A Case for TeX in India —
The Indian TeX Users Group

C. V. Radhakrishnan

Abstract

Unlike in other countries, TeX migrated to India as a medium of typesetting for the Western publishing world. With its vast human potential and the cheapness of its cost, India enticed the publishing giants like Elsevier, Academic Press, Springer-Verlag, etc., for their pre-press work and with that the TeX language found its way into this subcontinent. Its meager presence in the very many higher academic institutions and its pronounced absence from ordinary institutions strengthens the paradigm that TeX usage in India is primarily a gut-oriented phenomenon rather than an author driven one. This is further demonstrated by the clear absence of TeX-related research, newer macro development, font generation for the multitude of Indian scripts, etc. The Indian academy, in contrast to its Western counterparts, pays scant regard for such things or seldom considers it as a necessity. Therefore, the problems of TeX usage in India are diametrically different from those in other parts of the world. It is not strange that the recently formed Indian TeX Users Group faces the constraints of lack of research and economic issues of the users as well (quite strange!) since the vast majority of its members are from the typesetting industry who chose TeX language as a means of their livelihood. These and related issues are described in this article.

A general overview of TeX users in India

There can be a vertical split when we consider the general users of the TeX language in India, one from the typesetting industry and the other from the higher institutions of learning. The former category may outnumber the latter. Except for the Indian Institutes of Technology (a chain of institutes spread all over India, noted for its academic excellence and standards) and certain specific scientific institutions like the Indian Institute of Science, Inter University Center for Astronomy and Astrophysics, Tata Institute of Fundamental Research, etc., TeX is still alien to the academy or the researchers. The advent of WYSIWYG typesetting software has further pushed back the chances of TeX usage. Yet another paradoxical element you can observe here is that the Indian academy considers typesetting issues as the burden of publishing houses, and it is not the concern of the author to address the enigma of his own document presentation. This is the general philosophy of even the computer scientists working in various Universities in India. The limits of our document preparation skills are dictated by few Microsoft products. If any of the Microsoft product is incapable of presenting our document, we would resort to manual operations, thereby making it a childish doodle, for the present-day Indian mind is not at all carried away by refined and sophisticated presentation, in sharp contrast to the classic Indian aesthetic sensibilities.

This being the general attitude of the academics around this country, the quantum and the quality of TeX usage in the academy can be gauged by anybody. This may be the prime reason why India lagged behind in forming its user group when everyone else in the educated West went ahead with their user groups and made substantive contributions to the TeX language. India became a silent spectator, with a subdued longing for enjoying the fruits of TeX research in the West with an apparent resignation that is typical of a Hindu mind. The shape of things in the Indian typesetting industry is also not so bright. Due to lack of any meaningful research and development team, they solely depend on or unabashedly hire Western intelligence for the development of their in-house arsenal. Even in matters as simple as writing a filter for SGML to TeX or *vice versa*, they do get filters written by external agencies, present these to their clients as if developed by their own R&D team, and win huge contracts. In short, healthy usage of the TeX language is still a distant dream in any of these agencies.

Early work on Indian languages

One of the earlier work on TeX language that concerns Indian scripts are done by Avinash Chopde;[1] the package, called ITRANS, bundles lot of Indian scripts with LaTeX. You create an `.itx` file and run it through ITRANS to convert it to a `.tex` file. The commands are the same for Tamil, Sanskrit, Marathi, etc. His home page describes the system fully; it is available for UNIX and PC platforms.

There is also JTRANS (Sandeep Sibal), a Java program that enables you to see Sanskrit text in an HTML document. There is also an `Xdvng` font that if installed will display Sanskrit documents on the Web. All these are explained in detail in the file `index.html` available via anonymous `ftp` from `jaguar.cs.utah.edu` in the directory `private/sanskrit`. You will have fun with all these programs and Avinash can throw more light on all

---

[1] `http://www.paranoia.com/~avinash/itrans.html`

these topics, since he created the various ITRANS versions. He has also an ITRANS songbook that lists several thousand Hindi film songs in Devanagari script.

