Lecture 4 – Word Processing

Karl R. Wilcox

K.R.Wilcox@reading.ac.uk

Objectives

- **History of "word processing"**
- **The current state of play**
- **Issues with "word processing"**
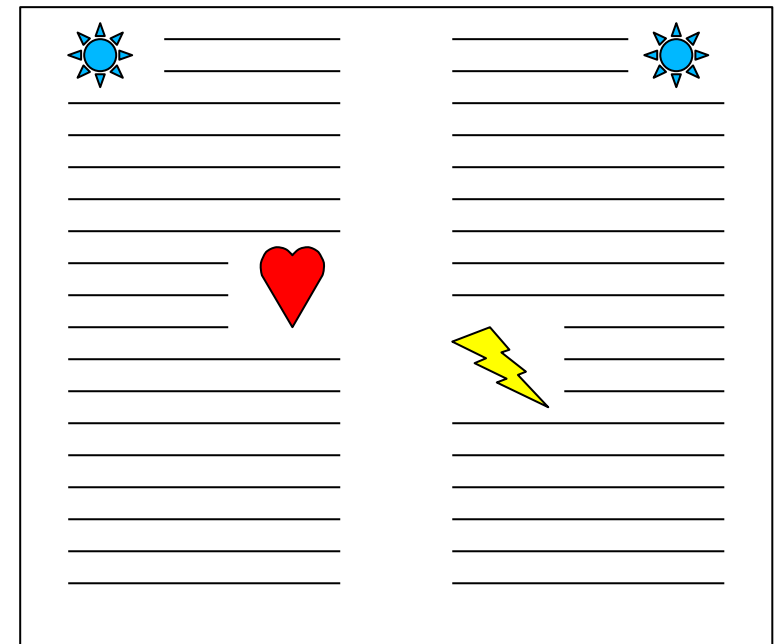
- **Today's practical**
  - Basic Microsoft Word use

Some History

- **Longhand / dictation / copy typing**
- **Single line memory typewriters**
- **Line printers, text editors & formatters**
- **Dedicated word processors (e.g. Amstrad PCW8512)**
- **Personal computers, text editors, text printers**
- **Dot matrix and laser printers (postscript)**
- **WYSIWYG (on paper) "word processors"**

- **(Typesetting is a different story, see next week)**

Word Processing is not…

- **Typesetting (more next week)**
- **Document management (more next week)**
- **Workflow / collaborative working**
- **Desktop Publishing**
  - Word processors support a single text flow with graphics and other objects inserted at fixed or floating locations
  - DTP defines various regions some of which may contain text flows
  - Try laying out a newspaper in Word (and changing it!)

The University of Reading

The State of the Art

- Absolute and relative positioning of text and graphics to a high degree of accuracy
- Tables of arbitrary organisation and format
- Dynamically updateable fields
- Automatic table of contents (if styles used correctly)

- Spell checking, grammar checking
- Automatic correction and automatic formatting
- Macro programming
- Change indications and annotations
- Mail merge

The University of Reading

But…

- **What is the purpose of the "words" being processed?**
  - **To be viewed on a screen?**
  - **To be printed and bound?**
  - **To be printed and displayed?**
  - **To be placed on a web site?**
  - **To form the content of an e-mail message?**
  - **To be archived as a matter of record?**
  - **To be re-published in a different medium?**
  - **To be used as data in a computer system?**

- **It is possible that any and all of the above could apply**

Word Processors Give Control

- **Word processors give very precise control over appearance of the document**

  - *But only if that document is printed on a particular size, shape and colour of paper!*

- **All other formats and purposes are usually compromised in some way**

Content Vs Format

- **Documents have content**
  - Text flows
  - Graphics
- **Content is _structured_**
  - Chapters
  - Paragraphs
  - Quotations
  - Captions

- **Format could depend on**
  - Output medium
  - User preference
  - Document purpose
- **Selection and order of content need not be fixed**
  - Summary or detailed
  - Skilled or novice
  - Conclusion first or last

Structure and Format by Styles

- **Most word processors combine structure and format into a single concept of a "style"**
  - **E.g. "Heading 1"**
    - **Should appear as a top level entry in the table of contents**
    - **Is 16 pt Arial bold, left aligned, 12pts before, keep with next…**

- **What if I want different types of chapters to have different formatting?**
  - **E.g. the chapters of a book are letters written by four friends, to distinguish them each chapter heading is formatted in a different fashion, but all appear in the table of contents**
    - **How would I do this in Word?**

Is a Style structure or format?

- **Why is a particular piece of text in an italic style?**

  - **Because it is in another language?**
    - **…the management finally acknowledged the *bona fides* of their dubious visitor.**
  - **For emphasis?**
    - **"You see, I *know* you posted a letter at Lower Swinstead."**
  - **Because it is a proper noun?**
    - **…after admitting to a very strange lapse of memory in missing *The Archers,* suddenly decided on a new line of enquiry.**

- **This would be important in a "text-to-speech" system**

Formatting for a purpose

- **Formatting a document for a particular purpose often precludes other purposes**
    - **Which you may not even have thought of!**
- **During the 1980's many newspapers replaced their microfilmed archives with Adobe PDF format files**
- **These files reproduced exactly the printed layout of the newspapers**
- **But the (original) PDF format did NOT maintain the reading order of elements on the page**
    - **I.e. there was no "electronic" link between headlines, stories and pictures**
- **On-line extraction of stories proved very difficult**

Separating Content from Format

- **The source document consists of text flows, graphics and other elements**
- **These are "marked up" according to the structure of the document and the role of each element**

- **Formatting is a separate process that depends on the type of output medium and the purpose of this document instance**
- **Formatting could include not just superficial appearance but selection and ordering of content**

The University of Reading

How to "Mark Up" Structure

- **Meta - XML**
  - XML is a "Meta-Language" that allows mark up languages to be defined

- **Generic - DOCBOOK**
  - Is an XML based language to describe the structure of books, technical reports and articles

- **Specific - ATA2100**
  - Is a industry specific standard for marking up the content of aircraft flying manuals

- **Others - LaTeχ, lout, nroff, lots of XML standards**

- **(HTML makes an even worse pig's ear of things…)**
  - Don't get me started….

The University of Reading

Coming up…

- **Today's practical**
  - **Basic use of Microsoft Word**

- **Next week**
  - **Document management**
  - **The production of large, complex, multi-authored, long life, evolving documents**