

Основы создания шрифтов



Форматы шрифтов

Лекция №5

Лингвистическое обеспечение АСОИУ

К.т.н., доцент Филиппович Анна Юрьевна

Формат шрифта

В отношении компьютерного шрифта слово **формат** (format) используется в двух смыслах:

- Во-первых, формат определяется платформой, для которой шрифтовой файл создан (например: Apple Macintosh или Windows PC).
 - Во-вторых, формат шрифтового файла отражает способ представления и организации собственно типографической информации.
 - **Форматом** представления цифрового шрифта называется способ (стандарт) представления цифровой информации, образующей шрифт.
-

Шрифтовая машина

- ❑ **Шрифтовая машина** - совокупность определенного формата представления шрифтов и средств воспроизведения шрифтов, заданных в этом формате.

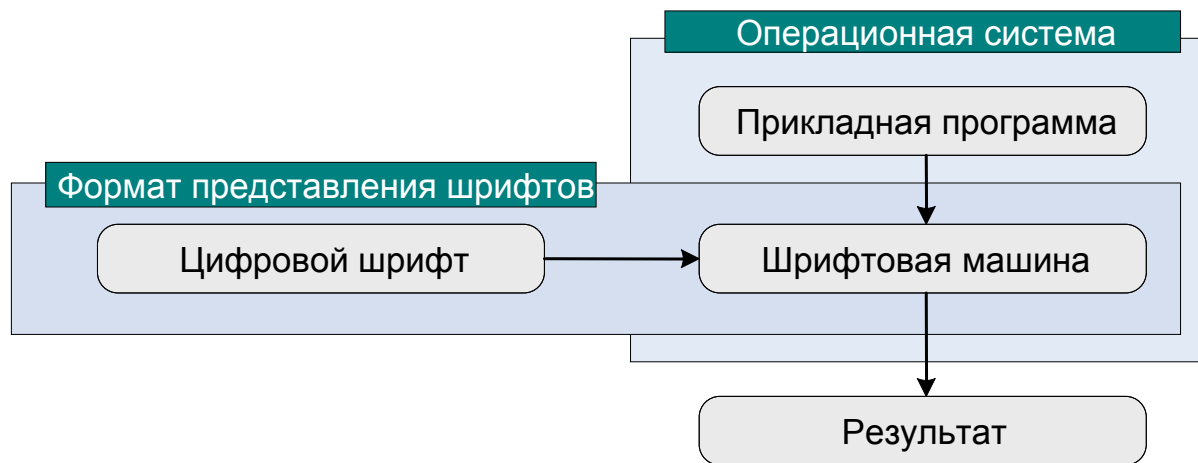


Схема реализации шрифтовой машины

Форматы контурных шрифтов



PostScript, Adobe Type 1



TrueType



OpenType

PostScript-шрифты

- Шрифты в этом формате **PostScript** основаны на языке описания страниц PostScript, и для их обработки и отображения требуется интерпретатор этого языка – RIP.

RIP

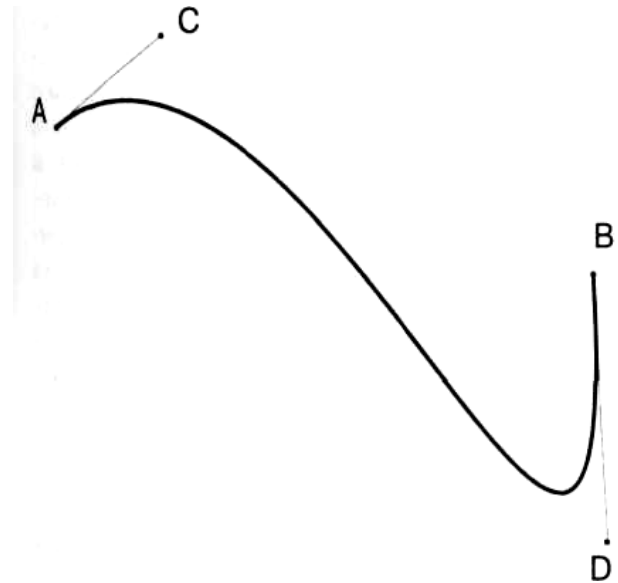
- ❑ **RIP** – raster image processor, растровый процессор.
 - ❑ У принтеров с высоким разрешением и фотонаборных автоматов он представляет собой отдельный процессор, предназначенный для преобразования PostScript-кодов в управляющие коды устройства.
 - ❑ Для устройств с низким разрешением, какими являются экран монитора и настольные офисные принтеры, PostScript-шрифты отображаются PostScript-интерпретатором, встроенным в операционную систему, или с помощью дополнительного приложения, например Adobe Type Manager (ATM).
-

PostScript-шрифты

- ❑ В настоящее время PostScript-шрифты стали стандартом в издательской отрасли.
 - ❑ Практически все устройства, используемые в издательствах снабжены растровыми процессорами. Естественно, такие процессоры лучше всего работают с PostScript-шрифтами.
-

PostScript-шрифты

- PostScript-шрифты построены с помощью кривых 3-го порядка, так называемых кривых Безье.



