469	10,940	29.20	5,248	14.0	0	0
7		49,128	12,400	25.24	6,456	13.1	0	0
8		39,307	9,480	24.12	5,340	13.6	0	0
9		36,790	9,860	26.80	5,527	15.2	0	0
10		10,467	3,670	35.06	1,628	15.6	0	0
11	신문 사설류	1,936	420	21.69	348	18.0	0	0
12		959	160	16.68	148	15.4	0	0
13		1,013	390	38.50	247	24.4	0	0
14		1,665	220	13.21	196	11.8	0	0
15		2,223	380	17.54	329	14.8	2	0
16		5,943	690	11.61	573	9.6	0	0
17		10,397	870	8.37	741	7.1	0	0
18		6,842	940	13.74	868	12.7	0	0
총 계		305,079	75,019	24.59	38,515	12.6	0	0

본 논문에서 역점역시 충돌을 피할 수 없는 경우에만 확장 점자 부호를 사용하는 경우의 실험에서도 역점역 오류가 발생하지 않음을 알 수 있다. 다만 15번 파일의 경우에 발생한 2개의 오류는 문서에 포함된 전각 문자로 인한 오류였으며, 이는 한글점자에서 사용되어서는 안되는 부호이기 때문이었다. 따라서 전각 문자로 인한 오류를 감안하더라도 본 논문에서 제안한 점자 확장 방법은 점자 파일을 한영문 파일로 정확하게 번역할 수 있게 하는 방안의 하나로 사용될 수 있음을 보여준다.

5. 결 론

한영문 문장의 역점역은 그 동안 큰 필요성을 느끼지 못하였으나 점자 컴퓨터를 사용하는 시각장애인이 늘어남에 따라 점자 파일을 한영문 파일로 자동 번역하는 기법의 필요성이 대두되어 왔다. 그러나 기존의 점자 체계 하에서는 6점의 점자 기호로 한글 자모와 특수 기호를 표현하는데 따른 점자 부호의 중복정의로 말미암아 역점역시의 오류율이 매우 높아 실질적인 자동 번역이 불가능하였다. 따라서 본 논문에서는 점자 파일을 한영문 파일로 자동 번역을 하였을 때에도 충분히 가독성 있는 파일을 생성할 수 있도록 하는 점자 부호체계를 제안하였다. 특히 점자 부호체계를 제안함에 있어서 기존의 점자 사용자에게 미치는 혼동을 최소화하기 위하여 확장하거나 변경되어야 하는 점자 부호의 수를 줄이도록 하였다. 또한 컴퓨터를 사용함에 따라 한영문 문장 내에서 사용하게 되는 아스키 코드에 대한 점자 부호도 추가하였다.

여러 가지 종류의 파일에 대한 실험 결과, 한영문 전환점자 부호를 명확히 규정하여 사용하는 것으로도 역점역의 오류를 크게 낮출 수 있었다. 또한, 보다 세밀한 역점역을 가능하게 하기 위해서는 기존의 점자 부호를 사용함에 따라 충돌을 야기하는 경우, 예를 들어 종성이 없는 음절 다음에 오는 문장부호가 한글 종성으로 번역될 수 있는 경우에는 확장 점자 부호를 사용하도록 하면 역점역시의 오류를 충분히 배제할 수 있었다. 앞으로는 고어, 수학 및 과학용 점자 부호를 사용하고 있는 점자문장을 역점역하는 경우에도 확장 점자 부호의 사용을 줄이면서도 완벽한 역점역이 가능한 방안 에 관한 연구를 수행할 예정이다.

참 고 문 헌

- [1] 서천석, "한글점자 교본," 연합 세계 선교회 점역 봉사실, 1992.
- [2] 김흥기, "통신을 지원하는 점자 코드를 KSC5602 코드로 변환하는 번역기의 설계," 통신학술 연구 보고서, 통신연구원, 1995. 3.
- [3] —, "소리문 사용자 교범," 청주맹학교, 1995. 12.
- [4] 임안수, "한글점자," 제10회 시각장애아교육연구회 발표집, pp. 71-89, 1994. 8.
- [5] 임안수, "시각장애자의 정보접근," 시각장애연구, Vol. 7, pp. 3-24, 한국시각장애연구회, 1991.
- [6] 이충한, 안광일, 김석일, "문서편집기와 연동하는 '한점변환기'의 설계 및 구현," 컴퓨터과학연구, Vol. 2, No. 1, pp. 1-13, 충북대학교, 1994. 3.
- [7] 정보윤, "점자를 한글코드로 번역할 때 발생하는 모호성 해소에 관한 연구," 충북대학교 컴퓨터과학과 학사논문, 1995. 2.
- [8] 임안수, "개정 한국 점자 통일안," 한국점자연구위원회, 1994. 10.
- [9] 김석일, "시각장애인 컴퓨터와 텔레마케팅," 제1회 시각장애인 기술세미나 자료집, pp. 40-60, 한국맹인복지연합회, 1996. 12.
- [10] 박성수, "우리 나라 점자의 변천사와 현 문제점 및 개선 방향," 제1회 시각장애인관련 기술세미나 자료집, 충북대학교 기초과학연구소, pp. 9-13, 1997.



길 태 영

1989년 단국대학교 특수교육학과 졸업 및 학사학위취득
 1989년~현재 청주맹학교교사로 재직중
 1998년 충북대학교 대학원 전자계산학과 석사과정입학
 관심분야: 전산교육, 시각장애인 사용자인터페이스

김 석 일

1975년 서울대학교 전기공학과
학사학위 취득

1975년~1990년 국방과학연구소
선임연구원으로 근무

1985년~1989년 미국 North Caro-
lina State University에서
공학박사 취득

1990년~현재 충북대학교 컴퓨터과학과 부교수로
재직중

관심분야: 병렬처리 컴퓨터 구조, 슈퍼컴퓨팅, 병렬처
리언어, 이기종분산처리, 시각장애인 사용
자 인터페이스 등임

김 흥 기

1961년 연세대학교 수학과 졸업,
학사학위취득

1985년 중앙대학교 대학원 응용
수학과졸업, 이학박사
학위취득

1986년~1987년 캘리포니아 주
립대학 교환교수로 활동

현재 충북대학교 컴퓨터과학과 교수로 재직중

관심분야: 퍼지 및 카오스이론, 계산이론, 신경회로
망, 유전진화알고리즘, 지능시스템, 수치해
석, 전산교육