

한글 문서를 텍으로: hmltotex

한글로 작성된 문서를 \LaTeX 으로 변환하는 프로그램

김환철

한양대학교

February 16, 2019

목차

- 1 일격필살팀 소개
- 2 이 프로그램을 만들게 된 계기
- 3 구현한 기능
- 4 수식의 변환
- 5 구현 예정인 기능
- 6 응용 가능성
- 7 프로그램 실행 방법

일격필살팀 소개

- 수능 수학 개념서와 문제집을 제작하는 팀 (Since 2012)

일격필살팀 소개

- 수능 수학 개념서와 문제집을 제작하는 팀 (Since 2012)
- 2015년까지 한글을 사용하다 2016년부터 \LaTeX 도입

일격필살팀 소개

- 수능 수학 개념서와 문제집을 제작하는 팀 (Since 2012)
- 2015년까지 한글을 사용하다 2016년부터 \LaTeX 도입
- 2018년 \LaTeX 으로 제작한 수학 참고서 출간
(『일격필살 기하와 벡터』, 『일격필살 N제』)

- 1 일격필살팀 소개
- 2 이 프로그램을 만들게 된 계기
- 3 구현한 기능
- 4 수식의 변환
- 5 구현 예정인 기능
- 6 응용 가능성
- 7 프로그램 실행 방법

이 프로그램을 만들게 된 계기

- \LaTeX 으로 넘어오면서 기존의 한글 문서들을 모조리 \LaTeX 으로 옮겨야 했음.

이 프로그램을 만들게 된 계기

- \LaTeX 으로 넘어오면서 기존의 한글 문서들을 모조리 \LaTeX 으로 옮겨야 했음.
- 한글 문서를 \LaTeX 으로 변환해주는 'hml-equation-parser' 라는 프로그램을 발견!

이 프로그램을 만들게 된 계기

- \LaTeX 으로 넘어오면서 기존의 한글 문서들을 모조리 \LaTeX 으로 옮겨야 했음.
- 한글 문서를 \LaTeX 으로 변환해주는 'hml-equation-parser' 라는 프로그램을 발견!
- 하지만...

$\sqrt{3x \over 4} + 1 \over 4$
 $\# \sqrt{3x \over 4} + 1 \over 4$

변환 전 : $\text{vec } \{a\} + \text{vec } \{d\}$

변환되지 않음!

변환 전 : $\text{vec } a$
 변환 후 : $\text{HULKVEC } a$

변환 전 : $123 \over 33$
 변환 후 : $123 \over 33$

변환 전 : $\{123\} \over \{33\}$
 변환 후 : $\# \text{frac} \{ 123 \} \{ 33 \}$

그러나

- *.HML 파일은 구조가 상당히 간단함!

그러나

- *.HML 파일은 구조가 상당히 간단함!
- 이 파일을 분석하면 변환 프로그램을 만들 수 있지 않을까?

그러나

- *.HML 파일은 구조가 상당히 간단함!
- 이 파일을 분석하면 변환 프로그램을 만들 수 있지 않을까?
- 직접 만들어보자!

개발 환경

- Visual Studio 2010, C++

HML 파일의 구조

- 일반 텍스트 : `<CHAR> ... </CHAR>`

HML 파일의 구조

- 일반 텍스트 : `<CHAR> ... </CHAR>`
- 수식 : `<SCRIPT> ... </SCRIPT>`

HML 파일의 구조

- 일반 텍스트 : `<CHAR> ... </CHAR>`
- 수식 : `<SCRIPT> ... </SCRIPT>`
- 문단 나눔 : `<P>`, `</P>`

HML 파일의 구조

- 일반 텍스트 : `<CHAR> ... </CHAR>`
- 수식 : `<SCRIPT> ... </SCRIPT>`
- 문단 나눔 : `<P>`, `</P>`
- 탭 문자 : `<TAB/>`
Shift+Enter : `<LINEBREAK/>`
Alt+Space : `<FWSPACE/>`
⋮

