LATEX Symbol Search

Kerry Kairui Zhang

11 June 2015

2014403073

Search

Tools

TABLE OF CONTENTS

INTRODUCTION

BACKGROUND Problem description

Data acquisition
Data sources

Index, Query, Rank Indexing Querying Ranking

Comparison

Tools

Why use LATEX?

- ► Platform independent
- ► Academics
 - ▶ Mathematic
 - ▶ Non-latin alphabets
- ► Markup: separation of content and style
- ► Free
- ► Automatic referencing
- ▶ Beautiful

Tools

- ► Platform independent
- ► Academics
 - Mathematics
 - ► Non-latin alphabets
- ► Markup: separation of content and style
- ► Free
- ► Automatic referencing
- ▶ Beautiful

- ► Platform independent
- ► Academics
 - ► Mathematics
 - ► Non-latin alphabets
- ► Markup: separation of content and style
- ► Free
- ► Automatic referencing
- ▶ Beautiful

- ► Platform independent
- ► Academics
 - ► Mathematics
 - ► Non-latin alphabets
- ► Markup: separation of content and style
- ► Free
- ► Automatic referencing
- ▶ Beautiful

- ► Platform independent
- ► Academics
 - ► Mathematics
 - ► Non-latin alphabets
- ► Markup: separation of content and style
- ► Free
- ► Automatic referencing
- ▶ Beautiful

- ► Platform independent
- ► Academics
 - ► Mathematics
 - ► Non-latin alphabets
- ► Markup: separation of content and style
- ► Free
- ► Automatic referencing
- ▶ Beautiful

- ► Platform independent
- ► Academics
 - ► Mathematics
 - ► Non-latin alphabets
- ► Markup: separation of content and style
- ▶ Free
- ► Automatic referencing
- Beautiful

- ► Platform independent
- ► Academics
 - ► Mathematics
 - ► Non-latin alphabets
- ► Markup: separation of content and style
- ► Free
- ► Automatic referencing
- ▶ Beautiful

- ► Steep learning curve
- ► Complicated workflow
- Many commands
 - Packages, sectioning, referencing etc.
 - Symbol commands

- ► Steep learning curve
- ► Complicated workflow
- Many commands
 - Packages, sectioning, referencing etc
 - Symbol commands

- Steep learning curve
- Complicated workflow
- ► Many commands
 - ► Packages, sectioning, referencing etc.
 - Symbol commands

- Steep learning curve
- Complicated workflow
- ► Many commands
 - ► Packages, sectioning, referencing etc.
 - Symbol commands

- Steep learning curve
- Complicated workflow
- ► Many commands
 - ► Packages, sectioning, referencing etc.
 - Symbol commands

α	\alpha	θ	\theta	0	0	au	\tau
β	\beta	1 9	\vartheta	π	\pi	\boldsymbol{v}	\upsilon
γ	\gamma	γ	\gamma	$\boldsymbol{\varpi}$	\varpi	φ	\phi
δ	\delta	κ	\kappa	ρ	\rho	φ	\varphi
ϵ	\epsilon	λ	\lambda	Q	\varrho	χ	\chi
ε	\varepsilon	μ	\mu	σ	\sigma	ψ	\psi
ζ	\zeta	ν	\nu	ς	\varsigma	ω	\omega
η	\eta	ξ	\xi		_		_
Г	\Gamma	Λ	\Lambda	Σ	\Sigma	Ψ	\Psi
Δ	\Delta	Ξ	\Xi	Υ	\Upsilon	Ω	\Omega
Θ	\Theta	Π	\Pi	Φ	\Phi		J

