

Лекция 8. Работа с текстом. Часть 2

Абзацы

Перед началом нового абзаца выполняем одно из двух действий:

- пропускаем пустую строку,
- выдаем команду `\par` (от слова paragraph).

Обычно в начале абзаца автоматически делается абзацный отступ. Его величину можно отрегулировать в преамбуле командой `\parindent`:

```
\parindent 20mm
```

Если абзацный отступ не нужен, то его можно подавить командой `\noindent` (без параметров).

Абзацы (продолжение)

В американской типографской традиции принято не делать абзацный отступ после заголовка раздела (т.е. в первом абзаце).

В русской традиции это не так.

Чтобы сделать абзацный отступ в начале раздела можно включить стилевой файл `indentfirst`:

```
\usepackage{indentfirst}
```

Другой способ: сделать в начале абзаца фиктивный пробел: `\ .`

Клей

В TeXе при формировании абзацев и страниц используется «клей» (`glue`).

Чтобы сделать текст красивым, TeX автоматически растягивает или сжимает горизонтальные и вертикальные интервалы в пределах заданных допусков.

Для расстояний между словами используется горизонтальный клей, для расстояний между абзацами — вертикальный клей.

Клей (продолжение)

Если уложиться в границы клея не удалось, то выдается сообщение `overfull` (переполнение) или `underfull` (слишком разреженный текст).

- Сообщение `overfull` (обычно для горизонтального клея) означает, что длина строки превысила заданные размеры. Указывается размер превышения в пунктах. Переполнение 2-3 pt обычно вполне допустимо.
- В сообщении `underfull` (обычно для вертикального клея) указывается параметр `badness` (плохость) в диапазоне 1000 – 10000.

Просмотр сообщений: файл протокола (log-файл).

Чтобы увидеть нарушение горизонтального клея:

```
\documentclass[draft]...
```

Выход в правое поле отмечается марашкой ■

Переносы

Информацию о правилах переноса слов TeX хранит в таблицах переносов.

При установке L^AT_EX создаются форматные файлы, в которые в частности закладываются эти таблицы.

В современных версиях MiKTeX в форматы автоматически включаются таблицы для основных языков, в том числе английского и русского.

Для русского языка используется таблица переносов Дмитрия Вулиса, City College of New York

Расстановка переносов вручную

При желании можно помочь TeXу, расставляя переносы вручную.

Возможные места переносов отмечают символами \-:

об\-ласть

функ\-ци\-о\-наль\-но-диф\-фе\-рен\-ци\-аль\-ный

Если слово повторяется часто, его можно включить в преамбулу:

\hyphenation{об\-ласть}

Если задать

творог_- или \mbox{творог}

это слово вообще не будет переноситься.

Разрывы строк

В случаях, когда разрывать подряд идущие слова нежелательно, можно использовать неразрывный пробел:

1999~г. Г.~М.~Фихтенгольц №~12

(получаем № 12)

Принудительно закончить строку, не начиная при этом нового абзаца, можно двумя способами:

- \linebreak — строка разрывается и растягивается с выравниванием по правому краю страницы. При этом она может получиться сильно разреженной.
- \ или \newline — строка разрывается, но не растягивается.

Разрывы строк (продолжение)

Для задания расстояний между строками или абзацами можно использовать команды

\[длина] \smallskip \medskip \bigskip

Например:

\[5mm]

Другой способ:

\vskip5mm

Разрывы страниц

Для начала новой страницы можно использовать команды:

- `\pagebreak` — растягивает страницу по вертикали,
- `\newpage` — просто обрывает.

Для запрета разрыва страницы используют команду `\nopagebreak`.

Центрирование и выравнивание строк

Для центрирования одной строки можно использовать команду

```
\centerline{...}
```

Для нескольких строк удобнее использовать окружение

```
\begin{center}  
.....  
\\  
.....  
\end{center}
```

Для выравнивания вправо используют команду `\rightline` или окружение `flushright`.