- 1 일격필살팀 소개
- 2 이 프로그램을 만들게 된 계기
- 3 구현한 기능**
- 4 수식의 변환
- 5 구현 예정인 기능
- 6 응용 가능성
- 7 프로그램 실행 방법

구현한 기능(수식)

- 분수, 제곱근, 지수, 밑첨자가 겹겹이 포함된 식을 (거의) 완벽하게 변환

구현한 기능(수식)

- 분수, 제곱근, 지수, 밑첨자가 겹겹이 포함된 식을 (거의) 완벽하게 변환
- 로만체와 이탤릭체, 선분, 벡터가 혼합된 수식을 간략화된 명령어를 이용하여 변환

구현한 기능(수식)

- 분수, 제곱근, 지수, 밑첨자가 겹겹이 포함된 식을 (거의) 완벽하게 변환
- 로만체와 이탤릭체, 선분, 벡터가 혼합된 수식을 간략화된 명령어를 이용하여 변환
- 인수가 중괄호로 둘러싸였든, 둘러싸이지 않았든, 인수의 범위를 (거의) 정확하게 인식하여 \LaTeX 코드로 변환

구현한 기능(수식)

- 분수, 제곱근, 지수, 밑첨자가 겹겹이 포함된 식을 (거의) 완벽하게 변환
- 로만체와 이탤릭체, 선분, 벡터가 혼합된 수식을 간략화된 명령어를 이용하여 변환
- 인수가 중괄호로 둘러싸였든, 둘러싸이지 않았든, 인수의 범위를 (거의) 정확하게 인식하여 \LaTeX 코드로 변환
- 쓸모없는 중괄호를 (거의) 제거한 코드로 변환

- 1 일격필살팀 소개
- 2 이 프로그램을 만들게 된 계기
- 3 구현한 기능
- 4 수식의 변환**
- 5 구현 예정인 기능
- 6 응용 가능성
- 7 프로그램 실행 방법

수식의 변환(1)

- e^{x+1}

수식의 변환(1)

- $e^{\hat{x}+1}$
→ e^{x+1}

수식의 변환(1)

- $e^{\hat{x}+1}$
→ e^{x+1}
- $e^{\hat{x}} + 1$

수식의 변환(1)

- e^{x+1}

→ e^{x+1}

- $e^x + 1$

→ $e^x + 1$

수식의 변환(1)

- e^{x+1}
→ e^{x+1}
- $e^x + 1$
→ $e^x + 1$
- $\sqrt{\frac{3x}{4} + \frac{1}{4}}$

수식의 변환(1)

- e^{x+1}

$$\rightarrow e^{x+1}$$

- $e^x + 1$

$$\rightarrow e^x + 1$$

- $\sqrt{\frac{3x}{4}} + \frac{1}{4}$

$$\rightarrow \frac{\sqrt{3x}}{4} + \frac{1}{4}$$

수식의 변환(1)

- e^{x+1}

$$\rightarrow e^{x+1}$$

- $e^x + 1$

$$\rightarrow e^x + 1$$

- $\sqrt{\frac{3x}{4}} + \frac{1}{4}$

$$\rightarrow \frac{\sqrt{3x}}{4} + \frac{1}{4}$$

- e^{x+ta}

수식의 변환(1)

- e^{x+1}

$$\rightarrow e^{x+1}$$

- $e^x + 1$

$$\rightarrow e^x + 1$$

- $\sqrt{\frac{3x}{4}} + \frac{1}{4}$

$$\rightarrow \frac{\sqrt{3x}}{4} + \frac{1}{4}$$

- e^{x+ta}

$$\rightarrow e^{x+ta}$$

수식의 변환(1)

