Haruhiko Okumura

Haruhiko Okumura promotes and supports the use of TEX in Japan.

[Interview completed 4 June 2007.]

Dave Walden, interviewer: Please tell me a bit about yourself personally, independent of the TEX world.

Haruhiko Okumura, interviewee: My background was in particle physics and gravitation theory, but in those days I couldn't get a job at a university so I was teaching math at high schools. During that period I bought my first microcomputer, read Knuth's *The Art of Computer Programming*, and began studying algorithms. I got especially interested in data compression. I designed the algorithm behind the once-famous compression archiver LHA, assembly-coded by Haruyasu Yoshizaki. My basic idea, that combines sliding dictionary and entropy coding, is still used in zip and gzip.

DW: I was interested to read your history of such compression archivers on your web site (`http://oku.edu.mie-u.ac.jp/~okumura/compression/history.html`).

How did you first become involved with TEX?

HO: I wrote my first book, on statistical algorithms, in 1986 and my second book, on computer algorithms, in 1987. I used a Japanese word processor but I couldn't typeset math properly. I learned about Knuth's TEX but it was not designed for multibyte characters. Then a Japanese publisher made a version of TEX for Japanese, named pTEX. (There was another one named JTEX, but pTEX was superior.) I was quite excited and began writing books with it. I studied the traditional Japanese art of typesetting, worked with a professional printing expert to improve the TFM files for Japanese, and wrote style files. My third book on algorithms in C, typeset with pTEX, was selected as one of the best 100 computer books (including translations) ever published in Japan. But people began to look at me as a TEX guru, not as an algorithm master! I was asked to write many books on how to typeset beautiful books with pTeX.

DW: I see from your web site (`http://oku.edu.mie-u.ac.jp/~okumura/`) that you are now a professor of computer education in the education department of Mie University. I gather that you have moved completely away from physics and are now training new school teachers like you yourself once were. How did this transition come about, and what sorts of things do you teach now?

HO: In 1960's, as a child I was hooked on the Japanese anime series named Astro Boy (`http://en.wikipedia.org/wiki/Astroboy`), the story of a robot kid supposed to be

created in 2003. When I bought my first microcomputer in late 1970's, I really thought that Astro Boy was coming soon. But the world did not change that fast. In 2003, at long last, we had a computer class in every Japanese high school. But it turned out that high-school computer teachers were quite unfamiliar with programming or the science underlying computing; many of them only teach how to use Windows and Microsoft Office. During 2003–2004 I was appointed to work part-time (confidentially) for the National Center for University Entrance Examination where I prepared SAT-like tests in computing. But since computing was optional for university admission, only a few hundred students chose the subject. All in all, I think our computer education is in trouble. I agree with Alan Kay who maintains that the present computer education only reflects pop culture, and with Edward Tufte who says PowerPoint is ruining science as well as education. I'm willing to dedicate my final ten years of professorship to improving the situation.

DW: Have you written other books in addition to those you mentioned on statistical algorithms, computer algorithms, and algorithms in C? I assume well-known volumes like Robert Sedgewick's *Algorithms in C* are translated into Japanese. Does your book on algorithms in C address the audience for a different audience or in a particularly Japanese way?

HO: My Algorithms in C book (and its Java version compiled later) is very limited in scope compared to Sedgewick's volumes. It is an alphabetical compilation of small but interesting algorithms, much like HACKMEM (I hadn't read this fine memorandum when I wrote my book). My newest book is a small textbook entitled "Computer Literacy" for college freshmen.

DW: HACKMEM (`ftp://publications.ai.mit.edu/ai-publications/pdf/AIM-239.pdf`) is such a wonderful document. (I worked with Mike Beeler, who was a co-author of HACKMEM, for many years at Bolt Beranek and Newman, one of the places he worked after MIT. He was still doing recreational math when I worked with him. I think the last time I saw Mike in person was at Knuth's set of six lectures at MIT on "Things a Computer Scientist Rarely Talks About" (`http://www-cs-faculty.stanford.edu/~knuth/things.html`), based substantially on his earlier book *3:16 Bible Texts Illuminated* (`http://www-cs-faculty.stanford.edu/~knuth/316.html`). Of course, another of the HACKMEM authors, Bill Gosper, worked with Knuth for a while at Stanford in the mid-1970s.)

Please tell me something about the traditional art of Japanese typesetting, how that tradition may differ from the English typesetting traditions, and how improving the Japanese TFM files may also be different than the parallel work for Latin alphabet fonts.

HO: Japanese letters are basically square in shape. You can break lines almost anywhere. But non-letter marks (punctuation, parentheses, etc.) need not be square; you cannot break lines before the closing marks and after the starting marks. Also, juxtaposed marks need pair kerning. The detailed rules were compiled in Japan Industrial Standard JIS X 4051:1995, "Line composition rules for Japanese documents". Although the original configuration of pTEX didn't conform to these rules, pTEX was so versatile that if we rewrote the TFM files and set penalties and parameters properly, we could arrive at a fairly decent approximation to the traditional rules. (See `http://oku.edu.mie-u.ac.jp/~okumura/texfaq/japanese/ptex.html`.)