Для выравнивания влево можно использовать окружение `flushleft`.

Использование русского языка

Для поддержки русского языка лучше всего использовать пакет `babel` (по-немецки — Вавилон) разработки Йоханнеса Брамса.

Когда речь идет о русификации, можно выделить уровни:

1. Кириллические шрифты.
2. Система русских переносов.
3. Набор стилевых файлов, которые позволяют оформлять документ по правилам, принятым в России.
4. Русифицированные варианты специальных утилит, используемых TeXом.
Например, средства работы с библиографическими базами данных (BiBTeX) и средства создания указателей (индексов) — MakeIndex.

Babel решает первые две задачи — шрифты и переносы.

Особенности преамбулы для работы с русским языком

В преамбуле (после команды `documentclass`) нужно указать следующие параметры

1. Тип внешней кодировки
2. Тип внутренней кодировки
3. Подключить `babel` с указанием используемых языков
4. Если нужны дополнительные шрифты, подключить их.

Например:

```
\usepackage[cp1251]{inputenc}
\usepackage[T2A]{fontenc}
\usepackage[english,russian]{babel}
\usepackage{type1ec}
```

Внешняя кодировка

Внешняя кодировка — кодировка, которую использует операционная система.

Она определяется загруженной кодовой страницей (code page).

Задается командой:

```
\usepackage[опция]{inputenc}
```

Возможные опции: cp1251 (Windows), cp866 (DOS), koi8-r (Unix), massys (Mac) и некоторые другие.

Внутренняя кодировка

Это внутренняя кодировка системы TeX. Она близка к внешней, но отличается от нее в деталях (например, при работе с русским языком — кодировка буквы ё).

Задается командой:

```
\usepackage[опция]{fontenc}
```

Внутренняя кодировка (продолжение)

TeX использует два вида внутренних кодировок символов — текстовые и математические.

Нас будут интересовать текстовые кодировки — их имена начинаются с префикса «Т» или «ОТ».

Первоначально Кнут использовал 7-битную кодировку OT1, с помощью которой можно закодировать только символы латинского алфавита. Впоследствии были введены 8-битные кодировки для символов национальных языков.

В настоящее время имеются три внутренних кодировки для символов кириллицы (разработка Ольги Лапко и Александра Ходулева — LH): T2A, T2B и T2C.

Все они поддерживают русский алфавит, а также национальные алфавиты, основанные на кириллице. Кроме того, пользователи могут вводить местные (локальные) кодировки, например LCY.

Подключение пакета babel

При подключении babel указывают опции языка

```
\usepackage[english,russian]{babel}
```

их можно указывать также в команде documentclass:

```
\documentclass[english,russian]{article}
\usepackage{babel}
```

Язык, который используется по умолчанию, указывается *последним*.

Переключение языков

Переключение языков выполняется командами *в тексте документа*:

```
\selectlanguage{english}  
\selectlanguage{russian}
```

Если выбран русский язык, то английский текст набирать можно, но при этом не работает система английских переносов.

Если выбран английский язык, работа с русским языком невозможна.

Подключение дополнительных шрифтов

Дополнительные шрифты нужно подключать после выбора кодировок и подключения babel.

Например,

```
\usepackage{type1ec}
```

обеспечивает работу со шрифтами cm-super (Postscript Type 1).

Кроме cm-super в пакет MiKTeX входят Postscript шрифты гарнитуры «Литературная» — пакет literat (лицензионно чистый шрифт).

Некоторые пакеты можно найти в Интернете (psyr, FontC) — «серые шрифты».

Буквальное воспроизведение текста

Иногда желательно воспроизвести текст буквально, без дополнительного форматирования. Например, если требуется напечатать текст программы на алгоритмическом языке, или требуется напечатать название команды ТЭХа. Буквальное воспроизведение выполняется шрифтом пишущей машинки. Его можно оформить двумя способами.