- e^{x+1}
→ e^{x+1}
- $e^x + 1$
→ $e^x + 1$
- $\sqrt[4]{3x} + \frac{1}{4}$
→ $\frac{\sqrt[4]{3x}}{4} + \frac{1}{4}$
- e^{x+ta}
→ e^{x+ta}
- $e^{x+\tan}$

수식의 변환(1)

- e^{x+1}
 $\rightarrow e^{x+1}$
- $e^x + 1$
 $\rightarrow e^x + 1$
- $\sqrt[4]{3x} + \frac{1}{4}$
 $\rightarrow \frac{\sqrt[4]{3x}}{4} + \frac{1}{4}$
- e^{x+ta}
 $\rightarrow e^{x+ta}$
- $e^{x+\tan}$
 $\rightarrow e^{x+\tan}$

수식의 변환(2)

- \overline{AB} = “`rmbarAB`”, “`bar{rmAB}`”, “`bar\AB`”, “`rmbar\AB`”

수식의 변환(2)

- \overline{AB} = “`rmbarAB`”, “`bar{rmAB}`”, “`bar\AB`”, “`rmbar\AB`”
→ `\ovr{AB}` (`\overline{\mathrm{AB}}`의 줄임)

수식의 변환(2)

- \overline{AB} = “`rmbarAB`”, “`bar{rmAB}`”, “`bar\AB`”, “`rmbar\AB`”
→ `\ovr{AB}` (`\overline{\mathrm{AB}}`)의 줄임)
- \overrightarrow{AB} = “`rmvecAB`”, “`vec{rmAB}`”, “`vec\AB`”, “`rmvec\AB`”

수식의 변환(2)

- \overline{AB} = “`rmbarAB`”, “`bar{rmAB}`”, “`bar\AB`”, “`rmbar\AB`”
→ `\ovr{AB}` (`\overline{\mathrm{AB}}`의 줄임)
- \overrightarrow{AB} = “`rmvecAB`”, “`vec{rmAB}`”, “`vec\AB`”, “`rmvec\AB`”
→ `\vrm{AB}` (`\vec{\mathrm{AB}}`의 줄임)

수식의 변환(2)

- \overline{AB} = “`rmbarAB`”, “`bar{rmAB}`”, “`bar\AB`”, “`rmbar\AB`”
→ `\ovr{AB}` (`\overline{\mathrm{AB}}`의 줄임)
- \overrightarrow{AB} = “`rmvecAB`”, “`vec{rmAB}`”, “`vec\AB`”, “`rmvec\AB`”
→ `\vrm{AB}` (`\vec{\mathrm{AB}}`의 줄임)
- 줄임 표현은 `customDef.sty` 파일에 정의되어 있음 (ikps 문법)

수식의 변환(3)

- `{pmatrix{3&4#2&1}}`

수식의 변환(3)

- `{pmatrix{3&4#2&1}}`

$$\rightarrow \begin{pmatrix} 3 & 4 \\ 2 & 1 \end{pmatrix}$$

수식의 변환(3)

- `{pmatrix{3&4#2&1}}`

$$\rightarrow \begin{pmatrix} 3 & 4 \\ 2 & 1 \end{pmatrix}$$

- `{sqrt {{1} over {x-1}}} over {{1} over {x}} + {sqrt {x}} over {x-1}}`

수식의 변환(3)

- $\{\text{pmatrix}\{3\&4\#2\&1\}\}$

$$\rightarrow \begin{pmatrix} 3 & 4 \\ 2 & 1 \end{pmatrix}$$

- $\{\text{sqrt}\{\{1\}\ \text{over}\ \{x-1\}\}\}\ \text{over}\ \{\{1\}\ \text{over}\ \{x\}\} + \{\text{sqrt}\{x\}\}\ \text{over}\ \{x-1\}$

$$\rightarrow \frac{\sqrt{\frac{1}{x-1}}}{\frac{1}{x} + \frac{\sqrt{x}}{x-1}}$$

$\backslash\text{dfrac}\{\backslash\text{sqrt}\{\backslash\text{dfrac}\{1\}\{x-1\}\}\}\{\backslash\text{dfrac}\{1\}\{x\}+\backslash\text{dfrac}\{\backslash\text{sqrt}\{x\}\}\{x-1\}\}$

