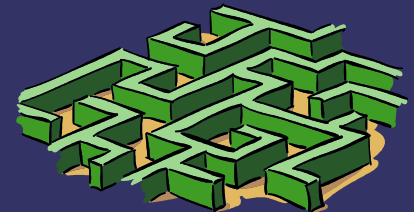


Towards Perfect Indic Rendering on Gnome

For Gnome Asia Summit,2011
2-3 Apr 2011, Bangalore

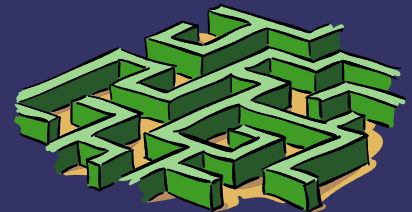
Arjuna Rao Chavala
Independent Software Technologist
Arjunaraoc [at] gmail.com
Copyright: CC-BY-SA-3.0

(Exceptions Unicode Telugu Table Image and image from Amir
Wali's arabic text flow map)



Agenda

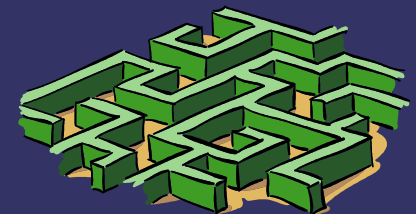
- ⇒ Timeline of indic computing
- ⇒ Gnome Text rendering stack
- ⇒ Issues in Indic rendering
- ⇒ Unicode Table(Telugu)
- ⇒ Open type font (Lohit-te)
- ⇒ Bugs in Pango
- ⇒ Harfbuzz-ng timeline and features
- ⇒ Harfbuzz-ng indic potential contributors
- ⇒ Summary
- ⇒ References



Timeline of Indic Computing*

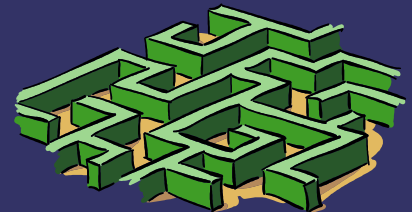
- ⇒ 2000-Windows OS support indic without special HW
- ⇒ 2005-Firefox build to render Indian languages
- ⇒ 2005-Padma Firefox plugin to display sites which use non uni-code fonts
- ⇒ 2006-Localisation of Desktop
- ⇒ 2007-Localisation of Office Suite
- ⇒ 2009-Commercial companies start offering Indic computing support(Google/Yahoo/Rediff/Microsoft)
- ⇒ 2005- Cell phones start sporting Indic languages

*Indicative Timeline

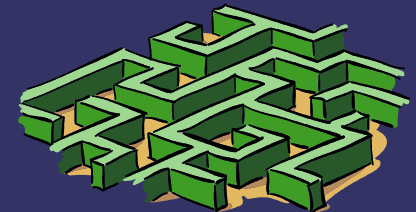
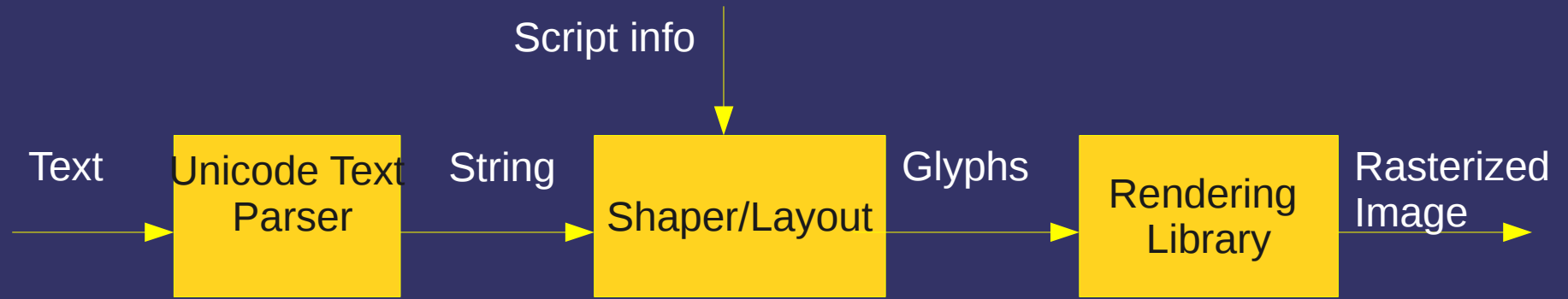