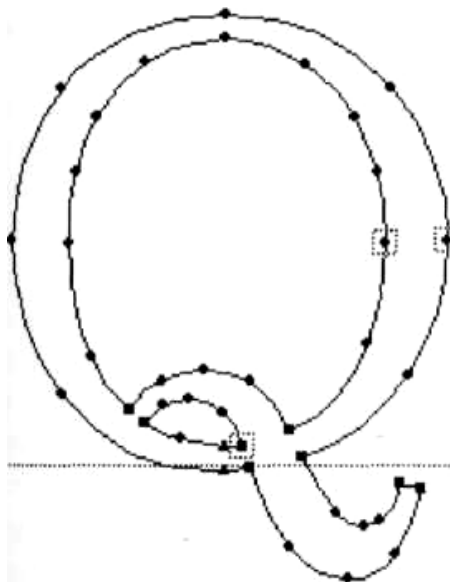
*Кривая Безье или
кубическая парабола*

$$R(t) = P_0(1-t)^3 + P_1(1-t)^2 + P_2 t^2(1-t) + P_3 t^3,$$

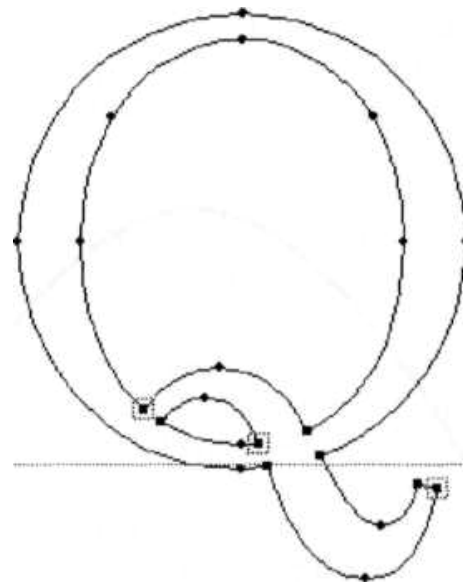
где $0 \leq t \leq 1$

PostScript-шрифты

- За счет использования кривых 3-го порядка достигается большая гладкость контуров и компактность шрифтовых файлов PostScript-шрифтов.



Литера TrueType-шрифта



Литера PostScript-шрифта

PostScript-шрифты

- Формат PostScript-шрифтов обычно содержит только стандартный комплект знаков (256) со стандартными параметрами.
-

Файлы PostScript-шрифта

Полноценный PostScript-шрифт обычно состоит из двух файлов:

- шрифтового (с расширением **PFB** или **PFA**);
- метрического - **AFM**-файла.

В некоторых случаях кроме AFM-файла требуется **INF**-файл, содержащий дополнительную информацию для регистрации шрифта в некоторых программах.

Файлы PostScript-шрифта

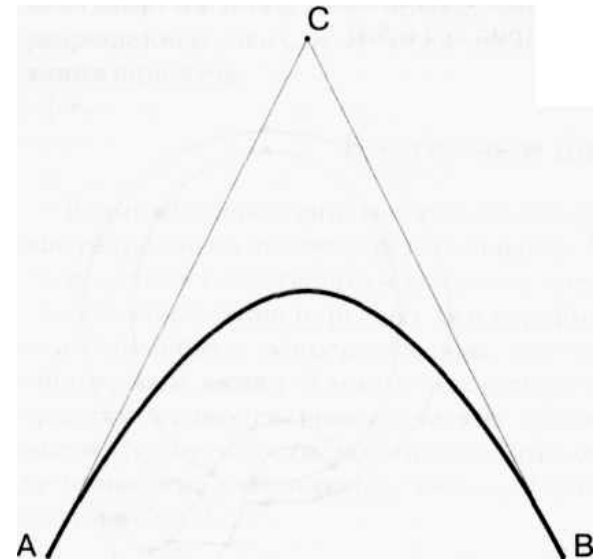
- ❑ **PFA** – PostScript Font File (шрифтовой файловый формат).
 - ❑ **PFB** – Print Font Binary (шрифтовой файловый формат, содержит информацию о контурах).
 - ❑ **PFM** – Print Font Metrics (шрифтовой файловый формат, содержит информацию о метрике).
 - ❑ **AFM** – Adobe Font Metrics (шрифтовой файловый формат, содержит информацию о метрике).
 - ❑ **INF** – файл, содержащий дополнительную информацию для регистрации шрифта в некоторых программах.
-

TrueType-шрифты

- ❑ TrueType-шрифты обычно распространяется без создаваемых вручную экранных (растровых) вариантов.
 - ❑ Экранное представление шрифта генерируется непосредственно из контура знака.
 - ❑ Формат TrueType допускает размещение более широкого комплекта знаков. В нем найдется место для альтернативных форм знаков и возможность контекстной замены знаков.
-

TrueType-шрифты

- TrueType-шрифты построены на кривых 2-го порядка.
- Каждый участок контура задается двумя точками – границами и направлением линии на каждой из границ



*Элементарная кривая
в шрифтах TrueType
(парабола второго порядка)*

OpenType-шрифты

- ❑ Формат **OpenType** является гибридным, он создан компаниями Adobe и Microsoft и сглаживает различия двух форматов, позволяя им сосуществовать в одном шрифтовом файле.
 - ❑ Он также дает возможность один и тот же шрифтовой файл использовать в обеих операционных системах Macintosh и Windows.
 - ❑ Шрифт OpenType может содержать данные формата TrueType, данные формата PostScript или (теоретически) обоих форматов.
 - ❑ ОС сама сортирует данные шрифта OpenType и выбирает только те из них, которые ее устраивают.
-

OpenType-шрифты

- ❑ Формат OpenType может содержать от 256 до 65 536 знаков. В один шрифтовой файл можно поместить все национальные кодировки, экспертные комплекты символов и практически любое количество дополнительных и альтернативных знаков.
 - ❑ Предлагает широкий набор дополнительных параметров, которые могут включаться, а могут и не включаться в каждый конкретный шрифтовой файл.
 - ❑ Одно из главных преимуществ новой технологии — поддержка расширенных типографских возможностей (т.н. OpenType features). Помимо собственно знаков шрифт OpenType может содержать правила использования этих знаков — позиционирование и подстановку одних знаков вместо других при определенных обстоятельствах.
-

OpenType-шрифты

- ❑ Главное в реализации расширенной типографики OpenType лежит в разделении числового кода знака (character) и его графемы (glyph).
 - ❑ Знак — это кодированная единица, упорядоченная в соответствии со стандартом Unicode, представляющая минимальную семантическую единицу языка, например букву.
 - ❑ Глиф — это графический образ знака.
 - ❑ Один знак может соответствовать нескольким глифам или один глиф может соответствовать комбинации нескольких знаков.
-

Файлы TrueType -шрифта

- TrueType-шрифт имеет расширение **ttf** или **ttc**.
 - OpenType-шрифты, содержащие данные PostScript имеют расширение **otf**, а шрифты на основе TrueType — расширение **ttf**.
-

