

9

## Text and Typography

**Based on material from  
*Digital Multimedia*, 3rd edition  
published by John Wiley & Sons, 2009  
© 2009 Nigel Chapman and Jenny Chapman**

**These lecture slides © 2009  
Nigel Chapman and Jenny Chapman**

**All figures © MacAvon Media Productions**

# Character Sets

**The content of text consists of the characters that make up the words, punctuation, symbols, and so on which convey the meaning or message.**

**The appearance of text comprises its visual attributes, such as the shape and size of characters, and their layout on the page.**

**An abstract character may have many different graphic representations.**

**A character set maps the abstract characters in its character repertoire to their code values in a set of integer code points.**

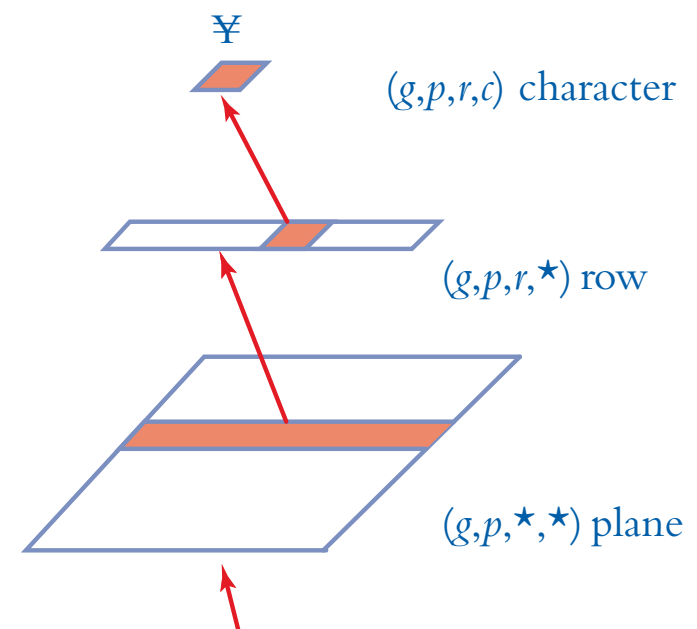
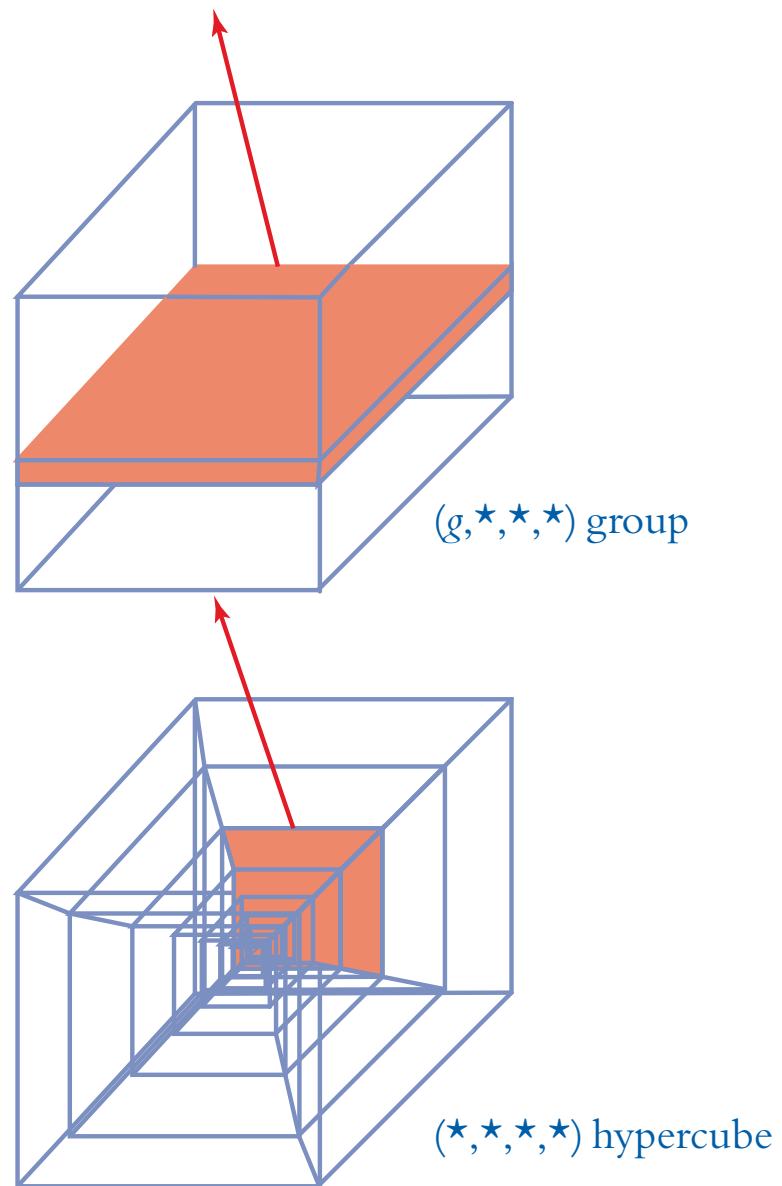
**ASCII is a 7-bit character set, providing 95 printable characters and some control characters. It is only adequate for a few languages, including English.**