Table 1: Greek Letters

INTRODUCTION

```
≤
     \leq
                                                  \equiv
                                                                   \models
                            \geq
≺
     \prec
                            \succ
                                                  \sim
                                                                   \perp
                                            \sim
≺
                      ≿
≫
     \preceq
                            \succeq
                                                  \simeq
                                                                   \mid
                                            \simeq
     \11
                                                                   \parallel
                            \gg
                                                  \asymp
\subset
     \subset
                      ٦
                                                                   \bowtie
                            \supset
                                                  \approx
                                            \approx
                                                             ×
⊆
     \subseteq
                            \supseteq
                                            \cong
                                                  \cong
                                                             \bowtie
                                                                   \Join^b
\sqsubset<sup>b</sup>
                            \sqsupset<sup>b</sup>
                                            ¥
                                                  \neq
                                                                   \smile
⊑
                                            ÷
     \sqsubseteq
                            \sqsupseteq
                                                  \doteq
                                                                   \frown
\in
     \in
                      ∋
                            \ni
                                            α
                                                  \propto
                                                             =
\vdash
                      \dashv
     \vdash
                            \dashv
                                             <
                                                 <
                                                             >
                                                                   >
```

Wikipedia: list of mathematical symbols

Wikipedia: list of mathematical symbols

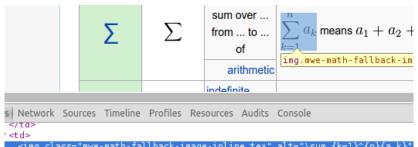
Basic symbols [edit]

Symbol	Symbol	Name				
-		Read as	Explanation	Examples		
in HTML in TeX		Category				
		addition				
		plus; add	4 + 6 means the sum of 4 and 6.	2 + 7 = 9		
		arithmetic				
+	+	disjoint union				
	the disjoin union of and		$A_1 + A_2$ means the disjoint union of sets A_1 and A_2 .	$A_1 = \{3, 4, 5, 6\} \land A_2 = \{7, 8, 9, 10\} \Rightarrow$ $A_1 + A_2 = \{(3, 1), (4, 1), (5, 1), (6, 1), (7, 2), (8, 2), (9, 2), (10, 2)\}$		
		set theory				
		subtraction				
		minus; take; subtract	9 – 4 means the subtraction of 4 from 9.	8 - 3 = 5		
		arithmetic				
	negative sign					
-	_	negative; minus; the opposite of	-3 means the negative of the number 3.	-(-5) = 5		
		arithmetic				

Symbol command in < img alt="\command" >

```
sum over ... from ... to ... \sum_{k=1}^{n} a_k \operatorname{mes}
                                 img.mwe-math-fallback-image-inli
                                              arithmetic
                                          indofinito
 Network Sources Timeline Profiles Resources Audits Console
'>...
'>...
٠,
td rowspan="1" bgcolor="#D0F0D0" align="center">...
td rowspan="1" bgcolor="#FFFFFF" align="center">
 <img class="mwe-math-fallback-image-inline tex" alt="\sum"</pre>
 eba63aafcecd3b92bd8ac4cd7dc0bca8.png">
1 de 14 de
```

Equation command in < img alt="\command" >



<img class="mwe-math-fallback-image-inline tex" alt="\sum {k=1}^{n}{a k}"</pre> be57a1f112b142efa33f2f7cb04d4b6e.png">

► Wikipedia list is not comprehensive

- ► Wikipedia list is not comprehensive
- ▶ Add Wikibooks

List of Mathematical Symbols [edit]

All the pre-defined mathematical symbols from the \TeX\ package are listed below. More symbols are available from extra packages.