Indic rendering software

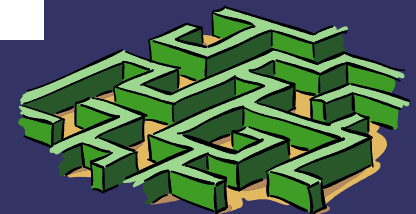
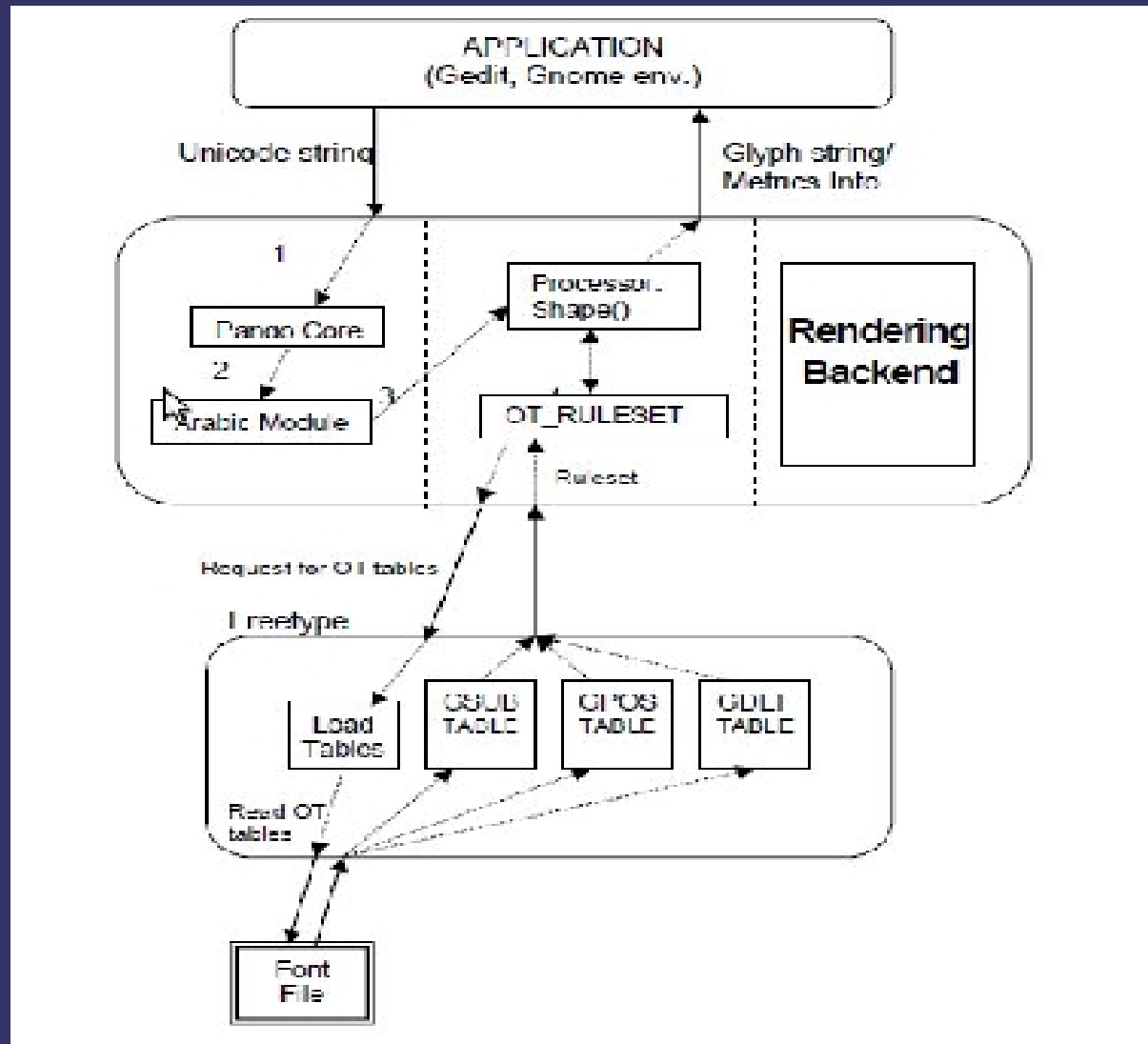
- ⇒ Pango
 - All Gnome applications
- ⇒ QT
 - All KDE applications
- ⇒ ICU
 - OpenOffice/Liberoffice



Gnome Text Rendering stack

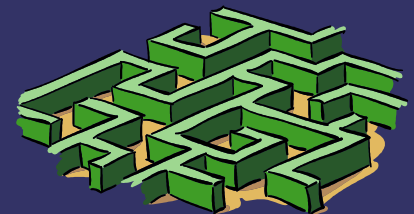


Text rendering flow (eg:Arabic)



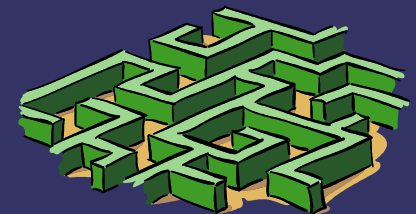
Issues in Indic rendering

- ⇒ Too many softwares
- ⇒ Single code base for each software.
 - Potential for bug fixes to introduce new bugs
- ⇒ Font encoding diversity
 - Unicode



Unicode Telugu

	0C0	0C1	0C2	0C3	0C4	0C5	0C6	0C7
0	ఱ	ఱ	ఱ	ఱ	ఱ	ఱ	ఱ	ఱ
1	ఱ	ఱ	ఱ	ఱ	ఱ	ఱ	ఱ	ఱ
2	ఱ	ఱ	ఱ	ఱ	ఱ	ఱ	ఱ	ఱ
3	ఱ	ఱ	ఱ	ఱ	ఱ	ఱ	ఱ	ఱ
4	ఱ	ఱ	ఱ	ఱ	ఱ	ఱ	ఱ	ఱ
5	ఱ	ఱ	ఱ	ఱ	ఱ	ఱ	ఱ	ఱ
6	ఱ	ఱ	ఱ	ఱ	ఱ	ఱ	ఱ	ఱ
7	ఱ	ఱ	ఱ	ఱ	ఱ	ఱ	ఱ	ఱ
8	ఱ	ఱ	ఱ	ఱ	ఱ	ఱ	ఱ	ఱ
9	ఱ	ఱ	ఱ	ఱ	ఱ	ఱ	ఱ	ఱ
A	ఱ	ఱ	ఱ	ఱ	ఱ	ఱ	ఱ	ఱ
B	ఱ	ఱ	ఱ	ఱ	ఱ	ఱ	ఱ	ఱ
C	ఱ	ఱ	ఱ	ఱ	ఱ	ఱ	ఱ	ఱ
D	ఱ	ఱ	ఱ	ఱ	ఱ	ఱ	ఱ	ఱ
E	ఱ	ఱ	ఱ	ఱ	ఱ	ఱ	ఱ	ఱ
F	ఱ	ఱ	ఱ	ఱ	ఱ	ఱ	ఱ	ఱ



Open Type Telugu Font (Lohit)

Names

General
Layers
PS UID
PS Private
OS/2
TTF Names
StyleSet Names
Grid Fitting
TeX
Size
Comment
FONTLOG
Mark Classes
Mark Sets
Lookups
Mac
Mac Features
Dates
Unicode Ranges

Fontname: Lohit-Telugu

Family Name: Lohit Telugu

Name For Humans: Lohit Telugu

Weight: Book

Version: 2.4.3

Same as Fontname

Base filename:

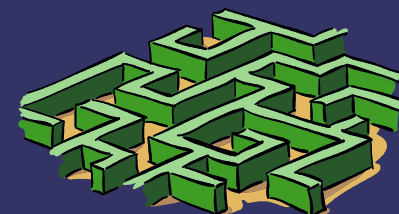
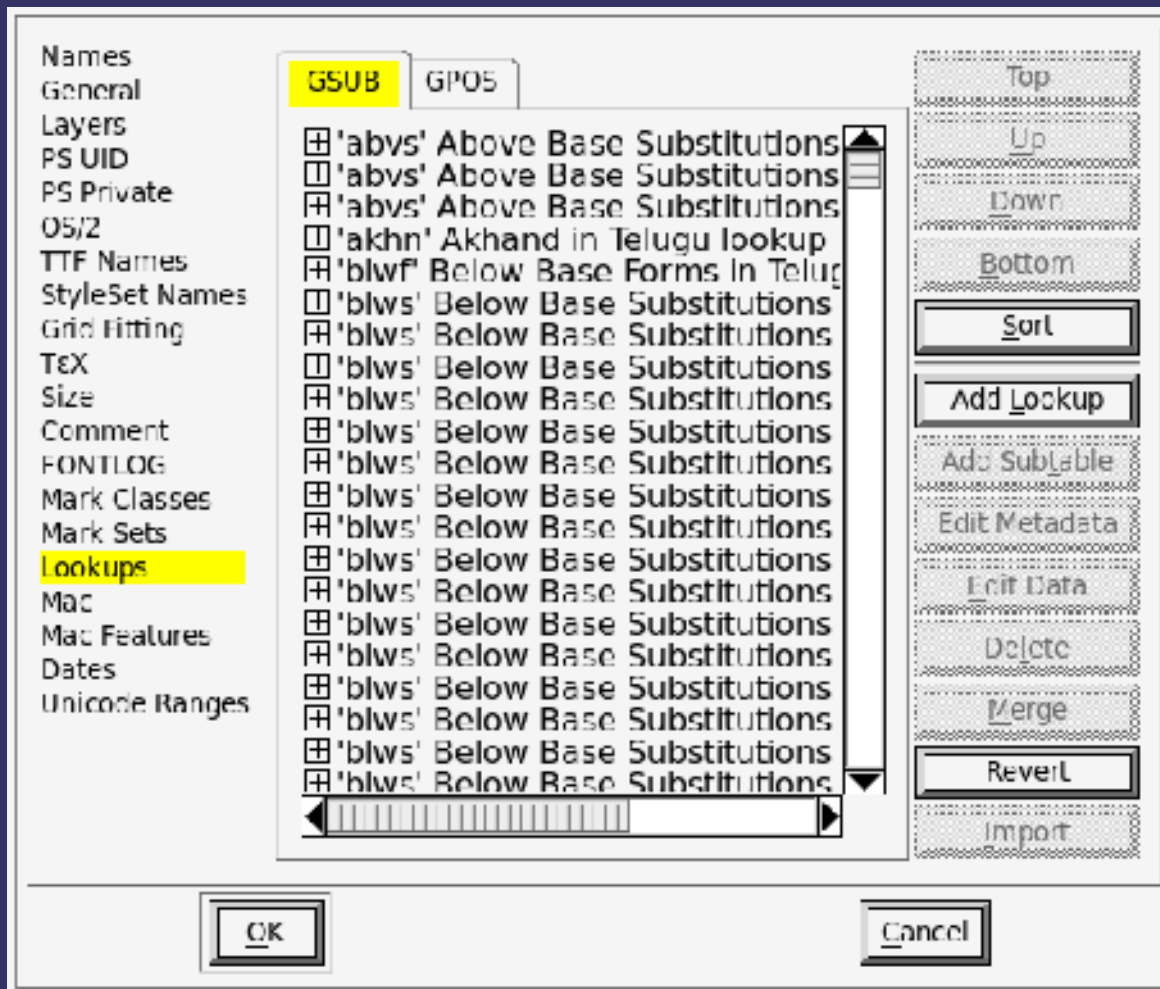
Copyright:

```
Copyright (c) 2003, Automatic Control  
Equipments, Pune, INDIA.  
Copyright 2006 Modular Infotech Pvt Ltd.  
Copyright 2009 Red Hat, Inc.  
  
Licensed under the GNU General Public License,  
version 2 (see file COPYING) along with the  
following exception:  
  
As a special exception, if you create a
```

