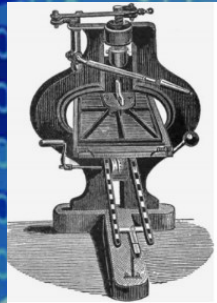


Evolutionary Changes in Persian and Arabic Scripts to Accommodate the Printing Press, Typewriting, and Computerized Word Processing

1600



1700



1800



1900



2000



Behrooz Parhami

University of California, Santa Barbara

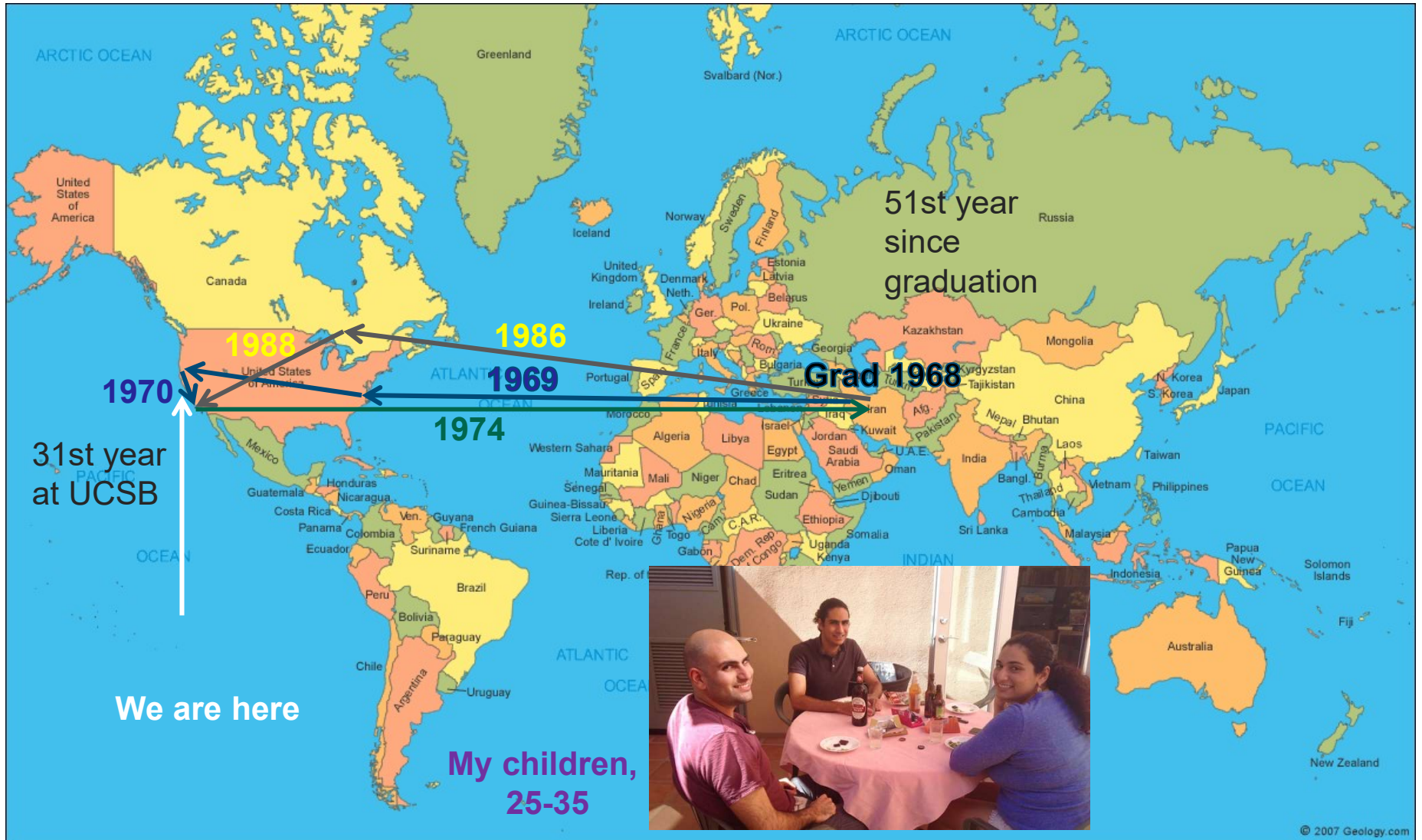
About This Presentation

This slide show was first developed in fall of 2017 for talks delivered as part of UCLA Bilingual Lecture Series on Iran. The two talks were presented on the UCLA campus, on Sun., Nov. 19, Dodd Hall 121, 4:00 PM (Persian), and Mon., Nov. 20, Humanities Bldg. 365, 2:00 PM (English). The current version was prepared for presentation at TUG Annual Conference, Palo Alto, CA, Aug. 9-11, 2019. All rights reserved for the author. ©2019 Behrooz Parhami