Кодировка шрифта

- ❑ Сейчас для кодировки шрифтов используется международный стандарт Unicode. Он расширяет кодовую схему, включая знаковые комплекты для нелатинских алфавитов.
 - ❑ Большинство шрифтовых файлов стандарта Unicode двухбайтовые и могут содержать до 65000 знаков.
 - ❑ У первых шрифтовых файлов имелись ячейки только для 256 знаков, и данный комплект знаков остается стандартом для большинства шрифтовых файлов.
 - ❑ Иногда, даже если формат шрифта поддерживает 65000 знаков, он включает только 256 стандартных.
-

Различия в кодировках

- ❑ До того как Macintosh и Windows стали поддерживать стандарт Unicode, они использовали разные кодовые таблицы.
 - ❑ Таблицы совпадали в основной части комплекта ASCII, но различались в так называемые знаках старших разрядов знаках (high-bit).
-

Различия в кодировках

Различные ОС применяли разные подмножества комплекта Latin 1 в качестве своих стандартных комплектов знаков:

- Комплект системы Macintosh (MacRoman)
 - Комплект системы для Windows (Win ANSI).
-

Структура шрифтового формата

- ❑ Область заголовка.
 - ❑ Область описания метрических параметров.
 - ❑ Область описания общих элементов.
 - ❑ Область описания системы кодирования.
 - ❑ Область описания разметки символов.
 - ❑ Область описания символов.
-

Управление растеризацией СИМВОЛОВ

- Фундаментальной особенностью контурных шрифтов является отделение информации о форме символов от процесса их воспроизведения на растровом выводном устройстве.
-

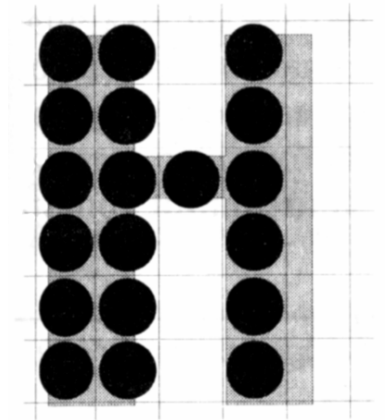
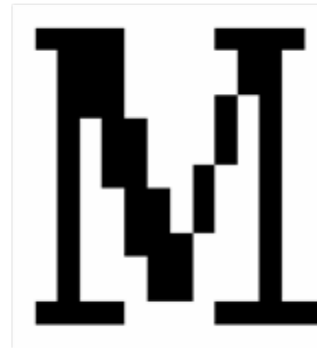
Растеризация символов

При воспроизведении каждого символа на растровом устройстве необходимо решить две задачи:

1. масштабировать контур символа до необходимого размера;
 2. активизировать все точки, попавшие во внутренние области этого контура, заполнить контур.
-

Проблемы растеризации

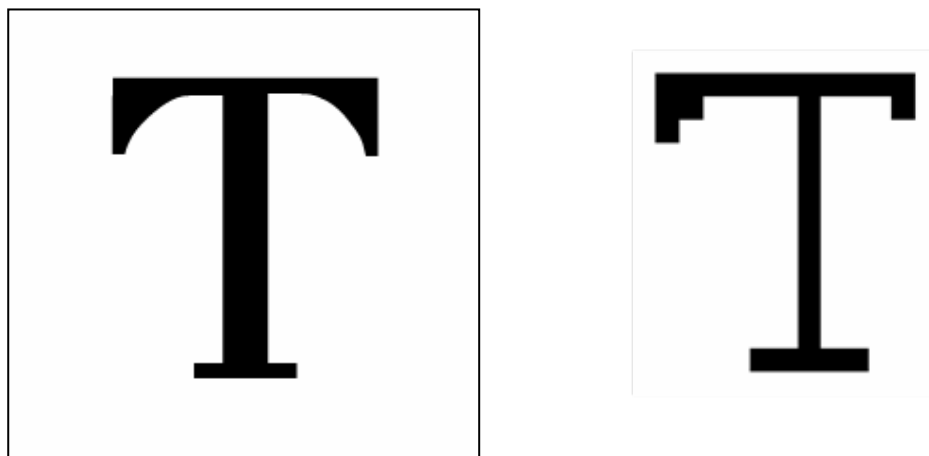
- Нарушение пропорций символа



Пример нарушения пропорций

Проблемы растеризации

- Нарушение симметричности некоторых символов



Пример нарушения симметричности

Проблемы растеризации

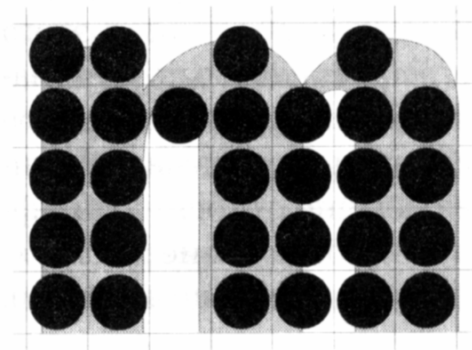
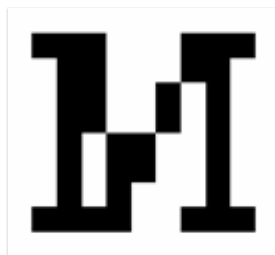
- ❑ Нарушение единства символов



Пример нарушения единства символов

Проблемы растеризации

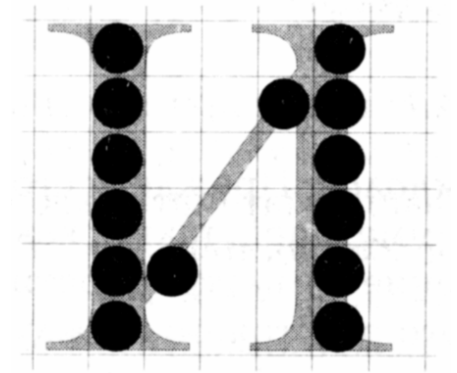
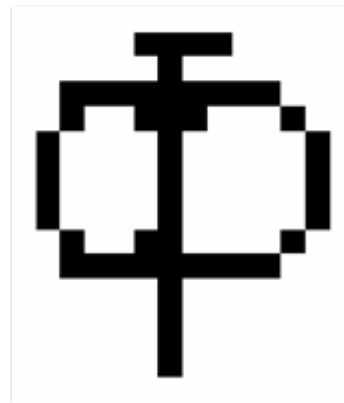
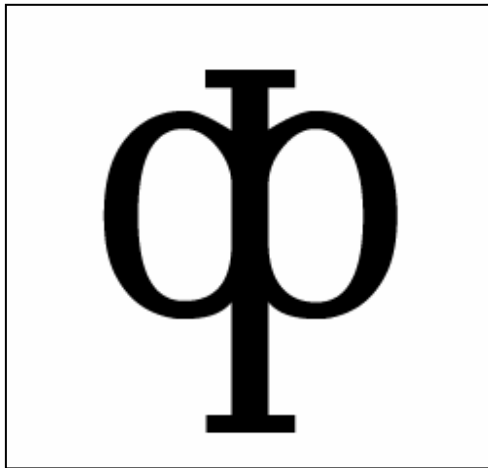
□ Смыкание штрихов



