

Latex and Sweave

BCB 504: Applied Bioinformatics

Matt Settles

University of Idaho
Bioinformatics and Computational Biology Program

February 8, 2012

Outline

- 1 LaTeX
 - What is LaTeX
 - History
 - LaTeX Resources
- 2 BibTeX
- 3 Sweave
- 4 Reproducible Research

What is Latex

LaTeX is a high-quality typesetting system; it includes features designed for the production of technical and scientific documentation. LaTeX is the de facto standard for the communication and publication of scientific documents.

History - TeX

TeX (= tau epsilon chi, and pronounced similar to "blecch"),
current version 3.1415926

In the late 1970s, Donald Knuth was revising the second volume of his multivolume magnum opus "The Art of Computer Programming, got the galleys, looked at them, and said (approximately) "blecch"! Around the same time, he saw a new book by Patrick Winston that had been produced digitally, and ultimately realized that typesetting meant arranging 0's and 1's (ink and no ink) in the proper pattern, and said (approximately), "As a computer scientist, I really identify with patterns of 0's and 1's; I ought to be able to do something about this", so he said out that were the traditional rules for typesetting math, what constituted good typography, and as much as he could about type design. A project he originally thought would take 6 months ended up taking almost 10 years.

LaTeX

LaTeX is a set of macros build on top of TeX

LaTeX is based on the idea that authors should be able to focus on the content of what they are writing without being distracted by its visual presentation. In preparing a LaTeX document, the author specifies the logical structure using familiar concepts such as chapter, section, table, figure, etc., and lets the LaTeX system worry about the presentation of the structures. It therefor encourages the separation of layout from content while still allowing manual typesetting adjustments where needed. This is similar to how HTML with CSS works.

LaTeX Resources

- Main site: www.latex-project.org/
- TeX User Group (tug): <http://www.tug.org/>
- Comprehensive TeX Archival Network: <http://www.ctan.org/>
- LaTeX for Biologists:
<http://www.lecb.ncifcrf.gov/toms/latex.html>

BibTeX

BibTeX is a reference management software for formatting lists of references. The BibTeX tool can be used together with the LaTeX document preparation system.

BibTeX makes it easy to cite sources in a consistent manner, by separating bibliographic information from the presentation of this information.

Sweave and Reproducible Research

Sweave is a function in the statistical programming language R that enable integration of R code into LaTeX. The purpose is "to create dynamic reports, which can be updated automatically if data or analysis change")

Because the Sweave files together with any external R files that might be sourced from them and the data files contain all the information necessary to trace back all steps of the data analysis, Sweave also has the potential to make research more transparent and reproducible to others.