

release note : hangul-ucs 4.0 무엇이 달라졌나

도은이아빠

제 1 절 도입 (Introduction)

올해 (2007년) 들어 한국텍학회 (Korean \TeX Society, KTS)가 창립되는 등 KTUG의 활동은 새로운 단계로 접어들었다. 이와 더불어, hangul-ucs 패키지는 4.0 버전을 릴리스하게 되었는데, 기존의 3.2.2 버전까지와는 전혀 다른, 완전히 새로운 코드로 쓰다시피하여 일신된 면모를 갖추게 되었다. 이 글은 새 버전에서 달라진 점을 일반 사용자에게 소개하는 것이 목적이다. 즉, dhucs의 코드 해설이 목적이 아니며, 최종 사용자 입장에서 필요한 사항을 위주로 기술한다. 이 글은 공식 문서가 아니라, 개인적인 메모이다. 더 자세한 사항은 hangul-ucs의 패키지 문서를 참조하라.

우리는 이 버전을 릴리스하는 것을 매우 자랑스럽게 생각한다. 새 버전을 준비한 김도현 교수께 깊이 감사드린다.

제 2 절 배경

새로운 버전이 태동하게 된 데는, 역시 “모임과 식사”가 있다. 이전에도 그러했거니와 여러 사용자가 만나서 모임을 이끌어가는 것은 새로운 기능에 대한 영감이나 의견을 교환하는 데 더없이 좋은 기회가 되었다. 아마도 finemath가 구체화되는 결정적 계기는 어느 토요일 공주에서 있었던 모임에 참석하기 위해 몇몇 분이 조진환 교수의 차를 타고 차안에서 나누었던 대화가 아니었던가 하는 전언도 있는 것이다.

The Asian Journal of \TeX 을 준비하는 과정에서 hangul-ucs의 기능을 검토하고 개선할 충분한 기회를 가질 수 있었다. 원래 소박하게 시작했던 finemath 옵션은 특별히 거창한 것이 아니라 AJT에 쓰인 수식의 처리를 구현하려는 시도에서 출발한 것이었다. 조진환 박사가 일찍부터 이 문제에 깊은 관심을 기울였는데, 실상 이 “간격”에 대한 논의는 김강수의 “ \LaTeX 으로 책 만들기”라는 글에서부터 그 가능성과 필요가 검토되기 시작했던 것이다. 수식과 한글 사이에 약간의 간격을 주자는 의견을 기존의 dhucs에서 구현한 간단한 코드 snippet이 제시된 것은 올 3월이었고, AJT 조판과정에서 dhucs에 대한 개선이 가능한지를 테스트해보았다. 그러다가, (완전하지

않더라도) 이것을 패키지에 옵션의 하나로 포함시키자는 논의가 나와서 그 이름 (finemath)을 서울의 어느 어두운 거리에서 이주호, 조진환, 김강수 셋이 모여서 결정했던 것이다. 이 건의를 김도현 교수가 받아들여서 이런저런 시행착오를 거치면서 마침내 hangul-ucs의 중심적 feature로 만들어놓았다. 한편, 김강수가 AJT에 제출한 기고논문 “한글 문장부호의 조판 관행에 대하여”라는 글 역시 finemath 기능의 확장과 깊은 관계가 있다.

어쨌든, 한글 조판에 있어서 마이크로타이포그래피의 영역이 도입된 것은 아마도 한글 L^AT_EX 10년 역사에 중대한 전환점으로 기록될 가치가 충분하다고 생각한다.

제 3 절 hangul-ucs의 새로운 면모

3.1 ϵ -T_EX, utf8.def

hangul-ucs 패키지는 매우 복잡한 여러 프로젝트를 총칭하는 말이다. 그 가운데 dhucs는 핵심적인 한글 조판 엔진에 해당하는 스타일이다. hangul-ucs의 새 버전은 dhucs의 (전면적인) 변모와 더불어, 폰트 스펙 파일의 도입 및 oblivoir/memhangul의 버전업으로 이루어졌다.

