

L^AT_EX Symbol Search

Kerry Kairui Zhang
2014403073

11 June 2015

TABLE OF CONTENTS

INTRODUCTION

BACKGROUND

Problem description

Data acquisition

Data sources

Index, Query, Rank

Indexing

Querying

Ranking

Comparison

Tools

PROBLEM DESCRIPTION

Why use L^AT_EX?

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

PROBLEM DESCRIPTION

Why use L^AT_EX?

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

PROBLEM DESCRIPTION

Why use L^AT_EX?

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

PROBLEM DESCRIPTION

Why use L^AT_EX?

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

PROBLEM DESCRIPTION

Why use L^AT_EX?

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

PROBLEM DESCRIPTION

Why use L^AT_EX?

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

PROBLEM DESCRIPTION

Why use L^AT_EX?

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

PROBLEM DESCRIPTION

Why use L^AT_EX?

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

PROBLEM DESCRIPTION

Problems with L^AT_EX

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

PROBLEM DESCRIPTION

Problems with L^AT_EX

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

PROBLEM DESCRIPTION

Problems with L^AT_EX

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

PROBLEM DESCRIPTION

Problems with L^AT_EX

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

PROBLEM DESCRIPTION

Problems with L^AT_EX

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

α	<code>\alpha</code>	θ	<code>\theta</code>	o	<code>o</code>	τ	<code>\tau</code>
β	<code>\beta</code>	ϑ	<code>\vartheta</code>	π	<code>\pi</code>	υ	<code>\upsilon</code>
γ	<code>\gamma</code>	γ	<code>\gamma</code>	ϖ	<code>\varpi</code>	ϕ	<code>\phi</code>
δ	<code>\delta</code>	κ	<code>\kappa</code>	ρ	<code>\rho</code>	φ	<code>\varphi</code>
ϵ	<code>\epsilon</code>	λ	<code>\lambda</code>	ϱ	<code>\varrho</code>	χ	<code>\chi</code>
ε	<code>\varepsilon</code>	μ	<code>\mu</code>	σ	<code>\sigma</code>	ψ	<code>\psi</code>
ζ	<code>\zeta</code>	ν	<code>\nu</code>	ς	<code>\varsigma</code>	ω	<code>\omega</code>
η	<code>\eta</code>	ξ	<code>\xi</code>				
Γ	<code>\Gamma</code>	Λ	<code>\Lambda</code>	Σ	<code>\Sigma</code>	Ψ	<code>\Psi</code>
Δ	<code>\Delta</code>	Ξ	<code>\Xi</code>	Υ	<code>\Upsilon</code>	Ω	<code>\Omega</code>
Θ	<code>\Theta</code>	Π	<code>\Pi</code>	Φ	<code>\Phi</code>		

Table 1: Greek Letters

\leq	<code>\leq</code>	\geq	<code>\geq</code>	\equiv	<code>\equiv</code>	\models	<code>\models</code>
\prec	<code>\prec</code>	\succ	<code>\succ</code>	\sim	<code>\sim</code>	\perp	<code>\perp</code>
\preceq	<code>\preceq</code>	\succeq	<code>\succeq</code>	\simeq	<code>\simeq</code>	\mid	<code>\mid</code>
\ll	<code>\ll</code>	\gg	<code>\gg</code>	\asymp	<code>\asymp</code>	\parallel	<code>\parallel</code>
\subset	<code>\subset</code>	\supset	<code>\supset</code>	\approx	<code>\approx</code>	\bowtie	<code>\bowtie</code>
\subseteq	<code>\subseteq</code>	\supseteq	<code>\supseteq</code>	\cong	<code>\cong</code>	\Join^b	<code>\Join^b</code>
\sqsubset^b	<code>\sqsubset^b</code>	\sqsupset^b	<code>\sqsupset^b</code>	\neq	<code>\neq</code>	\smile	<code>\smile</code>
\sqsubseteq	<code>\sqsubseteq</code>	\sqsupseteq	<code>\sqsupseteq</code>	\doteq	<code>\doteq</code>	\frown	<code>\frown</code>
\in	<code>\in</code>	\ni	<code>\ni</code>	\propto	<code>\propto</code>	$=$	<code>=</code>
\vdash	<code>\vdash</code>	\dashv	<code>\dashv</code>	$<$	<code><</code>	$>$	<code>></code>
$:$	<code>:</code>						