DW: Were the style files you mentioned developing related to use of the Japanese fonts or were they for other aspects of typesetting within the conventions for Japanese?

HO: My class files, jsarticle.cls (and friends), are equivalent to pTEX's jarticle.cls, which is in turn equivalent to article.cls, except for two exceptions: (1) jsarticle.cls discards pTEX's original font metrics and loads the improved TFM files; (2) it sets penalties and other parameters to better reflect traditional Japanese typesetting rules.

DW: When I trace the various links on your web site, I see a TEX Wiki, TEX questions and answers, a blog, and other ways you communicate with and support the use of TEX and other systems. I also do not see a Japanese TEX Users group when I look at the TUG web site.

HO: The Japanese TEX Users Group was discontinued long ago, because web sites, web forums (`http://oku.edu.mie-u.ac.jp/~okumura/texfaq/qa/`), and Wikis (`http://oku.edu.mie-u.ac.jp/~okumura/texwiki/`) turned out to be more appropriate for the impatient Japanese users. My site is getting so comprehensive that no one needs to buy my books!

DW: Have you written books in Japanese about TEX?

HO: Yes, I've written three or seven books, depending on how you count them. Two of them are out of date; four out of seven are the 1st through 4th editions of my LATEX 2_ε book (I rewrite it every three years). One out of seven is actually written by my colleagues under my supervision.

DW: What set of tools do you use when working with TEX (operating system, TEX engine, TEX format, classes, editor, etc.)?

HO: I've used MS-DOS, BSD, SunOS, Solaris, Linux, Windows, and Mac. Now I use Linux for servers and batch jobs and Mac OS X for client-side computing. My favorite editor is Emacs with AUCTEX. I use pTEX (pLATEX) and dvipdfmx to generate PDFs, because there's no pdfpTEX yet. Since pTEX for Unicode is now being developed and XETEX is acquiring pTEX-like versatility, next year I'll be using either the new pTEX or XETEX.

DW: You said earlier that a Japanese publisher created pTEX. Is pTEX a commercial product? If not, who is maintaining it now? Is it part of the TEX Live collection of software or does TEX in Japan exist sort of independently of that distribution which is jointly done by several TEX user groups?

HO: pTEX was created by ASCII Corporation (`http://www.ascii.co.jp/`), and is now distributed (`ftp://ftp.ascii.co.jp/pub/TeX/ascii-ptex/`) as a patch to teTEX 3.0. Its license is BSD-style. It is maintained by some engineers from the corporation with the help of a handful of outside volunteers. I wish it was incorporated in TEX Live, but I don't know how that can happen.

DW: Jin-Hwan Cho from Korea believes that TEX technologies are highly developed in Japan compared to Korea and China. Can you describe to what extent you believe this is true and why?

HO: I don't know much about Korean and Chinese TEX, but, yes, Japan has enjoyed 20 years of high TEXnology. We have had pTEX for 16-bit characters since 1987, and in 1990 Hisato Hamano wrote an article "Vertical typesetting in TEX" (`http://tug.org//TUGboat/Articles/tb11-3/tb29hamano.pdf`) in *TUGboat*.

DW: What sorts of projects is TEX used for in Japan beyond typesetting academic math? Is it widely used in the general publishing business, or are commercial systems like InDesign and QuarkXPress used by most Japanese publishers as is the case in the United States?

HO: TEX is used here for typesetting computer books as well as math and physics books. ASCII Corporation created a frontend to pTEX, called Editor's Work Bench (`http://www2.ascii.co.jp/ascii/EWB/`), mainly for the company's internal use. I hear it is much easier to use for non-techie editors. TEX is also used as a free substitute to PDFlib for Web-based systems. My university uses TEX to generate paper syllabi books from online syllabus system. But on the whole the trend is toward QuarkXPress and InDesign for publishers and Microsoft Word for authors. I don't think we can change the trend.

DW: You are listed as an editorial board member of the *Asian Journal of TEX*. Do you expect to participate heavily in that? Also, I am a little puzzled how a journal which apparently will have articles written in Korean, Japanese, and Chinese can work. I didn't think a reader of one of these languages can necessarily read another of them, or is the idea that readers of each language will find articles he or she can read in his or her native language?

HO: The CJK languages are so different from Latin ones that a journal dedicated to CJK typesetting is worth publishing. I'd like to participate in *AJT* as heavily as I can, but my time is quite limited.

DW: To what extent, if any, does your university encourage your involvement with TEX, and do you teach TEX to your students?

HO: I wish my university ever encouraged my involvement with TEX at all! I teach TEX, XHTML, CSS, C, Java, and all that my students should know, but almost all of them use Word to write essays and theses.

DW: Thank you for taking the time to participate in this interview and for educating me about some of the issues relating to TEX use in your country.