종래, ver 3.2.2까지의 dhucs는 Dominique Unruh 씨의 unicode 패키지가 제공하는 utf8x.def라는 유니코드 정의 파일과 그 처리방법에 전적으로 의존하였다. 사실상 이 패키지와 유니코드 솔루션은 hangul-ucs가 처음 나올 수 있었던 배경이기도 하다. unicode 패키지(latex-ucs)는 대단히 훌륭하게 동작하였으나, 골치아픈 PrerenderUnicode 문제라든가, 근본적인 한계도 없지 않아 있었다.

dhucs 4.0은 이 unicode 패키지를 버리고, L^AT_EX 표준이라고 할 수 있는 babel utf8.def로 나아갔다. 그러나 애시당초 utf8.def를 사용할 수 없었던 문제점, 즉 유니코드 문자들이 moving arguments에서 모두 풀려버려서 에러가 난다든가 하는 점을 해결해야 했다. 이를 위하여 과감하게 ϵ -T_EX primitives를 도입하게 된 것이다. 즉, dhucs 4.0은 ϵ -T_EX 엔진이 아닌 오리지널 T_EX 엔진에서는 동작하지 않는다. 다행히 오늘날 거의 모든 T_EX 배포판의 엔진이 pdf- ϵ -T_EX이므로, 엔진의 교체가 가져올 혼란은 크지 않다고 생각한다. 엔진의 교체와 같은 것은 아마도 일반 사용자에게 큰 의미가 없을 것이다. 그러나 이를 통하여 이전보다 더 세련된 코드를 구현하게 되었다는 정도만 적어두기로 하자.

L^yX 1.5.0 유저에게는 이 유니코드 정의 파일의 변화가 중요한데, 말하자면 L^yX의 Document Setting–Language 설정에서 이전까지 utf8x를 선택하던 것을 utf8로 바꾸는 정도만 신경쓰면 될 것이다.

oblivoir와 memhangul은 dhucs의 변화에 맞추어 동작하도록 개선되었다. 약간의 새로운 타이포그래피 요소도 함께 도입되었다.

3.2 부가 패키지의 변화

새 버전의 hangul-ucs에서 다음과 같은 변화가 생겨났다.

1. dhucs-ucshyper 패키지가 삭제되었다. 이 패키지는 hyperref의 버전업과 함께 의미가 없어진 지 오래 되었다. 원래 한글 pdf bookmark를 구현하기 위해 필요했던 것이지만, 지금은 hyperref이 알아서 잘 해준다. 이전 버전에서 작성한 문서에 이 스타일을 엮은 것이 있었다면, hyperref을 부르는 코드로 바꾸어주면 될 것이다. 대략 다음과 같이 하면 된다.

```
% \usepackage{dhucs-ucshyper}
\usepackage[unicode,pdftex,<other options>]{hyperref}
%% pdftex과 같은 driver 지정을 반드시 할 것.
```

2. memhangul과 oblivoir가 hangul-ucs에 통합되었다.
3. [finemath]라는 새로운 옵션이 생겨났다. 이에 대해서는 절을 바꾸어서 상론한다. memhangul과 oblivoir도 이 옵션을 받아들이게 되어 있다.
4. 사용자 인터페이스 수준에서 자주 접근할 일이 없기는 하나, lucenc.dfu라는 정의 파일과 hfontspec.XXXX라는 폰트 스펙 파일이 도입되었다. 자간과 미세간격 및 폰트 명칭 등을 하나의 fontspec으로 취급한다. 판면에 변화를 주고자 하는 디자이너라면 이 폰트 속성을 적절하게 조절할 수 있을 것이지만, 아주 미세한 값을 다루는 것이므로 default를 그대로 쓰는 쪽이 안전하다고 생각한다. 기본 글꼴인 은글꼴을 위한 폰트스펙 파일의 명칭은 hfontspec.un이다.
5. dhucs-interword의 \interhchar 명령이 의미가 바뀌었다. 따라서 이전의 자간 조절을 위해 사용하던 \interhchar는 \interHchar로 바뀌어야 한다. 그보다는, 자간을 폰트 속성에 해당하는 것으로 다루게 되었으므로 사용자가 임의로 바꾸지 않는 것이 더 좋은 방법이다. 이제부터 dhucs-interword는 그 이름대로 단어간격을 제어하는 역할만을 하게 될 것이다.
6. 새로운 개행 알고리즘을 도입하여, 이전 버전보다 더 지능적으로 행나눔을 실현한다.

제 4 절 한글 타이포그래피의 구현

4.1 간격

hangul-ucs가 새롭게 채택한 “간격”에 다음 몇 종류가 있다.

