The LaTeX3 Project

12 January 1999

Abstract

This article describes the motivation, achievements and future of the LaTeX3 Project, which was established to produce a new version of LaTeX, the widely-used and highly-acclaimed document preparation system. It also describes how you can help us to achieve our aims.

For Archive maintainers, Authors, Publishers and Distributors: The project team request that, whenever possible, you include this article in any of the following:

- Books about TeX and LaTeX.
- Instructions for authors on using LaTeX.
- The printed documentation of CD-ROM collections that contain LaTeX.
- On-line collections that include a significant proportion of documents encoded in LaTeX.

Outline

The purposes of the LaTeX3 system can be summarized thus: it will greatly increase the range of documents which can be processed; and it will provide a flexible interface for typographic designers to easily specify the formatting of a class of documents.

The LaTeX3 Project Team is a small group of volunteers whose aim is to produce this major new document processing system based on the principles pioneered by Leslie Lamport in the current LaTeX.

The major visible work of the team before 1997 was the development of the current *standard* version of LaTeX. This was first released in 1994 and has since then been actively maintained and enhanced by extensions to that core system. They will continue to develop and maintain this system, releasing updated versions every six months and recording these activities in the LaTeX bugs database (see below).

Although LaTeX may be distributed freely, the production and maintenance of the system does require expenditure of reasonably large sums of money. The LaTeX3 Project Fund has therefore been set up to channel money into this work. We know that some users are aware of this fund as they have already contributed to it—many thanks to all of them! If you want to know more about how you can help the project, see Page 6—and thanks in advance for your generosity in the future.

Background

With TEX, Knuth designed a formatting system that is able to produce a large range of documents typeset to extremely high quality standards. For various reasons (e.g. quality, portability, stability and availability) TEX spread very rapidly and can nowadays be best described as a world-wide de facto standard for high quality typesetting. Its use is particularly common in specialized areas, such as technical documents of various kinds, and for multi-lingual requirements.

The TEX system is fully programmable. This allows the development of high-level user interfaces whose input is processed by TEX's interpreter to produce low-level typesetting instructions; these are input to TEX's typesetting engine which outputs the format of each page in a device-independent page-description language. The LATEX system is such an interface; it was designed to support the needs of long documents such as textbooks and manuals. It separates content and form as much as possible by providing the user with a generic (i.e. logical rather than visual) mark-up interface; this is combined with style sheets which specify the formatting.

Recent years have shown that the concepts and approach of LATEX are now widely accepted. Indeed, LATEX has become the standard method of communicating and publishing documents in many academic disciplines. This has led to many publishers accepting LATEX source for articles and books; and the American Mathematical Society now provides a LATEX package making the features of AMS-TEX available to all users of LATEX. Its use has also spread into many other commercial and industrial environments, where the technical qualities of TEX together with the concepts of LATEX are considered a powerful combination of great importance to such areas as corporate documentation and publishing. This has also extended to on-line publishing using, for example, PDF output incorporating hypertext and other active areas.

With the spreading use of SGML-compliant systems (e.g. Web-based publishing using HTML or XML) TEX again is a common choice as the formatting engine for high quality typeset output: a widely used such system is *The Publisher* from ArborText, whilst a more recent development is the object-oriented document editor Grif. The latter is used for document processing in a wide range of industrial applications; it has also been adopted by the Euromath consortium as the basis of their mathematician's workbench, one of the most advanced of the emerging project-oriented user environments. Typeset output from SGML-coded documents in these systems is obtained by translation into LATEX, which will therefore soon also be a natural choice for the output of DSSSL-compliant systems.

Because a typical SGML Document Type Definition (DTD) uses concepts similar to those of LATEX, the formatting is often implemented by simply mapping document elements to LATEX constructs rather than directly to 'raw TEX'. This enables the sophisticated analytical techniques built into the LATEX software to be exploited; and it avoids the need to program in TEX.

Motivation

This increase in the range of applications of LATEX has highlighted certain limitations of the current system, both for authors of documents and for designers of formatting styles.

In addition to the need to extend the variety of classes of document which can be processed by LaTeX, substantial enhancements are necessary in, at least, the following areas:

- the command syntax (attributes, short references, etc);
- the layout specification interface (style design);
- the level of robustness (error recovery, omitted tags);
- the extendibility (package interface);
- the layout specification of tabular material;
- the specification and inclusion of graphical material;
- the positioning of floating material, and other aspects of page layout;
- the requirements of hypertext systems.

Further analysis of these deficiencies has shown that some of the problems are to be found in LaTeX's internal concepts and design. This project to produce a new version therefore involves thorough research into the challenges posed by new applications and by the use of LaTeX as a formatter for a wide range of documents, e.g. SGML documents; on-line PDF documents with hypertext links.

This will result in a major re-implementation of large parts of the system. Some of the results of such rethinking of the fundamentals are already available in Standard LaTeX, notably in the following areas:

- Font declaration and selection;
- Font and glyph handling within mathematical formulas;
- Handling multiple font glyph encodings within a document;
- Allowing multiple input character encodings within a document;
- A uniform interface for graphics inclusion;
- Support for coloured text;
- Building and interfacing new classes and extension packages.

Description

The strengths of the present version of LaTeX are as follows:

- excellent standard of typesetting for text, technical formulas and tabular material;
- separation of generic mark-up from visual formatting;
- ease of use for authors;

- portability of documents over a wide range of platforms;
- adaptability to many languages;
- widespread and free availability;
- reliable support and maintenance by the LaTeX3 project team.

These will be preserved and in many cases greatly enhanced by the new version which is being developed to fulfill the following requirements.

- It will provide a syntax that allows highly automated translation from popular SGML DTDs into LaTeX document classes (these will be provided as standard with the new version).