Edition	Released	Revised	Revised	Revised
First	Nov. 2017	Aug. 2019		

File: http://www.ece.ucsb.edu/~parhami/pres_folder/parh19-tug-evolution-persian-arabic-slides.pdf

My Personal Journey and Career



Aug. 2019

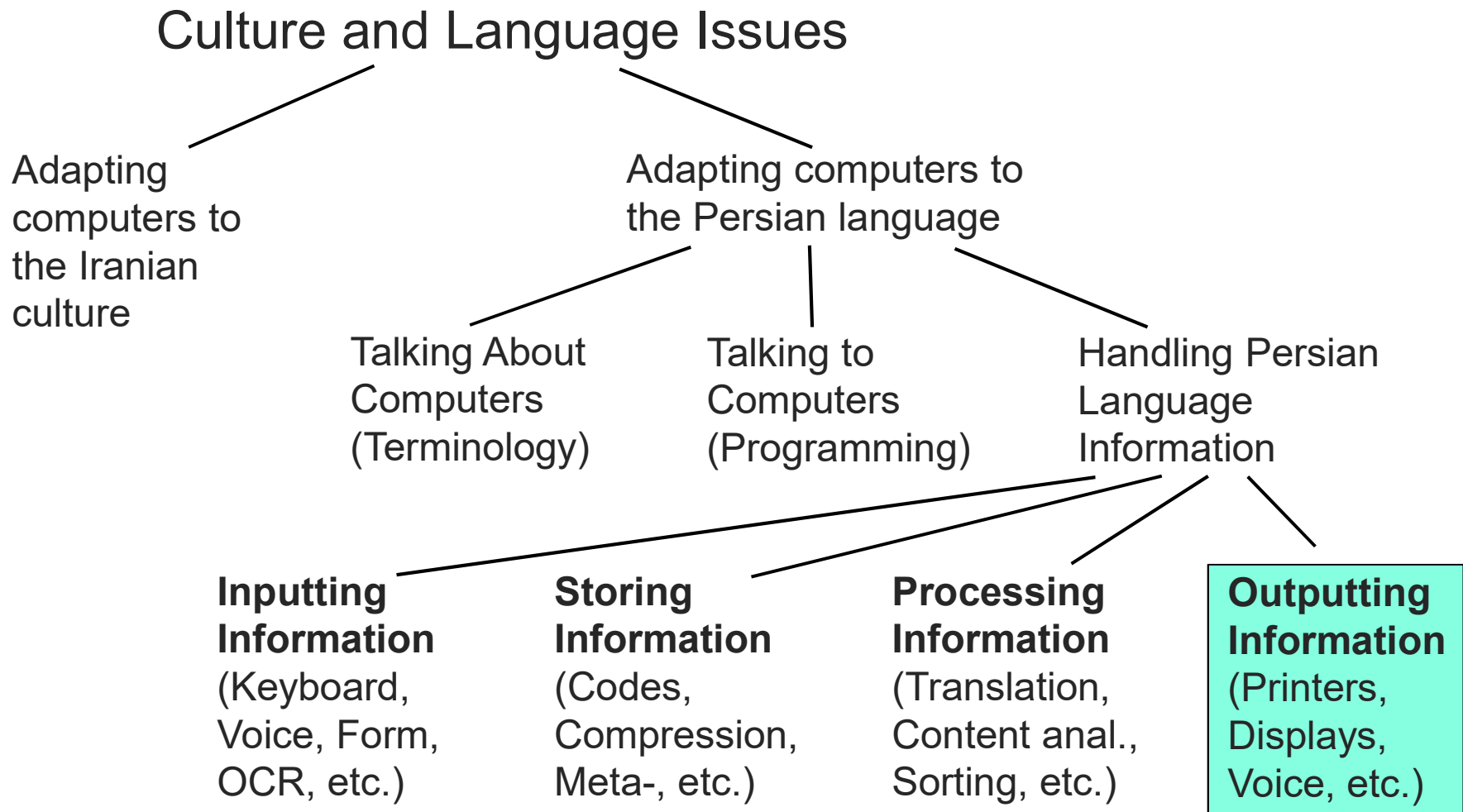


Evolutionary Changes in Persian & Arabic Scripts



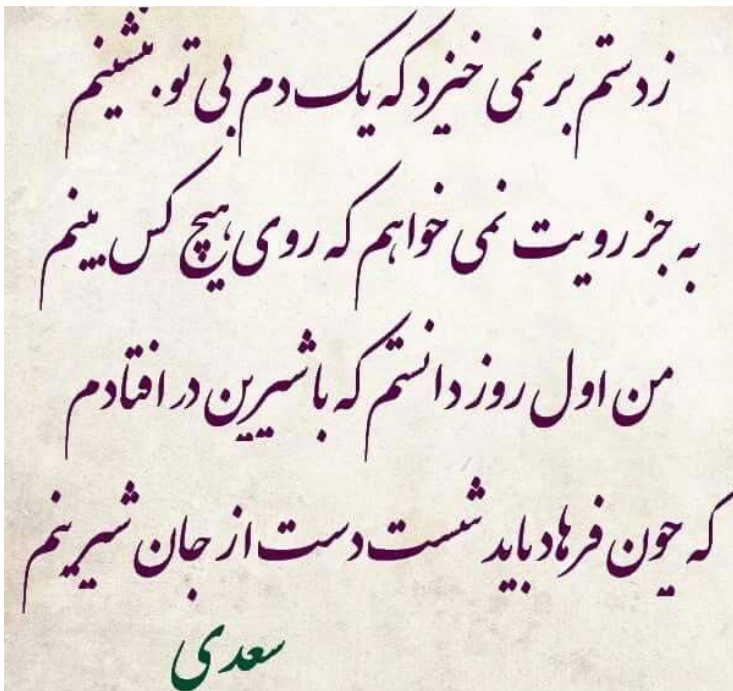
Slide 3

Computers and the Persian Language: Technology Transplantation (not Transfer)



Let's Start at the Very Beginning: A Very Good Place to Start

Nastaliq script was used
by scribes and
calligraphic artists



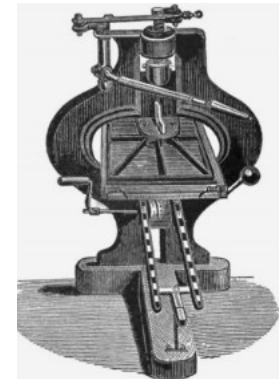
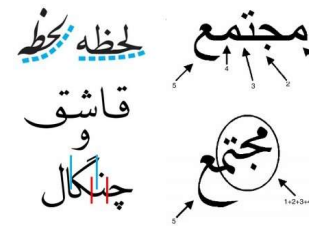
Text by
unknown
scribe



Art by
Farokh
Mahjoubi

Persian Script's Evolutionary Timeline

Year	Events Affecting the Development of Persian Script
1600	<ul style="list-style-type: none"> - Printing press arrives in Iran; little/no use early on - Armenian press established in Jolfa, Isfahan
1700	<ul style="list-style-type: none"> - Limited print runs; mostly on poetry and religion - Persian books published in Calcutta
1800	<ul style="list-style-type: none"> - First Stanhope hand-press arrives; printing spreads - Presses open in multiple cities; use of lithography - Technical books appear; newspapers flourish