- Большие участки текста можно поместить в окружение verbatim: `\begin{verbatim}`
...`\end{verbatim}`
- Короткий текст печатают с помощью команды `verb`. После команды указывается любой символ, который выступает как символ-ограничитель, затем следует текст, а затем вновь тот же символ. Эта команда удобна для воспроизведения команд ТЭХа:
`\verb#\ [4mm]#` дает `\ [4mm]`

Дефис и тире

Следует различать: дефис, короткое тире и длинное тире. По-английски: hyphen, n-dash, m-dash. Обозначения:

- -
- --
- --- (в русском тексте лучше "---).

Например: Жан-Жак (дефис), годы 1960–1970 (короткое тире), теорема Кронекера—Капелли (длинное тире).

Стили их использования могут различаться в разных странах и разных издательствах!

Дефис

Дефис (короткая черточка, hyphen) обозначается -. Употребляется:

- в переносах;
- в сокращениях (пример: арх-р);
- в сложных словах (пример: физико-математический);
- в составных именах и фамилиях (пример: Миттаг-Лёфлер, Франсуа-Мари);
- с числительными (пример: 1-й)

Строго различаются: Мусин-Пушкин (один и тот же человек, дефис), лемма Якубовича—Калмана (разные люди, длинное тире).

Короткое тире

Короткое тире (n-dash) обозначается --. Употребляется для обозначения диапазонов, пробелами не отделяется. Пример: С. ~10--15 дает С. 10–15.

Длинное тире

Длинное тире (m-dash) в английском тексте обозначается ---, в русском (babel) как "---. Использование:

- между словами по правилам грамматики данного языка (русского, английского),
- в названиях, составленных из **разных** фамилий. Примеры: теорема Больцано—Вейерштрасса, функционал Ляпунова—Красовского.

Замечание. Не допускается, чтобы строка начиналась со знака тире, например, — такой текст недопустим. Исключение: передача прямой речи. Например: — Ну и что?

Длинное тире: английский язык

Сложный вопрос — пробелы вокруг тире, он может решаться по-разному. В английском есть разные варианты:

слово---слово	дает:	слово—слово (Chicago style)
слово --- слово	дает:	слово — слово (Associated Press style)
слово -- слово	дает:	слово – слово (европейский стиль)

Длинное тире: русский язык

Тонкий набор русского текста (требуется пакет babel).

Основное тире: слово "--- слово Получаем: слово — слово

Тире в случае двух фамилий: Больцано"--~Коши Получаем: Больцано—Коши

Тире в начале предложения (прямая речь): "--* Вот так Получаем: — Вот так

Кавычки

В разных языках кавычки пишутся по-разному.

Вариант "текст" обычно считается неграмотным. Иногда его используют в научных отчетах.

В **английском языке**: ‘ ‘двойные’ ’ дает: “двойные”
 ‘одинарные’ дает: ‘одинарные’

В **русском языке** используют французские кавычки (ёлочки) и немецкие кавычки (лапки): <<ёлочки>> дает: «ёлочки»
 " ‘лапки’ " дает: „лапки“

В других языках есть другие варианты: >>шведские ёлочки>> дает: »шведские ёлочки«

Американские и британские кавычки

- В американском стиле основные кавычки — двойные, точки и запятые стоят внутри кавычек. В британском стиле основные кавычки — одинарные, точки и запятые стоят вне кавычек (если завершающая точка входит в цитату, ее ставят внутри кавычек).

Например: “Alas, poor Yorick,” he said по-американски ‘Alas, poor Yorick’, he said. по-английски

- Вложение кавычек: в американском — одинарные внутри двойных, в английском — двойные внутри одинарных.

Многоточия в тексте

Для многоточий в тексте используют команды: `\dots` и `\ldots`, которые эквивалентны.

Например: и так далее `\dots` дает: и так далее ...
 и так далее `\ldots` дает: и так далее ...