

ویژگی‌های یک قلم فارسی استاندارد

نویسنده: سعید رسولی

مقدمه‌ای بر انکودینگ

به الگوریتم تبدیل سلسله‌ای از نمادهای نوشتاری (کاراکترها) به سلسله‌ای از بایت‌ها (برای ذخیره‌سازی و انتقال در فضای دیجیتال) انکودینگ (encoding) می‌گویند (که طبیعتاً شامل الگوریتم عکس این عمل یعنی decoding هم می‌شود) انکودینگ‌ها شامل یک جدول از کاراکترها (character table) یا مجموعه کاراکترها (character set) هستند (به همین دلیل اصطلاح charset گاهی بجای اصطلاح encoding به کار می‌رود) انکودینگ‌های مختلفی وجود دارند که اغلب آن‌ها تنها از یک یا دو زبان یا در بهترین حالت از خانواده‌ای از زبان‌ها پشتیبانی می‌کنند. اما تنها انکودینگ‌های خانواده یونیکد هستند که از تمام زبان‌ها (و حتی بیشتر از آن!) پشتیبانی می‌کنند.

ساده‌ترین و پایه‌ای‌ترین انکودینگ، ASCII (اسکی) است که تنها شامل حروف و علائم زبان انگلیسی است.

انکودینگ arabic windows (که windows 1256 نام دیگر آن است)، شامل علائم اسکی و حروف زبان عربی و اغلب حروف فارسی است (غیر از «ی» و «ک» فارسی که با «ی» و «ک» عربی متفاوت است) که تنها مزیت آن کم‌حجم بودن متن خروجی است. یعنی اگر متنی به زبان انگلیسی و فارسی باشد، هر کاراکتر تنها شامل یک بایت است

انکودینگ‌های خانواده یونیکد، همه یک مجموعه مشترک و کامل از نمادها را پشتیبانی می‌کنند که شامل حروف تمام زبان‌های انسانی و یک سری علائم و نمادهای دیگر (مثل فلش و اشکال هندسی و علائم ریاضی و حتی علائم راهنمایی و رانندگی) می‌شود. اما روش تبدیل یا نگاشت این نمادها به بایت‌ها، بین انکودینگ‌های مختلف خانواده یونیکد متفاوت است. یک زیرمجموعه از این نمادها به نام BMP مخفف Basic Multilingual Plane (صفحه چندزبانه پایه‌ای) تعریف شده است که شامل 65536 نماد است (که قابل ذخیره‌سازی در ۲ بایت هستند) این نمادها اغلب کاراکترهای رایج بیشتر زبان‌های دنیا (از جمله فارسی و عربی) را شامل می‌شود.

انکودینگ UTF-8 از خانواده یونیکد بوده و سازگار با اسکی است (یعنی هر متن اسکی، با تبدیل به UTF-8 هیچ تغییری نمی‌کند) و کاراکترهای فارسی و عربی ۲ بایت اشغال می‌کنند (کاراکترهای بعضی زبان‌ها مثل زبان‌های آسیای شرقی ممکن است ۳ بایت اشغال کند). ولی با توجه به این که از تمام زبان‌ها (از جمله فارسی) به بهترین و استانداردترین نحو ممکن پشتیبانی می‌کند، بهترین گزینه برای ذخیره‌سازی در فایل است.

انکودینگ UTF-16 نیز از خانواده یونیکد بوده، ولی تفاوتش با UTF-8 این است که کاراکترهای فارسی و انگلیسی همه ۲ بایت اشغال می‌کنند و کار پردازش متن بسیار راحت‌تر می‌شود. به طور دقیق‌تر، کاراکترهای مجموعه BMP (که قبل‌تر گفته شد) ۲ بایت، و سایر کاراکترها (که به ندرت استفاده می‌شوند) ۴ بایت اشغال می‌کنند مشکلش هم این است که سازگار با اسکی نیست و برای ذخیره‌سازی فایل‌های متنی مناسب نیست اما برای تعریف نمادها در قلم‌های مدرن و چندزبانه، بهترین و پرکاربردترین گزینه است

اگر شما هم مثل من عاشق یونیکد شده‌اید، با من این ترانه را بخوانید :-)

unicode, unicode
uni-uni-unicode
with you to the end of road
in the international mode
everything is made of code