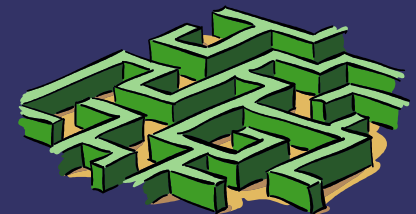
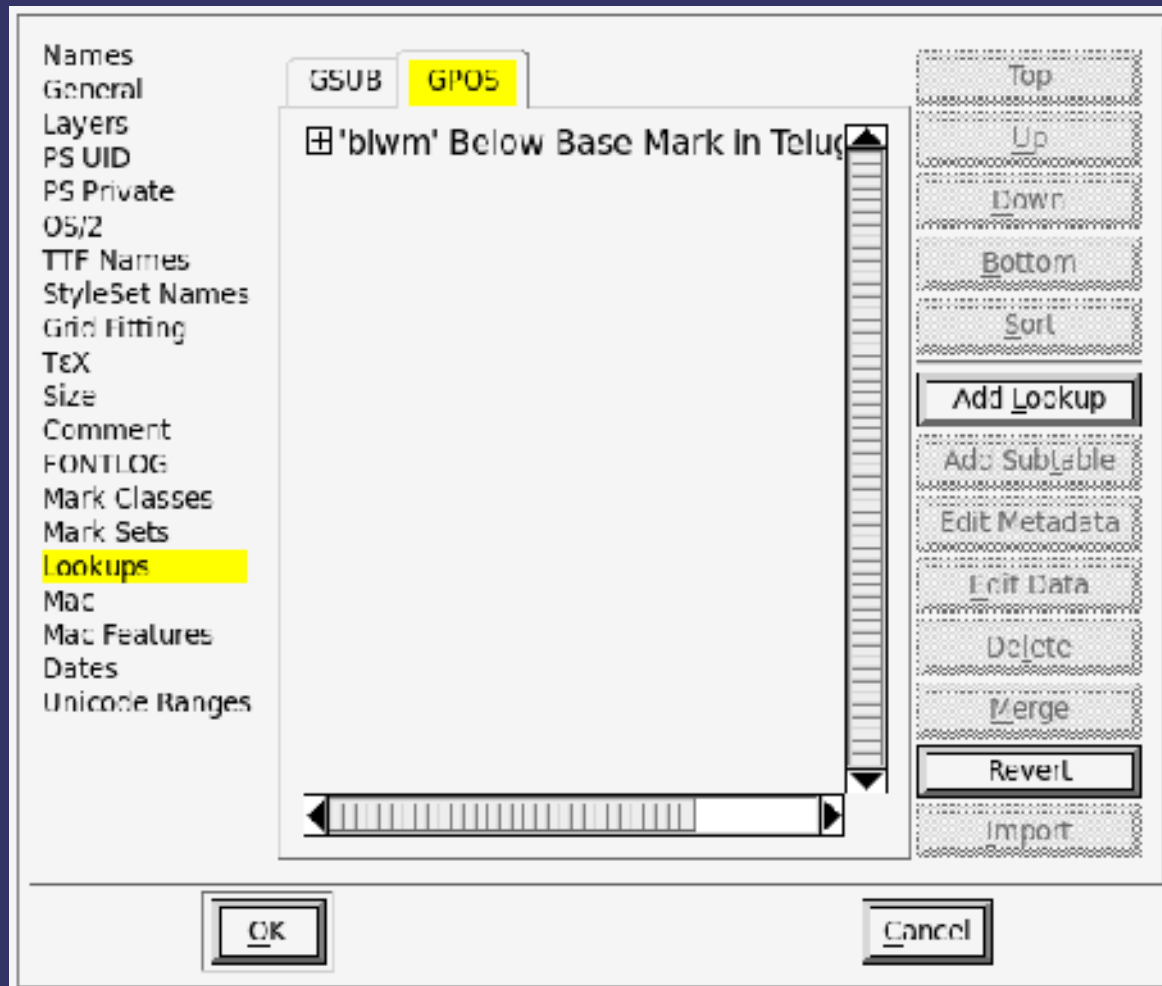
OK **Cancel**



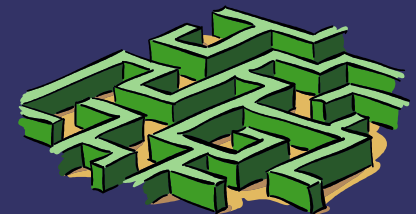
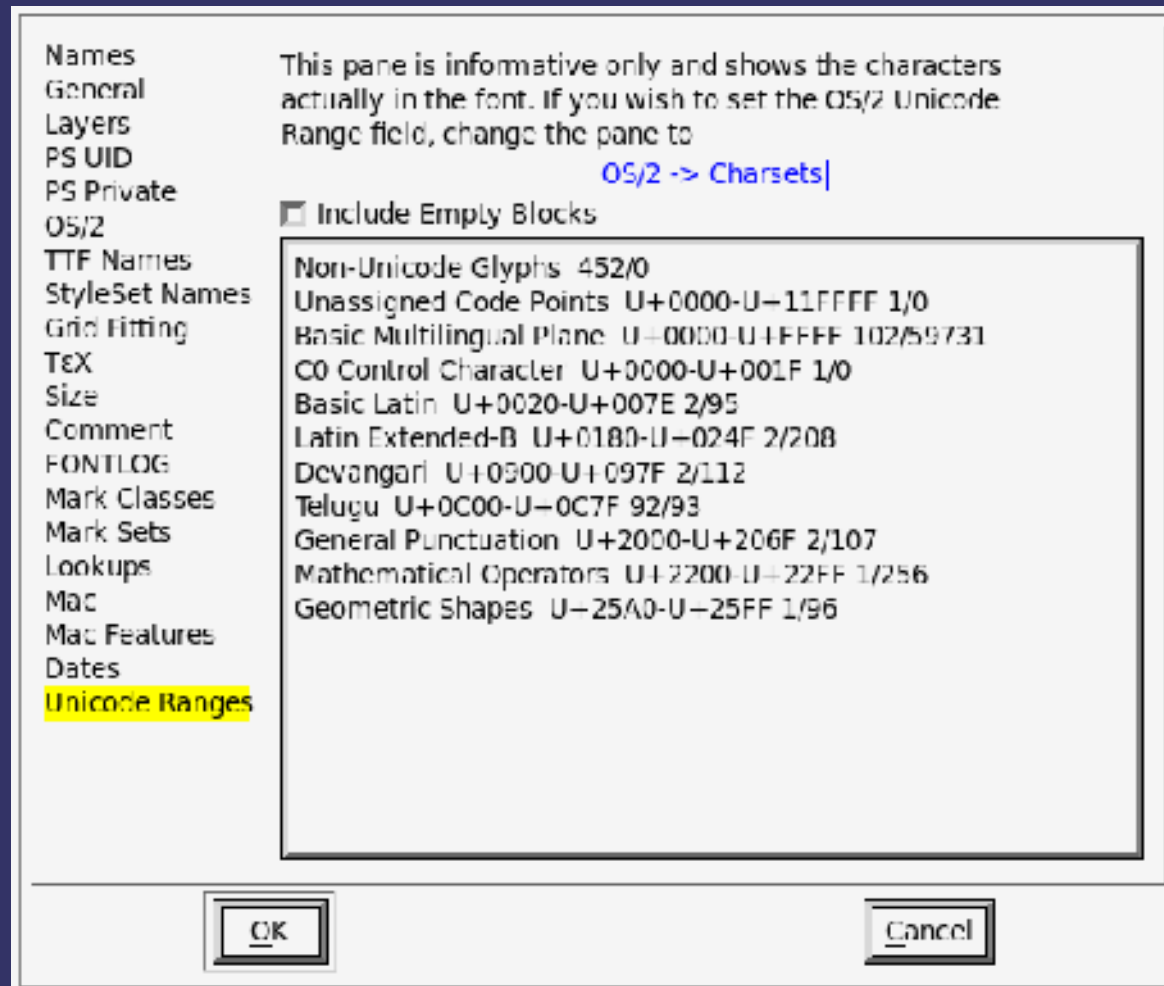
Open Type Telugu Font (Lohit)



