

Основи на L^AT_EX: ръководство и справочник за българи

Бойко Банчев

Институт по математика и информатика, БАН

ПИСАНЕ НА КИРИЛИЦА И БЪЛГАРСКИ ЕЗИК

Командите на L^AT_EX за оформяне на специфично именувани части на текста като таблици, фигури, библиография, изброени с буквени означения списъци и др. подразбират английски език и в частност латиница. Същото се отнася и до сричкопренасянето. За да се използва друг език или два или повече езика съвместно се прибегва до макропакета **babel**, като наред с това се уточнява и какво кодиране се използва за представяне на съответната азбука или азбуки, ако те са различни от латиницата или са нейни изменени варианти. Нужните за всичко това команди се поставят в преамбюла на документа.

За да се пише грамотно на даден език е нужно да се познават и средствата на L^AT_EX, които дават възможност да се възпроизвеждат в документа графично особените елементи на писмения език, като кавички, тирета, букви с ударения или надредни знаци, специфични за езика знаци и др.

По-долу давам кратки сведения за особеностите на L^AT_EX във връзка с употребата преди всичко на български език, но отчасти и сведения, отнасящи се до други езици.

Команди в преамбюла на документа

За да пишем на български или на български и английски можем да включим следните команди:

```
\usepackage[cp1251]{inputenc}
\usepackage[english,bulgarian]{babel}
```

които е най-добре да се поставят в началото на преамбюла на документа – непосредствено след `\documentclass`. Първата от тях се задава, за да съобщим на L^AT_EX според коя *кодова таблица* да тълкува входния текст – в случая това е таблицата *cp1251*, стандартна у нас за кирилица в Windows. Ако използваме друга кодова таблица, да речем *mik*, *koi8-r* или *utf8*, вместо *cp1251* посочваме нея.

Чрез втората команда въвеждаме в употреба пакета **babel** и избираме езиците, които ще използваме в документа. Последният от тях става *текущо използван* от началото на документа до края му или докато не бъде изрично променен (вж. по-долу). В примера за начален текущо използван език е избран българският.

Може да се окаже нужно да се уточни и *шрифтовата таблица*, която L^AT_EX трябва да използва с кирилицата, T2A. Тази шрифтова таблица би трябвало да се подразбира, но понякога това не се получава и тогава следва да използваме командата

```
\usepackage[T2A]{fontenc}
```

която поставяме също в началото на преамбюла.

Команди и други особености в тялото на документа

Избор на език. Преминаването към друг текущо използван език става с командата `\selectlanguage`, например `\selectlanguage{english}` или `\selectlanguage{bulgarian}`.

Смяна на езика е нужна, за да „знае“ \LaTeX правилата на кой език да използва при пренасяне на думи и поставяне на различни означения. При смяна на езика се отчитат и други особености, включително променя се множеството от достъпни команди. В частност, ако както в примера променим езика на английски, за \LaTeX престава да бъде видима дори кирилицата и ако тя се среща в текста, компилирането завършва неуспешно.

Не е нужно да променяме езика, когато в текста на български се вмъкват само една-две думи от друг. Латиницата е видима и когато избраният език е български, а пренасяне при малък брой чужди думи едва ли се налага.

Кавички. Означенията за кавички са различни в различните езици. Ето как изглеждат те в „българския и немския“, в „английския“ и в «руския и френския» езици. Като използвате `babel` с български език като текущ можем, където е уместно, да поставяме *всеки от видовете кавички* и дори всички видове кавички в един и същ документ. Това може да стане като въвеждаме двойки от знаци, с които разполага всяка клавиатура и които визуално наподобяват желанния резултат. По-точно , , `` (две запетаи и две тъпи ударения) за първия вид, `` ' (две тъпи ударения и два апострофа) за втория и << >> (два знака „по-малко“ и два „по-голямо“) за третия вид кавички.

В някои случаи в българския език вместо обичайните кавички се употребяват т. нар. *горни запетаи*. Лявата и дясната са еднакви и се задават със знака ` , който въвеждаме като ` .

Ударения. Ударенията¹ се поставят с помощта на обичайните за \TeX средства – като команда към съответната буква. В българския език ударенията са леви и се задават с \` пред гласната, напр. за да получим *парà* пишем `пар\`а`.

(В други езици, напр. руски, ударенията са десни и се задават с \', напр. пишем \'уже, за да получим *уже*. В някои езици се използват и двата вида ударения.)

Въпросът с поставяне на различни видове ударения и други знаци над буквите е пълноценно решен в \TeX още със създаването на програмата преди около четвърт век. В някои т. нар. текстообработващи програми като MS Word обаче и днес не са предвидени средства за поставяне на ударения. Почти никой от потребителите им няма представа как да преодолее това ограничение (то и не е лесно). Дори някак да успеят да направят това *въпреки програмата*, забелязва се, че ударението в избрания шрифт отсъства или изглежда зле – силно отместено спрямо буквата встрани, по вертикала или и двете. Само с два или три от стандартните за MS Windows шрифтове (поне до варианта XP вкл.) ударенията се изобразяват задоволително добре.

