نگارش فارسی با کامپیوتر: مروری بر دشواری‌ها و راهکارها

بهروز پرهامی[1]
دانشکدهٔ مهندسی برق و کامپیوتر، دانشگاه کالیفرنیا در سنتا باربارا

این مقاله در پی تدوین تاریخچه‌ای کامل و مستند از فعالیت‌های گسترده و ابداعاتی است که نگارنده، همکاران دانشگاهی او، و متخصصان بخش‌های دولتی و خصوصی

[1] صورت دیگری از این مقاله پیش از این در دو سخنرانی با عنوان "کامپیوتر و مشکلات نگارش فارسی" به زبان‌های فارسی در ۱۹ نوامبر و انگلیسی در ۲۰ نوامبر ۲۰۱۷ در سلسله سخنرانی‌های دوزبانهٔ برنامهٔ ایران‌شناسی دانشگاه یوسی‌ال‌ای عرضه شده است. از گردانندهٔ سخنرانی‌ها، خانم دکتر نیره توحیدی، و سایر دست‌اندرکاران و حامیان آن برنامه صمیمانه سپاس‌گزارم. نقش نگارنده در پژوهش‌ها و راهکارهایی که در این مقاله به آنها اشاره می‌رود عمدتاً طبقه‌بندی مسائل، جمع‌بندی، و گزارش بوده است. مشکلات مورد بحث بزرگ‌تر از آن‌اند که یک نفر به تنهایی قادر به حلشان باشد. همکارانم در دانشگاه صنعتی شریف، آقایان دکتر فرهاد مودّت و مهندس آرمن نهاپتیان، در ابتدای این راه نقش مهمی ایفا کردند. همکاران دیگر دانشگاه‌های ایران و پژوهشگران و مهندسان سازمان‌های دولتی و بخش خصوصی و دانشجویان دورهٔ کارشناسی ارشد دانشگاه صنعتی نیز در حل مسائل دست داشته‌اند. لازم به توضیح است که اشاره به مشکلات خط فارسی در این مقاله صورت‌های دیگر زبان فارسی، از جمله دَری، اُردو، و پَشتو را هم در بر می‌گیرد و مشکلات خط عربی هم عمدتاً شبیه فارسی است.

Behrooz Parhami, "Computers and Challenges of Writing in Persian," *Iran Namag*, Volume 4, Number 2 (Summer 2019),3-32.

بهروز پرهامی <parhami@ece.ucsb.edu> Behrooz Parhami (دانش‌آموختهٔ دکتری علوم کامپیوتر، دانشگاه یوسی‌ال‌ای، ۱۹۷۳) استاد دانشگاه کالیفرنیا در سنتا باربارا و معاون پیشین کالج مهندسی آن است که در حال حاضر به تدریس و پژوهش در رشتهٔ معماری و سخت افزار کامپیوتر اشتغال دارد. او عضو برجستهٔ انجمن بین‌المللی مهندسان برق و الکترونیک، انستیتوی مهندسی و تکنولوژی، و انجمن کامپیوتر بریتانیاست و جوایز متعددی دریافت کرده است. از ایشان ۶ کتاب و بیش از ۳۰۰ مقاله منتشر شده است. از ۱۹۷۴ تا ۱۹۸۶ در دانشگاه صنعتی شریف به مطالعهٔ زبان و خط فارسی برای استفاده در کامپیوتر و مسائل انتقال تکنولوژی و استانداردسازی مشغول بوده و در تأسیس انجمن انفورماتیک ایران و نشریهٔ آن انجمن با نام *گزارش کامپیوتر* مشارکت داشته است.

ایران برای تولید خط خوانا و زیبای فارسی صورت داده‌اند. در کنار ثبت تاریخچهٔ مشکلات و ابداعات اولیه، تصویری نیز از وضع کنونی نمایش و چاپ خط فارسی در کامپیوترهای امروزی و فهرستی از مسائل و مشکلات باقی‌مانده پس از تلاش‌های چشمگیر ۵۰ ساله ارائه خواهد شد.

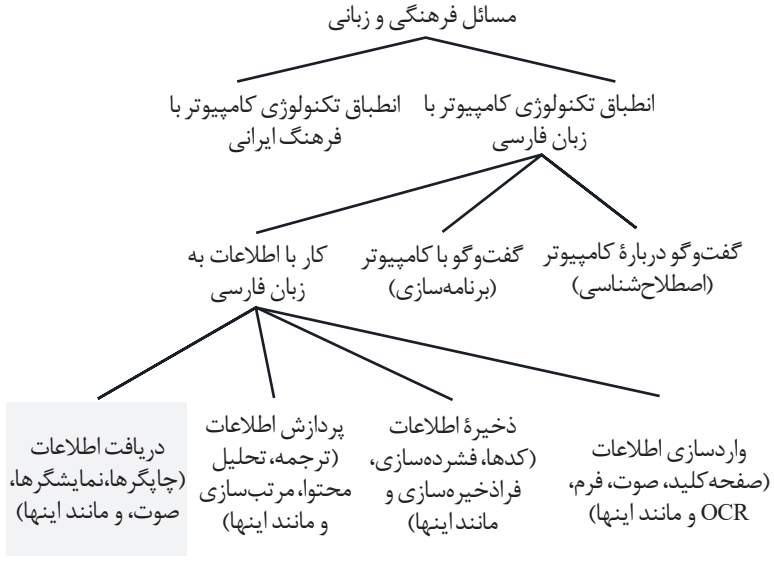
در واقع، عنوان پیشنهادی اولیه برای سخنرانی‌های دوزبانهٔ دانشگاه یوسی‌ال‌اِی، که مبنای این مقاله را تشکیل می‌دهند، "پنجاه سال بدخطی: چگونه کامپیوترها به زحمت نگارش فارسی را یاد گرفتند" بود و پنجاه سال از اواخر دههٔ ۱۹۶۰ آغاز می‌شود. در آن دوران، دانشگاه تهران به تازگی یک دستگاه کامپیوتر آی‌بی‌اِم ۱۶۲۰ را در مرکز محاسبات واقع در جنب سالن ورزش و خارج از محوطهٔ اصلی دانشگاه نصب کرده بود تا از آن برای داده‌پردازی و سیستم‌های مدیریت دانشگاه استفاده کند و برخی امکانات محاسباتی را نیز در اختیار دانشجویان قرار دهد. همزمان با این اقدامات دانشگاه، بانک‌های مهم و سازمان‌های بزرگ دولتی نیز به کامپیوتری کردن سیستم‌های اطلاعاتی و پرداخت حقوق مشغول بودند و در این راه به پردازش اطلاعات به زبان فارسی برای نام و نشانی کارمندان یا مشتریان خود نیاز داشتند.

در طول این پنج دهه رویدادهایی مهم و پیشرفت‌هایی چشمگیر را شاهد بوده‌ایم. آشنایی نگارنده با کامپیوتر در سال سوم دانشگاه از راه درس برنامه‌سازی فُرترن صورت گرفت که با معیارهای امروزی پیش‌پاافتاده و در عین حال کُند و دست‌وپاگیر بود: برنامه روی کارت‌های ۸۰ ستونی منگنه می‌شد و نتایج اجرای آن روی کاغذهای طومارگونه به کاربران برمی‌گشت و این مراحل تا هنگام رفع همهٔ اشکالات برنامه و رسیدن به جواب درست تکرار می‌شد. دانشجویان به دستگاه‌های نمایشگر کامپیوتری دسترسی نداشتند و چاپ اطلاعات فارسی هم در کار نبود. بانک‌ها و سازمان‌های دولتی هم که با اطلاعات فارسی سروکار داشتند ناچار به تحمل خطی ناخوانا و نازیبا بودند. اما امروز به یُمن سال‌ها تلاش متخصصان در ایران و برخی از کشورهای عرب‌زبان و جلب سرمایه و تخصص به سبب گسترش چشمگیر بازار کامپیوتر در منطقه، کیفیت خط حاصل از کامپیوترها بسیار بهتر شده است.

امید که این مقاله در نشان دادن دامنهٔ مشکلات اولیه و نیز سخت‌کوشی و ابتکار متخصصان بومی که با این مشکلات در مراحل متفاوت دست‌وپنجه نرم کردند و خود را با تغییرات و پیشرفت‌های بی‌امان تکنولوژی کامپیوتر تطبیق دادند، راهگشا و آموزنده باشد.

کامپیوتر و زبان فارسی

یکی از زمینه‌های پژوهشی نگارنده انتقال تکنولوژی بوده است. به عبارت دقیق‌تر، به اتفاق همکارانم علاقه‌مند به تطبیق تکنولوژی نسبتاً جدید کامپیوتر با فرهنگ، اقتصاد، و زبان کشورمان بوده‌ام.[۲] در واقع، عبارت "پیوند تکنولوژی" برای این فعالیت‌ها مناسب‌تر است، چرا که همانند پیوند اعضای بدن، احتمال مقاومت و دفع تکنولوژی جدید را هم در بحث منظور می‌کند. این مقوله با مبحث تکنولوژی متناسب،[۳] که شرایط اجتماعی و اقتصادی و محیط زیستی را در انتخاب و به‌کارگیری تکنولوژی مد نظر قرار می‌دهد، رابطه‌ای تنگاتنگ دارد.



شکل ۱. پیوند تکنولوژی: جنبه‌های فرهنگی و زبانی در تطبیق تکنولوژی کامپیوتر.

چنان‌که در شکل ۱ دیده می‌شود، منظور کردن عوامل فرهنگی و زبانی در حوزهٔ کامپیوتر زمینه‌ها و رشته‌های گوناگونی را در بر می‌گیرد. در خصوص عوامل فرهنگی فقط به ذکر این نکته اکتفا می‌کنیم که طراحی برنامه‌های آموزشی بومی نقش

[2]Farhad Mavaddat and Behrooz Parhami, "Informatics in Iran: Problems and Prospects," *Proc. Int'l Conf. Computer Applications in Developing Countries* (Thailand: Bangkok, August 1977), 121-133.

[3]Peter D. Dunn, *Appropriate Technology: Technology with a Human Face* (London: The Mac-Millan Press, 1979).

مهمی را در این زمینه ایفا می‌کند.[۴] در بدو امر، برای تطبیق کامپیوتر با زبان و خط فارسی، یعنی آنچه در سمت راست شکل ۱ دیده می‌شود، به واژگان مناسبی برای توصیف و تبادل نظر دربارهٔ تکنولوژی و کاربردهای کامپیوتر در قالب زبان فارسی نیازمندیم.[۵] ابداع واژگان جدید که هم از نظر علمی و هم از جنبهٔ زبان‌شناسی صحیح باشند و مقبول جامعه نیز قرار گیرند کار آسانی نیست. چه بسیار واژه‌های دقیق و صحیح که مقبول نیفتادند و در نتیجه، مورد استفاده قرار نگرفتند که از آن جمله "رایانه" است. حال آنکه واژه‌های دیگری مانند "ریزپردازشگر" خوب جا افتادند. از دیدگاه عملی، در برخورد با برخی واژه‌های جاافتادهٔ خارجی، مثل "الکترونیک،" نباید سختگیری به خرج داد.