Open Type Telugu Font (Lohit)

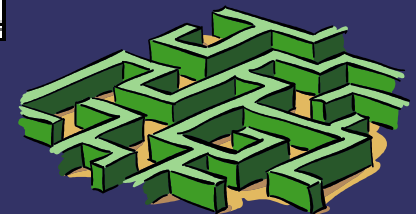


Open Type Telugu Font (Lohit)



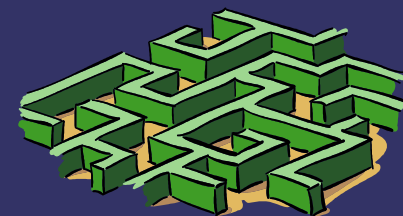
Open Type Telugu Font (Lohit)

File Edit Element Tools Hints Encoding View Metrics CID MM Window Help												
040 0x02101 0x02102 un:010 1ATTN CAPITAL LETTER YOGI												
0c00	0c01	0c02	0c03	0c04	0c05	0c06	0c07	0c08	0c09	0c0a	0c0b	
	ఠ	ఠ	ఠ		అ	ఆ	ఇ	ఈ	ఉ	ఊ	ఋ	
0c0c	0c0d	0c0e	0c0f	0c10	0c11	0c12	0c13	0c14	0c15	0c16	0c17	
	ృ	వ	వి	ఐ		ఋ	ౠ	ౡ	క	ఖ	గ	
0c18	0c19	0c1a	0c1b	0c1c	0c1d	0c1e	0c1f	0c20	0c21	0c22	0c23	
	ఘ	బ	చ	ఛ	జ	ఝ	ట	ఠ	డ	ఢ	ణ	
0c24	0c25	0c26	0c27	0c28	0c29	0c2a	0c2b	0c2c	0c2d	0c2e	0c2f	
	త	థ	ద	ధ	న		ప	ఫ	బ	భ	మ	య
0c30	0c31	0c32	0c33	0c34	0c35	0c36	0c37	0c38	0c39	0c3a	0c3b	
	ర	ల	ల	ళ		వ	శ	ష	స	హ		
0c3c	0c3d	0c3e	0c3f	0c40	0c41	0c42	0c43	0c44	0c45	0c46	0c47	
		ౌ	ౌ	ౌ	ృ	ృ	ృ	ృ		ౄ	ౄ	
0c48	0c49	0c4a	0c4b	0c4c	0c4d	0c4e	0c4f	0c50	0c51	0c52	0c53	
	౅	౅	౅	౅	౅							
0c54	0c55	0c56	0c57	0c58	0c59	0c5a	0c5b	0c5c	0c5d	0c5e	0c5f	
	ె	ె		ె	ె							
0c60	0c61	0c62	0c63	0c64	0c65	0c66	0c67	0c68	0c69	0c6a	0c6b	
	ఋ	ౠ	ౡ			ౢ	ౣ	౤	౥	౦	౧	
0c6c	0c6d	0c6e	0c6f	0c70	0c71	0c72	0c73	0c74	0c75	0c76	0c77	
	౨	౩	౪									



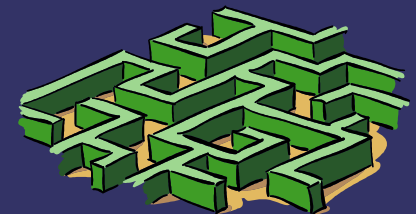
Open Type Telugu Font (Lohit)

fff0	fff1	fff2	fff3	fff4	fff5	fff6	fff7	fff8	fff9	fffa	fffb
				10000	10001	10002	10003	10004	10005	10006	10007
				L	□	□	□	□	□	□	□
10008	10009	1000a	1000b	1000c	1000d	1000e	1000f	10010	10011	10012	10013
□	□	□	□	□	కా	ఖా	గా	ఘా	జా	చా	
10014	10015	10016	10017	10018	10019	1001a	1001b	1001c	1001d	1001e	1001f
ఛా	జా	ఝా	ఞా	టా	ఠా	డా	ణా	తా	థా	దా	
10020	10021	10022	10023	10024	10025	10026	10027	10028	10029	1002a	1002b
ధా	నా	పా	ఫా	బా	భా	మా	యా	రా	లా	ళా	
1002c	1002d	1002e	1002f	10030	10031	10032	10033	10034	10035	10036	10037
వా	శా	షా	సా	హా	క్షా	కా	ఖా	గా	ఘా	జా	చా
10038	10039	1003a	1003b	1003c	1003d	1003e	1003f	10040	10041	10042	10043
ఛా	జా	ఝా	ఞా	టా	ఠా	డా	ణా	తా	థా	దా	
10044	10045	10046	10047	10048	10049	1004a	1004b	1004c	1004d	1004e	1004f
ధా	నా	పా	ఫా	బా	భా	మా	యా	రా	లా	ళా	
10050	10051	10052	10053	10054	10055	10056	10057	10058	10059	1005a	1005b
చా	ఛా	జా	ఝా	ఞా	టా	ఠా	డా	ణా	తా	థా	దా
1005c	1005d	1005e	1005f	10060	10061	10062	10063	10064	10065	10066	10067
ఛా	జా	ఝా	ఞా	టా	ఠా	డా	ణా	తా	థా	దా	