مقدمه‌ای بر گرافیک برداری

در تصاویر برداری، پیکسل‌ها ذخیره نمی‌شوند. در عوض، مرز اشکال به صورت تعدادی نقطه (با مختصات مشخص)، تعدادی پاره‌خط و تعدادی خم بیزیر (Bézier) تعریف شده است. خم Bézیر یک منحنی درجه سوم است (که می‌توان به صورت تعمیمی از قطاع دایره و خط در نظر گرفت) گر چه از لحاظ تئوری، نمی‌توان همه اشکال را به این صورت تعریف کرد، ولی عملاً تمام منحنی‌ها را می‌توان با تعدادی خم «بیزیر» طوری تعریف کرد که تفاوتی توسط چشم انسان تشخیص داده نشود. نرم‌افزارهای خاصی برای تولید و ویرایش تصاویر برداری وجود دارد که Corel Draw بین کاربران ویندوز معروف است، ولی ما هم نرم‌افزار آزاد Inkscape را داریم که حقیقتاً کم از Corel نداشته و حتی از جهاتی بهتر است.

البته نرم افزار LibreOffice Draw نیز قابل ذکر است.
نرم افزار Inkscape از فایل فرمت SVG استفاده می کند که فایل SVG خروجی آن نه تنها در Draw بلکه در FontForge (که نرم افزاری برای طراحی قلم است) هم ابل import است، پس در طراحی قلم هم SVG و Inkscape (که رابط بسیار بهتری از FontForge دارد) به شما کمک بسیاری خواهند کرد
نرم افزار Inkscape به قدری خوب و جذاب است که حتی گاهی ترجیح می دهیم بجای لیبره آفیس از آن استفاده کنیم

دسته بندی قلم ها

قلم ها را بر اساس خصوصیات مختلف آن های می توان دسته بندی کرد، مهم ترین این خصوصیات:

۱- فرمت فایل

قلم ها را بر اساس فرمت یا پسوند فایل می توان به دسته های مختلف تقسیم کرد. بعنوان مثال:

- TrueType (.ttf)
- Adobe PostScript Type 1 (.pda, .pdb)
- OpenType (.otf, .ttf, .ttc)
- .pcf (xserver, bitmap)
- .psf (console, bitmap)
- .fon (windows)

ما فعلا در مورد تک تک این فرمت ها صحبت نخواهیم کرد، ولی به بعضی از آن ها که در گنو/لینوکس کاربرد دارد، در قسمت های بعدی اشاره می کنیم.
در همین حد بدانید که بهترین فرمت مدرن و آزاد که برای فارسی قابل استفاده است، TrueType است که پسوند ttf دارد.

۲- یونیکد و غیر یونیکد

قلم های فارسی که ده ها سال پیش برای سیستم عامل های ویندوز و مک طراحی می شدند، یونیکد نبودند. خوشبختانه امروز اغلب این قلم ها به یونیکد تبدیل شده اند. اما همان طور که می دانید، در محیط گرافیکی گنو/لینوکس، همه چیز بر پایه یونیکد است.

در واقع، قلمی که یونیکد نباشد، برای ما ارزش کاربردی ندارد (و احتمالا با توجه به بحث مجوز و تاریخچه نامعلوم آن، هیچ گونه ارزشی ندارد)

البته این در مورد قلم هایی بود که قرار است در محیط گرافیکی استفاده شود. و گرنه همین الان تعدادی قلم غیر یونیکد روی توزیع های گنو/لینوکس وجود دارد (مثل قلم های psf) که فقط در کنسول یا ترمینال های غیر گرافیکی استفاده می شوند. همان جایی که وقتی Alt+Ctrl+F1 می زنید می روید! آیا می دانستید با تعیین قلم مناسب (که البته یونیکد نیست) می توانید آن جا هم فارسی ببینید؟

نمونه ای دیگر از قلم های غیر یونیکد، قلم های PostScript نوع ۱ شرکت آدوبی است که در گنو/لینوکس احتمالا در چنین مسیری تعدادی از آن ها را می بینید:
usr/share/fonts/Type1/

به هر حال، چون هدف ما در این مقاله، محیط گرافیکی و سیستم عامل گنو/لینوکس است و به پشتیبانی از فارسی هم نیاز داریم، از این پس فقط در مورد قلم های یونیکد صحبت خواهیم کرد.