Пример нарушения пропорций

Проблемы растеризации

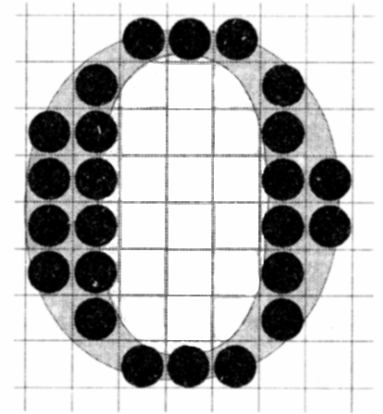
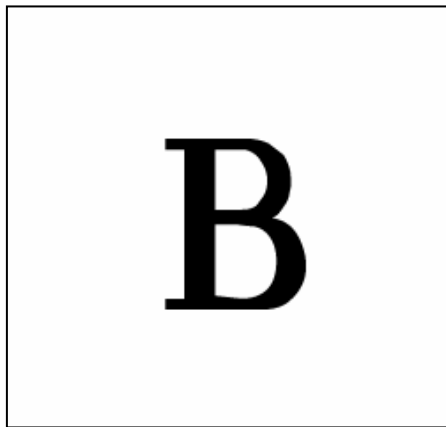
□ Выпадение точек



Пример выпадения точек

Проблемы растеризации

- ❑ Нарушение формы округлых букв



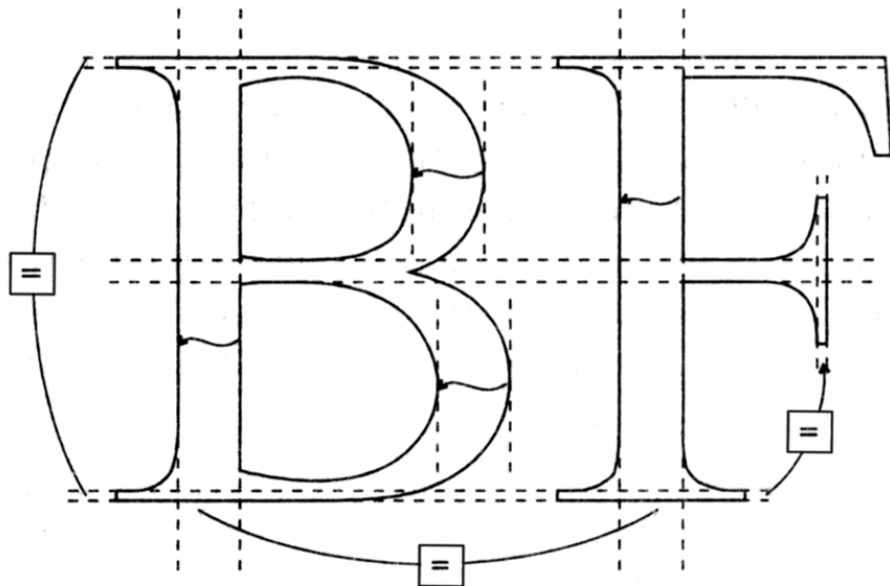
Пример нарушения округлых форм

Разметка символов

- *Разметкой* мы будем называть описание символов, их элементов и шрифта в целом, призванное улучшить качество растеризации символов.
 - Иногда разметку называют хинтингом (от англ. *hint* – подсказка), но этот термин обычно относят к шрифтам в формате Type 1 (для TrueType шрифтов используют понятие инструкций), поэтому мы считаем необходимым ввести новый, более общий, термин.
-

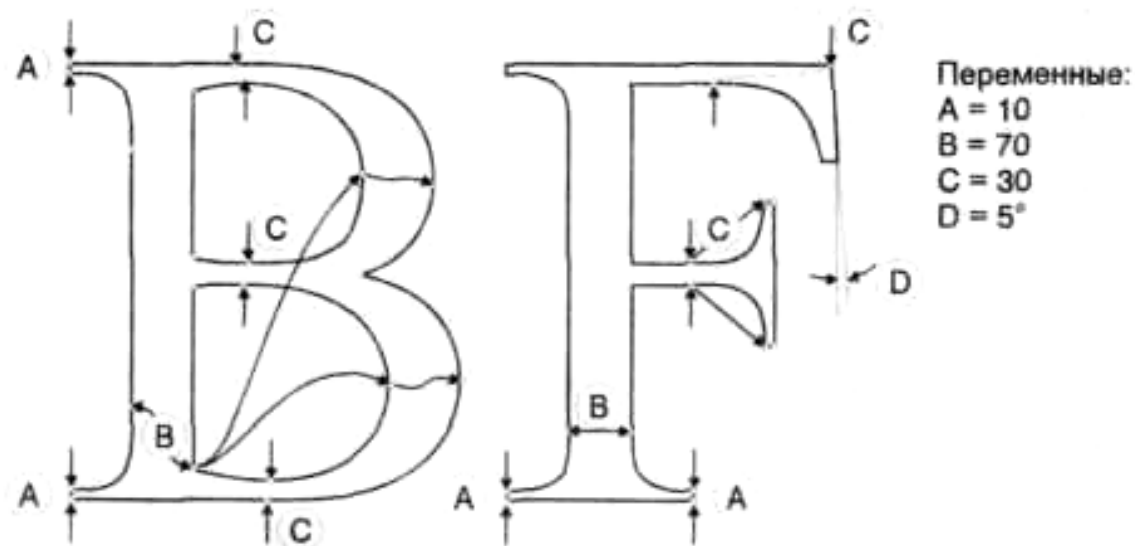
Методы разметки символов

- ❑ **Декларативный метод разметки** основан на описании особенностей символа при помощи их декларирования отдельно от описания контура.



