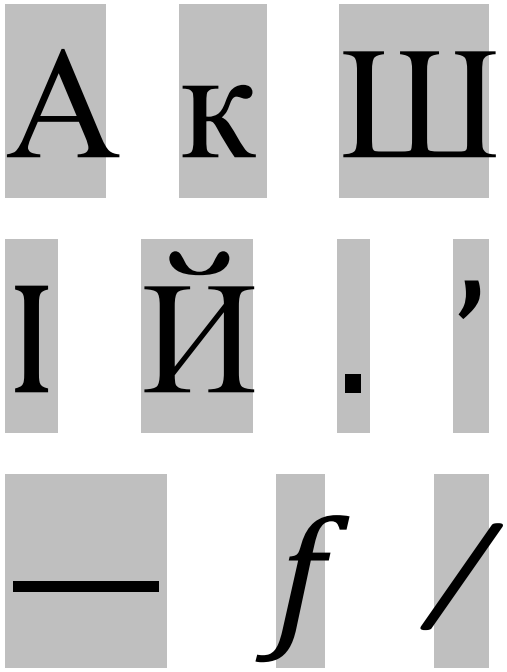
# III. Символы

- Принципы отображения текста
- Компьютерные шрифты
- Unicode
- Кодировки символов
- Специальные символы
- Подстановки символов
- Пробелы и переводы строк
- Оформление шрифта в CSS

# Принципы отображения текста




Образец шрифта

Образец шрифта

Образец шрифта

Образец шрифта

*Образец шрифта*

# Компьютерные шрифты

- Шрифт (font) — комплект векторных форм
- OpenType (.otf), TrueType (.ttf)
- Шрифты принадлежат операционной системе (все программы используют общий набор шрифтов)
- Способы получения шрифтов:
  - поставляются с операционной системой;
  - поставляются с прикладными программами;
  - устанавливаются вручную

# Классификация шрифтов

- Категории шрифтов:
  - по назначению: текстовые — декоративные
  - по ширине: пропорциональные — моноширинные
  - по наличию засечек: с засечками — без засечек
- Семейства шрифтов в CSS:
  - serif — пропорциональные с засечками
  - sans-serif — пропорциональные без засечек
  - monospace — моноширинные
  - cursive — декоративные, имитирующие пропись
  - fantasy — декоративные

# Unicode

- Консорциум Unicode: www.unicode.org
- Каталог символов Unicode:
  - алфавитные символы: www.unicode.org/charts
  - прочие символы: www.unicode.org/charts/symbols.html
- Около 100 000 символов
- У каждого символа — уникальный шестнадцатеричный номер и название
- Стандарт Unicode 5.0: www.unicode.org/versions/latest
  - классификация и свойства символов
  - кодировки символов
  - алгоритмы обработки символов и строк

# Некоторые диапазоны Unicode

| | | |
|---|---|---|
| основная латиница | U+0020…U+007E | A z 0 1 ! + @ & |
| расширенная латиница | U+00A0…U+00FF | Æ ó Ü ß Ø Þ |
| греческий | U+0370…U+03FF | α π Σ φ Ψ Ω |
| кириллица | U+0400…U+04FF | АяЄїЖґўЂѳҺҬ |
| пунктуация | U+2000…U+206F | – — ’ … ′ ″ |
| стрелки | U+2190…U+21FF | → ↑ ↗ ↳ ⇒ ↻ |
| математические символы | U+2200…U+22FF | ∫ ≤ ≈ ∀ ∈ ∞ |
| геометрические формы | U+25A0…U+25FF | ■ ▶ ◊ ◎ ◐ ◈ |
| пиктограммы | U+2600…U+26FF | ☺ ♣ ☮ ♋ ♛ ♯ |

# Кодировки символов

- Кодировка (charset, encoding) —
  способ представления символов в двоичном коде

| символ | код символа в Unicode | двоичный код в кодировке UTF-8 |
|---|---|---|
| Ф | ⇆ U+0424 | ⇆ 1101000010100100 |