DATA SOURCES

Wikipedia: list of mathematical symbols

DATA SOURCES

Wikipedia: list of mathematical symbols

Basic symbols [\[edit\]](#)

Symbol in HTML	Symbol in TeX	Name	Explanation	Examples
		Read as		
+	+	addition	4 + 6 means the sum of 4 and 6.	2 + 7 = 9
		plus; add		
		arithmetic	A ₁ + A ₂ means the disjoint union of sets A ₁ and A ₂ .	A ₁ = {3, 4, 5, 6} ∧ A ₂ = {7, 8, 9, 10} → A ₁ + A ₂ = {(3,1), (4,1), (5,1), (6,1), (7,2), (8,2), (9,2), (10,2)}
		disjoint union		
		the disjoint union of ... and ...		
−	−	set theory	9 − 4 means the subtraction of 4 from 9.	8 − 3 = 5
		subtraction		
		minus; take; subtract		
		arithmetic	−3 means the negative of the number 3.	−(−5) = 5
		negative sign		
		negative; minus; the opposite of		
		arithmetic		

DATA SOURCES

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

		sum over ... from ... to ...	$\sum^n a_k$ mea
		arithmetic	
		indefinite	

Network Sources Timeline Profiles Resources Audits Console

```
>...</tr>
>...</tr>
>
  |
```

DATA SOURCES

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

	Σ	Σ	sum over ... from ... to ... of	$\sum_{k=1}^n a_k$ means $a_1 + a_2 +$
			arithmetic	<code>img.mwe-math-fallback-im</code>
			indefinite	

Network Sources Timeline Profiles Resources Audits Console

```

</td>
<td>
<img class="mwe-math-fallback-image-inline tex" alt="\sum_{k=1}^n{a_k}"
be57a1f112b142efa33f2f7cb04d4b6e.png">
  
```

DATA SOURCES

- ▶ Wikipedia list is not comprehensive

DATA SOURCES

- 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	Symbol	Script	Symbol	Script	Symbol	Script	Symbol	Script
$<$	<code><</code>	$>$	<code>></code>	$=$	<code>=</code>	\parallel	<code>\parallel</code>	\nparallel	<code>\nparallel</code>
\leq	<code>\leq</code>	\geq	<code>\geq</code>	\doteq	<code>\doteq</code>	\asymp	<code>\asymp</code>	\bowtie	<code>\bowtie</code>
\ll	<code>\ll</code>	\gg	<code>\gg</code>	\equiv	<code>\equiv</code>	\vdash	<code>\vdash</code>	\dashv	<code>\dashv</code>
\subset	<code>\subset</code>	\supset	<code>\supset</code>	\approx	<code>\approx</code>	\in	<code>\in</code>	\ni	<code>\ni</code>
\subseteq	<code>\subseteq</code>	\supseteq	<code>\supseteq</code>	\cong	<code>\cong</code>	\smile	<code>\smile</code>	\frown	<code>\frown</code>
\subsetneq	<code>\subsetneq</code>	\supsetneq	<code>\supsetneq</code>	\simeq	<code>\simeq</code>	\models	<code>\models</code>	\notin	<code>\notin</code>
\sqsubset	<code>\sqsubset</code>	\sqsupset	<code>\sqsupset</code>	\sim	<code>\sim</code>	\perp	<code>\perp</code>	\mid	<code>\mid</code>
\sqsubseteq	<code>\sqsubseteq</code>	\sqsupseteq	<code>\sqsupseteq</code>	\propto	<code>\propto</code>	\prec	<code>\prec</code>	\succ	<code>\succ</code>
\preceq	<code>\preceq</code>	\succeq	<code>\succeq</code>	\neq	<code>\neq</code>	\sphericalangle	<code>\sphericalangle</code>	\measuredangle	<code>\measuredangle</code>