  The syntax of the new LaTeX user-interface will, for example, support the SGML concepts of 'entity', 'attribute' and 'short reference' in such a way that these can be directly linked to the corresponding SGML features.

- It will support hypertext links and other features required for on-line structured documents using, for example, HTML and XML.

- It will provide a straightforward style-designer interface to support both the specification of a wide variety of typographic requirements and the linking of entities in the generic mark-up of a document to the desired formatting. These two parts of the design process will be clearly separated so that it is possible to specify different layouts for the same DTD.

  The language and syntax of this interface will be as natural as possible for a typographic designer. As a result, this language could easily be interfaced to a visually-oriented, menu-driven specification system.

  This interface will also support DSSSL specifications and style-sheet concepts such as those used with HTML and XML.

- It will provide an enhanced user-interface that allows expression of the typesetting requirements from a large range of subject areas. Some examples are listed here.

  - The requirements of technical documentation (e.g. offset layout, change bars, etc).
  - The requirements of academic publishing in the humanities (critical text editions, etc).
  - The requirements of structural formulas in chemistry.
  - Advanced use of the mathematical-typesetting features of TeX.
  - The integration of graphical features, such as shading, within text.
  - the integration of hypertext and other links in on-line documents using systems such as HTML, XML and PDF.

  Special care will be taken to ensure that this interface is extensible: this will be achieved by use of modular designs.

- It will provide a more robust author-interface. For example, artificial restrictions on the nesting of commands will be removed. Error handling will be improved by adding a more effective, interactive help system.

- It will provide access to arbitrary fonts from any family (such as the PostScript and TrueType fonts) including a wide range of fonts for multilingual documents and the specialist glyphs required by documents in various technical and academic areas.

- The new interfaces will be documented in detail and the system will provide extensive catalogues of examples, carefully designed to make the learning time for new users (both designers and authors) as short as possible.

- The code itself will be thoroughly documented and it will be designed on modular principles. Thus the system will be easy to maintain and to enhance.

The resulting new LaTeX will, like the present version, be usable with any standard TeX system (or whatever replaces it) and so will be freely available on a wide range of platforms.

LaTeX documentation

A complete description of Standard LaTeX can be found in:

LaTeX: A Document Preparation System Leslie Lamport,
Addison Wesley, 2nd ed, 1994.

The LaTeX Companion Goossens, Mittelbach and Samarin,
Addison Wesley, 1994.

A recent addition to the publications closely associated with the project is:

The LaTeX Graphics Companion Goossens, Mittelbach and Rahtz,
Addison Wesley, 1997.

This LaTeX distribution comes with documentation on several aspects of of the system. The newer features of the system are described in the following documents:

LaTeX 2ε for authors describes the new features of LaTeX documents, in the file `usrguide.tex`;

LaTeX 2ε for class and package writers describes how to produce LaTeX classes and packages, in the file `clsguide.tex`;

LaTeX 2ε font selection describes the new features of LaTeX fonts for class and package writers, in the file `fntguide.tex`.

For further contacts and sources of information on TeX and LaTeX, see the addresses on Page 6.

The LaTeX3 Project Fund

Although LaTeX may be distributed freely, the production and maintenance of the system does require expenditure of reasonably large sums of money. There are many necessities that need substantial financing: examples are new or enhanced computing equipment and travel to team meetings (the volunteers come from many different countries, so getting together occasionally is a non-trivial exercise).

This is why we are appealing to you for contributions to the fund. Any sum will be much appreciated; the amount need not be large as small contributions add up to very useful amounts. Contributions of suitable equipment and software will also be of great value. This appeal is both to you as an individual author and to you as a member of a group or as an employee: please encourage your department or your employer to contribute towards sustaining our work.

We should like to see funded projects that make considerable use of LaTeX (e.g. conferences and research teams who use it to publish their work, and electronic research archives using it) include contributions to this fund in their budgets.

We are also asking commercial organisations to assess the benefits they gain from using, or distributing, a well-supported LaTeX and to make appropriate contributions to the fund in order that we can continue to maintain and improve the product. If you work for, or do business with, such an organisation, please bring to the attention of the relevant people the existence and needs of the project.

In particular, we ask that all the large number of organisations and businesses that distribute LaTeX, within other software or as part of a CD-ROM collection, should consider pricing all products containing LaTeX at a level that enables them to make regular donations to the fund from the profit on these items. We also ask all authors and publishers of books about LaTeX to consider donating part of the royalties to the fund.

Contributions should be sent to one of the following addresses:

TeX Users Group, P.O. Box 2311, Portland, OR 97208-2311 USA
Fax: +1 503 223 3960 Email: `tug@tug.org`

UK TUG, 1 Eymore Close, Selly Oak, Birmingham B29 4LB UK
Fax: +44 121 476 2159 Email: `uktug-enquiries@tex.ac.uk`

Cheques should be payable to the user group (TUG or UKTUG) and be clearly marked as contributions to the LaTeX3 fund. Many thanks to all of you who have contributed in the past and thanks in advance for your generosity in the future.

Contacts and information

In addition to the sources mentioned above, LaTeX has its home page on the World Wide Web at:

`http://www.latex-project.org/`

This page describes LaTeX and the LaTeX3 project, and contains pointers to other LaTeX resources, such as the user guides, the TeX Frequently Asked Questions, and the LaTeX bugs database.

More general information, including contacts for local User Groups, can be accessed via:

`http://www.tug.org/`

The electronic home of anything TeX-related is the Comprehensive TeX Archive Network (CTAN). This is a network of cooperating ftp sites, with over a gigabyte of TeX material:

```
ftp://cam.ctan.org/tex-archive/
ftp://dante.ctan.org/tex-archive/
ftp://tug.ctan.org/tex-archive/
```

For more information, see the LaTeX home page.