1. 행중수식 (in-line math) 과 한글 사이의 간격
2. 영문자/숫자와 한글 사이의 간격
3. 괄호, 가운뎃점, 느낌표, 물음표, 꺾쇠기호 등 문장부호와 한글 사이의 간격
4. 문장부호 가운데 일부에 베이스라인 이동을 적용한 수직 간격
5. `trivcj` 에 적용된 일어/중국어 문장에서의 간격

이 가운데 (5) 번 문제는 이 글에서 논외로 할 것이다. (1)에서 (4)에 이르는 간격 조절은 `[finemath]` 옵션으로 (모두) 활성화된다. (1)과 달리 (2), (3), (4)는 필요하다면 그 기능을 끌 수도 있다.¹⁾

4.2 행중수식의 간격

실제 사용자는 `finemath` 옵션을 선언하는 것으로 위의 “간격제어” 기능을 호출할 수 있으므로, 여기서는 그 결과 (효과) 만을 시각적으로 제시하기로 하자. 다음 문장은 어느 대학 수학교재에서 가져온 예제이다. 왼쪽은 `finemath` 기능을 off 한 결과이고, 오른쪽은 그것을 활성화한 것이다.

집합 \mathbb{N} 이 위로 유계라 가정하자. 그러면 완비성 공리에 의하여 최소상계 $\alpha \in \mathbb{R}$ 가 존재한다. 실수 $\alpha - 1$ 은 α 보다 작은 수이므로 \mathbb{N} 의 상계가 아니고, 따라서 부등식 $\alpha - 1 < n \leq \alpha$ 를 만족하는 자연수 $n \in \mathbb{N}$ 이 존재한다. 그러면 $n + 1 \in \mathbb{N}$ 임에도 불구하고 $n + 1 > \alpha$ 이기 때문에 α 가 \mathbb{N} 의 상계라는 데 모순이다.

집합 \mathbb{N} 이 위로 유계라 가정하자. 그러면 완비성 공리에 의하여 최소상계 $\alpha \in \mathbb{R}$ 가 존재한다. 실수 $\alpha - 1$ 은 α 보다 작은 수이므로 \mathbb{N} 의 상계가 아니고, 따라서 부등식 $\alpha - 1 < n \leq \alpha$ 를 만족하는 자연수 $n \in \mathbb{N}$ 이 존재한다. 그러면 $n + 1 \in \mathbb{N}$ 임에도 불구하고 $n + 1 > \alpha$ 이기 때문에 α 가 \mathbb{N} 의 상계라는 데 모순이다.

실제로 수식이 많은 문서에서 이 미세한 간격은 수식의 가독성을 높여준다고 생각된다. 숫자나 영문자와 한글 사이에도 `finemath` 간격이 들어간다.

4.3 문장부호

문장부호의 조판은 그 동안 소홀히 다루어져 왔으나, 실제 문서의 품위를 결정하는 아주 중요한 요소이다. 특히 대부분의 문장부호를 영문 폰트의 것을 가져다 쓰는 관행을 발전시켜온 `LATEX` 한글 조판에서는 더욱 그러하다. 한글 폰트 안에도 문장부호가 있으므로, 영문 폰트의 것을 쓸 것인지 한글 폰트의 것을 쓸 것인지에 대해서도 근본적인 검토가 필요한 시점이다. 이 문제에 대한 문제제기로 김강수의 글 (in AJT)을 참고할 수 있다.

1) 해제할 때는 `\disablehangulfontspec` 명령을 이용한다.

아무튼, 관행적인 영문 폰트의 문장부호를 쓴다 하더라도 그 위치를 조정하는 것이 이번에도 도입된 새로운 미세조정의 일부라고 한다. 이 위치 조정은 수직 위치와 수평 위치의 두 차원에서 이루어진다. 수직 위치는 베이스라인의 재조정이고 수평 위치는 간격을 부여하는 것이다.

괄호, 느낌표, 물음표 등에 대해서는 미세간격(finemath 간격)을 부여한다. 또한 전각 문장부호는 특별한 처리를 거쳐서 어떤 폰트를 사용하든 일정한 간격을 지니도록 설계하였다. 그리고 마침표(온점), 느낌표, 물음표에 대해서는 hfontspec에 지정된 값을 이용하여 수직 위치를 조절한다. 다음 문장부호의 사용을 비교해볼 수 있다. 왼쪽은 finemath를 활성화하지 않았을 때의 문장부호 식자 상태이고 오른쪽은 간격을 적용한 것이다.