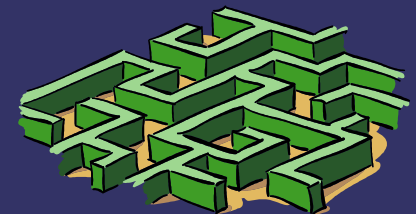
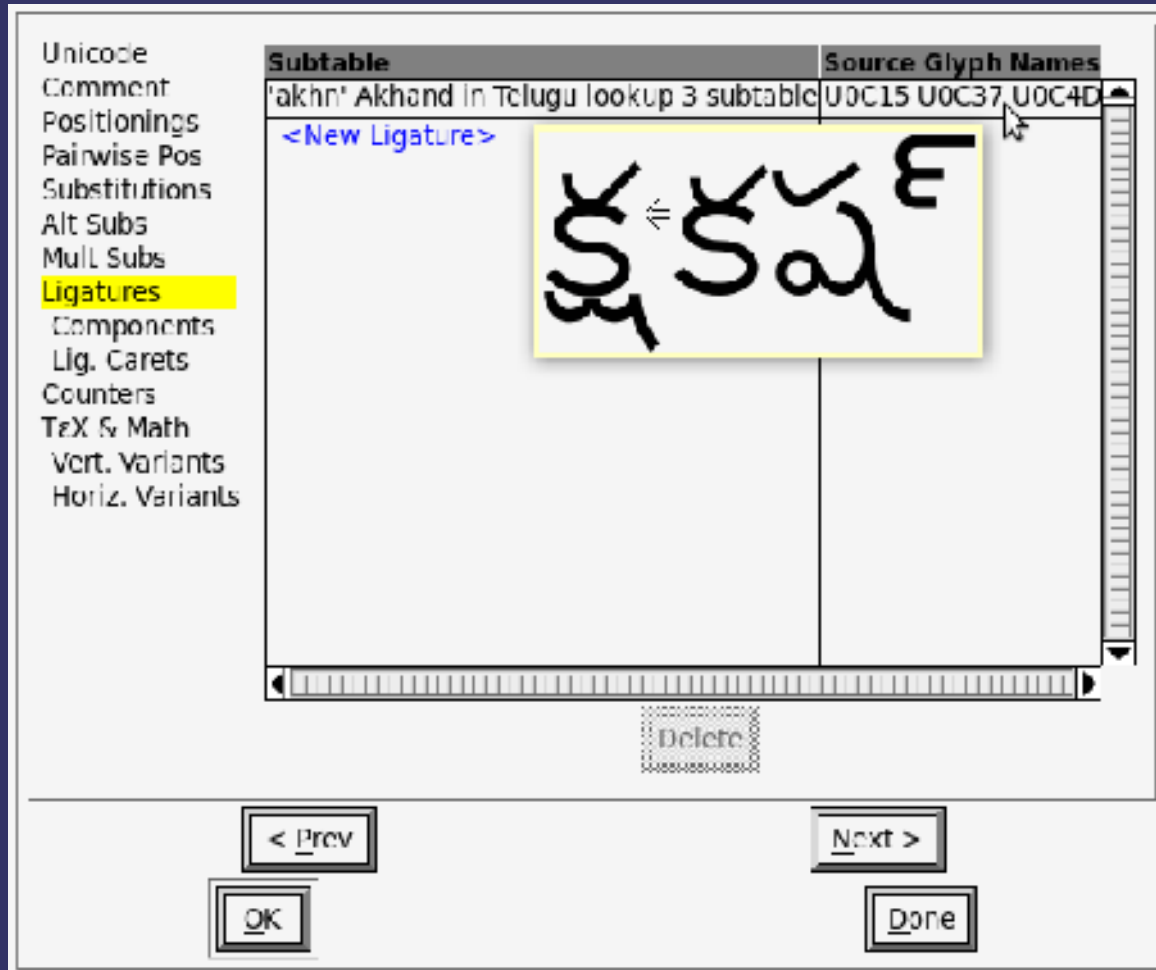


Open Type Telugu Font (Lohit)

File Edit Element Tools Hints Encoding View Metrics CID MM Window Help										
65922 (0x0182) 00000000 00000000 00000000 00000000										
1014d	1014e	1014f	10150	10151	10152	10153	10154	10155	10156	10157
వ	క	ష	న	హ	క్ష	జ	ఝ	వు	పూ	పు
10158	10159	1015a	1015f	1015c	1015d	1015e	1015f	10160	10161	10162
వు	పూ	ర	ఖ	ఁ	లు	జ	చ	చ	ఙ	ఱ
10164	10165	10166	10167	10168	10169	1016a	1016b	1016c	1016d	1016e
ఱ	ఱ	ఱ	ఱ	ఱ	ఱ	ఱ	ఱ	ఱ	ఱ	ఱ
10170	10171	10172	10173	10174	10175	10176	10177	10178	10179	1017a
ఱ	ఱ	ఱ	ఱ	ఱ	ఱ	ఱ	ఱ	ఱ	ఱ	ఱ
1017c	1017d	1017e	1017f	10180	10181	10182	10183	10184	10185	10186
ఱ	ఱ	ఱ	ఱ	ఱ	ఱ	ఱ	ఱ	ఱ	ఱ	ఱ
65922 0x0182 0-7777 *00000000_00000000_00000000	10186	10187	10188	10189	1018a	1018b	1018c	1018d	1018e	1018f
ఱ	ఱ	ఱ	ఱ	ఱ	ఱ	ఱ	ఱ	ఱ	ఱ	ఱ
10194	10195	10196	10197	10198	10199	1019a	1019b	1019c	1019d	1019e
ఱ	ఱ	ఱ	ఱ	ఱ	ఱ	ఱ	ఱ	ఱ	ఱ	ఱ
101a0	101a1	101a2	101a3	101a4	101a5	101a6	101a7	101a8	101a9	101aa
ఱ	ఱ	ఱ	ఱ	ఱ	ఱ	ఱ	ఱ	ఱ	ఱ	ఱ
101ac	101ad	101ae	101af	101b0	101b1	101b2	101b3	101b4	101b5	101b6
ఱ	ఱ	ఱ	ఱ	ఱ	ఱ	ఱ	ఱ	ఱ	ఱ	ఱ
101b8	101b9	101ba	101bb	101bc	101bd	101be	101bf	101c0	101c1	101c2
ఱ	ఱ	ఱ	ఱ	ఱ	ఱ	ఱ	ఱ	ఱ	ఱ	ఱ



