TeX-free LaTeX, an overview
&
Standards for LaTeX documents and processors
OR
Wither LaTeX? (The Language

Chris Rowley
Open University, UK

TUG Meeting
Notre Dame, July 2009

Abstract

The LaTeX *is* the message

As some of you will be aware, and all should be, LaTeX code, possibly with some variations, extensions or simplifications, has for a long time been used, raw and unprocessed, as a lingua franca for communicating mathematics via text files in computers. [I have even seen it used on napkins and coffee tables.]

This led to a proliferation of LaTeX-like input systems for mathematical information and this in turn produced a reluctance by users of maths notation to adopt any other type of input. However, much of this math input is not intended (primarily) to ever be input to a TeX machine (It may get swallowed by a TeX-like system after, for example, some copy-paste actions).

More recently, systems are being developed to produce whole LaTeX-encoded documents that are to be processed by systems such as OMDoc or LaTeXML and so will not necessarily ever pass through a TeX-like engine. Sytems such as PlasTeX also belong in this category, despite using TeX as a helper utility in their implementation.

A very recent discovery surprised me more than a little: that many systems in the maths world are not only able to produce LaTeX output (e.g., computer algebra packages) but, currently at least, have LaTeX-maths as their only or primary output! This is because: it is wanted or preferred by mathematicians; it is widely accepted by other mathematical software; or simply that nothing else is known to be available for a consistent and familiar encoding of maths notation.

A more sophisticated reason put forward for the increasing ubiquity of LaTeX is that if you are looking for a user-friendly and flexible editor for structured documents, then there are no rivals to the various environments available for the production and editing of LaTeX documents (such as `auctex+(x)emacs`).

Standards. *What* standards?

It would be possible to make an exhaustive list of everything that is allowed to appear in a Standard Basic LaTeX document. But that would be both tedious, uncheckable and ignored.

It is currently much easier to pin down which parts of the LaTeX language are accepted by the various non-TeX-like processors of LaTeX (from the first part). Also, there are copora that can be aoutomatically studied to produce definitions of the subsets actually used by various communities.

Amongst those who handle mathematics in computers there is a growing demand to analyse these de facto standards, at least for LaTeX-math, and to produce reference standards in this area. These would be used to compare systems and communities and make recommendations for usage. This could possibly lead to some more formal standards and, most importantly, extension mechanisms so that, for example, general-purpose parsers can be used to read such code.