


# pT<sub>E</sub>X-ng

马起园

2014 年 11 月—12 月



pT<sub>E</sub>X-ng 作者近照 → 

pT<sub>E</sub>X-ng 是一个**编译器**  
T<sub>E</sub>X 是一种**动态语言**

# pT<sub>E</sub>X-ng 是什么

pT<sub>E</sub>X-ng 是下一代的 pT<sub>E</sub>X。在底层引擎上支持汉字处理、禁则处理、汉字和西文间距处理、汉字直排。目前只支持 UTF-8 编码，覆盖有中日韩三种语言支持。

pT<sub>E</sub>X 是日本 **ASCII** 公司开发的汉字处理引擎。在 2008 年和 2010 年经历合并以及收购之后，pT<sub>E</sub>X 的实际开发已经停止，已经转向由社区维护。

pT<sub>E</sub>X-ng 的开发最早可以追溯到 2012 年春节假期，最早在 LuaT<sub>E</sub>X 进行试验性质的修改，在 2013 年一度中断开发，正式版在今年（2014 年）10 月份开始发布。由于作者的时间有限，目前只做漏洞修补，下一部分大规模开发要拖到 2015 年春节。pT<sub>E</sub>X-ng 是自由软件，分发遵循 GPL 第二版许可。

# pT<sub>E</sub>X-ng 的开发概要

pT<sub>E</sub>X-ng 的前身是 Y&Y 公司的一个用 C 语言实现的带内存管理的 T<sub>E</sub>X。早期的源代码是使用 web2c 转换，不具可读性和可扩展性，这部分代码在 2014 年上半年进行了重写。pT<sub>E</sub>X-ng 的源代码建立在这部分代码的基础之上。

pT<sub>E</sub>X-ng 的汉字处理相关代码来源于 pT<sub>E</sub>X；Unicode 编码处理代码来源于 upT<sub>E</sub>X。对于  $\epsilon$ -T<sub>E</sub>X 的支持来源于 eupT<sub>E</sub>X。

pT<sub>E</sub>X-ng 只支持 PDF 文件输出，相关的代码来源于 dvipdfmx，没有任何来自 PDF<sub>T</sub><sub>E</sub>X 的代码，所以大量 primitive 不被支持。

# pT<sub>E</sub>X-ng 的将来

pT<sub>E</sub>X-ng 在将来会支持 OpenType 字体、绑定动态语言扩展宏编程。

- ① Lua, Lua 5.2/LuaJIT
- ② JavaScript, MuJS/V8/SpiderMonkey
- ③ Ruby, mruby
- ④ Scheme, GNU Guile

# pT<sub>E</sub>X-ng 的 catcode

0	escape character	10	space
1	beginning of group	11	letter
2	end of group	12	other character
3	math shift	13	active character
4	alignment tab	14	comment character
5	end of line	15	invalid character
6	parameter	16	kanji
7	superscript	17	hiragana, katakana, alphabet
8	subscript	18	cjk symbol codes
9	ignored character	19	hangul codes



# pT<sub>E</sub>X-ng 的 primitive

- `\jfont` 和 `\tfont`，这两个和 `\font` 语义一致。pT<sub>E</sub>X 中的 TFM 有两种，一种是 T<sub>E</sub>X82 定义了的 TFM 格式；另一种是 pT<sub>E</sub>X 定义的专门用于重载字符宽度的 JFM（后文会专门介绍 JFM）。在 JFM 中，已经定义了对应的文字方向，所以无论使用上述三种那个都不会产生问题。另外，还支持级数和齿数，即 Q 和 H，在大小上为 0.25cm，在传统意义上只有前者可用于定义字体大小，后者用来支配其他非字体元素的尺寸。
- `\font\tenmin=upjisr-h % mincho(KANJI)`  
`\font\sevenmin=upjisr-h at 7pt`  
`\font\fivemin=upjisr-h at 5pt`

# pT<sub>E</sub>X-ng 的 primitive

- `\kanjiskip` 和 `\xkanjiskip`, 前者会被自动插入到汉字之间, 后者会自动插入到非汉字与汉字之间。在 T<sub>E</sub>X82 上, 可以使用 `ex` 和 `em`。而在 pT<sub>E</sub>X 中, 可以使用 `zw` 和 `zh`, `z` 表示 *zenkaku* (全角, ぜんかく), 前者为全角宽度, 后者为全角高度。
- `\kanjiskip=0pt plus .4pt minus .4pt`  
`%\xkanjiskip=2.5pt plus 1pt minus 1pt`  
`\xkanjiskip=.25zw plus 1pt minus 1pt`

# pT<sub>E</sub>X-ng 的 primitive

- `\tbaselineshift` 和 `\ybaselineshift`, 控制基线的偏移值。前者为直行文本的偏移量, 请注意直行的汉字的基线位垂直方向的中心; 后者为横行文本的基线偏移量。
- `\tbaselineshift=3pt`  
`\ybaselineshift=3pt`
- `林``hayasi``林`, `林``hayasi``林`

# pTeX-ng 的 primitive

- `\ngbanner`, 输出默认 banner。
- This is pTeX-ng, Version 3.14159265
- `\ngostype`, 输出当前系统名称。
- Win64

# 特别感谢

罗心澄	苏杰	齐亮	刘海洋	高虎	孙亮	徐腾飞
夏晓昊	黄晨成	黄金泽	梁海宇	孙伟	都龙山	王侠兵
李国宝	胡玉进	杨海宇	周浩	曹梦迪	李任之	卢焱
吕文博	翟羽	钟毓	肖智博	罗希睿	李清	张康
王政	戴唯思	李天池	乔崇智	林坤	朱文俊	张麒
王昭礼	常金龙	冷俊园	高学远	张伟文	杨笑生	张潇
王昊	林科	朱焕杰	李润泽	王璐	杨林恒	陈甫
孙文全	荣健欣	汤进伟	唐俊荣	郝运	罗晨星	朱翀
何春奇	林懿伦	刘思江	金晨羽	陈浩	刘中阳	王冠
杨文清	贾泊	王昊	朱宁	朱佳宁	陈欣	冷轩
鲁尚文	韩勇杰					

# 休憩