Методы разметки символов

- ❑ **Программируемый метод разметки** основан на точном определении в шрифте всех действий, которые должен выполнять растеризатор.



Общая структура шрифта в формате Type 1 (PostScript)

1. Открытая часть:

- Обозначение шрифта
- Заголовок шрифта
- Кодовая таблица шрифта
- Уникальный идентификатор шрифта

2. Закрытая часть:

- Область глобальной разметки
 - Область глобальных подпрограмм
 - Область подпрограмм разметки и контурных подпрограмм
 - Область описаний символов
-

Обозначение шрифта

%! PS-AdobeFont-1.0: TimeRoman 001.1
%%CreationDate: Wed Oct 20 17:08:26 1993
%%Creator: FontLab(c) for Windows v2.5

Типичный заголовок Type 1-шрифта

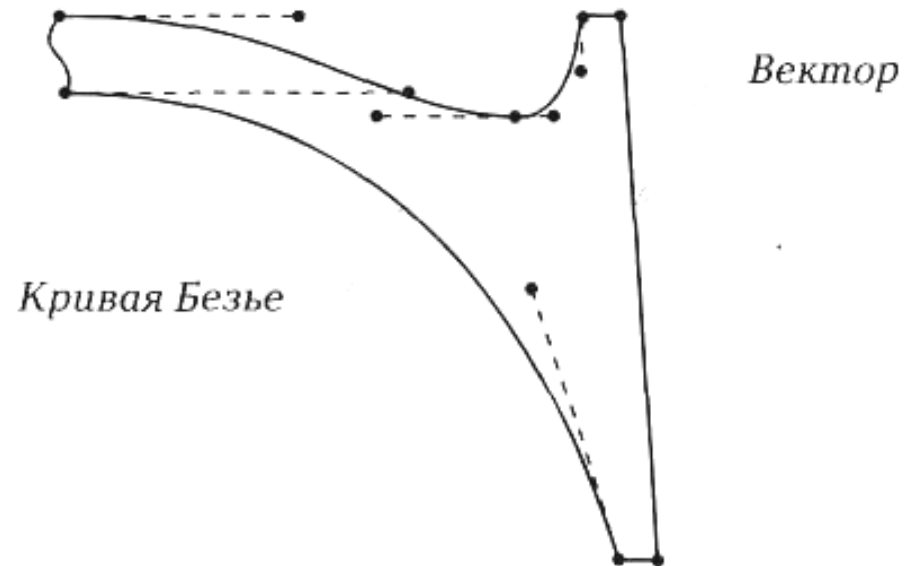
```
/FontInfo 9 dict dup begin
/FullName (Times New Roman) readonly def
/FamilyName (Times) readonly def
/version (001. 1) readonly def
/Weight (Normal) readonly def
/Notice ((c) Copyright Monotype, 1990) readonly def
/ItalicAngle 0 def
/isFixed Pitch false def
/UnderlinePosition -100 def
/UnderlineThickness 50 def end readonly def
/FontName /TimesNewRoman def
/PaintType 0 def
/FontType 1 def
/FontMatrix [ 0. 001 0 0 0. 001 0 0 ] readonly def
/FontBBox { -63 -231 1148 882 } readonly def
```

Уникальный идентификатор шрифта

- Уникальный идентификатор шрифта – 24-разрядное число (от 0 до 16777215).

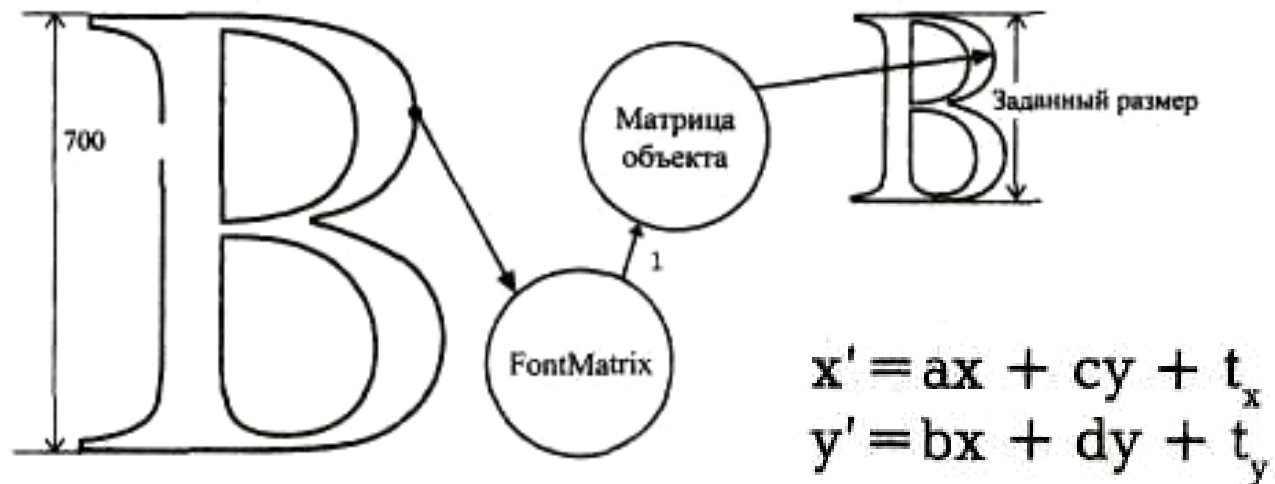
Построение символов PostScript Type 1-шрифтов