۳- برداری یا غیر برداری

با مقدمه ای که گفته شد، می توانید حدس زده باشید که اغلب قلم های مدرن از نوع برداری هستند. علت آن هم واضح است. چون ما متن را در اندازه های مختلف ممکن است ببینیم. اگر قرار باشد یک قلم نمادها را بصورت پیکسلی یا bitmap تعریف کرده باشد، باید بزرگترین اندازه ممکن را در نظر گرفته باشد (چون اگر بزرگتر از آن را بخواهیم، لبه های متن، مات خواهد شد و عملا غیر قابل استفاده است). حتی در آن صورت هم، کوچک کردن نمادهایی که بصورت تصاویر پیکسلی بزرگ تعریف شده اند، بار زیادی روی سیستم وارد می کند. یک قلم خوب اگر می خواهد bitmap هم باشد باید نمادها را در اندازه های مختلف قلم تعریف کرده باشد.

البته ممکن است یک قلم برداری، برای سازگاری بیشتر، تصاویر bitmap نمادها را هم در چند اندازه مختلف در کنار تصاویر برداری در خودش گنجانده باشد (کاری که در قلم ترافیک شرکت فارسی وب انجام شده)
ولی تقریبا هیچ قلم مدرنی (برای محیط گرافیکی) نیست که فقط بر اساس bitmap باشد.

لازم به ذکر است، قلم‌های psf برای ترمینال غیر گرافیکی (که پیش‌تر ذکر شد)، و قلم‌های pcf (مخفف Portable Compiled Format) برای xserver (که مثلاً در ترمینال xterm استفاده می‌شود) دو دسته از قلم‌های قدیمی به ارث مانده از یونیکس هستند که همه bitmap و غیربرداري هستند. تعداد زیادی از قلم‌های pcf را می‌توانید در گنو/لینوکس در چنین مسیری ببینید: usr/share/fonts/75dpi/

۴- عرض متغیر و عرض ثابت (مونو)

در هر قلمی، هر نماد در یک مستطیل فرضی تعریف شده است. که هنگام render کردن متن توسط قلم، این مستطیل‌ها در کنار هم قرار می‌گیرند (مگر در موارد خاص مثل حرکت‌ها) ارتفاع این مستطیل‌ها معمولاً یکسان است. اما عرض آن‌ها در اکثر قلم‌ها متغیر است. چون این خاصیت زبان‌های معمول نوشتاری است که عرض حروف با یکدیگر متفاوت است، و یکی کردن عرض این مستطیل‌های فرضی، به معنی گذاشتن فاصله اضافی در دو طرف حروف کم‌عرض‌تر است (و یا عریض‌تر کردن خود حروف)، که این خوانایی متن را پایین می‌آورد. مگر آنکه واقعاً نیاز به این ثابت بودن عرض نمادها داشته باشیم. مثلاً هنگام برنامه‌نویسی، ثابت بودن عرض نمادها باعث می‌شود که نمادهایی که در جایگاه‌های یکسان در خطوط متوالی هستند، زیر هم نشان داده شود. ضمن این که بیشتر بودن عرض کاراکتر فاصله باعث نمایش بهتر indenting (دندانه‌گذاری) نیز می‌شود (که در زبانی مثل پایتون، حیاتی است). اگر برنامه‌نویس باشید نیاز به توضیح بیشتر ندارید، و می‌دانید که هیچ برنامه‌نویسی از یک قلم عرض-متغیر (مثل تاهوما یا Arial) در ویرایشگر یا IDE محبوبش استفاده نمی‌کند. همین‌طور در خط فرمان (ترمینال / کنسول) - چه گرافیکی و چه غیر گرافیکی - به دلیل ذات و منطق ترمینال، و ابزارهای نمایش/ویرایش متن در آن، فقط می‌توان از قلم‌های عرض ثابت استفاده کرد. در واقع می‌توان گفت تنها استفاده عملی قلم‌های عرض ثابت (که با نام Mono یا Monospace هم شناخته می‌شوند) در خط فرمان و ویرایشگرهای برنامه‌نویسی است اما در سایر قسمت‌های محیط گرافیکی، مثل میز کار، مرورگر، واژه‌پرداز و... از قلم‌های عرض-متغیر استفاده می‌شود و ما هم از این پس در مورد این قلم‌ها صحبت خواهیم کرد