در مرحلهٔ بعد به برقراری ارتباط با کامپیوتر نیاز داریم تا رهنمودها و دستورات لازم را برای اجرای کارهای مورد نظرمان به آن بدهیم. در این حوزه زبان نقشی جانبی دارد. اگر بخواهیم که کامپیوتر سلسله محاسباتی را تکرار کند، مادام که مقدار متغیر x مثبت است، مشکل ما درک عبارت انگلیسی "while x > 0 do" نیست، چون یاد گرفتن چند واژهٔ به‌کاررفته در یک زبان برنامه‌سازی در مقایسه با مفاهیم و ساختارهای برنامه‌سازی کاری آسان است. با این حال، چون در دههٔ ۱۹۷۰ زبانی با نام کوبال برای برنامه‌سازی سیستم‌های تجاری به کار می‌رفت و از عبارات و جملات انگلیسی زیاد استفاده می‌کرد، گروهی به فکر طراحی یک زبان برنامه‌سازی فارسی افتادند. این روش برقراری ارتباط با کامپیوتر به زبان فارسی، که پس از چندی کنار گذاشته شد، امروز به علت رواج یافتن ارتباط صوتی و گویشی دوباره مطرح است.

برای پردازش متن فارسی باید اول آن را از راه صفحه‌کلید،[۶] گویش، فرم‌های پُرشده یا مدارک واردشده به دستگاه حروف‌خوان نوری به کامپیوتر داد.[۷] سپس به کُدهایی

---

[4]Behrooz Parhami, "University Education in Computer Science and Technology: The New Iranian Plan," *Proc. IFIP 4th World Conf. on Computers in Education* (Norfolk, August 1985), 923-930; Behrooz Parhami, "Computer Science and Engineering Education in a Developing Country: The Case of Iran," *Education and Computing*, 2:4 (1986), 231-242.

[5]See Behrooz Parhami and Vida Daie, "Glossary of Computers and Informatics: English/Persian," *Informatics Society of Iran* (May 1981).

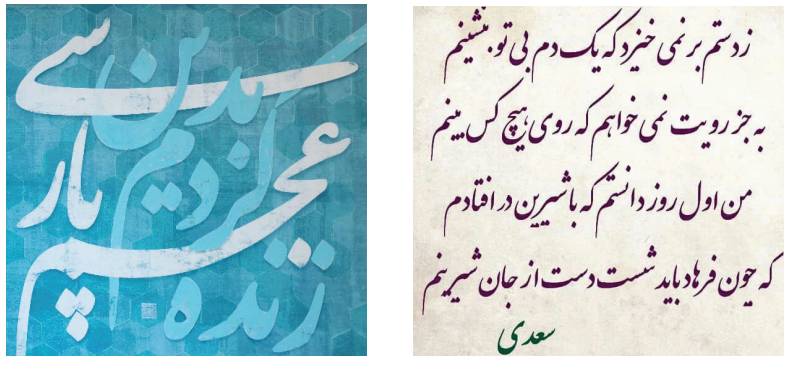
[6]Behrooz Parhami, "Standard Farsi Information Interchange Code and Keyboard Layout: A Unified Proposal," *Journal of Institution of Electrical and Telecommunications Engineers*, 30:6 (1984), 179-183.

[7]Behrooz Parhami and Mahmood Taraghi, "Automatic Recognition of Printed Farsi Texts," *Pattern Recognition*, 14:1-6 (1981), 395-403.

برای ذخیرهٔ اطلاعات و احتمالاً ابزار فشرده‌سازی داده‌ها نیاز داریم تا در کاربرد فضای حافظه و امکانات مجراهای تبادل داده‌ها صرفه‌جویی شود.[۸]

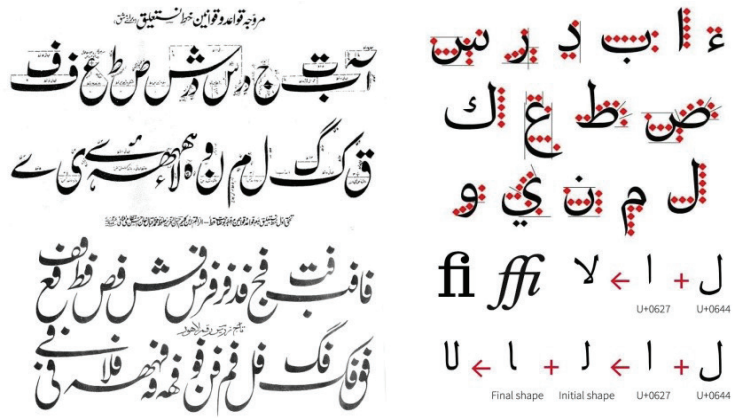
عمل پردازش اطلاعات در کامپیوتر نیازمند نرم‌افزارها و الگوریتم‌هایی است که برای هر زبان متفاوت‌اند. مثال خوبی در این زمینه مرتب کردن نام‌های فارسی به ترتیب الفبایی است که به سبب وجود صورت‌های مختلف حروف و تفاوت‌های املایی، به‌ویژه در نام‌های مرکب، مشکل‌زاست. مطالعاتی از نوع تجزیه و تحلیل محتوا و سبک متون نیازمند همکاری متخصصان رشته‌های دیگر مانند انسان‌شناسی، حقوق، ادبیات، و جامعه‌شناسی است. مسائل ترجمهٔ متون زبان‌های دیگر هم از همین دست است.

آخرین مرحله در این روند تولید و ارایهٔ اطلاعات خروجی با کامپیوتر است. زبان در جایی که کامپیوتر از راه کنترل مستقیم دستگاه‌ها و فرایندها با دنیای خارج ارتباط برقرار می‌کند نقشی ندارد، اما هنگامی که خروجی به صورت چاپی، نمایشی یا گویشی ارائه می‌شود، نقش زبان مهم است. در حال حاضر، کاربرد خروجی گویشی به فارسی بسیار محدود است، ولی طی چند سال آینده اهمیت این نوع خروجی افزایش خواهد یافت.



شکل ۲. راست: دست‌خط نستعلیق از خطاطی ناشناس، برگرفته از
https://i.pinimg.com/originals/cc/0c/f6/cc0cf6edb8bb6851d91c8ce2edbc2c76.jpg
چپ: نمونهٔ خط تزیینی از فرخ محجوبی، برگرفته از
https://festiveart.com/persian-calligraphy-farrokh-mahjoubi.html#.XS1ut-hKiUk/.

[8] Iranian Plan and Budget Organization, *Final Proposal for the Iranian National Standard Information Code (INSIC), Persian and English versions*, 1980.



شکل ۳. قوانین نوشتن حروف و ترکیبات برای تولید خط زیبای فارسی. راست بالا: فونت خط نسخ، برگرفته از
https://www.rock-cafe.info/posts/arabic-calligraphy-fonts-naskh-617261626963.html
چپ بالا و پایین: برگرفته از
http://www.handwriting.pk/calligraphy.html/.

تاریخچهٔ مختصر خط فارسی

پیشینهٔ خط نوین فارسی، که بر اساس خط عربی بنا شده است، به ۱۲۰۰ سال پیش بازمی‌گردد.[9] زبان فارسی اما قدیمی‌تر و پیش از آن به صورت فارسی کهن و فارسی میانه در ایران رایج بوده است. در بخش عمده‌ای از این ۱۲۰۰ سال، خط فارسی با دست نوشته می‌شد و کسانی که می‌خواستند آثارشان ثبت و پخش شود خطاطانی را استخدام می‌کردند که متونشان را به‌گونه‌ای خوانا و زیبا با یکی از خطوط رایج در آن زمان می‌نوشتند (شکل ۲، راست). کتابی که به این ترتیب تولید می‌شد منحصربه‌فرد بود و هر نسخه‌اش می‌بایست جداگانه نوشته شود. نوعی روش چاپ بسیار ابتدایی هم شامل حکاکی متن روی سنگ یا چوب و آغشته کردنش به مرکب و قرار دادنش روی کاغذ یا پوست وجود داشت که استفاده از آن مشکل و وقت‌گیر بود.

اهمیتی که ایرانیان برای زیبایی خط قائل بودند به پیدایش و گسترش سریع خطاطی هنری، که در آن زیبایی ناشی از تناسب و تقارن و رنگ‌آمیزی متون بیشتر از خوانایی خط مدّ نظر بود (شکل ۲، چپ)، انجامید. پیشرفت این رشتهٔ هنری نوعی خط‌های

[9] Gilbert Lazard, "The Rise of the New Persian Language," *The Cambridge History of Iran: Period from the Arab Invasion to the Saljuqs* (Cambridge: Cambridge University Press, 2008), vol. 4, 566-594.

ویژه مانند کوفی مربع به وجود آورد که پادشاهان و معماران از آنها برای تزیین مساجد، کاخ‌ها و دیگر بناهای مهم استفاده می‌کردند. در این خطوط، متن مورد نظر با کنار هم گذاشتن کاشی‌های سیاه و سفید یا رنگی ایجاد می‌شد.

از نظر تطبیق با تکنولوژی، خط فارسی در سه مرحلهٔ مهم متحول شده است که متناظر با ورود صنعت چاپ، استفاده از ماشین تحریر، و رواج کاربرد کامپیوتر است. هر یک از این مراحل ویژگی‌هایی را از مرحلهٔ قبلی به ارث برد، به طوری که خط ماشین تحریر گونه‌ای تغییریافته از خط چاپی بود و خط کامپیوتری هم از خط ماشین تحریر مشتق شد. باید به خاطر داشت که مشکلات تولید خط عربی عمدتاً با فارسی مشترک است.

پیش از دوران نیاز به تطبیق با تکنولوژی، چند نوع خط و از جمله خطوط پُرطرفدار نستعلیق و نسخ در نوشته‌ها به کار می‌رفت. قوانین خط‌نویسی، مانند اندازهٔ حروف بر حسب تعداد نقطه‌های معادل در ابعاد افقی و عمودی و شکل‌های متفاوت هر حرف، از استاد به شاگرد منتقل می‌شد (شکل ۳، راست). هر استاد خط‌نویس سبک ویژهٔ خود را داشت که بنا بر شهرت و اعتبارش طرفدارانی داشت. برای هرچه زیباتر شدن خط، برخی ترکیبات از حروف به صورت یک‌جا خطاطی می‌شدند، چنان‌که در خط لاتین هم ترکیب‌هایی مانند fi و ffi علامت‌های ویژه‌ای دارند که با کنار هم چیدن حروف تشکیل‌دهنده‌شان فرق دارند. مثال بارز این امر در خط فارسی ترکیب "لام-الف" یا "لا" است. چنین ترکیب‌هایی در خط نستعلیق بیشتر به چشم می‌خورند (شکل ۳، چپ).

به علت تطبیق بهتر با تکنولوژی‌های نمایش و چاپ کامپیوتری، امروز خط نسخ از نستعلیق بیشتر رایج است و به گونه‌های متفاوت عرضه می‌شود. اما خط نستعلیق هنوز هم جای ویژه‌ای در دل فارسی‌زبانان سراسر دنیا دارد. مثلاً نفیس‌ترین دیوان‌های شعر با خط دست‌نویس نستعلیق چاپ می‌شوند. در برگه‌های تبلیغاتی، پوسترها و بروشورهایی که عمدتاً با خط نسخ چاپی تهیه می‌شوند، گاه برای عناوین اصلی یا جلب توجه بیشتر خط نستعلیق را به کار می‌گیرند. تولید خط نستعلیق با کامپیوتر ممکن است و پیشرفت‌هایی هم در این زمینه صورت گرفته است، ولی کیفیت خط حاصل هنوز به حد دست‌خط استادان خط‌نویس نرسیده است.