Open Type Telugu Font (Lohit)



Fix of Font rendering bugs for Telugu, Kannada

- Reph handling for Telugu
- Halanth ending of conjugates
- Telugu akhand problem
 - (కా+ప్) correct: కాప్ wrong: క్షా
- Kannada akhand problem

సంస్కృతము
సంస్కృతము

ఫైర్‌ఫాక్స్ లోహిత్ తెలుగు ఫాంటు
ఫైర్‌ఫాక్స్ పోతన ఫాంటు

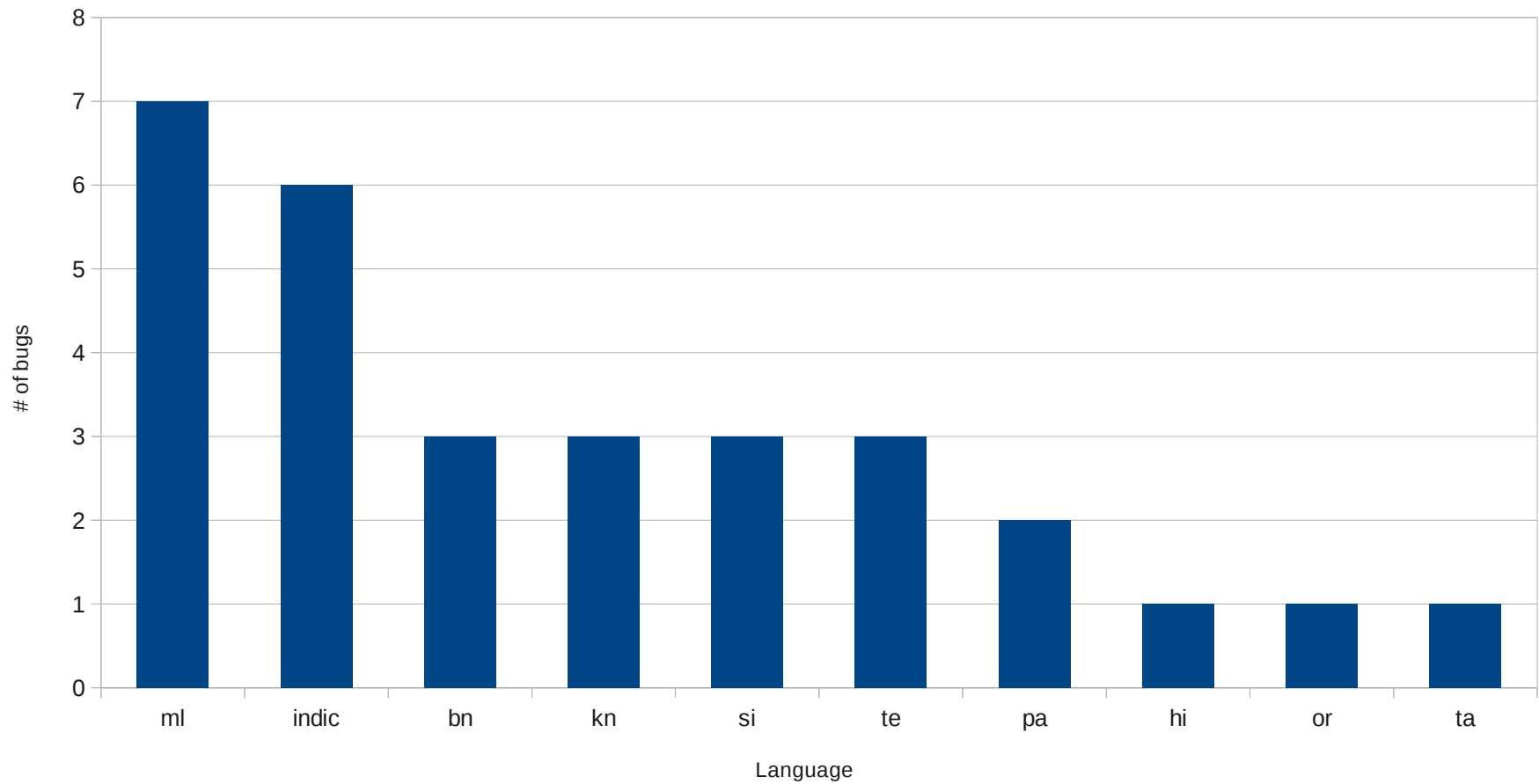
Actual: ಕ್ಷ
Correct: ಕ್ಷ
Key sequence :
Inscript: k + d + k + <
Rawcode: 0c95+0ccd+0c95+0cb7