**ISO 8859 defines a collection of 8-bit character sets, each one covering a set of related languages.**

**ISO 8859-1 or ISO Latin1 covers most Western European languages. It is identical to ASCII for code points 0 to 127.**



**ISO 10646 defines a 32-bit Universal Character Set (UCS), arranged in 256 groups, each of which consists of 256 planes accommodating 65,536 characters each.**



*The structure of ISO 10646*

**The UCS-4 encoding uses four bytes to hold the full 32-bit code value for any character.**

**The UCS-2 encoding uses just two bytes, to hold 16-bit values for characters on the (0,0,\*,\*) plane (the Basic Multilingual Plane).**

**UCS-2 is identical to Unicode, and provides code values for all the characters used to write contemporary major languages.**

**ISO Latin1 is the 8-bit code equivalent to the (0,0,0,\*) row of ISO 10646.**

**UTF-8 allows any ISO 10646 or Unicode value to be encoded as a sequence of 8-bit bytes, such that ASCII values are left unchanged in a single byte.**

**UTF-16 is an extension mechanism which provides Unicode with access to an extra 15 planes of the full ISO 10646 character set.**

**UTF-8 is the default character set used on the World Wide Web.**

# Fonts



**A glyph is a graphic representation of a character's shape. A character may be represented by many different glyphs.**

q q q q q q q q q q  
q q q q q q q q q

*19 glyphs for a single letter*

**A font is a collection of glyphs.**

**Fonts may be embedded in SWF or PDF files.**

**Web browsers usually access fonts on the user's system, but may download them from a remote server.**

**Fonts may be classified by their spacing (monospaced or proportional); the presence or absence of serifs (serifed or sans serif); their shape (upright, italic or slant); stretch (extended or condensed); and weight (bold, normal or light).**

Monospaced Font: Courier  
Each letter occupies the same amount of horizontal space, so that the text looks as if it was typed on a typewriter.

*A monospaced font*

Proportional Font: Bembo  
Each letter occupies an amount of horizontal space proportional to the width of the glyph, so that the text looks as if it was printed in a book.

*A proportional font*



### Sans Serif Font: Univers

The letters of a sans serif (or sanserif) font lack the tiny strokes known as serifs, hence the name. They have a plain, perhaps utilitarian, appearance.

*A sans serif font*

*Italic Font: Bembo Italic*

*The letters of an italic font slope to the right, and are formed as if they were made with an italic pen nib. Italics are conventionally used for emphasis, and for identifying foreign words and expressions.*

***An italic font****Slanted Font: Lucida Bright Oblique*

*The letters of a slanted font share the rightward slope of italic fonts, but lack their calligraphic quality. Slanted fonts are sometimes used when a suitable italic font is not available, but may also be preferred to italics when a more modern look is wanted.*

***A slanted font***



**Fonts are grouped into families, consisting of related versions of a typeface.**

Text that is set in Bembo goes well  
with *Bembo italic* and ***Bembo bold italic***  
but it looks quite wrong mixed  
with Cheltenham and *Cheltenham*  
*italic*. (All this text is 12 pt.)

***Combining fonts from different families***

**Text fonts are intended for extended passages of text, so they must be readable and unobtrusive. Usually they are upright and serified.**

**Display fonts are intended for short pieces of isolated text, such as headlines and slogans. They should be eye-catching and are often unconventional.**