ورود صنعت چاپ به ایران

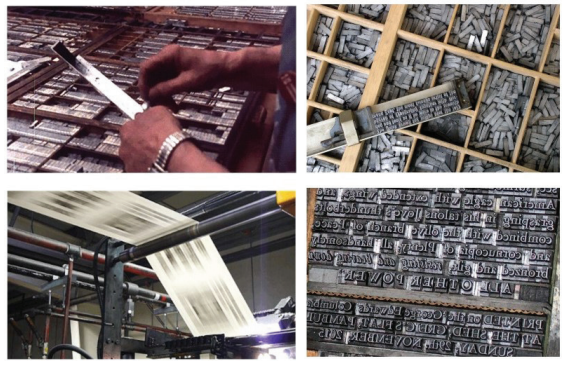
ماشین چاپ چهار قرن پیش به ایران وارد شد. در سال ۱۶۱۸، شاه‌عباس اول از وجود حروف چاپی عربی و فارسی آگاه شد و و تصمیم گرفت آنها را به کشور وارد کند.[۱۰] در سال ۱۶۲۹، دستگاه چاپ و حروف لازم به اصفهان ارسال شدند، اما هیچ مدرکی دال بر استفاده از این تجهیزات بر جا نمانده است. کاربرد عملی صنعت چاپ به صورت گسترده حدود سه قرن در ایران سابقه دارد. اختراع دستگاه چاپ دستی استانهوپ (Stanhope) در سال ۱۸۰۰ صنعت چاپ را در جهان دگرگون کرد، چون هم اندازه‌اش نسبتاً کوچک بود و هم عملکرد ساده‌ای داشت. در سال ۱۸۱۶، ایرانیانی که به اروپا و روسیه سفر می‌کردند این دستگاه را به تبریز و چند سال پس از آن به اصفهان و تهران آوردند و برای نشر کتاب‌های گوناگون، و نه فقط معدودی از کتب مذهبی، به کار گرفتند.[۱۱] یکی از دلایل این تأخیر مشکلات اولیهٔ تطبیق صنعت چاپ با خط فارسی بود. دستگاه‌های چاپ وارداتی بر اساس روش حروف انتقال‌پذیر بنا شده بودند، به این مفهوم که بلوک‌های فلزی حاوی حروف الفبا و علامت‌های دیگر کنار هم چیده می‌شدند و حروف تنها یا دنباله‌ای از حروف از یک نقطه در متن به نقاط دیگر قابل انتقال بودند. متن از ردیف کردن حروف و علامت‌ها بر روی خطوط تشکیل می‌شد (شکل ۴). حروفچین در مقابل سینی قسمت‌بندی‌شده‌ای که حاوی حروف متفاوت بود می‌ایستاد و اگر متن شامل زبان‌ها یا خط‌های گوناگون بود، چند سینی لازم داشت. حروف یک‌به‌یک در قاب کوچکی که در دست حروفچین بود قرار می‌گرفتند و هنگامی که آن قاب با چند خط از متن پر می‌شد، متن چیده شده به قاب بزرگ‌تری انتقال می‌یافت و چیدن قطعهٔ بعدی آغاز می‌شد. با تکرار این مراحل، نهایتاً ۸ صفحه از متن برای چاپ بر روی برگهٔ بزرگی از کاغذ و ۸ صفحهٔ دیگر هم برای چاپ در پشت برگه آماده می‌شد و ترتیب صفحات آن هم مطابق شکل ۵ بود.

در مرحلهٔ بعدی، برگه‌های چاپ شدهٔ ۱۶ صفحه‌ای تا می‌شد و در کنار برگه‌های تاشدهٔ دیگر قرار می‌گرفت تا کتاب کامل را، که تعداد صفحاتش همواره مضربی از

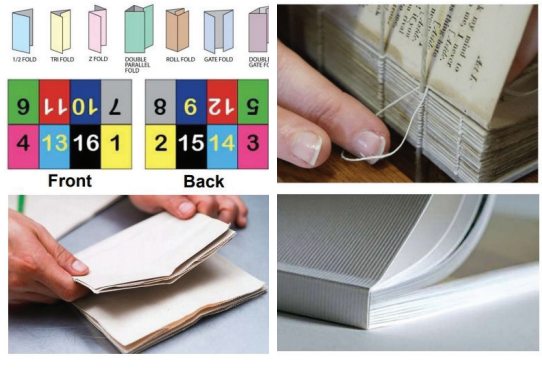
[10]Willem M. Floor, "Čāp," in *Encyclopedia Iranica*, vol. I/7, 760-764; available online at http://www.iranicaonline.org/articles/cap-print-printing-a-persian-word-probably-derived-from-hindi-chapna-to-print-see-turner-no

[11]Nile Green, "Persian Print and the Stanhope: Industrialization, Evangelicalism, and the Birth of Printing in Early Qajar Iran," *Comparative Studies of South Asia, Africa, and the Middle East*, 30:3 (December 2010), 473-490.

۱۶ بود و بنابراین احتمالاً صفحات سفیدی در انتها داشت، تشکیل دهد. سپس، گروه‌های ۱۶ صفحه‌ای به هم دوخته می‌شدند، لبه‌های به‌هم‌چسبیدهٔ صفحات با یک دستگاه بزرگ برش از هم جدا می‌شدند و کتاب با جلد مناسبی صحافی می‌شد. به مرور ماشین‌های دوخت و صحافی جای کار دستی وقت‌گیر را گرفتند و استفاده از چسب به جای دوختن این مرحله از کار را آسان‌تر، ارزان‌تر، و سریع‌تر کرد. برش لبه‌های کتاب گاه ناکامل بود و خواننده می‌بایست با دست لبه‌های چسبیده را از هم باز کند.



شکل ۴. حروف‌چینی دستی با حروف انتقال‌پذیر، قاب حاوی متن چیده شده، و ماشین چاپ. تصویر راست بالا برگرفته از
https://franciscojaviertostado.files.wordpress.com/2013/12/metal_movable_type.jpg
تصویر راست پایین برگرفته از
https://www.abc.net.au/news/image/7189966-3x2-460x307.jpg/.



شکل ۵. ترتیب انتقال متن بر روی برگهٔ کاغذ، تا کردن آن، دوختن، و صحافی کتاب. تصاویر صحافی برگرفته از
https://neilpaints.com/wp-content/uploads/2018/09/coloring-book-printing-and-binding-marvelous-manhattan-perfect-bound-book-of-coloring-book-printing-and-binding.jpg/.

شکل ۶. حروف انتقال‌پذیر فارسی در سینی قسمت‌بندی‌شدهٔ حروفچین. برگرفته از https://blog.29lt.com/.

اولین مسئله در چاپ فارسی تولید بلوک‌های فلزی حامل حروف و علایم برجستهٔ لازم بود (شکل ۶). چون برخلاف خط لاتین که حروف آن جدا هستند و با فاصلهٔ کوچکی کنار هم قرار می‌گیرند، حروف فارسی در بسیاری از موارد به هم می‌چسبند. بنابراین، تولید حروف لازم مشکل‌تر بود، ضمن آنکه در خط متداول آن زمان حروف همیشه افقی کنار هم قرار نمی‌گرفتند، بلکه گاه به صورت عمودی روی هم سوار می‌شدند. مثلاً در واژهٔ "مجتمع،" در پایین شکل ۷، حروف اول و دوم (م، ج) و نیز حروف سوم و چهارم (ت، م) را نمی‌توان در راستای افقی از هم جدا کرد.

گونهٔ جدید واژهٔ "مجتمع،" که در بالای شکل ۷ دیده می‌شود، این مشکل را با تغییر روش خطاطی از بین می‌برد، به طوری که حروف به‌کاررفته در راستای افقی تجزیه‌پذیر می‌شوند. مشکلات مشابهی نیز در برخی حروف پهن یا بلند فارسی وجود داشت. مثلاً در واژهٔ "چنگال" (شکل ۷)، سرکش حرف "گ" به محدودهٔ حرف "ن" امتداد یافته است که جلوی جداسازی حروف و قرار دادن هر حرف روی بلوکی جداگانه را می‌گیرد. همهٔ این مشکلات با تغییراتی در قوانین خط‌نویسی و گاه با قربانی کردن زیبایی و خوانایی خط به منظور تجزیه‌پذیری آن حل شد.

به مصداق مثل "عیب می جمله بگفتی، هنرش نیز بگو،" یکی از ویژگی‌های خط فارسی کار حروفچین را آسان می‌کند و آن امکان کشیدن حروف است. مثلاً در حروفچینی شعرهای فارسی، که معمولاً به شکل دو ستونی (مصرع اول و مصرع

دوم) است، رسم بر این است که ستون‌ها پهنای ثابتی داشته باشند. در سمت چپ شکل ۷ می‌بینیم که در واژهٔ "آفریدم،" حروف "ف" و "ی" کشیده شده‌اند و واژه به صورت "آفـــریـــدم" درآمده تا پهنای آن مصرع با مصرع‌های دیگر یکسان شود. این نوع کشیدن حروف با استفاده از علامت "ـــ" امکان‌پذیر است که بین حروف چسبیده قرار می‌گیرد، چون چسباندن حروف همواره روی یک خط افقی با نام محور اتصال صورت می‌گیرد.

علاوه بر مشکلات ناشی از به‌هم‌چسبیدگی و روی‌هم‌افتادگی حروف فارسی که به آنها اشاره شد، شکل و اندازهٔ حروف نیز مشکل‌زاست. حروف لاتین شکل‌های هندسی و تقریباً هم‌اندازه دارند (شکل ۸، بالا)، حال آنکه حروف فارسی بیشتر قوس دارند و از نظر پهنا و بلندی بسیار متغیرند. به این علل و نیز به سبب حذف برخی از حروف صدادار، خط فارسی از نظر افقی فشرده‌تر از لاتین و در راستای عمودی، به سبب نیاز به فاصلهٔ بیشتر بین خطوط، جاگیرتر است. در نتیجه، بلوک‌های حاوی حروف فارسی با اندازه‌های بزرگ و کوچکشان کار حروفچینی را دشوارتر می‌کنند.

آن‌گونه که در بحث ماشین‌های تحریر خواهیم دید، استفاده از حروف با پهنای یکنواخت به کیفیت خط فارسی صدمهٔ چشمگیری می‌زند. در قسمت پایین شکل ۸، چند بیت از یک شعر فارسی با حروف دارای پهنای یکنواخت و متغیر حروفچینی شده است که اثر آن بر کیفیت خط حاصل به خوبی مشهود است. همهٔ مشکلاتی که برای خط فارسی برشمردیم در خط عربی هم موجود است. تفاوت فقط در این است که فارسی ۴ حرف اضافی دارد که بسیار به برخی حروف موجود در عربی شبیه‌اند. یکی از راه‌های فهماندن این مشکلات به کسانی که با فارسی آشنا نیستند، تشبیه آنها به مشکلات موجود در خطاطی (خط دست‌نویس) در لاتین است که در آن هم حروف به هم می‌چسبند و اندازه‌هایشان بسیار متفاوت است.