- Образец контура символа



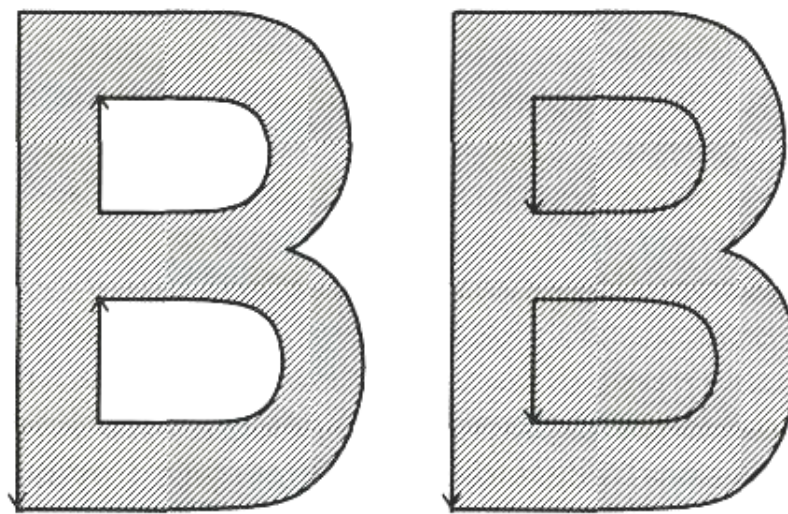
Построение символов PostScript Type 1-шрифтов

- ❑ Схема преобразования символа



Построение символов PostScript Type 1-шрифтов

- Правило заполнения контуров

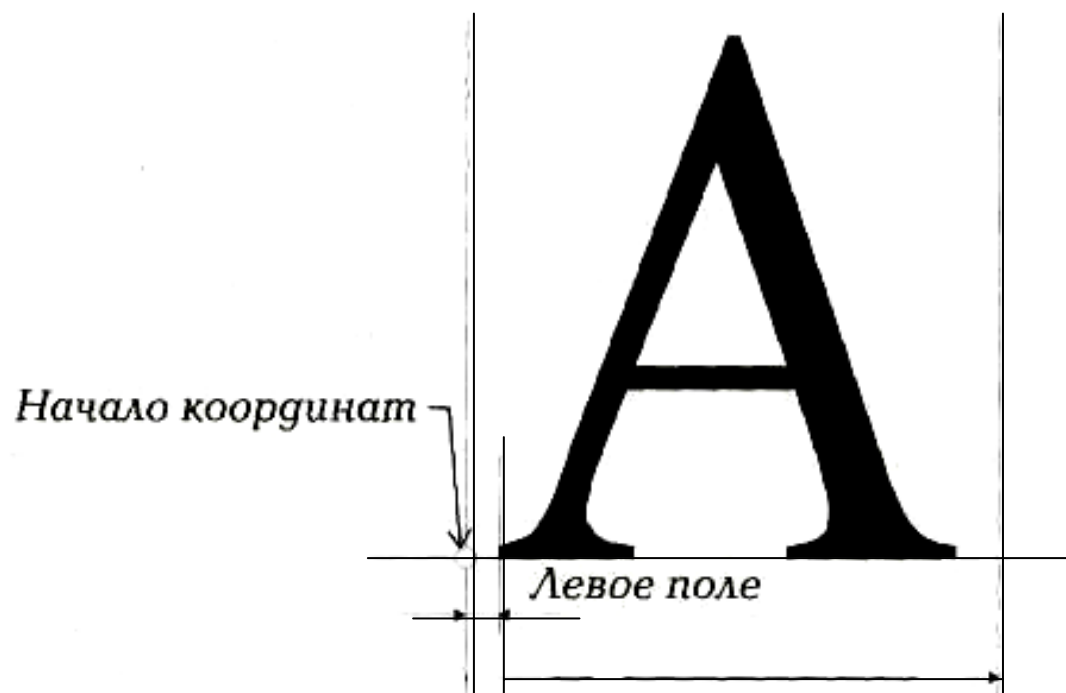


Правильно

Неправильно

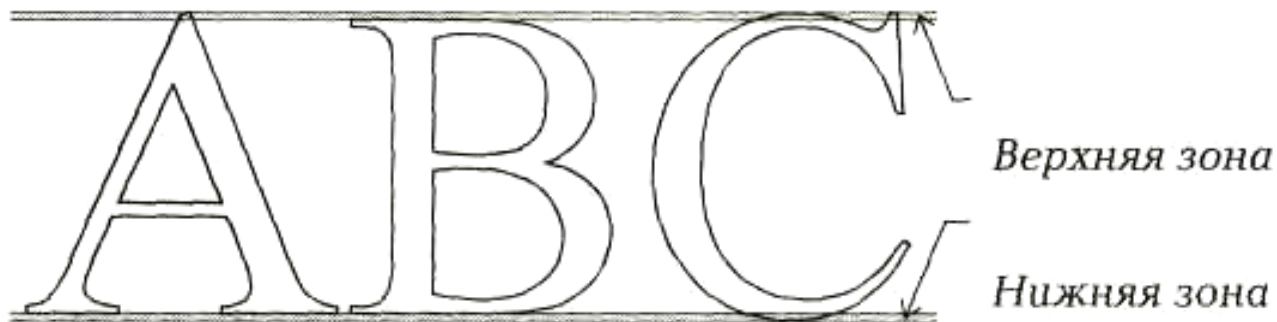
Построение символов PostScript Type 1-шрифтов