تا این‌جا، خصوصیات نسبتاً بدیهی قلم‌ها را توضیح دادیم، بگذارید به قلم‌های TrueType یونیکد برداری عرض متغیر، بگویم قلم‌های معمولی. با این تعریف، اغلب قلم‌هایی که امروزه در محیط گرافیکی بطور موثری استفاده می‌شوند قلم‌های معمولی هستند و این حداقل خواسته ما از یک قلم خوب فارسی است. حال به خصوصیات دیگر قلم‌ها که چندان بدیهی نیستند می‌پردازیم

۵- مجوز

قلم‌ها را با توجه به نوع مجوز آن‌ها می‌توان به سه دسته تقسیم کرد:

- ۱- غیر قابل توزیع
- ۲- قابل توزیع
- ۳- قابل توزیع، ویرایش و توزیع مجدد

عبارت متن‌باز بیشتر در مورد نرم‌افزار مصداق دارد. مثلاً در مورد قلم، اصلاً کد منبع (سورس کد) معنی ندارد. حتی عبارت منبع یا سورس هم به سختی مصداق دارد، چون قلم‌های ttf شامل اشکال برداری نمادها هستند و از لحاظ فنی قابل ویرایش هستند.

پس بحث آزاد بودن یا نبودن بیشتر مفهوم قانونی (یا اخلاقی) دارد تا فنی. و شاید هم (با توجه به ضعف قانون کپی‌رایت در کشور) به همین دلیل است که قلم‌های آزاد فارسی خیلی کم هستند. از چند ده سال پیش تا به حال متأسفانه افراد و گروه‌های زیادی، قلم‌های دیگران را (که براساس مجوز، قابل ویرایش هم نبودند) ویرایش می‌کردند و حتی دوباره به نام خود منتشر می‌کردند. که باعث ایجاد هرج و مرج در زمینه قلم‌های فارسی شده است. قلم‌های زیادی با نام یکسان یا مشابه وجود دارند که گاهی تفاوت زیادی از لحاظ فنی با هم دارند و هر یک به نام گروه یا شرکت خاصی کپی‌رایت شده‌اند و کسی نمی‌داند تاریخچه هر قلم چیست و چند بار دست به دست شده است.

قلم‌های آزاد فارسی:

در حال حاضر تنها قلمی که می‌توانم با اطمینان بگویم آزاد است قلم FreeFarsi است که از سازندگان آن واقعاً تشکر می‌کنم، گرچه به نظر بنده، از لحاظ خوانایی و حتی زیبایی توان رقابت با قلم‌های غیر آزادی مثل ترافیک و میترا را ندارد. برای دانلود قلم FreeFarsi از این آدرس استفاده کنید: <http://slmd.ir/2h>

البته شرکت فارسی‌وب هم ۷ قلم واقعاً خوب منتشر کرده (شامل الهام، هما، کودک، نازلی، رویا، ترافیک و تیترا) که به نظر می‌آید ۳ تا از آنها (هما، نازلی و تیترا) آزاد و تحت مجوز GPL است. باقی آن‌ها قرار است (یا بود) وقتی به ورژن 1.0 برسند، به صورت آزاد (احتمالاً GPL) منتشر شوند. آخرین نسخه بسته این قلم‌ها 0.5 هست که به سال‌ها پیش برمی‌گردد.

در مورد سه قلمی که تحت GPL هستند (با توجه به این که همانام با چند قلم دیگر است) هم تاریخچه مشخصی وجود ندارد، ولی متن کپی‌رایت که درون این قلم‌ها نوشته شده می‌گوید:

Copyright (c) 2003 by Sharif FarsiWeb, Inc. All Rights Reserved.

Copyright (c) 1999 by (URW)++ Design & Development

You may redistribute, change or rename this font according to GPL.

نگارنده هم سه قلم آزاد عربی به نام‌های KacstBook و KacstOne و KacstQurn که توسط موسسه تحقیقاتی KACST تحت مجوز GPL منتشر شده است را به کمک نرم‌افزار FontForge و Inkscape از عربی به فارسی تبدیل کرده و با نام‌های FaKacstBook و FaKacstOne و FaKacstQurn منتشر کرده است. تنها اشکال فنی این سه فونت، عربی بودن ارقام ۵ ۴ ۶ فارسی است (چون از ارقام عربی کپی شدند) من در هنگام نوشتن این مقاله، به این مسئله پی بردم و این مشکل را در قلم FaKacstBook (که خواناتر و زیباتر از ۲ تای دیگر است) برطرف کردم. اکنون می‌توان گفت FaKacstBook یکی از استاندارددترین قلم‌های فارسی است، که البته آزاد نیز هست. فایل‌هایی ttf این سه قلم و نیز راهنمای تبدیل قلم‌های عربی به فارسی را (که البته قدیمی است و در سال ۲۰۰۸ ساخته شده) می‌توانید از این صفحه دانلود کنید <https://github.com/ilius/kacst-fonts-fa>