مثال‌هایی که از خط چاپی فارسی امروز در شکل ۹ دیده می‌شوند حاوی صفحهٔ اول چند روزنامه با عنوان‌های درشت و متن‌های ریزتر از سال ۲۰۱۷ است. ممکن است خواننده تصور کند که تضمین خوانایی عناوین روزنامه‌ها به علت درشتی خط کار دشواری نیست، اما باید توجه کرد که این عناوین باید از فاصله‌ای دورتر خوانا باشند و جلب توجه کنند و بنابراین، همان مسائل دربارهٔ آنها هم وجود دارد.

شکل ۷. تطبیق خط فارسی با صنعت چاپ و حروف انتقال‌پذیر آن. تصویر راست برگرفته از
https://www.researchgate.net/figure/Demonstration-of-Arabic-typeface-characteristics-A-A-sample-of-a-five-letter-word_fig1_302059929
تصویر وسط برگرفته از
https://arabictattoo.wordpress.com/2009/02/03/spoon-and-fork-in-naskh/.

شکل ۸. مقایسهٔ حروف و خط لاتین با حروف و خط فارسی و عربی. تصویر حروف عربی برگرفته از
https://ya-webdesign.com/transparent250_/arabic-alphabet-png-4.png/.

شکل ۹. نمونه‌هایی از خط چاپی فارسی امروزی در روزنامه‌های ایران. برگرفته از
www.theatlantic.com/.

ماشین تحریر فارسی

ماشین تحریر حدود ۱۲۰ سال پیش به ایران وارد شد، ولی مانند صنعت چاپ کاربرد وسیع آن بلافاصله آغاز نشد؛ این امر تغییراتی در خط فارسی می‌طلبید. برای چاپ، هر حرف فارسی به چهارگونه ساخته می‌شد که متناظر با شکل‌های تنها، ابتدایی، میانی، و پایانی آن حرف در هر کلمه بود. چهار شکل برخی از حروف، مانند "ع" (قسمت بالای شکل ۱۰) بسیار متفاوت‌اند و می‌بایست چهار کلید متفاوت ماشین تحریر یا دو کلید همراه با کلید "تغییر مکان" به آنها اختصاص یابد. البته تهیهٔ صفحه‌کلیدی با ۶۴ کلید حرفی (۳۲ حرف الفبا ضرب در ۲) و تعداد دیگری کلید برای ارقام و علامت‌های متداول اصلاً عملی نبود. خوشبختانه، برای بیشتر حروف فارسی گونه‌های ابتدایی و میانی و نیز گونه‌های تنها و پایانی تا حدودی مشابه‌اند و ادغام آنها صدمهٔ زیادی به زیبایی و خوانایی خط حاصل نمی‌زند. برای مثال، حروف "م" و "ب" فقط دو گونه لازم دارند (شکل ۱۰) و برخی از حروف مانند "ه" هم، با صدمه‌ای نسبتاً بیشتر به خوانایی و زیبایی خط، می‌توانند فقط دو گونه داشته باشند. روی هم رفته، به استثنای دو حرف "ع" و "غ،" بقیهٔ حروف را می‌توان با دو گونهٔ کوچک و بزرگ–مانند حروف کوچک و بزرگ لاتین–کارسازی کرد.

با به بازار آمدن ماشین‌های تحریر هوشمند، تحول بزرگی در وارد کردن اطلاعات فارسی به واژه‌پردازها و کامپیوترها حاصل شد. این گونه دستگاه‌های ورودی هوشمند و نرم‌افزارهای آنها فقط به یک گونه از هر حرف فارسی نیاز داشتند، چون می‌توانستند گونهٔ تایپ‌شده یا نمایش‌داده‌شده را با تجزیه و تحلیل متن ورودی تعیین کنند. مثلاً برای وارد کردن واژهٔ "کمتر" (شکل ۱۰، چپ)، کاربر ابتدا حرف "ک" را وارد می‌کند که به سبب نامعلوم بودن حرف بعدی به صورت تنها روی نمایشگر ظاهر می‌شود. در مرحلهٔ دوم، با وارد کردن حرف "م،" واژه‌پرداز می‌فهمد که حرف قبلی، یعنی "ک،" باید به صورت ابتدایی نمایش داده شود و آن را تغییر می‌دهد. این روش تا برخورد به انتهای خط یا انتهای کلمه (فشار دادن کلید بازگشت یا فاصله) ادامه می‌یابد. این‌گونه تغییر پس از چاپ در ماشین تحریر ممکن نیست و بنابراین، ماشین‌های تحریر هوشمند متن ورودی را با یک حرف تأخیر چاپ می‌کردند.

ماشین‌های تحریر فارسی در ابتدا صفحه‌کلیدهایی مانند آنچه در سمت چپ شکل ۱۱ دیده می‌شود داشتند. برای آنکه کاربران ماشین‌های تحریر بتوانند به سادگی با ماشین‌های تحریر مختلف کار کنند، استانداردی برای ترتیب کلیدها تدوین و

تصویب شد.[۱۲] مکانیسم چاپ، که در ابتدا شامل چکش‌های کوچکی بود که با برخورد به نوار مرکب و کاغذ شکل حروف را بر روی کاغذ ایجاد می‌کردند، آرام‌آرام بهبود یافت.

ماشین تحریری که خط فارسی‌اش کیفیت بسیار بالایی داشت آی‌بی‌ام سلکتریک (IBM Selectric) بود که مکانیسم چاپ آن گویی پلاستیکی به شکل توپ گلف بود (شکل ۱۱، راست). با فشار دادن هر کلید، چرخش و خم شدن گوی چاپ حرف مورد نظر را به نقطهٔ چاپ می‌آورد و با ضربهٔ گوی به نوار مرکب و کاغذ آن را چاپ می‌کرد. در آوردن گوی و جایگزین کردن آن با گویی دیگر بسیار آسان بود و به این ترتیب، حروف با شکل‌ها و اندازه‌های متفاوت، علایم ریاضی، و حتی حروف زبان‌های دیگر به سادگی قابل ترکیب بودند. نگارنده چنین ماشین تحریری را برای آماده کردن کتاب *آشنایی با کامپیوتر* خود به کار گرفت که نمونهٔ نتایج حاصل از آن در شکل ۱۲ دیده می‌شود.[۱۳]

ماشین تحریر آی‌بی‌ام سلکتریک همچنین دارای حافظهٔ موقتی بود که می‌توانست محتوی یک سطر کامل از متن را در خود ذخیره کند. با فشار دادن کلید "پس‌رفت،" حرف قبلی از حافظه خوانده می‌شد و گوی چاپ دوباره همان حرف را با ضربه روی نوار سفید و کاغذ چاپ می‌کرد. این ویژگی کارِ ماشین‌نویس‌ها را بسیار ساده می‌کرد، چون دیگر نیازی به استفاده از مایع پاک‌کن سفید و مشکل انتظار برای خشک شدن آن پیش از ادامهٔ کار نداشتند.

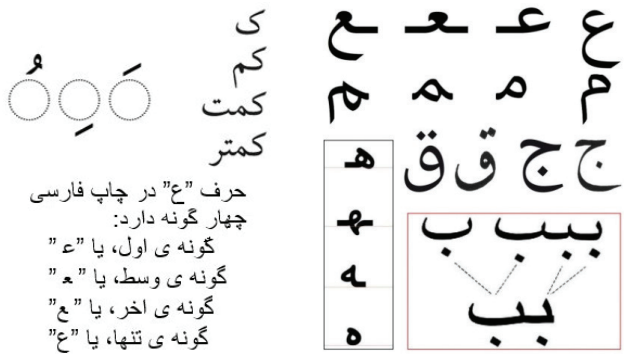
در اواسط دههٔ ۱۹۷۰، تلاش‌هایی برای طراحی و تصویب استانداردی واحد برای صفحه‌کلید در ماشین‌های تحریر و کامپیوترها صورت گرفت. پیشنهاد حاصل، که به علت نابسامانی‌های ادارات دولتی در یکی دو سال قبل از انقلاب هرگز به تصویب و اجرا نرسید، در سمت چپ شکل ۱۳ دیده می‌شود.[۱۴] برای این طرح استاندارد، به خاطر کلیدهای واقع در ردیف پایین صفحه‌کلید، نام "زودگذر" انتخاب شد و امید بر آن بود که استاندارد چندان زودگذر نباشد. استانداردی هم برای ذخیره و تبادل

[12]Institute of Standards and Industrial Research of Iran, *Character Arrangement on Keyboards of Persian Typewriters* (Tehran: ISIRI, 1976).

[۱۳]بهروز پرهامی، *آشنایی با کامپیوتر* (تهران: طلوع آزادی، ۱۳۶۳).

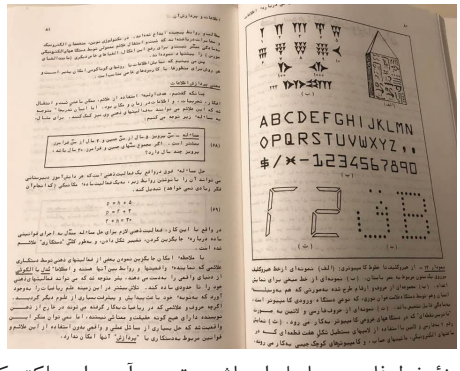
[14]See Parhami, "Standard Farsi Information Interchange Code and Keyboard Layout."

اطلاعات فارسی بر مبنای کُد سازمان بین‌المللی استانداردها پیشنهاد شد.[15] آن هم مانند استاندارد صفحه‌کلید هرگز تصویب و ابلاغ نشد.



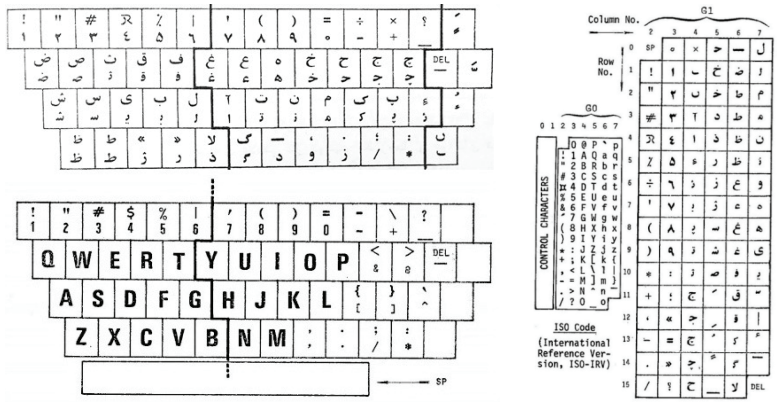
شکل ۱۰. حروف فارسی در چاپ چهارگونه، در ماشین تحریر دوگونه و در واژه‌پردازهای هوشمند یک گونه دارند.

شکل ۱۱. صفحه‌کلید یکی از ماشین‌های تحریر قدیمی فارسی و گوی حروف در ماشین تحریر برقی آی‌بی‌ام سلکتریک.