If you visit `http://www.concentric.com/~Dchand/jaguar`, click on Processing Tools, where several packages for processing Sanskrit on the net are described with pointers to ITRANS, JTRANS, and others. Currently ITRANS supports Devanagari (Sanskrit/Hindi/Marathi), Tamil, Telugu, Kannada, Bengali, Gujarati, and Romanized Sanskrit script output.

The input text to ITRANS is in a transliterated form. Each letter in an Indian Script is assigned an English equivalent, and the English letters are used to construct what will eventually print out in the Indian Language Script.

ITRANS offers a choice of two input encodings: ITRANS, and CS/CSX. ITRANS encoding is a 7-bit ASCII encoding, while the CS/CSX encoding is an 8-bit encoding. The ITRANS encoding requires multi-character English code be used to represent each Indic Script letter, while the CS/CSX encoding uses a one-character code to represent each Indic script letter.

Other meaningful work undertaken in TEX-related areas includes various fonts created using METAFONT or tools like that. Some of the work has been undertaken by non-Indians too.

1. ItxGuj, a Gujarati font, and ItxBeng, a Bengali font, were added to ITRANS. These fonts have been donated to ITRANS by Shrikrishna Patil, and are available in PostScript Type 1 and TrueType formats, so they can be used for printing as well as for display on WWW browsers such as Netscape 3.0 (or later).

2. Though a lot of improvisation is needed, KannadaTEX (developed by the Central Institute of Indian Languages, Maysore) is a commendable work in the right direction. For the Kannada font from the KannadaTEX package, ITRANS support was added by Raghunath K. Rao. This is a font in METAFONT format, so it can be used only with TEX.

3. Devanagari font: `Xdvng`, by Sandeep Sibal is available in PostScript Type 1 and TrueType formats, so it can be used for printing as well as for display on WWW HTML browsers such as Netscape 3.0 (or later). The `Xdvng` font is a derivative of the `Devnag` font developed by Frans Velthuis.

4. Romanized Devanagari fonts: `CSUtopia`, by Dominik Wujastyk,[2] and Washington Indic Roman by Thomas Ridgeway; both in Classical Sanskrit Roman encoding (CS/CSX encoding).

5. Malayalam font: by Jeroen Hellingman,[3] a commendable work for both the traditional and reformed scripts. This is complete except for METAFONT sources; instead a range of pre-compiled sizes is included for the main font, and is available at CTAN. This system comes with two pre-processors, `patc` and `mm`.

The `malyalam.sty`[4] package is an interface to malayalamTEX, for use with LATEX 2_ε. It works by loading Hellingman's macro files `mmmacs.tex` and `mmtrmacs.tex` to interpret the TEX macros generated by the `patc` and `mm` pre-processors.

Note that these macro files are *not* provided as part of malayalam.sty, but must be collected separately from CTAN or elsewhere.

Problems of TEX in India

As you can see, except for baseline research on some font generation, nothing substantive is forthcoming from the Indian TEX world. TEX has never percolated into the local publishing industry. As such, document preparation (especially technical documents) in the regional languages suffers considerably and its current status is deplorably poor. It has an indirect effect on the development of scientific document presentation in local languages. People quite simply are forced to believe that our languages are not fit for scientific document presentation and console themselves that it is a cherished domain of European languages. For instance, the State Languages Institute of Kerala (Kerala is one of the Indian States where the literacy rate has surpassed 95%), the official body for the production of school and University text books in Malayalam (the language of Kerala), finds it difficult to produce advanced scientific books with a quality comparable to English language text books, though intellectual resources are abundant.

Secondly, with a very healthy and vibrant literature, the Indian regional languages publishing is one of the richest industries in the country. But electronic digitizing and archiving of the multitude of books released in a variety of languages (both officially recognized and otherwise) is a distant dream for us. No effort has been invested to address the

---

[2] `d.wujastyk@ucl.ac.uk`

[3] `etmjehe@etm.ericsson.se`

[4] `http://ctan.tug.org/tex-archive/language/malayalam/`

problems of archiving of text data and its retrieval. SGML (Standard Generalized Markup Language) is still alien to Indian languages. A vast heritage of Indian Literature still thrives on paper, raising a multitude of issues relating to storage and retrieval of information. The advent of InterNet and the WWW has prompted very many Indian regional periodicals to enter into the world of electronic magazines. Without proper fonts and an encoding scheme fit for the WWW browsers, most of them are still wallowing in the primitive world of presenting images of whole text pages, which becomes highly unpopular among Indian viewers where the poor dialup line speed prevents easy browsing.