۶- بر اساس شکل حروف

قلم‌های Serif به قلم‌هایی گفته می‌شود که در آن‌ها در انتهای حروف لاتین یک زائده وجود دارد که باعث زیبایی حرف می‌شود. گرچه این اصطلاح ریشه در لاتین دارد، ولی با کمی اغماض در تعریف، به قلم‌های فارسی هم قابل تعمیم است. وقتی که نوشته را با قلمی مانند قلم خوشنویسی (که سطح مقطعی باریک دارد) می‌نویسید، ضخامت نوشته در نقاط مختلف متفاوت است (گرچه زاویه قلم، به زبان و خط بستگی دارد) قلمی که شبیه نوشته با چنین قلمی باشد را می‌توان گفت Serif. قلم‌های Sans-Serif که به اختصار به آن‌ها Sans هم گفته می‌شود، عموماً ضخامت یکسانی در تمام نقاط حروف دارند. مثل نوشته‌ای که با خودکار (یا ماژیک سرگرد) نوشته شده باشد، بدون تغییر میزان فشار بر کاغذ. کلمه Sans در زبان فرانسوی به معنای «بدون» است. و کلمه Serif از کلمه Schreef در زبان هلندی آمده که به معنی خط یا تکان دادن قلم (در انتهای حروف) است. بطور خلاصه Sans-Serif یعنی «بدون دندانه انتهای حروف» تفاوت‌های قلم‌های Serif و Sans-Serif را می‌توان در چنین جدولی خلاصه کرد:

مشخصه	Serif	Sans-Serif
وجود دندانه (در لاتین)	دندانه‌دار	بدون دندانه
ضخامت خط	ضخامت خط متغیر	ضخامت خط ثابت (Monoline)
شکل هندسی	شکل هندسی پیچیده و هنری	شکل هندسی ساده و منظم
قابل نوشتن توسط...	قلم خوشنویسی	قلم سرگرد و عمودی با فشار ثابت
خوانایی / زیبایی	خوانایی کمتر، زیبایی بیشتر	خوانایی بیشتر، زیبایی کمتر

می‌توان هر یک از این چند فاکتور را بعنوان تعریف Serif و Sans-Serif در نظر گرفت. ولی من ترجیح می‌دهم ترکیب این فاکتورها را برای تعریف و دسته‌بندی فونت‌ها استفاده کنم. با چنین تعریفی، این دو اصطلاح را (که از لاتین نشئت گرفته‌اند) می‌توان به فارسی هم تعمیم داد. قلم‌های Sans-Serif کمتر جنبه هنری کمتری داشته و به خصوص در صفحات نمایش کامپیوتر/تبلت/موبایل که بخاطر ذات پیکسلی بودنشان جزئیات ظریف و گوشه‌های تیز را به خوبی نشان نمی‌دهند، قلم‌های Sans خوانایی بیشتری دارند و به همین خاطر بیشتر استفاده می‌شوند. گرچه قلم‌های Serif هم بخاطر زیبایی‌شان طرفداران خود را دارند، بخصوص اگر قصد پرینت داشته و یا اهل استفاده از کتابخوان‌های الکترونیک باشید. تعدادی از قلم‌ها هم هستند که کاملاً جنبه فانتزی یا گرافیکی دارند ((بسیار هنری‌تر از قلم‌های Serif)). این قلم‌ها بیشتر در کارهای گرافیکی استفاده شده و تقریباً کسی برای خواندن متون از آن‌ها استفاده نمی‌کند.

۷- خوانایی و زیبایی

شاید نتوانیم قلم‌ها را بر اساس شکل حروف بطور دقیقی دسته‌بندی کنیم، ولی می‌توانیم طبق سلیقه خودمان، به خوانایی و زیبایی قلم، دو نمره متفاوت بدهیم. یک قلم ممکن است خوانایی و زیبایی را همزمان در حد قابل قبولی داشته باشد، ولی عملاً نمی‌تواند هر دوی آن‌ها را در حد کمال داشته باشد. پس باید مصالحه کرد و دید خوانایی و زیبایی هر کدام چقدر برای ما اهمیت دارد. این در نهایت به سلیقه

و نیاز شما برمی گردد. ولی با توجه به نیاز امروز کاربران کامپیوتر و زیاد بودن قلم‌های هنری و کم بودن یک قلم خوانای خوب و کم‌اشکال، ما فعلا به خوانایی اهمیت بیشتری می‌دهیم.