Relation Symbols Symbol Script Script Symbol Symbol Script Symbol Script Symbol Script \parallel \nparallel > < \leq \bowtie \geq \dotea \asymp \bowtie **\11** >> \gg \equiv \vdash \dashv \subset \supset \approx \in \in € \ni \subseteq \supseteq \cong \smile \frown ⊅ \nsubsetea \nsupsetea \simea ⊨ \models ₫ \notin \sqsubset \sqsupset \perp \mid \sim \sqsubseteq \sqsupseteq \propto \prec \prec \succ \propto \preceq \neq \sphericalangle \preceq \succeq \neq ⋖ \measuredangle

Binary Operations										
Symbol	Script	Symbol	Script		Symbol	Script		Symbol	Script	
\pm	\pm	Π	\cap		\$	\diamond		\oplus	\oplus	
	\mp	U	\cup		Δ	\bigtriangleup		Θ	\ominus	

▶ Wikibooks

- ► Only symbol commands
- ► No description or examples
- ► Restrict to pre-defined symbols in the TeXpackage

- Wikibooks
 - ► Only symbol commands
 - ► No description or examples
 - ► Restrict to pre-defined symbols in the TEXpackage

- Wikibooks
 - Only symbol commands
 - ► No description or examples
 - ► Restrict to pre-defined symbols in the TFXpackage

- Wikibooks
 - Only symbol commands
 - ► No description or examples
 - ► Restrict to pre-defined symbols in the TEXpackage

Wikipedia

Each row is treated as a document

Country I	T-V	Name		
Symbol	TEX	Read as	Explanation	Examples
in <u>TEX</u>	Category Category			
		multiplication		
		times; multiplied by	3×4 or $3\cdot 4$ means the multiplication of 3 by 4.	7 · 8 = 56
		arithmetic		
×	\times	dot product		
		dot	$\mathbf{u}\cdot\mathbf{v}$ means the dot product of $\underline{\text{vectors}}\;\mathbf{u}$ and \mathbf{v}	$(1,2,5) \cdot (3,4,-1) = 6$
	\cdot	linear algebra		
		placeholder		
		(silent)	A · means a placeholder for an argument of a function. Indicates the functional	•
		functional	nature of an expression without assigning a specific symbol for an argument.	
		analysis		

- ▶ Command
- ▶ Kevword
- ▶ Content



WIKIPEDIA

Each row is treated as a document

Complete I	Name Name				
Symbol to T-V	_	Read as	Explanation	Examples	
<u>167</u>	in TEX Command Categor				
		multiplication			
		times; multiplied by	3×4 or $3 \cdot 4$ means the multiplication of 3 by 4.	7 · 8 = 56	
		arithmetic			
×	\times	dot product			
		dot	$\mathbf{u} \cdot \mathbf{v}$ means the dot product of <u>vectors</u> \mathbf{u} and \mathbf{v}	$(1,2,5) \cdot (3,4,-1) = 6$	
	\cdot	linear algebra			
		placeholder (silent) A · means a placeholder for an argument of a function. Indicates the functional			
			•		
		functional	nature of an expression without assigning a specific symbol for an argument.	11 . 11	
		analysis			

- ▶ Command
- ▶ Keyword
- ▶ Content



Wikipedia

Each row is treated as a document

C	T-V	Name		
Symbol to T-V	TEX	Read as	Explanation	Examples
ш <u>трх</u>	in TeX Command Category			
		multiplication		
		times; multiplied by	3×4 or $3\cdot4$ means the multiplication of 3 by 4.	$7 \cdot 8 = 56$
		arithmetic		
×	\times	dot product		
		dot	$\mathbf{u}\cdot\mathbf{v}$ means the dot product of $\underline{\text{vectors}}\;\mathbf{u}$ and \mathbf{v}	(1,2,5) · (3,4,-1) = 6
	\cdot	linear algebra		
	pla	placeholder		
		(silent)	A · means a placeholder for an argument of a function. Indicates the functional	•
		functional	nature of an expression without assigning a specific symbol for an argument.	
		analysis		