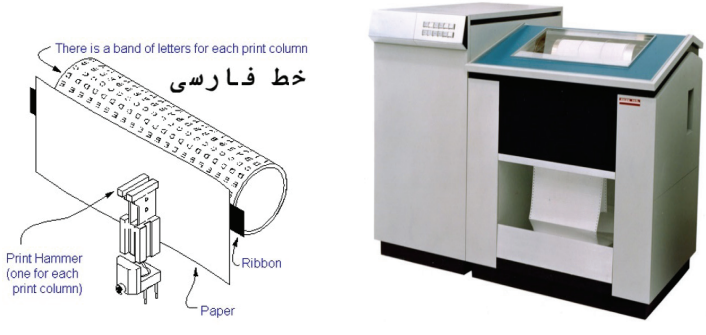


شکل ۱۲. نمونهٔ خط فارسی حاصل از ماشین تحریر آی‌بی‌ام سلکتریک در تدوین کتاب *آشنایی با کامپیوتر* نگارنده.

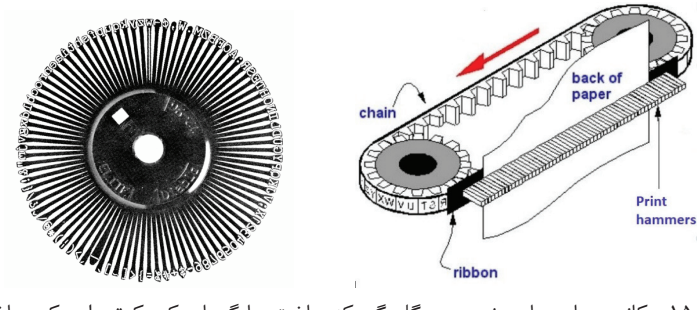
[15] See Parhami, "Standard Farsi Information Interchange Code and Keyboard Layout;" Iranian Plan and Budget Organization, *Final Proposal for the Iranian National Standard Information Code*.



شکل ۱۳. استانداردهای پیشنهادی برای ترتیب کلیدها در صفحه‌کلید و کُد نمایش و تبادل اطلاعات فارسی.



شکل ۱۴. چاپگر طبله‌ای، نمونهٔ کوچکی از خط حاصل از آن، و نوعی دستگاه چاپ سطری.
تصویر راست برگرفته از IBM و تصویر چپ برگرفته از
PC Magazine Encyclopedia.



شکل ۱۵. مکانیسم‌های چاپ زنجیری و گلبرگی که ساخت چاپگرهای کوچک‌تر را ممکن ساختند.
تصویر چپ برگرفته از پایگاه اینترنتی Computer Sc IT & Management و تصویر راست برگرفته از
PC Magazine Encyclopedia.

چاپگرها و نمایشگرهای اولیهٔ کامپیوتر

قدیمی‌ترین دستگاه‌های چاپ کامپیوتری چاپگرهای سطری بودند که یک سطر کامل از متن را بر روی صفحات پهن و طومارگونه یک‌جا چاپ می‌کردند (شکل ۱۴، راست). مکانیسم چاپ در این چاپگرها طبله‌ای فلزی بود که به ازای هر یک از ستون‌های چاپ ــ معمولاً ۱۳۲ ستون ــ یک حلقه از حروف و علایم مورد نیاز را شامل می‌شد (شکل ۱۴، چپ). وقتی که طبله با سرعت زیاد حول محورش می‌چرخید، همهٔ حروف قابل چاپ به نوبت از موضع چاپ رد می‌شدند و در هر ستون، درست هنگامی که در محل سطر چاپ قرار می‌گرفت، چکش کوچکی بر روی حرف مورد نظر ضربه می‌زد و به کمک نوار مرکب آن را روی کاغذ چاپ می‌کرد. چاپ کردن هر سطر نیازمند یک گردش طبله بود که فقط چند هزارم ثانیه طول می‌کشید.

یکی از عیب‌های چاپگرهای طبله‌ای برای چاپ خط فارسی فواصلی بود که به سبب فاصلهٔ بین چکش‌ها و نیز بین حلقه‌های روی طبله بین حروف مجاور ایجاد می‌شد و کیفیت خط را پایین می‌آورد. مثال‌هایی از این فواصل در بالای شکل ۱۴ مشهودند. گاه به دلایلی چکش درست در وقت موعد به طبله نمی‌خورد و در نتیجه، حروف تشکیل‌دهندهٔ یک سطر کمی نسبت به یکدیگر بالا و پایین ظاهر می‌شدند که کیفیت خط را باز هم بدتر می‌کرد.

برای مدتی کوتاه، روش جالبی برای بهبود کیفیت خط فارسی به کار گرفته شد که بعدها به علت برخی اشکالات اضافی آن را کنار گذاشتند. در این روش، جهت سطرهای چاپ تغییر یافت و به جای آنکه حروف در ستون‌های متفاوت با چکش‌های گوناگون چاپ شوند، یک چکش به هر سطر اختصاص یافت تا حروف سطر را به نوبت چاپ کند. به این ترتیب، حروف بدون فاصله‌های ناخوشایند موجود در شکل ۱۴ چاپ می‌شدند، ولی اگر صفحه‌ای حاوی فقط یک یا دو سطر اطلاعات بود، چاپ کردن آن به اندازهٔ یک صفحهٔ کاملاً پُر وقت می‌گرفت.

مکانیسم‌های چاپ زنجیری و گلبرگی (شکل ۱۵) از طبله کوچک‌تر بودند و در نتیجه، چاپگرهایی که از آنها استفاده می‌کردند حجم و وزن کمتری داشتند. این دو نوع چاپگر هم از چکش‌های کوچکی برای ضربه زدن و تشکیل حروف بر روی کاغذ استفاده می‌کردند و از نظر کیفیت خط حاصل همان اشکالات چاپگرهای طبله‌ای را داشتند.

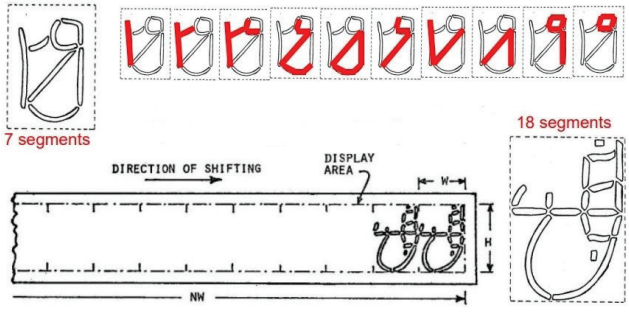
پیش از ادامهٔ بحث دربارهٔ بهبود کارایی چاپگرهای کامپیوتری، شرح مختصری دربارهٔ خروجی تصویری روی نمایشگرهای کامپیوتری لازم است. یکی از ابتدایی‌ترین روش‌های نمایش اطلاعات عددی شامل استفاده از نمایشگرهای هفت‌قطعه‌خطی بود، که در بالای شکل ۱۶ دیده می‌شود. خاموش یا روشن بودن هفت لامپ کوچک، که بعدها از یکسوکننده‌های منوّر ساخته شدند، شکل ارقام صفر تا ۹ را برای کاربردهایی چون ماشین حساب، ساعت، و دماسنج ایجاد می‌کرد. با گسترش کاربرد این نمایشگرها، انواع دیگری از آنها با ۹ یا ۱۴ یا تعداد دیگری از قطعه‌خط‌ها برای نمایش ارقام، حروف و علایم ویژه ساخته شدند که کیفیت نسبتاً خوبی در نمایش حروف بزرگ لاتین داشتند، ولی برای حروف کوچک و برخی علایم ویژه چندان کارا نبودند.

در ایران هم تلاش‌هایی برای کاربرد نمایشگرهای قطعه‌خطی در خط فارسی صورت گرفت. نتیجهٔ کار این بود که ارقام فارسی هم با ۷ قطعه خط قابل نمایش‌اند (شکل ۱۷، بالا)، ولی برای نمایش متون حد اقل ۱۸ قطعه‌خط لازم است (شکل ۱۷، پایین). هر دوی این طرح‌ها در حد پژوهش و نمونه‌سازی باقی ماندند و به مرحلهٔ تولید نرسیدند، چرا که نمایشگرهای قطعه‌خطی کم‌کم از دور خارج شدند و جای خود را به نمایشگرهای نقطه‌ای دادند. آرایه‌ای از چراغ‌ها با ۷ سطر و ۵ ستون برای نمایش خط لاتین در فهرست‌های پرواز فرودگاه‌ها، تابلوهای ثبت نتایج در میدان‌های ورزشی، و موارد مشابه دیگر کافی بود. هر چه تعداد چراغ‌های مورد استفاده برای هر حرف بالاتر می‌رفت، مثلاً ۹ سطر و ۷ ستون، و چراغ‌ها نزدیک‌تر به هم جاسازی می‌شدند، کیفیت خط حاصل بالاتر می‌رفت.

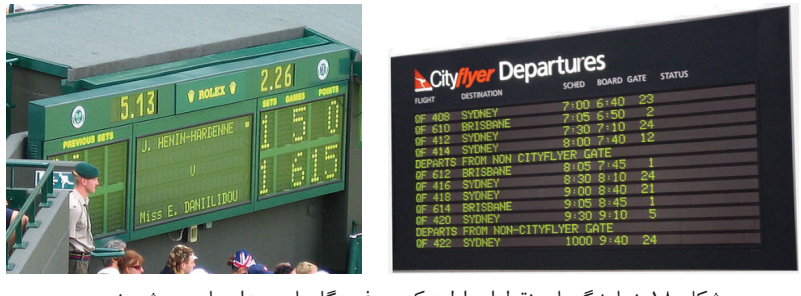
نمایشگرهای کاتُدی، که مانند تلویزیون‌های قدیمی شامل صفحهٔ حساس فسفری و شعاع الکترونیکی بودند که بر آن می‌تابید، حروف و علایم را با رسم خطوط مستقیم و منحنی یا با استفاده از سخت‌افزارهای ویژه‌ای موسوم به "علامت تولیدکن،" که شعاع الکترونیکی را برای ترسیم علامت مورد نظر هدایت می‌کردند، روی صفحهٔ نمایشگر می‌تاباندند. نهایتاً، طراحان این نمایشگرها به روش بُرداری روی آوردند که در آن صفحهٔ نمایش به صورت آرایه‌ای از صفر و یک (نقاط سیاه و سفید) منظور می‌شد و شعاع الکترونیکی در حال طی کردن ردیف به ردیف صفحهٔ نمایش، برای ایجاد نقاط سیاه و سفید خاموش و روشن می‌شد. نمایشگرهای رنگی در واقع سه شعاع جداگانه برای سه رنگ اولیهٔ قرمز، آبی، و زرد داشتند. نمایشگرهای کاتُدی در

نهایت با نمایشگرهای مسطح جایگزین شدند، ولی روش نقطه‌ای نمایش اطلاعات هنوز هم پابرجاست.

شکل ۱۶. نمایشگرهای قطعه‌خطی برای ارقام، حروف الفبای لاتین، و علایم ویژه. تصویر راست برگرفته از
https://upload.wikimedia.org/wikipedia/en/thumb/0/0a/14_Segment_LCD_characters.jpg/220px-14_Segment_LCD_characters.jpg/.



شکل ۱۷. نمایشگرهای قطعه‌خطی پیشنهادی برای ارقام و حروف الفبای فارسی.



شکل ۱۸. نمایشگرهای نقطه‌ای اولیه که در فرودگاه‌ها و میدان‌های ورزشی نصب می‌شدند. تصویر راست برگرفته از
http://computronics.biz/productimages/prodairport3.jpg/.