۸- چندزبانه‌بودن

علائم نوشتاری را بر اساس زبان می‌توان به چند دسته تقسیم کرد

- ۱- کاراکترهای اسکی (انگلیسی)
- ۲- کاراکترهای لاتین غیر اسکی
- ۳- حروف عربی
- ۴- حروف فارسی، اردو و کردی که در عربی نیستند
- ۵- حروف زبان‌های خانواده سیریلیک (مثل روسی)
- ۶- حروف زبان‌های آسیای شرقی (چینی، ژاپنی، کره‌ای)
- ۷- سایر زبان‌ها

تقریبا هیچ قلمی نیست که از تمام زبان‌ها پشتیبانی کند. اگر هم قلمی باشد که از بیشتر زبان‌ها (از جمله لاتین و فارسی) پشتیبانی کند، احتمالا اشکال فارسی‌اش مطابق با سلیقه ما نیست. پس ما همیشه به چندین قلم نیاز داریم. در سیستم‌عامل گنو/لینوکس، در بیشتر چهارچوب‌های گرافیکی / میزکارها و بیشتر برنامه‌ها این قابلیت وجود دارد، که اگر تعدادی کاراکتر در قلم دلخواه شما وجود نداشت، برای نمایش آن کاراکتر از یک قلم عمومی و پیش‌فرض (که از اغلب زبان‌ها پشتیبانی می‌کند) استفاده کند.

حتی اگر یک قلم فارسی برای برنامه تعیین کنید که حروف انگلیسی را نداشته باشد، حروف انگلیسی توسط همان قلم / قلم‌ها پیش‌فرض نمایش داده می‌شوند

گرچه به ندرت، برنامه‌ای مثل مرورگر فایرفاکس به شما اجازه می‌دهد، برای هر زبان یک قلم جداگانه تنظیم کنید. گرچه ما خیلی علاقه‌مندیم یک قلم خوب و آزاد فارسی داشته باشیم که از حروف انگلیسی (و حتی سایر زبان‌های لاتین) هم پشتیبانی کند تا هنگام خواندن متن‌های دوزبانه (شامل تمام متن‌های فنی) حروف فارسی و انگلیسی از لحاظ اندازه و شکل با هم هماهنگ باشند. چون اندازه قلم، مثل سایز پیراهن است و یک استاندارد کلی و جامع وجود ندارد. واضح‌تر بگوییم، حروف سایز ۱۴ قلم A ممکن است بزرگ‌تر از حروف سایز ۱۴ قلم B باشند. ضمن این که شکل یا style حروف هم بهتر است بین فارسی و انگلیسی هماهنگ و یکسان باشد.

پس دوست داریم است یک قلم داشته باشیم که تمام یا بیشتر کاراکترهای موردنیاز ما (فارسی، عربی، انگلیسی و حتی سایر زبان‌های اروپایی مثل آلمانی و فرانسوی) را پشتیبانی کند.

۹- حجم فایل

طبیعی است که مانند هر نوع فایل دیگری، حجم فایل قلم هم برای ما اهمیت دارد. گرچه فشرده‌سازی نقش زیادی در قلم ندارد و حجم بیشتر معمولا به معنای تعداد شکل بیشتر است (با این فرض که قلم برداری ttf است). گرچه دقت اشکال هم بی‌تأثیر نیست. تعداد شکل بیشتر هم معمولا به معنای تعداد کاراکتر بیشتر است (تفاوت این دو را خواهیم گفت) و تعداد کاراکتر بیشتر هم معمولا به معنای تعداد زبان‌های بیشتری است که پشتیبانی می‌کند. اما این جا لازم می‌دانیم معنای اصطلاح Ligature یا لیگاتور را بگوییم.

در فارسی، شکل یک کلمه عموما از به هم چسباندن شکل حروف ایجاد می‌شود که شکل حروف هم بستگی به چسبان و غیرچسبان بودن خود حرف و حرف قبل و بعدش، یکی از چهار حالت ممکن است که تمام ۴ حالت حرف (جدا، ابتدایی، میانی، انتهایی) در قلم تعریف شده است (گرچه ممکنه است ۲ یا ۳ تا از این ۴ حالت یکسان باشند).