1900	<ul style="list-style-type: none"> - First typewriter arrives in Iran - Typewriters begin to be used widely - Electric typewriters, Linotype, and computers arrive - Standards for information code and keyboard layout
2000	<ul style="list-style-type: none"> - Use of personal computers broadens - Computer-software and mobile-app industries thrive



Early Challenges with Print Typefaces: Decomposition and Added Uniformity

Naskh script was more suitable for adaptation to printing press, which came to Iran in the 1600s

لحظه بخط

قاشق

و چنگال

مجمع

مجمع

تخت جمشید

تخت جمشید یا پارسه (یا پرسپولیس، پرسه پلیس، هزارستون، صدستون و یا چهل منار) نام یکی از شهرهای باستانی ایران است که طی سالیان پیوسته، پایتخت باشکوه و تشریفاتی پادشاهی ایران در زمان امپراتوری هخامنشیان بوده است. در این شهر باستانی کاخی به نام تخت جمشید وجود دارد که در دوران زمامداری داریوش

چو شهر عشق من شهری ندیدم
ز نادانی بسی غربت کشیدم
هر آوازی که در عالم شنیدم
چو ماری بی پرو پا میخزیدم
چو گل بی حلق و بی لب می چشیدم
که من محنت سرائی **افریدم**
از آنجا آمدن هم می رمیدم
که من نزدیک، چون حبل الوری دم
فسون و عشوه او را خریدم

سفر کردم، بهر شهری دویدم
ندانستم ز اول قدر آن شهر
بغیر از عشق آواز دهل بود
در آن عالم که جان بودم به تنها
از آن باده که لطف و خنده بخشد
ندا آمد ز عشق: ای جان سفر کن
چنان کاکنون ز رفتن میگیریم
بگفت: ای جان برو هر جا که خواهی
فسون کرد و مرا بس عشوه ها داد