- Начало координат



Построение символов PostScript Type 1-шрифтов

- Зоны выравнивания в глобальной разметке



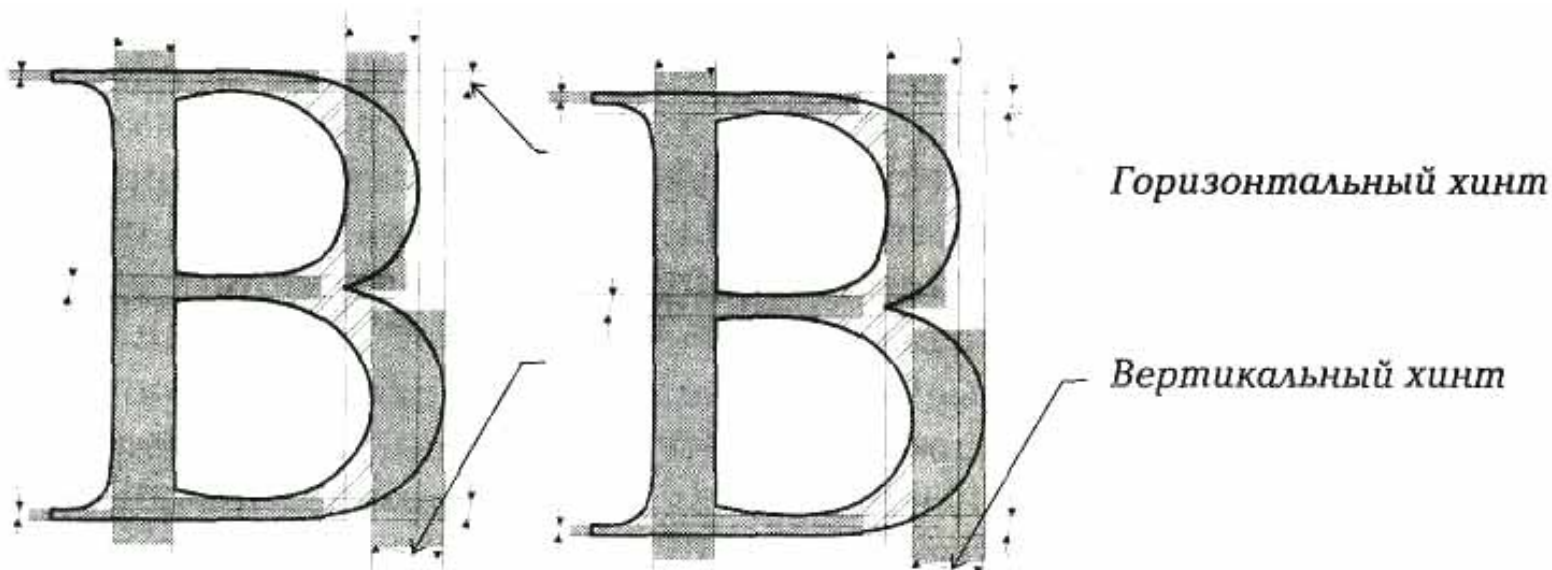
Построение символов PostScript Type 1-шрифтов

Метрическая информация (файл AFM) о шрифте состоит:

- Заголовок.
 - Таблица имен и метрик символа.
 - Описание пар кернинга.
 - Описание трекинга.
-

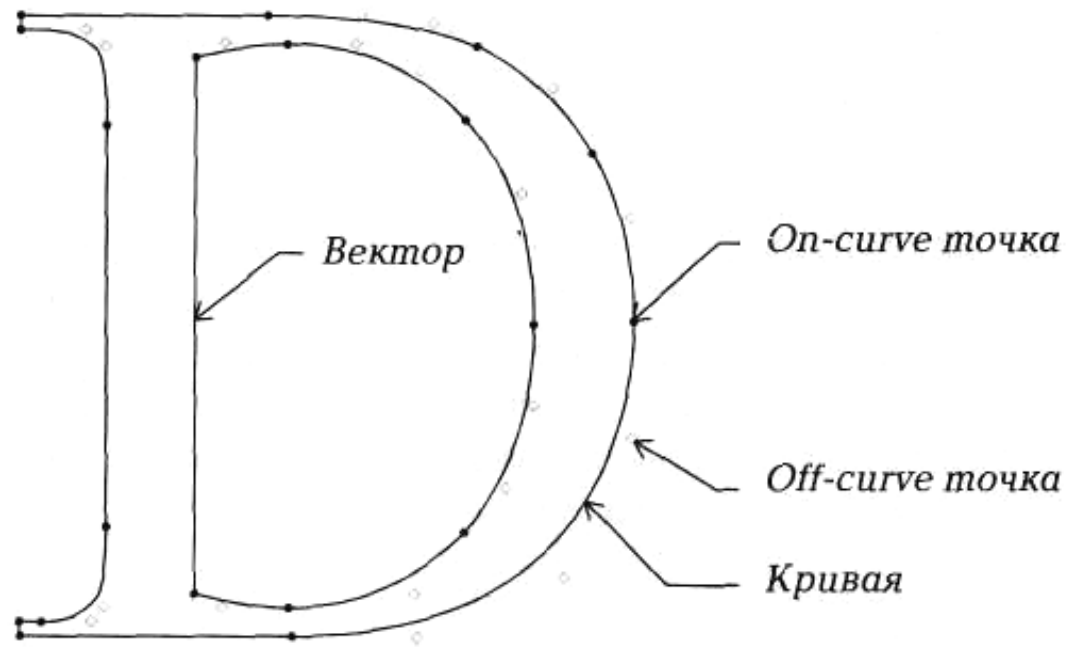
Построение символов PostScript Type 1-шрифтов