- ► Command
- ▶ Keyword
- ► Content



Wikipedia

Each row is treated as a document

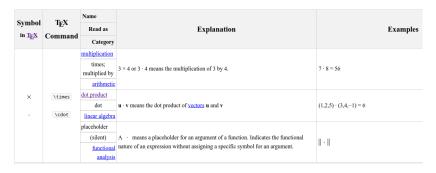
C	Name Name					
Symbol	Read	Read as	Explanation	Examples		
in <u>TEX</u>	Command Category					
		multiplication				
		times; multiplied by	3×4 or $3\cdot 4$ means the multiplication of 3 by 4.	$7 \cdot 8 = 56$		
		arithmetic				
×	\times	dot product				
		dot	$\boldsymbol{u}\cdot\boldsymbol{v}$ means the dot product of $\underline{vectors}\;\boldsymbol{u}$ and \boldsymbol{v}	(1,2,5) · (3,4,-1) = 6		
	\cdot	linear algebra				
		placeholder				
		(silent)	A · means a placeholder for an argument of a function. Indicates the functional	•		
		functional	nature of an expression without assigning a specific symbol for an argument.	11 - 11		
		analysis				

- ► Command
- ► Keyword
- ▶ Content



WIKIPEDIA

Each row is treated as a document

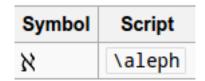


- ▶ Command
- ► Keyword
- ► Content



Wikibooks

Each row is treated as a document



Indexed fields

► Command

Wikibooks

Each row is treated as a document



Indexed fields

► Command

WIKIBOOKS

Each row is treated as a document



Indexed fields

► Command

Querying

Supports fuzzy search (maximum edit distance of 3)

QUERYING

Supports fuzzy search (maximum edit distance of 3)

LATEX Symbol Search

esplon	Search

Mouse over equations to display their TFX commands

Symbol in <u>T</u> <u>F</u> X	T _E X Command	Name Read as Category	Explanation	Examples
\mathbf{E} , ϵ and ε	E, \[\epsilon\]\ and \[\varepsilon\]			

QUERYING Can search content



QUERYING

Can search content

LATEX Symbol Search

the set of all real numbers Search

Mouse over equations to display their TFX commands

Symbol TEX in TEX Commi	ToV	Name		
	_	Read as	Explanation	
	Command	Category		
	real numbers			
\mathbb{R}		R;		
11/2		the (set of)	R means the set of real numbers.	
R		real numbers;	at means the set of real numbers.	
n (mathbright)	the reals			
		numbers		

► Okapi BM25F

- ▶ Probabilistic relevance model
- Document generative model

 - "probability of document being generated by query"

$$score(D, Q) = \sum_{i=1}^{n} IDF(q_i) \cdot \frac{f(q_i, D) \cdot (k_1 + 1)}{f(q_i, D) + k_1 \cdot (1 - b + b \cdot \frac{|D|}{avgdl})}$$

- ► Okapi BM25F
 - ► Probabilistic relevance model
 - Document generative model
 - $ightharpoonup \Pr(Q, D \mid R) = \Pr(D \mid Q, R) \Pr(Q \mid R)$
 - "probability of document being generated by query

$$score(D, Q) = \sum_{i=1}^{n} IDF(q_i) \cdot \frac{f(q_i, D) \cdot (k_1 + 1)}{f(q_i, D) + k_1 \cdot (1 - b + b \cdot \frac{|D|}{avgdl})}$$

- ► Okapi BM25F
 - ► Probabilistic relevance model
 - Document generative model
 - $Pr(Q, D \mid R) = Pr(D \mid Q, R) Pr(Q \mid R)$
 - "probability of document being generated by query"

$$score(D,Q) = \sum_{i=1}^{n} IDF(q_i) \cdot \frac{f(q_i,D) \cdot (k_1+1)}{f(q_i,D) + k_1 \cdot (1-b+b \cdot \frac{|D|}{avgdl})}$$

- ► Okapi BM25F
 - ► Probabilistic relevance model
 - Document generative model
 - $Pr(Q, D \mid R) = Pr(D \mid Q, R) Pr(Q \mid R)$
 - "probability of document being generated by query"

$$score(D, Q) = \sum_{i=1}^{n} IDF(q_i) \cdot \frac{f(q_i, D) \cdot (k_1 + 1)}{f(q_i, D) + k_1 \cdot (1 - b + b \cdot \frac{|D|}{avgdl})}$$