پیشرفت در تکنولوژی چاپ کامپیوتری

طی سال‌ها، روش‌های گوناگون چاپ کامپیوتری ابداع و آزمایش شدند، ولی همهٔ آنها نهایتاً جای خود را به روش چاپ نقطه‌ای دادند که شباهت‌هایی به خط کوفی مربع دارد؛ همان که کاشی‌کاران برای نوشتن متون در مساجد، کاخ‌ها، و ساختمان‌های مهم دیگر استفاده می‌کردند (شکل ۱۹، راست). همان‌گونه که یک کاشی‌کار می‌تواند همهٔ کاشی‌های یک ردیف را بچیند و بعد به سراغ ردیف بعدی برود، چاپگرهای نقطه‌ای اولیه دارای هفت سوزن در یک ردیف بودند که هر کدام می‌توانستند در یکی از دو حالت بالا و پایین قرار گیرند. فشردن این سوزن‌ها به نوار مرکب و کاغذ بین صفر تا ۷ نقطه را بر روی کاغذ چاپ می کرد. پس از تکمیل چاپ یک ردیف، یا سوزن‌ها به محل ردیف بعدی می‌رفتند یا کاغذ کمی در جهت عکس حرکت می کرد و این امر تا چاپ تمام صفحه تکرار می شد (شکل ۱۹، چپ). بعدها چاپگرهای نقطه‌ای با سوزن‌های بیشتر در یک یا چند ردیف به بازار آمدند. سوزن‌ها در ردیف‌های مختلف یا با هم مطابق بودند یا کمی نسبت به هم جابه‌جایی داشتند. سوزن‌های مکانیسم چاپ نهایتاً با مرکب‌پاش‌های کوچکی جایگزین شدند که مشکلات ناشی از حرکت مکانیکی سوزن‌ها را حل کرد.

با ایجاد انعطاف‌پذیری در طراحی خط، چاپگرهای نقطه‌ای کیفیت خط حاصل را بالا بردند و در ضمن قابلیت چاپ نقشه، شکل‌های هندسی و هر نقش دیگری را هم داشتند (شکل ۲۰). با کوچک‌تر شدن نقطه‌ها و افزایش آنها در هر اینچ مربع از سطح کاغذ یا صفحهٔ نمایشگر، کیفیت خط حاصل به‌تدریج بهبود یافت، چنان‌که امروز به زحمت می‌توان نقطه‌های به‌کاررفته در تشکیل حروف و علایم را از هم تشخیص داد. مشکلات ناشی از پیوستگی و اندازه‌های متفاوت حروف فارسی هم از بین رفتند. آنچه تا آن زمان مسئله‌ای سخت‌افزاری بود، به مفهومی نرم‌افزاری (مشخص کردن نقطه‌های تاریک و روشن برای تشکیل دادن هر حرف) تبدیل شد.

طبیعتاً در ایران هم طراحی حروف فارسی به صورت ماتریس‌های نقطه‌ای آغاز شد (شکل ۲۱). پس از طراحی مجموعهٔ حروف و علایم مورد نظر، آنها را روی کاغذ کپی می‌کردیم، می‌بریدیم و در ترکیبات مختلف کنار هم قرار می‌دادیم تا کیفیت خط حاصل را ارزیابی کنیم (شکل ۲۱، وسط). اگر از کیفیت خط راضی نبودیم، تغییراتی در شکل حروف و علایم می‌دادیم و ارزیابی را تکرار می‌کردیم. نهایتاً، برای ایجاد این ترکیبات برنامه‌ای نوشتیم و دیگر لازم نبود این کار را به صورت دستی انجام دهیم.

همزمان با بهتر شدن طرح‌های ما، اندازهٔ ماتریس‌های نقطه‌ای نیز تدریجاً بزرگ‌تر و دقت آنها بیشتر می‌شد که به بالا رفتن خوانایی و زیبایی خط کمک می‌کرد.

تأثیر اندازهٔ ماتریس نقطه‌ای بر کیفیت خط حاصل در سمت چپ شکل ۲۲ مشهود است که در آن، حرف R لاتین در ماتریس‌های ۷در۵ و ۱۶در۱۶ با نقاط هم‌پوشان به نمایش درآمده است. جالب آنکه خوانایی خط با سهولت تشخیص آن توسط کامپیوتر رابطه‌ای مستقیم دارد. در یک طرح پژوهشی برای تشخیصِ خودکارِ عناوین روزنامه‌ها–در آن زمان، دقت دستگاه‌های تصویربرداری برای خط معمولی روزنامه کافی نبود–مسئلهٔ تفکیک تکه‌ای از متن به حروف تشکیل‌دهندهٔ آن بر رسی و حل شد.[۱۶] در سمت راست شکل ۲۲ نتیجهٔ تفکیک بخش اول از واژهٔ "تهران،" شامل سه حرف "ت " و "ه " و "ر،" با کامپیوتر دیده می‌شود. نتایج حاصل از این طرح پژوهشی به خودی خود جالب و مفید بودند و به طراحی حروف نقطه‌ای خواناتر هم کمک کردند. طی نزدیک به چهار دهه که از اجرای طرح پژوهشی بالا می‌گذارد، رشتهٔ شناسایی کامپیوتری متون فارسی پیشرفت چشمگیری کرده است و نتیجهٔ جستجوی "شناسایی حروف فارسی" در گوگل بیش از ۶ میلیون صفحه است.

روش چاپ یا نمایش نقطه‌ای بالاخره مشکل خروجی فارسی از کامپیوتر را به نحو مطلوبی حل کرد، ضمن آنکه اجازهٔ طراحی برخی از مجموعه‌های علایم ویژه و تخصصی را هم به دست داد.[۱۷] در شکل‌های ۲۳ و ۲۴ دیده می‌شود که درجهٔ بهبود کیفیت خط فارسی از چاپگرهای ضربه‌ای و نقطه‌ای ساده در دههٔ ۱۹۷۰ تا چاپگرهای نقطه‌ای نسل دوم در دههٔ ۱۹۸۰ چشمگیر بوده است. در سه دههٔ اخیر کیفیت خط باز هم بهبود یافته و در نتیجه، روش چاپ و نمایش نقطه‌ای تثبیت شده است.

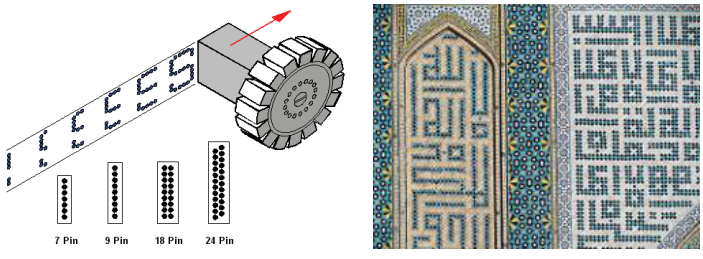
با رفع مشکلات چاپ و نمایش خط فارسی، بحث‌های گاه و بی‌گاه دربارهٔ متحول ساختن خط فارسی به منظور تطبیق آن با تکنولوژی و پیش‌گیری از "عقب افتادگی فرهنگی"[۱۸] هم کمتر شده است. این نتیجهٔ جانبی موجب خوشحالی است، چون

[16]See Parhami and Taraghi, "Automatic Recognition of Printed Farsi Texts."

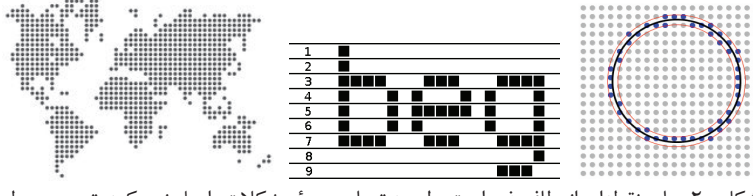
[17]Behrooz Parhami, "Optically Weighted Dot-Matrix Farsi and Arabic Numerals," *Proc. 3rd Jerusalem Conf. Information Technology 78* (North-Holland, 1978), 207-210.

[18]Maryam Borjian and Habib Borjian, "Plights of Persian in the Modernization Era," *Handbook of Language and Ethnic Identity: The Success-Failure Continuum in Language and Ethnic Identity Efforts* (Oxford: Oxford University Press, 2011), vol. 2, 254-267.

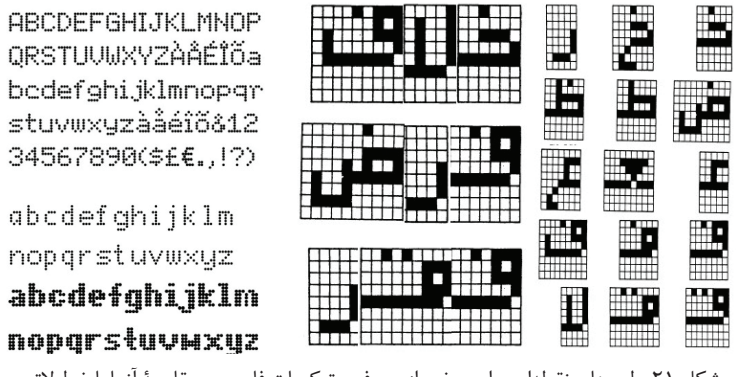
روش درست در پیوند تکنولوژی تطبیق فنون نوین با فرهنگ بومی است و نه تغییر فرهنگ به علت محدودیت‌های فنی. البته، زیاده‌روی در جهت عکس، یعنی اصرار بر خلوص زبان و مقاومت در برابر هر نوع تغییر و تحول، هم مشکل‌آفرین خواهد بود.



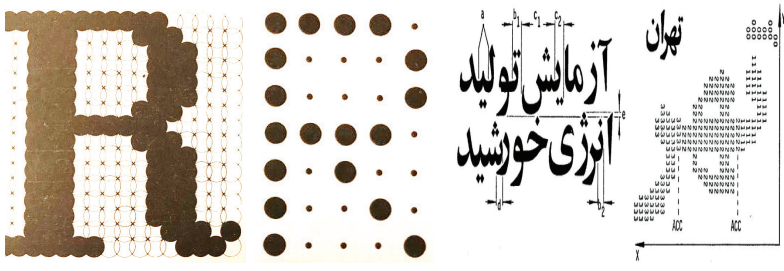
شکل ۱۹. خط کوفی مربع، که کاشی‌کاران ابداع کردند، و تکنولوژی ابتدایی چاپ نقطه‌ای. تصویر چپ برگرفته از PC Magazine Encyclopedia.



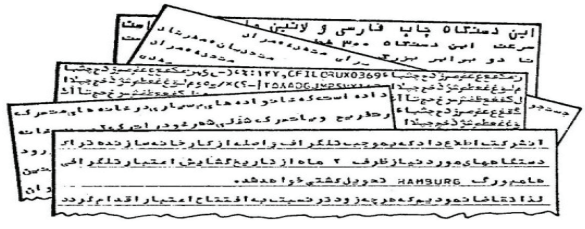
شکل ۲۰. چاپ نقطه‌ای انعطاف‌پذیر است، ولی به تنهایی همهٔ مشکلات را حل نمی‌کند. تصویر وسط برگرفته از
https://blog.gatunka.com/2009/11/03/japanese-computers-still-living-it-8-bit/.