- 1 일격필살팀 소개
- 2 이 프로그램을 만들게 된 계기
- 3 구현한 기능
- 4 수식의 변환
- 5 구현 예정인 기능**
- 6 응용 가능성
- 7 프로그램 실행 방법

구현 예정인 기능

- 특수문자를 수식창에 입력하여 기호를 표현하는 경우
ex) 수식창에 ANGLE 대신 특수문자 \angle 을 삽입하여 각을 표현

구현 예정인 기능

- 특수문자를 수식창에 입력하여 기호를 표현하는 경우
ex) 수식창에 ANGLE 대신 특수문자 \angle 을 삽입하여 각을 표현
- 수식에서 LEFT와 RIGHT의 짝이 맞지 않는 경우
(수식 뒤에 ****이 출력됨)

구현 예정인 기능

- 특수문자를 수식창에 입력하여 기호를 표현하는 경우
ex) 수식창에 ANGLE 대신 특수문자 \angle 을 삽입하여 각을 표현
- 수식에서 LEFT와 RIGHT의 짝이 맞지 않는 경우
(수식 뒤에 ****이 출력됨)
- 표와 그림이 포함된 경우

구현 예정인 기능

- 특수문자를 수식창에 입력하여 기호를 표현하는 경우
ex) 수식창에 ANGLE 대신 특수문자 \angle 을 삽입하여 각을 표현
- 수식에서 LEFT와 RIGHT의 짝이 맞지 않는 경우
(수식 뒤에 ****이 출력됨)
- 표와 그림이 포함된 경우
- 자동으로 display 수식임을 판단하기

- 1 일격필살팀 소개
- 2 이 프로그램을 만들게 된 계기
- 3 구현한 기능
- 4 수식의 변환
- 5 구현 예정인 기능
- 6 응용 가능성**
- 7 프로그램 실행 방법

응용 가능성

- 소수의 \LaTeX 편집자만 있으면 한글 사용자와 협업하여 \LaTeX 을 활용할 수 있음.

응용 가능성

- 소수의 \LaTeX 편집자만 있으면 한글 사용자와 협업하여 \LaTeX 을 활용할 수 있음.
- 한글에 익숙한 사용자가 \LaTeX 으로 자연스럽게 넘어올 수 있게 함.

응용 가능성

- 소수의 \LaTeX 편집자만 있으면 한글 사용자와 협업하여 \LaTeX 을 활용할 수 있음.
- 한글에 익숙한 사용자가 \LaTeX 으로 자연스럽게 넘어올 수 있게 함.
- 한글 수식 기반 프로그램을 만들 때 유용하게 사용할 수 있음.

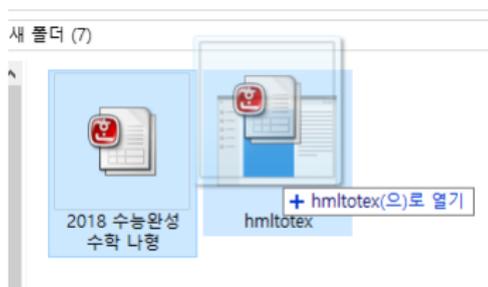
응용 가능성

- 소수의 \LaTeX 편집자만 있으면 한글 사용자와 협업하여 \LaTeX 을 활용할 수 있음.
- 한글에 익숙한 사용자가 \LaTeX 으로 자연스럽게 넘어올 수 있게 함.
- 한글 수식 기반 프로그램을 만들 때 유용하게 사용할 수 있음.
- 한글의 세계화(?)