**Text fonts designed for print may not work well on screen.**

**1 pt (point) is  $\frac{1}{72}$  of an inch.**

**1 pc (pica) is 12 pt.**

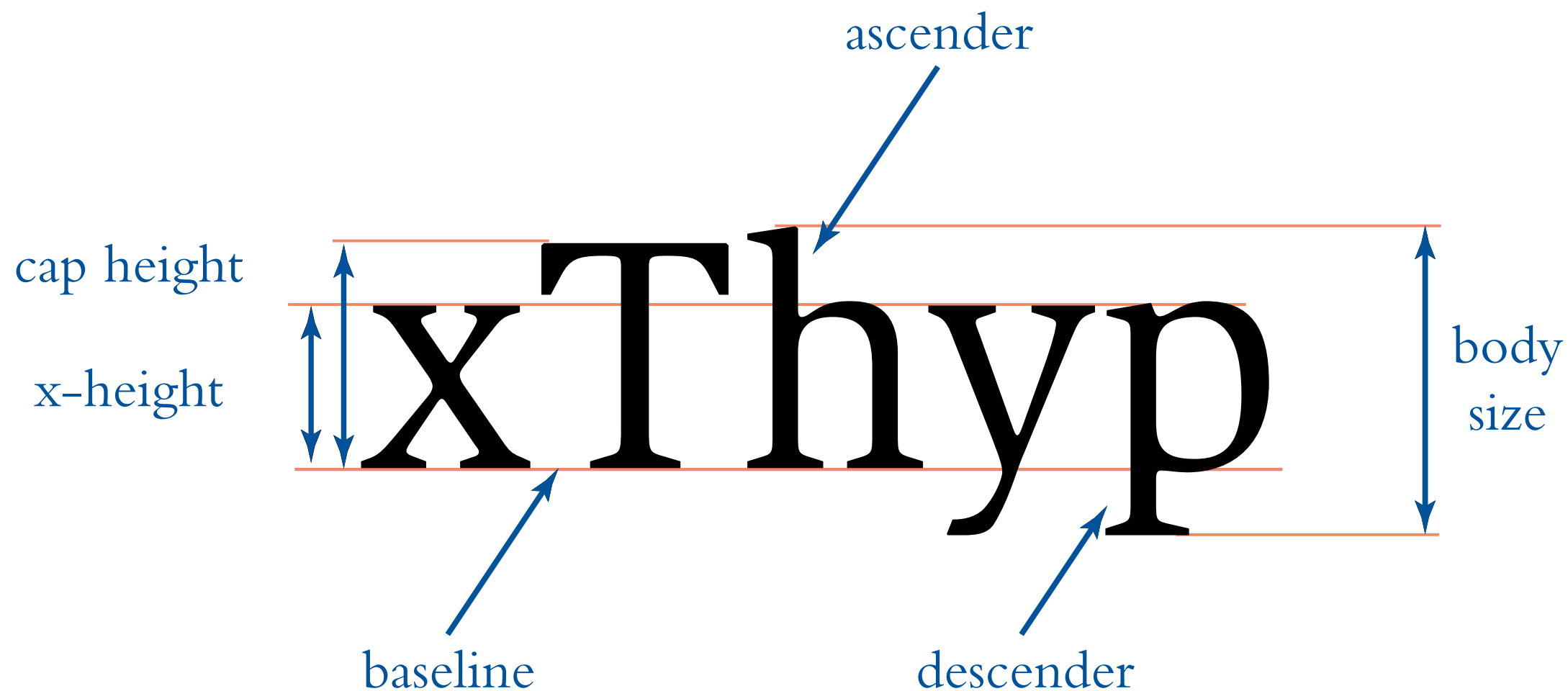
**The body size is the nominal size of a font, but is not usually equal to the height of any glyph in the font.**

**The leading is the distance between baselines.**



**The x-height is the height between the baseline and the top of a lower-case letter x.**

**Strokes that rise above the x-height are ascenders; those that extend below the baseline are descenders.**



*Some font terminology*

**Relative units are useful for measurement.**

**1 em is equal to the body size;**

**1 en is 0.5 em;**

**1 ex is equal to the x-height.**

**Kerning is the adjustment of the space between letter pairs (e.g. AV) to make them look more uniform.**

**Ligatures are composite single glyphs used to replace combinations of letters (e.g. fi) that don't look right next to each other.**

kerned    not kerned

**AV    AV**

**Ta    Ta**

*Kerning*

without    fine fluffy soufflés

with    fine fluffv soufflés

*Ligatures*

**Type 1 (PostScript), TrueType or OpenType are outline font formats.**

**Anti-aliasing is often applied to type to make it appear smoother.**



*Anti-aliased text*



# Layout and Formatting

**Markup consists of annotations that control a text document's layout and formatting or indicate its structure.**

**In WYSIWYG systems, markup is invisible and the effects of formatting commands are displayed immediately.**

**Markup may take the form of readable tags, which can be inserted using any text editor.**

**In visual markup, tags specify aspects of the text's appearance. In structural markup, they identify logical elements, such as paragraphs, lists or headings.**

**Structural markup allows global formatting changes to be made easily, permits the same markup to be used for different output media, allows a separation of concerns between content creation and design, and makes it easier for computer programs to analyse and process marked-up documents.**

**A set of styles can be collected into a stylesheet, which may be attached to many documents to ensure a consistent and unified appearance.**