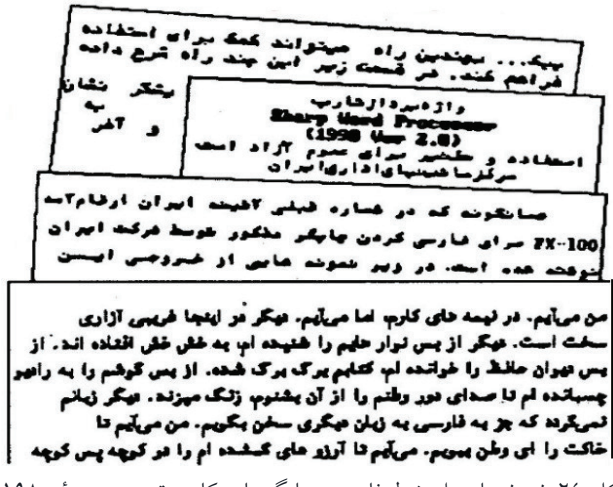
شکل ۲۱. طرح‌های نقطه‌ای برای برخی از حروف و ترکیبات فارسی و مقایسهٔ آنها با خط لاتین.



شکل ۲۲. ماتریس‌های کوچک (۷در۵) و بزرگ (۱۶در۱۶، با نقاط همپوشان) و مثالی از حروف پیوسته.



شکل ۲۳. نمونه‌هایی از خط فارسی چاپگرهای کامپیوتری در دهۀ ۱۹۷۰.



شکل ۲٤. نمونه‌هایی از خط فارسی چاپگرهای کامپیوتری در دهۀ ۱۹۸۰.

خطوط فارسی و عربی در کامپیوترهای امروزی

با گسترش ارتباطات الکترونیکی و شبکه‌های اجتماعی، پیشرفت‌های بسیاری در کیفیت چاپ و نمایش خط فارسی حاصل شده است. در حال حاضر، همۀ

سیستم‌های موجود از روش نقطه‌ای برای تولید خط استفاده می‌کنند. برای نمونه، باید از بهبود خط فارسی در فیس‌بوک یاد کرد، ولی خطِ بهتر غالباً فراگیر نیست و فقط در گوشه‌هایی از سیستم ظاهر می‌شود. مثلاً در سال ۲۰۱۷، پیشنهاداتی برای رفع برخی از مشکلات خط فارسی به فیس‌بوک ارائه دادم و در اواخر همان سال، خط تازه‌ای بر روی "سکوهای تلفن همراه" پیاده شد که کیفیت بهتری داشت. اما متأسفانه، سکوهای کامپیوتری از این پیشرفت‌ها بی بهره ماندند (شکل ۲۵، چپ). در سمت راست شکل ۲۵ نمونهٔ خط فارسی یک پایگاه خبری دیده می‌شود که کیفیت بسیار خوبی دارد. درسی که از این تحولات باید گرفت این است که کاربران باید همیشه هشیار باشند و در هر موقعیتی مشکلات و ایرادها را به مسئولان گوشزد کنند. با در دست بودن مجموعه‌های بزرگی از متون فارسی در شبکه‌های اجتماعی و سیستم‌های ارتباطی دیگر، هیچ بهانه‌ای برای ارزیابی نکردن کیفیت خط و فقدان نظر خواهی از کاربران برای بهبود آن پذیرفته نیست.

مشکلات تولید خط خوانای فارسی عمدتاً همان‌ها هستند که خواندن کامپیوتری خط را هم دشوار می‌کنند. عنوانی از یک روزنامه، که در شکل ۲۶ دیده می‌شود و چهار دهه قبل در یک طرح پژوهشی از آن استفاده شد،[19] به خوبی مشکلات تشخیص نقاط اتصال حروف (a)، حروف مشابه (b_1, b_2)، تفاوت‌های چشمگیر در پهنای حروف (c_1, c_2)، هم‌پوشی افقی حروف واقع در یک سطر (d)، و هم‌پوشی عمودی بین حروف بلند در سطرهای مجاور (e) را نشان می‌دهد. یکی از نتایج جانبی این پژوهش تعیین پهنای قلم چهار نقطه‌ای (شکل ۲۷، چپ) برای احراز دقت قابل قبول در شناسایی حروف بود. همین پهنای قلم چهار نقطه‌ای ضابطهٔ خوبی برای خوانا کردن خط است، چون پهنای چهار نقطه یا بیشتر باعث واضح‌تر شدن انحناها، سوراخ‌ها، و دیگر ویژگی‌های حروف فارسی می‌شود (شکل ۲۷).

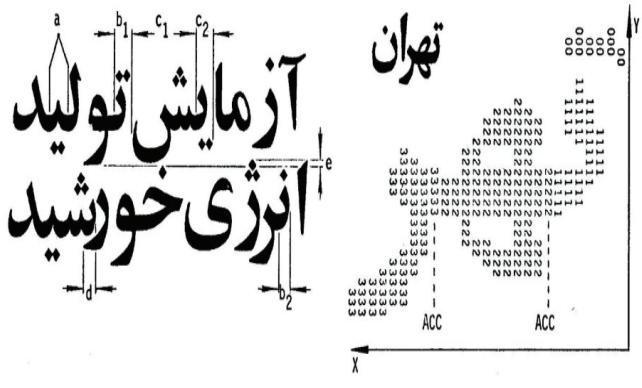
حال که حد پایینی اندازهٔ ماتریس نقطه‌ای را برای تولید خط زیبا و خوانای فارسی می‌دانیم، بد نیست اشاره‌ای هم به اندازهٔ ماتریس در تولید خط قابل قبول برای کاربردهایی کنیم که باید با هزینهٔ کم پیاده شوند.[20] نتایج شبیه‌سازی با ماتریس‌های ۷در۵، ۷در۲÷۹، و ۹در۲÷۹ در بخش وسطی شکل ۲۸ دیده می‌شوند. وقتی که یکی از ابعاد ماتریس را به صورت m÷۲ می‌نویسیم، منظور این است که m نیم‌ستون

[19]See Parhami and Taraghi, "Automatic Recognition of Printed Farsi Texts."

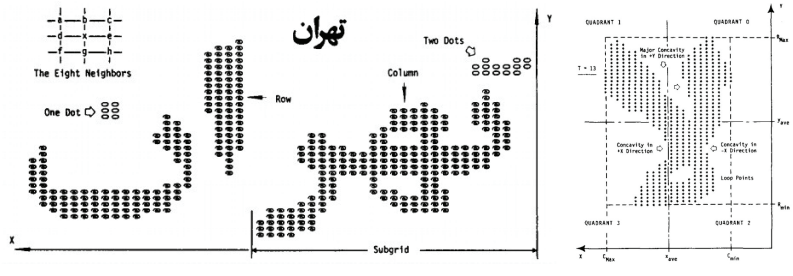
[20]Behrooz Parhami, "On Lower Bounds for the Dimensions of Dot-Matrix Characters to Represent Farsi and Arabic Scripts," *Proc. 1st Annual CSI Computer Conf.* (Tehran, December 1995), 125-130.

داریم که نقاط آنها نسبت به یکدیگر کمی جابه‌جا شده‌اند. این‌گونه ماتریس‌ها در نمایش خطوط مایل و منحنی بهتر از ماتریس‌های مستطیلی عمل می‌کنند.

شکل ۲۵. نمونهٔ خط نقطه‌ای فارسی از فیس‌بوک نگارنده (سمت چپ) و یک پایگاه خبری.



شکل ۲۶. برخی از مشکلات شناسایی متون چاپی فارسی.



شکل ۲۷. روش‌های به‌کاررفته و نتایج حاصل از شناسایی خودکار متون چاپی فارسی.

از آنجا که چاپ یا نمایش برخی از حروف قدبلند فارسی بیشتر از حروف دیگر مشکل‌زاست، شبیه‌سازی جداگانه‌ای برای این حروف صورت گرفت که پاره‌ای از نتایج آن در قسمت بالا و سمت راست شکل ۲۸ برای سه ماتریس ذکرشده دیده می‌شوند. در همان‌جا نتایج چاپ یا نمایش با ماتریس‌های دارای پهنای متغیر نیز

ثبت شده‌اند. نتیجهٔ نهایی این بود که ماتریس نقطه‌ای با ابعاد ۹در۲÷۹ کوچک‌ترین اندازه‌ای است که می‌توان برای حداقل خوانایی قابل قبول به کار برد.

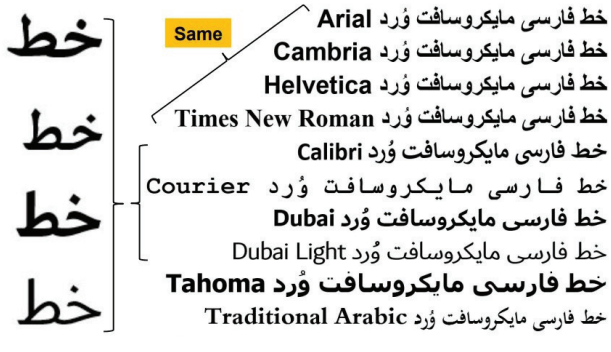
هنگامی که به خطوط ریزتر برای نمایش متن روی نمایشگرهای محدود تلفن‌های همراه و برخی دستگاه‌های الکترونیکی ارزان‌قیمت نیاز داریم، دو راه در پیش داریم. اول آنکه حروف کوچک‌تری را به کار بگیریم. مجموعه‌های حروف در اندازه‌های مختلف ارائه می‌شوند که به آن اندازهٔ حرف یا "پوینت" می‌گویند (شکل ۲۹، بالا). حروف با اندازهٔ ۲۰ معمولاً نصف حروف با اندازهٔ ۴۰ نیستند، چون افزایش خوانایی خط حکم می‌کند که برخی جزئیات در حروف با اندازهٔ کوچک‌تر با اغراق نمایش داده شود تا حروف مشابه از هم تمیز داده شوند. روش دوم کوچک کردن خطی حروف از راه فشرده‌سازی است. این کاری است که مثلاً در هنگام نمایش نقشه‌ای بزرگ بر روی صفحهٔ کوچک تلفن همراه صورت می‌گیرد و همهٔ عوامل نقشه، از جمله حروف و کلمات روی آن، به یک نسبت فشرده می‌شوند. نیمهٔ پایین شکل ۲۹ تفاوت‌های دو روش را نشان می‌دهد.

در پایان این بخش، اشاره‌ای هم به حروف فارسی مورد استفاده در برنامه‌های مایکروسافت و از جمله واژه‌پرداز وُرد می‌کنیم. شکل ۳۰ برخی از این مجموعه‌های حروف را نشان می‌دهد که متأسفانه در حال حاضر، چهار گونهٔ اول علی‌رغم نام‌های متفاوت از نظر شکل یکسان‌اند. شاید برنامه بر این باشد که این چهار گونه در آینده از هم متمایز شوند. می‌بینیم که حروف ارائه‌شده از نظر زیبایی و خوانایی متفاوت‌اند. شاید بتوان گفت که خطوط کالیبری و دوبی در مجموع بهترین‌اند-به خط‌های بزرگ‌شده در سمت چپ شکل ۳۰ توجه کنید. در خط کوری‌یر، اشکالات متداول حروف با پهنای یکسان به چشم می‌خورند، ولی این‌گونه حروف کاربردهایی هم دارند و گاه مفید واقع می‌شوند.