- ► Okapi BM25F
 - ► Probabilistic relevance model
 - Document generative model
 - $Pr(Q, D \mid R) = Pr(D \mid Q, R) Pr(Q \mid R)$
 - "probability of document being generated by query"

$$score(D,Q) = \sum_{i=1}^{n} IDF(q_i) \cdot \frac{f(q_i,D) \cdot (k_1+1)}{f(q_i,D) + k_1 \cdot (1-b+b \cdot \frac{|D|}{avgdl})}$$

- ► Combines search results, using priority order:
 - ► Exact match: commands and keywords
 - ► Or match: commands and keywords
 - ► Fuzzy: commands
 - ► Fuzzy: keywords
 - ► Fuzzy: content

- ► Combines search results, using priority order:
 - ► Exact match: commands and keywords
 - ▶ Or match: commands and keywords
 - ► Fuzzy: commands
 - ► Fuzzy: keywords
 - ► Fuzzy: content

- ► Combines search results, using priority order:
 - ► Exact match: commands and keywords
 - ► Or match: commands and keywords
 - Fuzzy: commands
 - ► Fuzzy: keywords
 - ► Fuzzy: content

- ► Combines search results, using priority order:
 - ► Exact match: commands and keywords
 - Or match: commands and keywords
 - Fuzzy: commands
 - ► Fuzzy: keywords
 - ► Fuzzy: content

- ► Combines search results, using priority order:
 - ► Exact match: commands and keywords
 - Or match: commands and keywords
 - ► Fuzzy: commands
 - ► Fuzzy: keywords
 - ► Fuzzy: content

- ► Combines search results, using priority order:
 - ► Exact match: commands and keywords
 - ► Or match: commands and keywords
 - ► Fuzzy: commands
 - ► Fuzzy: keywords
 - ► Fuzzy: content

LaTeX Symbol Search by Chris Clark

- ► Only uses keywords
- ► No fuzzy search
- ▶ No content search
- No explanation or examples

LaTeX Symbol Search

Don't know the LaTeX code for the symbol you need? Search for it in the box below!

You can enter the real mathematical name instead of the LaTeX code. For example, you can enter "gradient" to find the "\nabla" code.

=	=
≠	\neq
>	>
<	<
<u>≤</u>	\leq
\geq	\geq
~	\sim
«	\II
>>	\gg
<u>.</u>	\doted

LaTeX Symbol Search by Chris Clark

- ► Only uses keywords
- ► No fuzzy search
- ► No content search
- No explanation or examples

LaTeX Symbol Search

Don't know the LaTeX code for the symbol you need? Search for it in the box below!

You can enter the real mathematical name instead of the LaTeX code. For example, you can enter "gradient" to find the "\nabla" code.

epsion

LaTeX Symbol Search by Chris Clark

- ► Only uses keywords
- ► No fuzzy search
- ► No content search
- No explanation or examples

LaTeX Symbol Search

Don't know the LaTeX code for the symbol you need? Search for it in the box below!

You can enter the real mathematical name instead of the LaTeX code. For example, you can enter "gradient" to find the "\nabla" code.

the set of all real numbers

LaTeX Symbol Search by Chris Clark

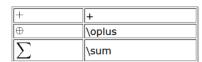
- ► Only uses keywords
- ► No fuzzy search
- ► No content search
- No explanation or examples

LaTeX Symbol Search

Don't know the LaTeX code for the symbol you need? Search for it in the box below!

You can enter the real mathematical name instead of the LaTeX code. For example, you can enter "gradient" to find the "\nabla" code.

sum



Tools

Pure Python

- ► BeautifulSoup: HTML parser
- ► Whoosh: search engine
- ► Flask: Web framework

Thank you!