Проблемът със задаване на ударението е породил напълно неуместната практика вместо думата „й“ (лично местоимение в дателен падеж или кратка форма на притежателно местоимение) да се пише „й̀“. Тази практика е толкова по-нелепа като се има предвид, че буквата „й“ (с или без надреден знак) е гласна, а „и кратко“ не само е различна, а и съгласна буква, която, тъй да се каже, случайно прилича по вид на „й̀“. Много по-приемливо е при невъзможност да се пише „й̀“ да оставим просто „и“.

Така или иначе, \TeX не ни принуждава да изкривяваме езика в никаква степен или форма – напротив, улеснява ни да пишем грамотно.

Тирета и чертици. В българския език се различават *тире* (–) и *чертица* (*малко тире*, -). В \TeX и \LaTeX второто можем да въведем със знака минус (-), а първото – с два такива знака последователно (--).

Пакетът `babel` с избран български език дава някои допълнителни възможности, свързани с използване на чертица и с пренасяне на думи. Командата “- може да се използва (както и универсалната за \TeX команда \-) за посочване на място в дума, където пренасянето е допустимо. Командите “= и “~ поставят чертица в двойни имена, полуслято написани сложни думи, прилагателни и наречия в сравнителна или превъзходна степен, съкращения като д-р и др., съответно с и без възможност за пренасяне след чертицата.

Руски букви, които не са български. Пишещият чрез \LaTeX може, като използва `babel` с български език, да пише и на руски, без да се налага изрично да задава и последния като

¹Езиковедите различават понятията *ударение* и *надреден знак с вид на ударение*, но използваният знак е един и същ

употребяван. В руската азбука има няколко букви, които не са в българската, но това не затруднява използването им наред с останалите.

Буквите \ddot{e} и \ddot{E} могат да се запишат с `\"e` и `\"E`, т.е. като използваме общите средства на \TeX за поставяне на [надбуквени означения](#).

В руската азбука присъстват също буквите υ , Υ , ε и ε , но те нямат свързани с получаването им в \TeX особености: стига да присъстват във входния текст, те се изобразяват в документа наред с буквите от българската азбука (и кодовите, и шрифтовата таблица за кирилица ги съдържат).

Ако нямаме удобен начин да въвеждаме специфичните руски букви в текста, можем да ги задаваме чрез поименни команди. За \ddot{e} , υ и ε използваме командите `\cyruo`, `\cyruery` и `\cyrerev`, а за \ddot{E} , Υ и ε – `\CYRYO`, `\CYRERY` и `\CYREREV`. (Поименни команди всъщност има за всички букви от всички варианти на кирилицата – руски, сръбски, украински и пр. \LaTeX заменя всяка буква от кирилицата със съответната $\grave{\text{h}}$ команда и едва след тази промяна на входния текст го подлага на същинско компилиране.)

Разбира се, за по-пълноценно използване на \LaTeX при писане на руски следва да се избира `babel` именно с този език, както и да се осигури възможност за непосредствено задаване на всички букви в текста.

Други знаци. Непосредствено писане. Някои знаци, които \LaTeX отпечатва с помощта на команди, се срещат и в кодовата таблица `cp1251`. Ако използваме именно нея, което е почти винаги за потребителите на \LaTeX с български език, тези знаци могат да се въвеждат не само с команди, а и непосредствено във входния за \LaTeX текст.

Следната таблица показва знаците и съответните им команди.

–	--	±	<code>\pm</code>	·	<code>\cdot</code>
—	---	¬	<code>\neg</code>	•	<code>\bullet</code> <code>\textbullet</code>
...	<code>\dots</code>	№	<code>\No</code>	°	<code>\textcircled{}</code>
‘	`	%	<code>\textperthousand</code>	†	<code>\dag</code> <code>\dagger</code>
’	’	§	<code>\S</code>	‡	<code>\ddag</code> <code>\ddagger</code>
„	„ (две запетаи)	¶	<code>\P</code>	©	<code>\copyright</code>
“	“	£	<code>\pounds</code>	®	<code>\textregistered</code>
”	”	¤	<code>\textcurrency</code>	™	<code>\texttrademark</code>
«	<<				
»	>>				

Въвеждането чрез команди е по-общ метод, тъй като не зависи от използваната кодова таблица, но прякото въвеждане прави текста по-нагледен и по-къс, а има и някои други предимства.

Разбира се, прякото въвеждане изисква текстовият редактор да има съответните възможности или те да бъдат предоставяни от клавиатурен драйвер на операционната система.²

Знаците `...`, `§`, `¶`, `†`, `‡`, `©`, `®` и `™` и произвеждащите ги команди от горната таблица могат да се използват където и да е в текста. Други знаци и команди имат особености по отношение на използването си, които е добре да имаме предвид.