اما این قاعده همیشه هم صادق نیست. مثلا تقریبا هیچ کس کلمه «لازم» را بصورت «**لازم**» نمی‌نویسد. چون چسباندن حرف لام به الف (با قاعده ذکر شده) ترکیب زشتی را می‌دهد. به همین دلیل در اغلب قلم‌های فارسی ترکیب این دو حرف بصورت مجزایی با شکل «لا» رسم شده است. (جالب است بدانید یونیکد یک کاراکتر مجزا با کد U+FEFB برای «لا» دارد) به این نوع شکل که مربوط به ترکیبی از حروف متوالی است، لیگاتور می‌گویند. خط نستعلیق را می‌توان پیچیده‌ترین و البته زیباترین خط فارسی دانست. شکل هر حرف در این خط، حالت‌های بسیار متعددی دارد و به عوامل مختلف بستگی دارد.

اگر قلم IranNastaliq را دیده باشید، حجم آن بیشتر از یک مگابایت است که برای یک قلم خیلی زیاد است. علت آن هم تعداد بسیار زیاد اشکال است که بیشتر آن‌ها هم مربوط به فارسی است. نه تنها تعداد لیگاتورها خیلی زیاد است (حتی کلمه بلندی مثل «شکستگی» هم بصورت یک شکل مجزا ترسیم شده!!) بلکه حالت‌های مختلفی از تک‌حرف‌ها هم هستند که لیگاتور نیستند و من واقعا نمی‌دانم با چه ابزاری و طبق چه قاعده‌ای تعریف شده‌اند.

گرچه این قلم آزاد نیست و به صورت قانونی قابل ویرایش نیست، ولی اگر بخواهید هم احتمالا نمی‌توانید آن را ویرایش کنید!!

در ضمن، با وجود این همه پیچیدگی، این قلم یک پیاده‌سازی کامل از خط نستعلیق نیست. خط نستعلیق (در حالت کامل) با یک قلم قابل پیاده‌سازی نیست! نیاز به نرم‌افزارهای تخصصی مثل «نامه‌نگار» دارد که فقط هم در ویندوز قابل اجرا هستند.

جدای از این که استفاده از خط نستعلیق در صفحات وب و مستندات باعث سخت شدن خواندن متن می‌شود، ضمناً این قلم هم یک سری اشکالات فنی دارد (که خب با توجه به پیچیدگی خط نستعلیق، طبیعی است و حتی باید سازنده آن را تحسین کرد) ولی این قلم برای کارهای گرافیکی، در نرم‌افزارهایی مثل GIMP و Inkscape (یا معادل‌های غیر آزاد آن‌ها) به شدت مفید خواهد بود.

۱۰- استاندارد بودن نمادهای فارسی و عربی

این استاندارد را می‌توان شامل مواردی دانست که در ادامه عنوان شده است:
۱- ارقام فارسی و عربی (جدای از ارقام لاتین)

English	Persian	Arabic
0	۰	۰
1	۱	۱
2	۲	۲
3	۳	۳
4	۴	۴
5	۵	۵
6	۶	۶
7	۷	۷
8	۸	۸
9	۹	۹

دقت کنید که ارقام ۴ و ۵ و ۶ بین فارسی و عربی متفاوت است همین طور بهتر است رقم صفر فارسی، توخالی باشد
تمام ارقام فارسی و عربی، حتی در صورت شبیه بودن، در استاندارد یونیکد، با هم متفاوت هستن (مثلاً ۲ فارسی معادل U+06F2 است و ۲ عربی معادل U+0662 است)

۲- حروف مشابه ولی متفاوت بین فارسی و عربی (مثل ک و ی)

۳- فاصله مجازی

۴- علائم نگارشی فارسی متفاوت با انگلیسی (مثل ویرگول، درصد و گیومه)