마침표, 물음표 등 《문장부호》의 위치가 조정된다. 영문자(라틴문자)의 뒤에서는 조정되지 않는다. not adjusted. 자! 가자!	마침표, 물음표 등 《문장부호》의 위치가 조정된다. 영문자(라틴문자)의 뒤에서는 조정되지 않는다. not adjusted. 자! 가자!
질문있어요? “질문 있나요?”	자! 질문있어요? “질문 있나요?”

매우 미세한 간격이라 언뜻 눈에 잘 들어오지 않으나, 한 페이지를 조판해두고 비교하면 판면의 느낌이 많이 다른 것을 알 수 있을 것이다.

4.4 행나눔과 ‘고아’ 회피

개행(line-breaking)은 이미 dhucs 2.7.1에서부터 L^AT_EX의 문제점을 수정해 왔다. 새로이 도입된 행나눔을 예시하는 것은 생략하려 하나, 적어도 부적절한 행나눔으로 인한 귀찮은 수작업은 획기적으로 줄어들었을 것이라고 생각한다.

행나눔과 관련하여 특기할 만한 것으로, 이른바 ‘고아’ 문제라고 불리던 문단 마지막의 한 글자 한 행을 회피하도록 하는 것이 어느 정도 가능해졌다는 것이다.

자동 고아 회피는, 문단의 마지막에 한 글자와 종지부가 오고 그 다음 문자가 문단 끝 문자(\par)일 경우 동작한다. 만약 문단 끝 문자가 오기 전에 다른 매크로 등이 놓인다면 동작하지 않는다. 그러므로 실제로는 문단이 끝나지만 코드 시퀀스에서 문단 끝 문자를 두지 않는 각주와 같은 곳에서는 자동 고아 회피가 동작하지 않는다. 또한 한 글자 다음에 문장부호가 둘 이상 잇대어(예를 들면 .”) 오는 경우에도 자동 고아 회피가 동작하지 않는다.

자동 고아 회피 이외에, 고아 회피를 강제하기 위한 매크로 ‘\d’가 도입되었다. 이것은 ‘\nobreak d’와 동일하게 동작한다. 문단 마지막 글자가 ‘다’가 아닌 경우도 많겠으나, 역시 ‘다’가 가장 대표적이므로 이 매크로 하나를 두어 편의를 도모하고자 했다. 어떤 이유에선가 자동 고아 회피가 동작하지 않으면 이 매크로를 활용하면 되겠다.

4.5 Micro Typography²⁾

pdfTeX의 저자인 Hàn Thê Thành의 박사학위논문 *Micro-typographic extensions to the T_EX typesetting system*³⁾은 pdfTeX이라는 엔진으로 margin kerning과 font expansion을 가능하게 해주는 마이크로 타이포그래피 개념을 T_EX 세계에 도입하였다.

마진 커닝(margin kerning)이란, 여백 공간이 가지런히 보이도록 하기 위해 글자들을 텍스트 블록의 마진 쪽으로 아주 조금 이동시켜주는 테크닉이다. 마진 커닝을 적용하지 않으면 일부 문자들이 마진 경계가 시작하기 전에 끝나기 때문에 오히려 울퉁불퉁하게 보인다. 이것은 문장부호 끌어내기(hanging punctuation)와 비슷한 것이지만 문장부호만이 아니라 일반적인 글자에도 적용할 수 있다. 이를 적용함으로써 텍스트 블록의 모양을 현저히 개선할 수 있다.

폰트 확장(font expansion)이란 단어 간격이 좀더 균일하게 보이도록 하기 위해 폰트의 폭을 약간 좁게 혹은 넓게 만드는 테크닉이다. ... 폰트를 넓게 혹은 좁게 만듦으로써 개행 엔진이 더 나은 행나눔을 할 수 있게 된다.

한글 문서에 있어서 행나눔을 결정할 때 부득이하게 ‘병병한’ 공백을 남기고 개행해야 하는 경우가 빈번했다. 또한 문장부호가 행 끝에 걸리는 경우 여백의 모양이 균일하지 않게 보이는 것을 어찌할 수 없었다. 마진 커닝과 폰트 확장이 매우 요긴하게 요청되는 바였지만, 이를 쉽게 구현하기 어려웠으며, 간간이 마진 커닝 정도가 쓰이는 데 그쳤다. dhucs 4.0은 pdfTeX의 폰트 확장을 활용할 수 있게 한다.⁴⁾

이를 위해서는 다음 조건이 갖추어져야 한다.