Знаците `±` и `¬` се допускат само във формулен режим и могат да се задават само посредством съответните им команди, така че ако искаме да напечатаме `±`, трябва да напишем `\pm`. Командите `\cdot` и `\bullet` също са само за формулен режим, но в текстов режим да пишем самите знаци `·` и `•`. Командата `\textbullet` е равнозначна на непосредствено записан знак `•`.

Дългото и много дългото тирета, както и кавичките следва да се използват само в текстов режим.³

²Един специфичен за редактора *Vim* вариант, независещ нито от наличието на драйвер за кирилица или др., нито от операционната система, е описан в <http://www.math.bas.bg/bantchev/vim/kbdbul.html>.

³Всъщност във формулен режим използваме `'`, `''` и т.н., но те се изобразяват съответно като `'`, `''` и т.н.

Командите за %₀ и ∅ и самите знаци могат да се използват само в текстов режим и то само където текущият език е български. Командата \No пък се използва и в текстов, и във формулен режим при избран български език, но знакът № – само в текстов.

Знакът ° (градус) се допуска и в текстов, и във формулен режим, но командата \circ – само във формулен. Освен това тя произвежда не °, а о; за да получим °, трябва да повдигнем и намалим знака, пишейки $\text{\textasciicircum}\circ\text{\textasciicircum}$.

Командите \dagger и \ddagger произвеждат варианти на знаците † и ‡ само за формулен режим, където те се тълкуват като операции – около тях се оставя празно място.

Във връзка със знаците за кавички трябва да обърнем внимание на следното. Често наричаме „лява и дясна единични кавички“ двата знака на клавиатурата, които се намират съответно отляво в горния край и отдясно в средата ѝ. Всъщност това са *тъпо ударение* (`) и *апостроф* ('). Т_ЕX заменя тези знаци от входния текст с истинските лява и дясна кавички, които виждаме отпечатани. Той обаче приема като вход и истинските кавички, за които няма назначени клавиши, но както другите особени знаци, също могат да се поставят в текста.

Използването непосредствено на истинските кавички е удобно с това, че във входния текст ги виждаме така, както и в печатния; по-специално, ‘ и ’ са симетрични една на друга, за разлика от техните заместители ` и '. Същото се отнася и за “...” и „...“ в сравнение с ``...`` и ,,...``.

От друга страна, там, където тъпото ударение и апострофът са част от команда на Т_ЕX/Л_АT_ЕX, напр. в [команда за поставяне на ударение](#), трябва да използваме именно тях – не можем да ги заменим с лява и дясна кавички.

Знакът €. Въпреки че присъства в кодовата таблица *cp1251*, този знак не може да бъде отпечатан непосредствено. Всъщност той дори не трябва да присъства във входния текст, защото това предизвиква грешка при компилирането: знакът не е включен в шрифтовата таблица T2A, нито в някоя друга.

Извеждането на знака € изисква особен подход, тъй като той се появи сравнително скоро и не е предвиден в стандартните шрифтове и команди. Налице са различни възможности, повечето изискващи използване на допълнителен макропакет за Л_АT_ЕX. Може би най-добрият вариант е пакетът **eurosym** – той е прост за употреба и получаваният знак напълно съответства на определения чрез стандарт рисунък,⁴ а и освен изписването на €, **eurosym** предоставя някои допълнителни функции.

За включване на макропакета **eurosym** поставяме `\usepackage{eurosym}` в преамбула, а знака € в текста получаваме с команда `\euro`. Командата `\EUR` служи за извеждане на € като обозначение на конкретна сума; например `\EUR{107,99}` отпечатва 107,99 €.

Имена на функции. Командите `\tan` и `\arctan` за [формулният режим на Т_ЕX](#) не изписват имената на тригонометричните функции *тангенс* и *аркустангенс* по приетия у нас начин, а именно „tg“ и „arctg“. Затова българският режим на макропакета **babel** дава на разположение командите `\tg` и `\arctg`. Имаме също `\ctg` и `\arcctg` – за *котангенс* и *аркускотангенс*.

Разни. Следните команди на Л_АT_ЕX могат да се използват и с кирилица, което е удобно напр. при оформяне на заглавия:

<code>\MakeUppercase{Текст}</code>	⇒	ТЕКСТ
<code>\MakeLowercase{Текст}</code>	⇒	текст

За автоматично поставяне на текуща дата се избира някоя от командите `\today` или `\todayRoman` – получаваме съответно, да речем, 1 януари 2010 г. и 1. I. 2010 г. .

Някои особености при писане на български език, свързани с избора на интервали за разделяне на различни части на печатния текст, се разискват в раздела [Mind the gap: особени интервали](#).

⁴Вж. напр. [вебстраницата на автора на пакета](#) или <http://evertype.com/standards/euro>