Thirdly, there is a wide gap between Indian TeX Users and the current status of TeX in the world. The old LaTeX 2.09 is still in popular usage among most of the typesetting houses and general users as well. Most users are afraid of LaTeX 2_ε. When the world is hopefully anticipating the arrival of LaTeX3, our users are still in the domain of the obsolete LaTeX 2.09. In the workshop held along with the inauguration of the Indian TeX Users Group, most of the participants had not heard of graphics inclusion programs like METAPOST, XYpic, PSTricks, etc. PostScript and its relationship with TeX is also still unfamiliar.

Formation and relevance of the Indian TeX Users Group

It is at this state of affairs, some of the TeXies in the southernmost part of India came up with the idea of forming a Users Group in India. Sebastian Rahtz[5] of UKTUG played a key role in its formation. A few academics and researchers from the University of Kerala, Trivandrum, scientists from the Space Center, programmers from Indian software and typesetting companies assembled together and launched the Indian TeX Users Group (short-named TUG*India*). The aims and objectives of TUG*India* do not differ much from that of the international TUG, aside from a special emphasis on extending TeX to Indian languages. As a first step towards this goal TUG*India* is associating with Yannis Haralambous[6] of the French TeX Users Group to build an Omega-Malayalam system. Preliminary work done so far gives quite encouraging results, and with these results the local education department has agreed in principle to finance projects relating to actualizing and perfecting the Omega Malayalam system that

can solve the problem of technical document preparation in Malayalam.

This is only a first step towards extending TeX to Indian languages. Slowly and steadily this mission would be spread to other parts of India to cover all the major languages. It may sound a little strange that a single user group in a vast country with diverse lingua and cultures can hold all the users with different identities together. Unlike other parts of the world, this is an amazing truth so far as India is concerned and the TUG*India* Board has decided that its secretariat would be shifting its location to different centers in India in a fixed periodical manner so that the current bias towards the south will be annulled.

To familiarize its members with the emerging trends in TeX research, TUG*India* will be holding periodical seminars and workshops, etc.; the first of these was conducted along with the inaugural ceremony. The main themes discussed were LaTeX to SGML conversion strategies, pdfTeX and related issues, Hypertext in TeX, graphics and color inclusion in LaTeX, and METAPOST and other graphics programs.

Miscellany

The majority of TeX users in India are from the typesetting companies and the TeX implementations are naturally the choice of their employers. Most of them are using Y&Y with Win95 operating platform except for one company (to my knowledge) which uses *Textures* for Mac. Still another company which has more than 100 terminals employs Novell Netware and DOS based TeX implementation too. Most of the computers used are Intel based PCs. Unlike this scenario, the academic institutions like the Indian Institute of Technology, Indian Institute of Science, Inter-Univ. Center for Astronomy and Astrophysics, etc., where TeX remains a leading document preparation medium, the operating platform is various flavors of UNIX. Most of these institutions have Sun workstations, DEC Alpha systems, or HP workstations, and TeX implementations for these systems are in use.

Epilogue

The Indian TeX Users Group would be publishing a journal *viz.*, TUG*India* Journal, every four months and the first issue is getting ready to be released within a fortnight. TUG*India* welcomes articles from all the TeXies interested in publishing an article in our journal. Various User Groups are also informed that TUG*India* will be happy to reprint each other's articles on a reciprocal basis. The address

---

[5] s.rahtz@elsevier.co.uk

[6] yannis@pobox.com

for communication: Indian TEX Users Group, TC 24/548, KRIPA, Sastha Gardens, Thycaud, Trivandrum 695014, India. Tel. +91 471 324341, Fax. +91 471 333186, e-mail: `tugindia@mailexcite.com`.

⋄ C. V. Radhakrishnan
River Valley Technologies
Software Technology Park
Trivandrum 695034
India
`river@earthling.net`