۱۱- هماهنگی بین شکل و اندازه حروف زبان‌های مختلف

مقایسه قلم‌های فارسی آزاد و قلم‌های معروف

نام قلم	مجوز	سبک (Style)	خوانایی (۰-۹)	زیبایی (۰-۹)	زبان	استاندارد بودن فارسی	حجم (KB)
FreeFarsi	GPL	Serif	۸	۶	فارسی + انگلیسی	۱۰۰	۱۱۳
B Mitra	نامشخص	Serif	۹	۸	فارسی	۵۵	۵۸
Roya	قابل کپی	Sans	۹	۸	فارسی	۹۵	۴۸
Terafik	قابل کپی	Sans	۹	۷	فارسی + انگلیسی	۹۵	۸۵
Tahoma	قابل کپی	Sans	۹	۵	کامل	۹۵	۶۹۸
Nazli	GPL	Serif	۷	۸	فارسی + انگلیسی	۹۵	۵۵
FaKacstBook	GPL	Serif	۷	۸	فارسی + انگلیسی	۱۰۰	۱۱۸
Koodak	قابل کپی	Sans	۹	۷	فارسی	۹۵	۸۲
Titr	GPL	Serif	۸	۷	فارسی	۹۵	۸۶
FaKacstQurn	GPL	Serif	۷	۷	فارسی + انگلیسی	۹۰	۱۲۹
Homa	GPL	Sans	۶	۷	فارسی	۹۵	۴۸
FaKacstOne	GPL	Serif	۸	۶	فارسی + انگلیسی	۹۰	۱۲۴
DejaVuSans	آزاد	Sans	۸	۵	کامل	۱۰۰	۷۴۱
Arial	نامشخص	Sans	۷	۶	کامل	۹۵	۲۷۵
Elham	قابل کپی	Sans	۵	۷	فارسی	۹۵	۴۲
DroidNaskh	Apache	Other	۹	۷	فارسی	۹۸	۸۹
FreeSans	GPL	Sans	۹	۷	انگلیسی و لاتین	۰	۱۵۶۰
FreeSerif	GPL	Serif	۸	۸	انگلیسی و لاتین	۰	۳۳۰۰
DroidSans	Apache	Sans	۹	۸	انگلیسی و لاتین	۰	۱۹۰
DejaVuSerif	آزاد	Serif	۸	۷	انگلیسی و لاتین	۰	۳۶۸

ستون دوم از چپ، درصد استاندارد بودن پشتیبانی از فارسی است که به این صورت محاسبه شده:

۳۰	حروف گ چ پ ژ
----	--------------

۲۰	پشتیبانی از ارقام یونیکد فارسی
۱۰	کاراکترهای نگارشی یونیکد فارسی
۱۵	انگلیسی بودن ارقام و کاراکترهای نگارشی اسکی
۱۰	متفاوت بودن ارقام ۴۵۶ فارسی از عربی
۱۰	داشتن حروف « ی ک » فارسی و متفاوت بودن آن‌ها از حروف عربی (ي ك)
۵	شکل صحیح فاصله مجازی (تهی) (نداشتن آن بهتر از یک خط عمودی است)

Free Farsi	نمونه‌ای از متن فارسی با قلم‌های مختلف
B Mitra	نمونه‌ای از متن فارسی با قلم‌های مختلف
Roya	نمونه‌ای از متن فارسی با قلم‌های مختلف
Terafik	نمونه‌ای از متن فارسی با قلم‌های مختلف
Tahoma	نمونه‌ای از متن فارسی با قلم‌های مختلف
Nazli	نمونه‌ای از متن فارسی با قلم‌های مختلف
FaKacstBook	نمونه‌ای از متن فارسی با قلم‌های مختلف
Koodak	نمونه‌ای از متن فارسی با قلم‌های مختلف
Titir	نمونه‌ای از متن فارسی با قلم‌های مختلف
FaKacstQum	نمونه‌ای از متن فارسی با قلم‌های مختلف
Homa	نمونه‌ای از متن فارسی با قلم‌های مختلف
FaKacstOne	نمونه‌ای از متن فارسی با قلم‌های مختلف
DejaVuSans	نمونه‌ای از متن فارسی با قلم‌های مختلف
Arial	نمونه‌ای از متن فارسی با قلم‌های مختلف
Elham	نمونه‌ای از متن فارسی با قلم‌های مختلف
DroidNaskh	نمونه‌ای از متن فارسی با قلم‌های مختلف

گرچه فرصت نشد در مورد Text Rendering و کارهای آقای «بهداد اسفهد» بنویسم، ولی همین قدر بدانید که اگر زحمات ایشان نبود، هر چقدر هم قلم خوبی داشتیم، باز هم هرگز نمی‌توانستیم متون فارسی را در گنو/لینوکس، به این خوبی مشاهده کنیم و بخوانیم.