1. pdf_latex으로 컴파일할 것
2. type 1 글꼴 혹은 트루타입 글꼴이 사용될 것
3. microtype 패키지에 의한 설정이 활성화될 것

그림 1은 폰트 확장을 적용하지 않은 경우와 그것을 적용한 경우의 행나눔과 스페이스 간격을 비교해본 것이다. 왼쪽 것이 font expansion을 적용하지 않은 것이고, 오른쪽 것이 그것을 적용한 것이다. 더 균일하고 아름다운 판면이 생성되는 것을 눈으로 확인할 수 있을 것이다.

샘플 문서의 소스 중에서 font expansion과 관련된 부분은 다음과 같다.

2) 이 절은 dhucsdoc에서 그 내용을 가져왔음.

3) Oct., 2000. Masaryk University. cf. <http://www.pragma-ade.com/pdf_latex/thesis.pdf>.

4) 김도현, “microtype을 아시나요,” <http://kts.ktug.kr/node/207>.

나는 여름 방학이 되어 집
에 내려 오면 한 번씩은 이
집을 찾는다. 이 집에는 나
보다 한 살 아래인, 열세 살
되는 누이뻘 되는 소녀가 있
었다. 실상 촌수를 따져 가
며 통내외까지 할 절척도 아
니지만, 서로 가깝게 지내는
터수라, 내가 가면 여간 반
가워하지 아니했고, 으레 그
소녀를 오빠가 왔다고 불러
내어 인사를 시키곤 했다. 소
녀의 몸매며 옷매무새는 제
법 색시꼴이 박히어 가기
시작했다. 그때만 해도 시골
서 좀 범절있다는 가정에서
는 열 살만 되면 벌써 처녀
로서의 예모를 갖추었고 침

나는 여름 방학이 되어 집
에 내려 오면 한 번씩은 이
집을 찾는다. 이 집에는 나
보다 한 살 아래인, 열세 살
이런 이야기 저런 이야기로
얼마 지체한 뒤에, 안 건넌
었다. 실상 촌수를 따져 가
며 통내외까지 할 절척도 아
니지만, 서로 가깝게 지내는
터수라, 내가 가면 여간 반
가워하지 아니했고, 으레 그
소녀를 오빠가 왔다고 불러
내어 인사를 시키곤 했다. 소
녀의 몸매며 옷매무새는 제
법 색시꼴이 박히어 가기 시
작했다. 그때만 해도 시골서
좀 범절있다는 가정에서는
열 살만 되면 벌써 처녀로
서의 예모를 갖추었고 침선

그림 1: Font Expansion

```
\usepackage[verbose=true]{microtype}
\DeclareMicrotypeSet{dhucsmicro}{encoding=LUC}
\UseMicrotypeSet[expansion]{dhucsmicro}
```

제 5 절 마치는 말

dhucs 4.0으로, 한글 L^AT_EX은 다른 어떤 워드 프로세서보다 나은 한글 문서를 작성할 수 있게 되었다. 마이크로 타이포그래피와 미세 간격을 도입한 것은 실로 획기적인 일이며, 이로써 한글 타이포그래피의 요구사항을 조금이나마 충족하게 된 것을 매우 기쁘게 생각한다. 이것은 KTUG 5년이 가져온 기술의 축적, 연구와 노력의 결과라고 할 것이다. KTS라는 공식적인 단체의 설립과 더불어 이러한 결과물을 내보일 수 있게 된 것도 특기해두어야 마땅할 것이다.

한글 문서 다운 한글 문서를 조판하게 하는 스타일을 제작하는 것은 실로 요원한 일이다. 이번 미세 타이포그래피의 도입은 그 한 걸음을 내디딘 것에 불과할 것이다. 개행과 같은 문제만 하더라도 언뜻 생각하면 별것 아닌 것처럼 보일는지 모르나, 그 동작의 경우를 분석하고 이를 구현하는 것이 만만한 문제는 아니었을 것이다. 이외에도 dhucs 새버전과 이에 기반한 memhangul 및 oblivoir의 세세한 변화가 많으나, 일일이 소개하는 것은 다른 문서로 미루자.

사용자 여러분의 다양한 의견과 테스트, 격려와 조언은 저자들에게 큰 기쁨과 힘이 된다.