- Кодировки Unicode:
  - UTF-8 — рекомендуется для веб-страниц;
    отключайте BOM (Byte Order Mark, Unicode Signature)
  - UTF-16, UTF-32
- Расширенные кодировки ASCII — устарели
  - windows-1251, koi8-r

# Сведения о кодировке

- В текстовом редакторе —
  при сохранении файла (для каждого файла)
  или в настройках текстового редактора (для всех файлов)
- В параметре charset элемента meta (не обязательно)
  - <meta http-equiv="Content-Type"
    content="text/html; charset=utf-8" />
- В конфигурации веб-сервера (.htaccess и др.)
  - AddDefaultCharset utf-8
- В объявлении XML
  - <?xml version="1.0" encoding="utf-8"?>

# Типографские знаки

| | | | |
|---|---|---|---|
| | неразрывный пробел | U+00A0 | Alt 0160 |
| « | левая кавычка | U+00AB | Alt 0171 |
| » | правая кавычка | U+00BB | Alt 0187 |
| × | знак умножения | U+00D7 | Alt 0215 |
| – | короткое тире | U+2013 | Alt 0150 |
| — | длинное тире | U+2014 | Alt 0151 |
| ’ | апостроф | U+2019 | Alt 0146 |
| … | многоточие | U+2026 | Alt 0133 |

# Редкие символы

- Чтобы символы отображались на веб-странице — задавайте в CSS шрифты, где есть эти символы
  - Lucida Sans Unicode (более 1700 символов)
  - Arial Unicode MS (более 50 000 символов)

```
<span class="unicode">ω ∈ ∅</span>

.unicode {
   font-family: 'Arial Unicode MS', 'Lucida Sans Unicode';
}
```

- Чтобы символы отображались в коде — используйте моноширинный шрифт, где есть эти символы
  - Everson Mono Unicode (около 5000 символов)

# Подстановки символов (character references)

- Обязательные подстановки:

  < <

  & &

- Необязательные подстановки:

  > >

  " "

  ' '

- Не используйте подстановки для символов, которые можно ввести непосредственно (  , — , &copy и т. д.)

# Пробелы и переводы строк

- Пробельные символы (white-space characters):
  - пробел (U+0020)
  - перевод строки, LF (U+000A)
  - возврат каретки, CR (U+000D)
  - горизонтальная табуляция (U+0009)
- перевод строки:
  - Windows: CR LF
  - Unix, Mac OS X: LF
- Пробельные символы сокращаются
- Автоматические переводы строк расставляются браузером
- Элемент <br /> — принудительный разрыв строки

# Свойство white-space — обработка пробельных символов

- Значения

| | пробельные символы | автоматические переводы строк |
|---|---|---|
| normal | сокращаются | есть |
| nowrap | сокращаются | нет |
| pre | сохраняются | нет |

- Элемент pre
  - по умолчанию white-space: pre;
  - поэзия, компьютерный код

# Свойство font-size — размер (кегль) шрифта

- Кегль шрифта — высота кегельной площадки
- Значения в процентах и em:
  - Для всех элементов, кроме html:
    1em = 100 % кегля шрифта родительского элемента
  - Для элемента html:
    1em = 100 % базового кегля шрифта
- Значения в других единицах — не рекомендуются (ухудшается масштабируемость шрифта)

# Оформление шрифта в CSS

- Свойства:
  - font-family — гарнитура
    font-family: гарнитура, семейство;
  - font-weight — насыщенность
    normal, bold, 100…900 (не поддерживаются, кроме 600)
  - font-style — наклон
    normal, italic
  - font-variant (не рекомендуется)
  - text-transform: uppercase; — верхний регистр
- Собирательное свойство font:
  - font: наклон насыщенность размер гарнитура; значения normal можно пропускать