- 1 일격필살팀 소개
- 2 이 프로그램을 만들게 된 계기
- 3 구현한 기능
- 4 수식의 변환
- 5 구현 예정인 기능
- 6 응용 가능성
- 7 프로그램 실행 방법

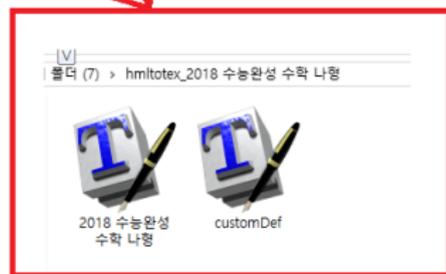
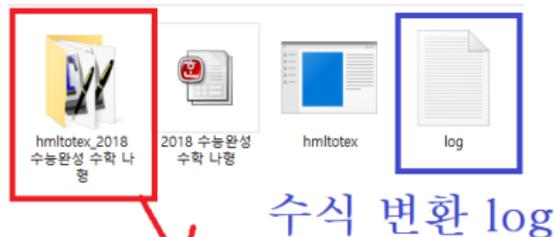
프로그램 실행 방법 (일괄 변경)

- 변환하고 싶은 hml파일을 exe 파일에 Drag-and-Drop



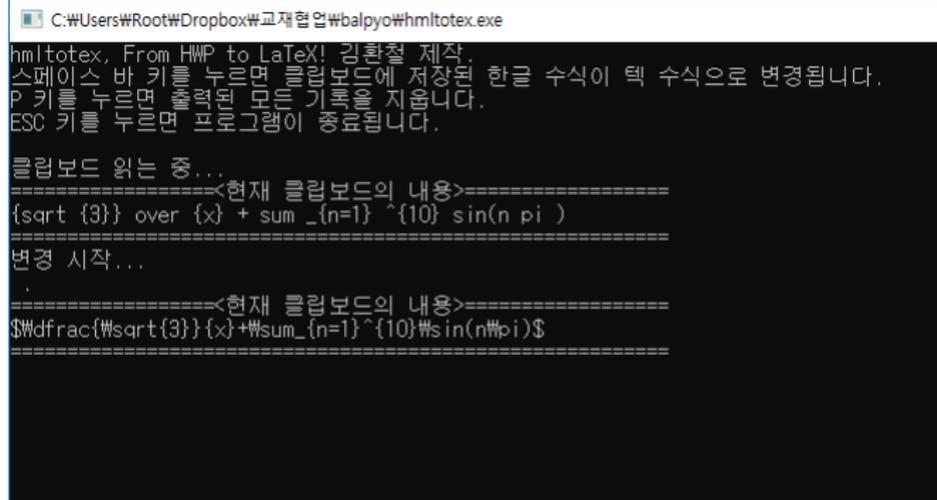
프로그램 실행 방법 (일괄 변경)

- "hmltotex_(파일이름)" 폴더 안에 tex파일 생성
exe파일과 같은 경로에 수식 변환 log 파일 생성



프로그램 실행 방법 (한 수식 변경)

- 클립보드에 한글 수식을 복사한 채로, 프로그램을 실행하고 스페이스 바 입력



```
C:\Users\Root\Dropbox\고재협업\balpyo\hmltotex.exe
hmltotex, From HWP to LaTeX! 김환철 제작.
스페이스 바 키를 누르면 클립보드에 저장된 한글 수식이 텍 수식으로 변경됩니다.
P 키를 누르면 출력된 모든 기록을 지웁니다.
ESC 키를 누르면 프로그램이 종료됩니다.

클립보드 읽는 중...
===== <현재 클립보드의 내용> =====
{\sqrt {3}} over {x} + sum _{n=1} ^{10} sin(n pi )
=====
변경 시작...
.
===== <현재 클립보드의 내용> =====

$$\frac{\sqrt{3}}{x} + \sum_{n=1}^{10} \sin(n\pi)$$

=====
```

- https://github.com/yeo32cjf/hmltotex_open

감사합니다.

Q & A