Pango Open bugs -Language

Pango Open bugs

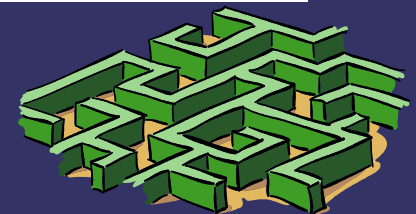
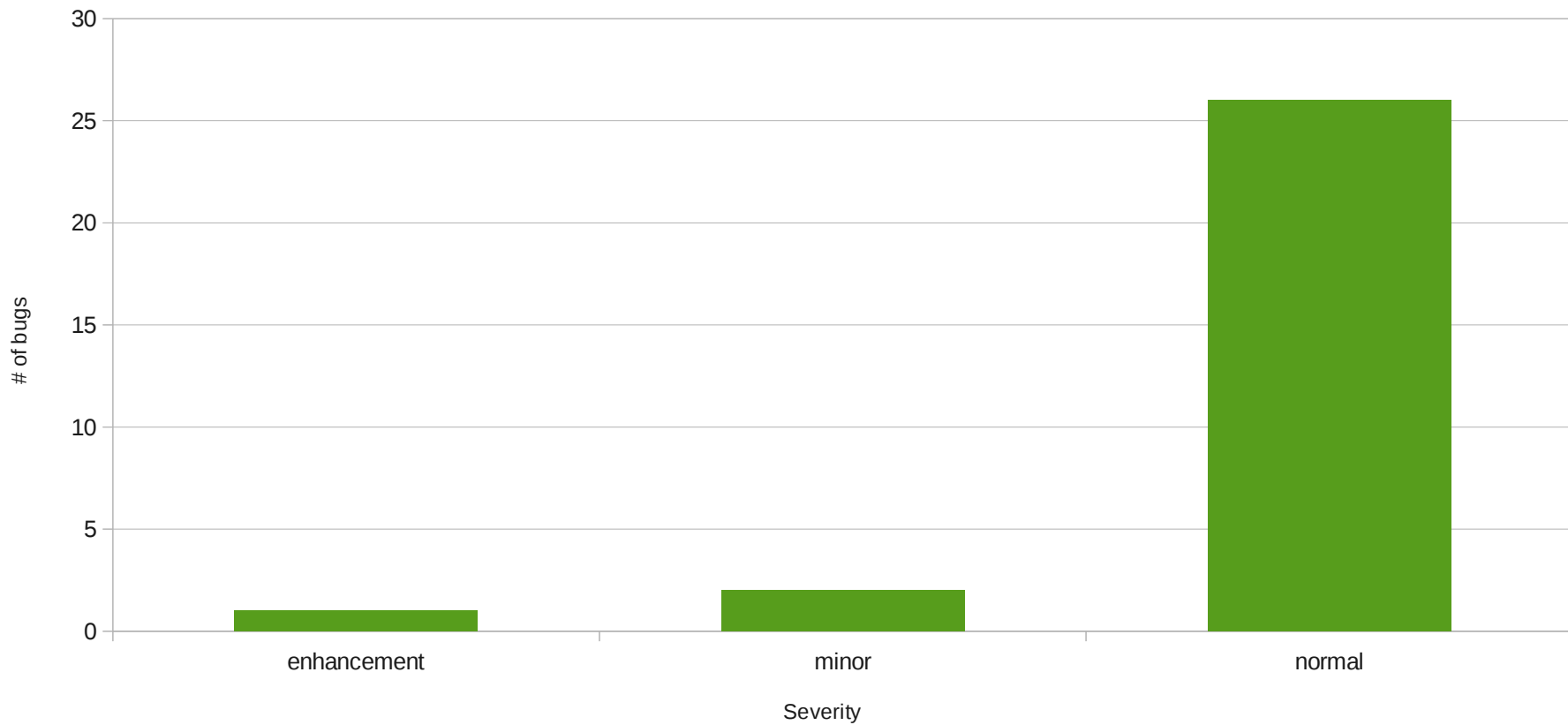
as on 2011 Mar 30



Pango Indic Bugs -Severity

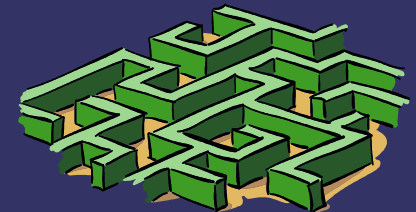
Pango Indic Open bugs

as on 2011 Mar 30



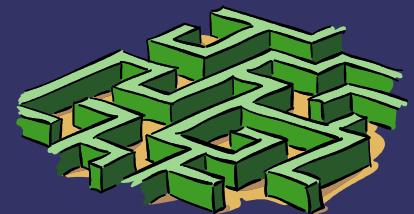
Harfbuzz-Ng Timeline

- ⇒ 2009-07-05: First public version of need for unifying rendering as presented at the Gran Canaria Desktop Summit by Behdad (See reference)
- ⇒ 2009-08-06 Harfbuzz-ng layout code supporting Latin shaping merged into Pango
- ⇒ 2010-10-13 Arabic/Syrian/N'Ko shaper working
- ⇒ 2011-03-22 Firefox 4 ships with Harfbuzz-ng
- ⇒ Future
 - 1.0 with Stable API and shaping for Latin, Arabic by Mar 2011
 - Hebrew, Hangul ... TBD
 - New Tai Lue, Indic ...TBD



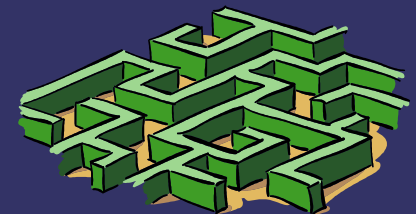
Harfbuzz-Ng API Features

- ⇒ Usability top priority (best practice from Cairo)
- ⇒ Threadsafe
- ⇒ Backend agnostic API (hb.h)
- ⇒ Open type specific API (hb-ot.h)
- ⇒ Call backs for font/ Unicode
- ⇒ Use of Font data directly



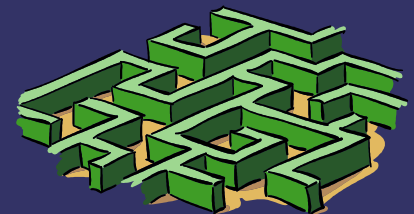
Harfbuzz-Ng Indic potential contributors

- Indic: Tom Hacoen
- Telugu, Arjuna Rao Chavala, Kesavardhana Gandla
- Hindi, Pravin Satpute
- Tamil, Muguntharaj Subramanian



Summary

- ⇒ Perfect rendering of Indic languages on Computers, Smart-phones critical for business and personal use
- ⇒ Need for Indians to contribute to the initiative.



References

- ➔ Wikipedia article on Pango
 - <http://en.wikipedia.org/wiki/Pango>
- ➔ State of Text Rendering
 - <http://behdad.org/text/>
- Unicode standard
 - <http://unicode.org/>
- Harfbuzz-ng API thread
 - <http://lists.freedesktop.org/archives/harfbuzz/2009-August/000359.html>
- ➔ Harfbuz-ng
 - <http://cgit.freedesktop.org/harfbuzz/>
 - <git://anongit.freedesktop.org/harfbuzz>
 - <ssh://git.freedesktop.org/git/harfbuzz>
- ➔ Mailing list
 - <http://lists.freedesktop.org/mailman/listinfo/harfbuzz>
 - <https://lists.sourceforge.net/lists/listinfo/indlinux-group>