Binary Operations

Symbol	Script	Symbol	Script	Symbol	Script	Symbol	Script
\pm	<code>\pm</code>	\cap	<code>\cap</code>	\diamond	<code>\diamond</code>	\oplus	<code>\oplus</code>
\mp	<code>\mp</code>	\cup	<code>\cup</code>	\bigtriangleup	<code>\bigtriangleup</code>	\ominus	<code>\ominus</code>

DATA SOURCES

► Wikibooks

- Only symbol commands
- No description or examples
- Restrict to pre-defined symbols in the $\text{T}_{\text{E}}\text{X}$ package

DATA SOURCES

- ▶ Wikibooks
 - ▶ Only symbol commands
 - ▶ No description or examples
 - ▶ Restrict to pre-defined symbols in the \TeX package

DATA SOURCES

- ▶ Wikibooks
 - ▶ Only symbol commands
 - ▶ No description or examples
 - ▶ Restrict to pre-defined symbols in the \TeX package

DATA SOURCES

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

INDEXING

WIKIPEDIA

Each row is treated as a document

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

Indexed fields

- ▶ Command
- ▶ Keyword
- ▶ Content

INDEXING

WIKIPEDIA

Each row is treated as a document

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

Indexed fields

► Command

► Keyword

► Content

INDEXING

WIKIPEDIA

Each row is treated as a document

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

Indexed fields

- Command
- Keyword
- Content

INDEXING

WIKIPEDIA

Each row is treated as a document

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

Indexed fields

- Command
- Keyword
- Content

INDEXING

WIKIPEDIA

Each row is treated as a document

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

Indexed fields

- Command
- Keyword
- Content

INDEXING

WIKIBOOKS

Each row is treated as a document

Symbol	Script
\aleph	<code>\aleph</code>

Indexed fields

- Command

INDEXING

WIKIBOOKS

Each row is treated as a document

Symbol	Script
\aleph	<code>\aleph</code>

Indexed fields

- Command

INDEXING

WIKIBOOKS

Each row is treated as a document

Symbol	Script
\aleph	<code>\aleph</code>

Indexed fields

- Command

QUERYING

Supports fuzzy search (maximum edit distance of 3)

QUERYING

Supports fuzzy search (maximum edit distance of 3)

L^AT_EX Symbol Search

Mouse over equations to display their T_EX commands

Symbol in <u>T_EX</u>	T _E X Command	Name	Explanation	Examples
		Read as		
		Category		
E, ϵ and ε	<code>E</code> , <code>\epsilonpsilon</code> and <code>\varepsilonpsilon</code>			

QUERYING

Can search content

QUERYING

Can search content

L^AT_EX Symbol Search

Mouse over equations to display their T_EX commands

Symbol in T_EX	T _E X Command	Name	Explanation
		Read as	
		Category	
\mathbb{R}	<code>\mathbb{R}</code>	real numbers	\mathbb{R} means the set of real numbers.
\mathbf{R}	<code>\mathbf{R}</code>	R; the (set of) real numbers; the reals	
		numbers	

RANKING

► Okapi BM25F

- Probabilistic relevance model
- Document generative model
 - $\Pr(Q, D | R) = \Pr(D | Q, R) \Pr(Q | R)$
 - “probability of document being generated by query”

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

RANKING

- ▶ 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”

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

RANKING

- ▶ 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”

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

RANKING

- ▶ 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”

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

RANKING

- ▶ 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”

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

RANKING

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

RANKING

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

RANKING

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

RANKING

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

RANKING

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

RANKING

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

COMPARISON

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.

Number of results: 273

=	=
≠	\neq
>	>
<	<
≤	\leq
≥	\geq
~	\sim
≪	\ll
≫	\gg
⋅	\cdot

COMPARISON

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.

Number of results: 0

COMPARISON

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 " ∇ " code.

Number of results: 0

COMPARISON

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.

Number of results: 3

$+$	<code>+</code>
\oplus	<code>\oplus</code>
Σ	<code>\sum</code>

TOOLS

Pure Python

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

Thank you!