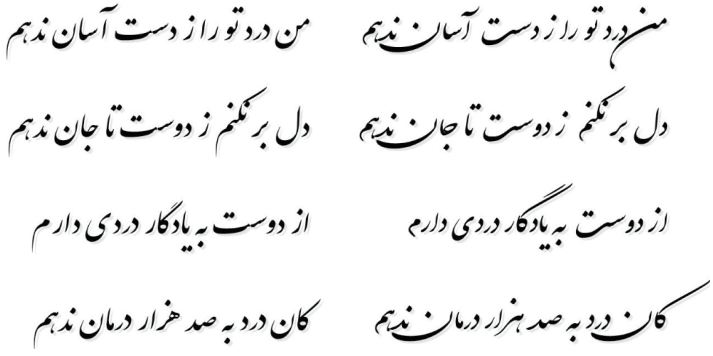
غیر از نمایش و چاپ اطلاعات به منظور خواندن متون، خطاطی با حروف هنری گوناگون نیز گاه مورد توجه است. پایگاه NastaliqOnline.ir امکاناتی را برای نمایش خطوط نستعلیق و شکسته و غیره در اختیار کاربران قرار می‌دهد. متن مورد نظر در یک جعبهٔ مستطیل شکل وارد شده و سپس فرمان "خطاطی کن" صادر می‌شود. نتیجهٔ خطاطی به صورت تصویر jpg به کاربر بازمی‌گردد که قابل کپی کردن و چسباندن در جاهای دیگر است. شکل‌های ۳۱ و ۳۲ حاوی مثال‌هایی در این زمینه‌اند.

شکل ۲۸. برخی از نتایج حاصل از پژوهشی با هدف تعیین حداقل اندازهٔ ماتریس‌های نقطه‌ای برای ارائهٔ خط فارسی.

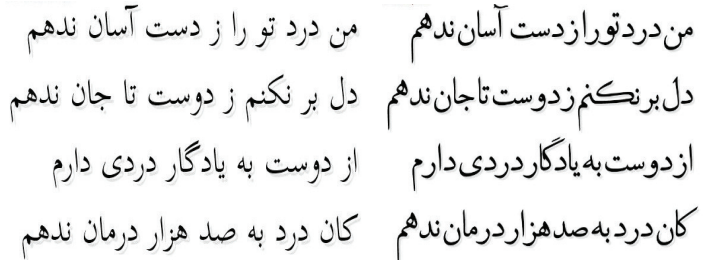
شکل ۲۹. حروف فارسی با اندازه‌های متفاوت و تأثیر کاهش اندازهٔ خط با کاربرد حروف کوچک‌تر یا فشرده‌سازی تصویری.



شکل ۳۰. مثال‌هایی از حروف فارسی واژه‌پرداز مایکروسافت وُرد و خطوط حاصله.



شکل ۳۱. نمونه‌های تولید خطوط نستعلیق (چپ) و شکسته (راست) در پایگاه NastaliqOnline.ir.



شکل ۳۲. نمونه‌هایی از دو خط دیگر که کامپیوتر تولید کرده است: عثمان طه (راست) و نیریزی (چپ). تولیدشده با ابزار NastaliqOnline.ir.

نتیجه‌گیری و بیان مشکلات باقی‌مانده

پس از ۵۰ سال تلاش و نوآوری، هنوز هم در مقولهٔ تولید خط زیبا و خوانای فارسی به پایان راه نرسیده‌ایم. البته افزایش سال‌به‌سال کیفیت و دقت نمایشگرهای کامپیوتری، که در نوع اچ‌دی حاوی ۱۹۲۰در۱۰۸۰ نقطه‌اند و در ماورای اچ‌دی ۳۸۴۰در۲۱۶۰ نقطه دارند، کمک بزرگی است. چاپگرهای امروزی هم بسیار دقیق‌اند و حتی ارزان‌ترین آنها که برای مصارف شخصی و خانگی ساخته می‌شوند توانایی چاپ حداقل ۶۰۰ نقطه در هر اینچ را دارند. بنابراین، می‌توان گفت که مشکل امروز ما فقدان امکانات سخت‌افزاری نیست، بلکه به طراحی حروف و الگوریتم‌های پردازش متن بازمی‌گردد. باید آنچه را که در این ۵۰ سال آموخته‌ایم با کاربرد روش‌های تجربیِ علمی و بهره‌گیری از نظرخواهی‌های گسترده به کار گیریم.

یکی از زمینه‌هایی که در آن به کار بیشتر نیاز داریم بررسی رابطهٔ تنگاتنگ بین زیبایی و خوانایی خط است. در این خصوص، به‌کارگیری روش مشارکت جمعی

و نظرخواهی‌هایی که در آن دو خط متفاوت کنار هم قرار داده شده و از کاربران درخواست می‌شود که آنها را از دید زیبایی و خوانایی ارزیابی کنند، کارساز خواهد بود. دیگر، مسائل قالب‌بندی متون است، به‌خصوص هنگامی که با متون دو یا چندزبانه سروکار داریم. یکی از دلایل مشکلات ترکیب خطوط فارسی و لاتین در یک متن جهت‌های متفاوت نگارش است. اگر مثلاً بخواهیم در میان یک متن فارسی عبارت English text را بیاوریم، بسته به آنکه این عبارت در کجای سطر قرار بگیرد (وسط یا نزدیک انتها که در نتیجه باید تمام یا بخشی از آن به سطر بعدی انتقال یابد)، احتمال بروز اشکال وجود دارد. علایم نقطه‌گذاری و انواع پرانتزها، که نوع باز و بسته‌شان در فارسی و لاتین عکس یکدیگرند، نیز گاه مسئله‌سازند.

گسترش چشمگیر تعداد کاربران چاپگرها و نمایشگرهای فارسی انگیزهٔ لازم را برای بهبود کیفیت خط فارسی ایجاد کرده است. این کاربران باید فعال باشند و از ابراز نظر نگریزند و از این راه به بهبود سیستم‌ها کمک کنند. طراحی یک صفحه‌کلید دو زبانهٔ فارسی-لاتین، که استفاده از آن در همهٔ محیط‌های کاربردی به سادگی ممکن باشد، و بهبود کیفیت سیستم‌های تشخیص خودکار حروف فارسی از دیگر گام‌های مهم اولیه‌اند. علاوه بر تلاش برای بهتر کردن سیستم‌های موجود ورودی و خروجی فارسی، باید نگاهی هم به آنچه که در راه است بیاندازیم. ورودی صوتی به زبان انگلیسی پیشرفت چشمگیری داشته است، ولی برای ورودی صوتی فارسی هنوز در ابتدای راهیم.[۲۱] همین امر در مورد خروجی صوتی هم صدق می‌کند.[۲۲]

غیر از مطالعات دربارهٔ چاپگرها و نمایشگرهای گوناگون، باید نگاهی هم به خود خط فارسی بیاندازیم تا ویژگی‌های آن و نحوهٔ برخورد با ضرورت‌های زیبایی و خوانایی آن را بهتر بشناسیم. در اینجا به ذکر دو نمونه از کارهایی که در زبان انگلیسی با خط لاتین صورت گرفته و می‌توان آنها را به زبان و خط فارسی تعمیم داد اکتفا می‌کنیم.

در کتابی که نگارنده برای یکی از درس‌های دورهٔ کارشناسی ارشد راجع به مقابله با خرابی‌ها و اشتباهات کامپیوتری به کار گرفته است،[۲۳] پیوست اول حاوی جدولی است که آمار اشتباهات تایپی را در وارد کردن دستی اطلاعات به کامپیوتر به

[21]Hossein Sameti et al., "Nevisa: A Persian Continuous Speech Recognition System," *Proc. 13th Int'l CSI Computer Conf.* (Berlin: Springer, 2008), 485-492.

[22]Mohammad Hadi Bokaei et al., "Niusha: The First Persian Speech-Enabled IVR Platform," *Proc. 5th IEEE Int'l Symp. Telecommunications* (2010), 591-595.

[23]Robert W. Bailey, *Human Errors in Computer Systems* (New Jersey: Prentice Hall, 1983).

دست می‌دهد. بر طبق این جدول، حروف و علایمی که کمترین اشتباه (کمتر از ۱ درصد) در ورود آنها صورت می‌گیرد عبارت‌اند از W،M، 3، 7، A، 9، E، C. حروفی که احتمال اشتباه در آنها پنج برابر بیشتر است (حدود ۵ درصد) عبارت‌اند از J، V، G، N، 0، 5. اشتباه‌برانگیزترین حروف در زبان انگلیسی Z (۱۳ درصد) و I (۲۵ درصد) هستند. واضح است که اشتباه‌برانگیز بودن و خوانایی خط و حروف رابطه‌ای نزدیک دارند، ولی به نظر نمی‌رسد که این رابطه در خط فارسی بررسی شده باشد.

در متون انگلیسی و برخی زبان‌های دیگر، معیار سهولت درک به تفصیل مطالعه شده است.[۲۴] سهولت درک، که با خوانایی فرق دارد، با کلمات مورد استفاده (کوتاه یا بلند، متداول یا نادر)، ساختار جملات (کوتاه و ساده یا بلند و پیچیده) و ویژگی‌های ظاهری متن (نوع حروف، اندازه، رنگ و غیره) ارتباط دارد. مفهوم خوانایی محدودتر است و به شکل حروف و علایم، فواصل بین آنها و شکل هندسی کلمات بستگی دارد. مثلاً برای شناسایی واژهٔ English، مغز انسان پیش از توجه به حروف موجود در آن به شکل ظاهری (بالا و پایین رفتن، انحناها و غیره) توجه می‌کند. ظاهراً مطالعاتی در این زمینه در خط فارسی صورت نگرفته است، ولی برای خط عربی دست‌کم یک مرجع موجود است.[۲۵]

امید می‌رود که این مقاله مکملی بر مقالات مروری پیشینی، که حدود چهار دهه قبل انتشار یافته‌اند،[۲۶] باشد و نقشی در ثبت و گسترش اطلاعات مرتبط با پایه‌گیری فناوری کامپیوتر در ایران ایفا کند.

---

[24]Mostafa Zamanian and Pooneh Heydari, "Readability of Texts: State of the Art," *Theory & Practice in Language Studies*, 2:1 (2012), 43-53.

[25]N. Chahine, "Reading Arabic: Legibility Studies for the Arabic Script" (Leiden University, Doctoral Dissertation, Faculty of the Humanities, 2012).

[26]Behrooz Parhami and Farhad Mavaddat, "Computers and the Farsi Language: A Survey of Problem Areas," *Information Processing 77: Proc. IFIP World Congress* (Amsterdam: North Holland, 1977), 673-676; Behrooz Parhami, "On the Use of Farsi and Arabic Languages in Computer-Based Information Systems," Proc. Symp. Linguistic Implications of Computer-Based Information Systems (New Delhi, India, November 1978), 1-15; Behrooz Parhami, "Impact of Farsi Language on Computing in Iran," *Mideast Computer*, 1 (September 1978), 6-7; Behrooz Parhami, "Language-Dependent Considerations for Computer Applications in Farsi and Arabic Speaking Countries," *System Approach for Development: Proc. IFAC Conf.* (Amsterdam: North-Holland, 1981), pp. 507-513.