□ Хинты локальной разметки



Построение символов TrueType шрифтов

□ Построение контуров символов



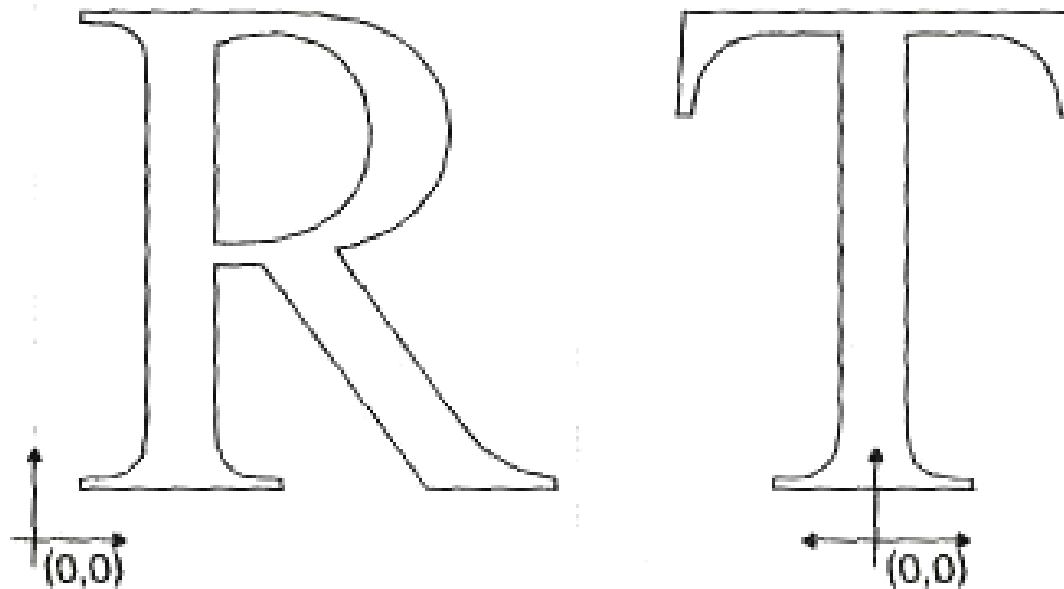
Построение символов TrueType шрифтов

- Пример заполнения контуров



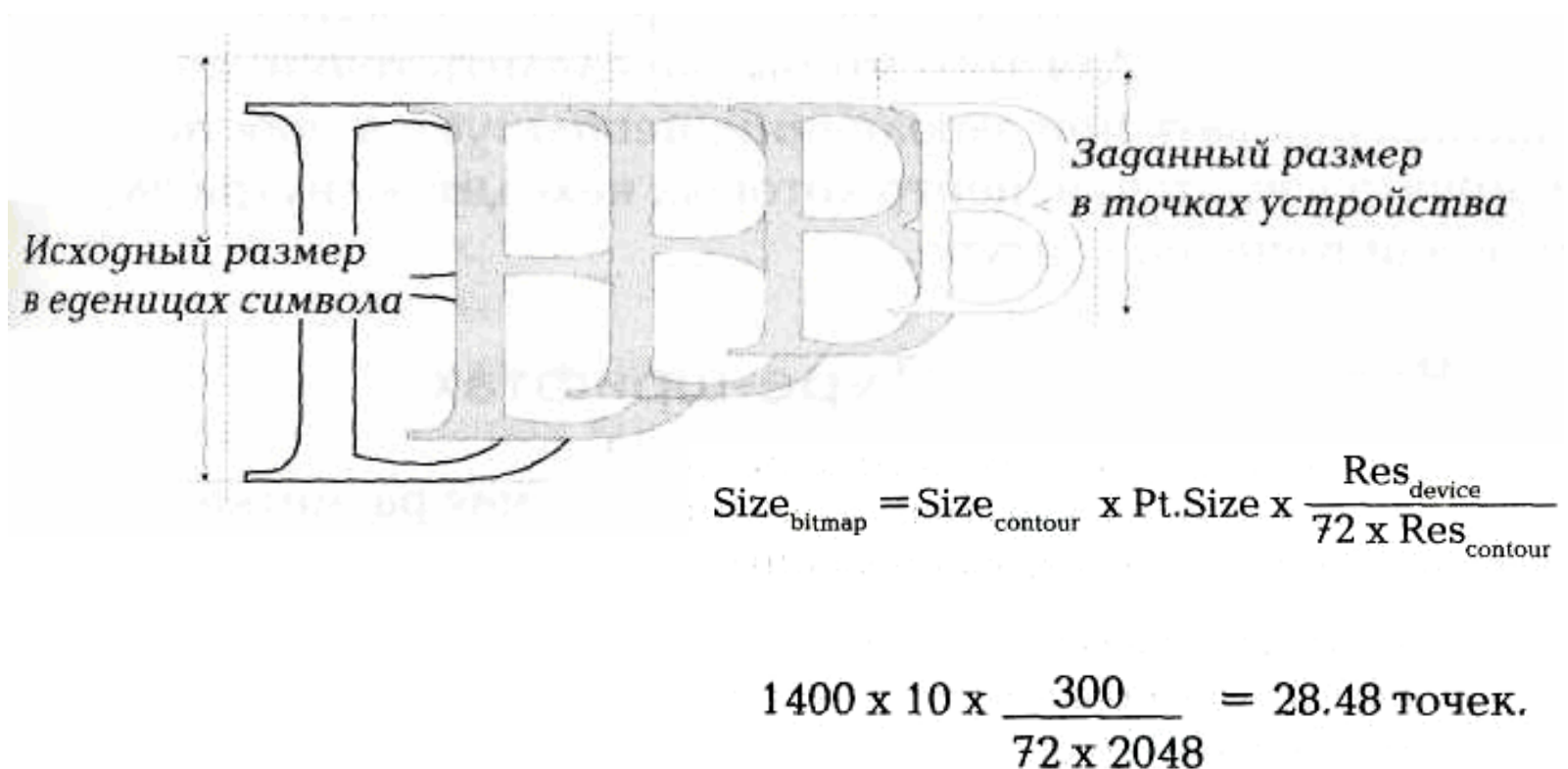
Построение символов TrueType шрифтов

- Положение точки начала координат в координатной системе



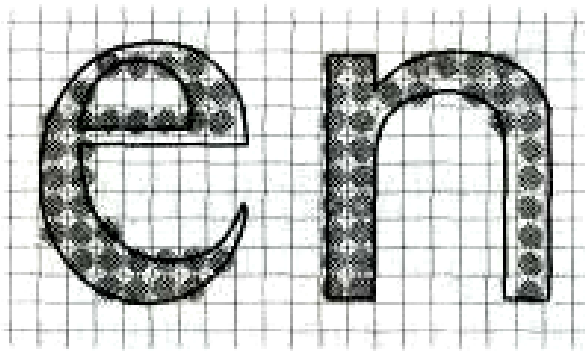
Построение символов TrueType шрифтов

- Схема процесса масштабирования

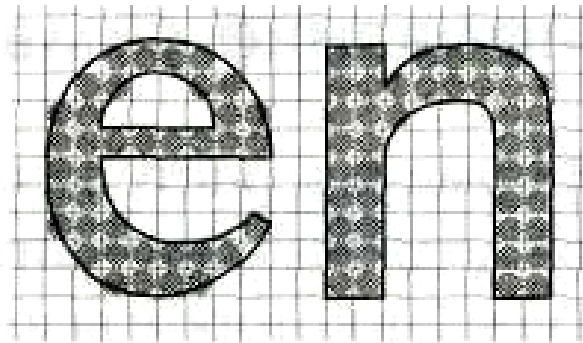


Построение символов TrueType шрифтов

- Изменение контура для улучшения качества воспроизведения (хинтинг)



Исходное состояние



После изменения контуров