ز راهم برد و آنجا هم بره کرد
کز آن ره من نرفتم من رهیدم

مولوی (مولانا جلال الدین محمد بلخی رومی)

The Printing Press and Movable Type

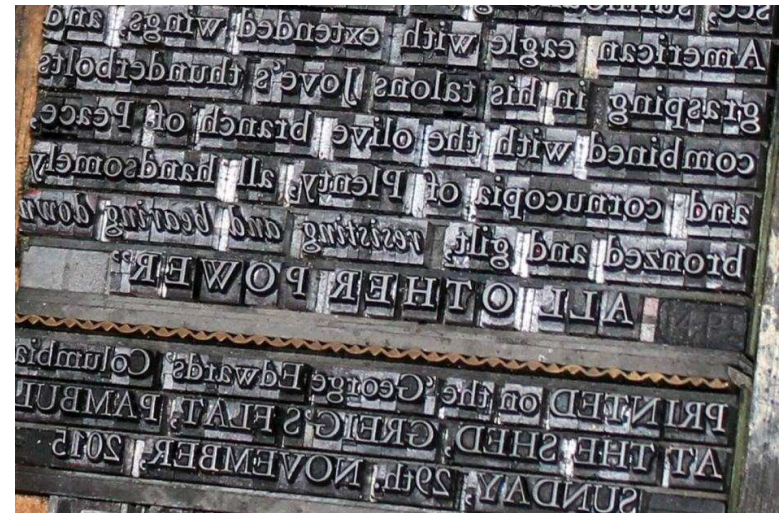
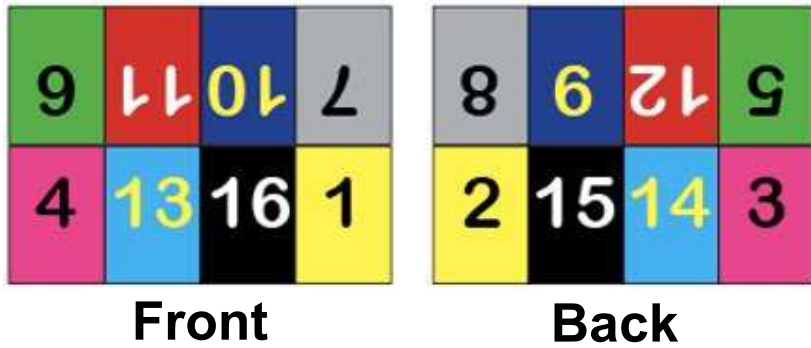
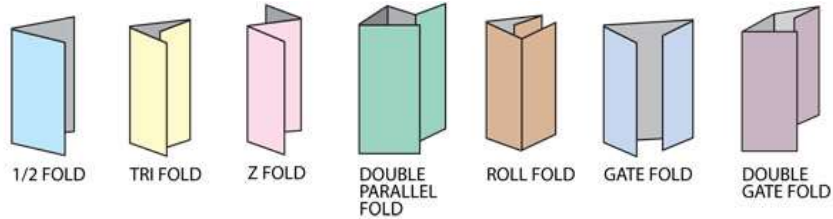


Image Transfer to Paper and Binding



Movable Persian Type Developed



Aug. 2019



Evolutionary Changes in Persian & Arabic Scripts



Slide 11

Persian Typefaces Developed into Modern Versions Now in Use



Aug. 2019

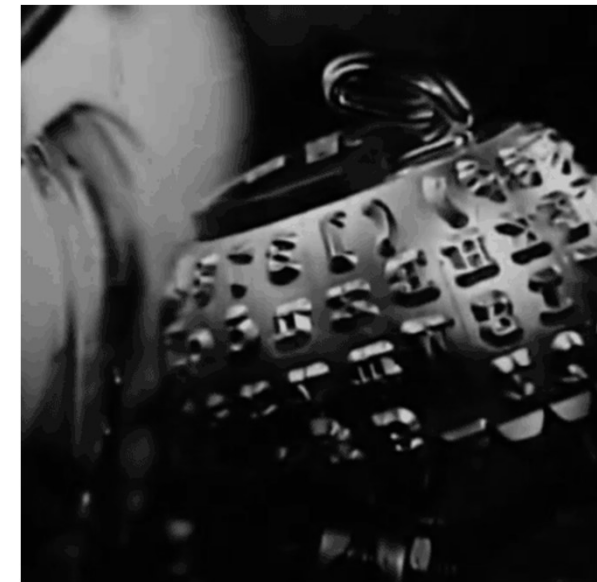


Evolutionary Changes in Persian & Arabic Scripts



Slide 12

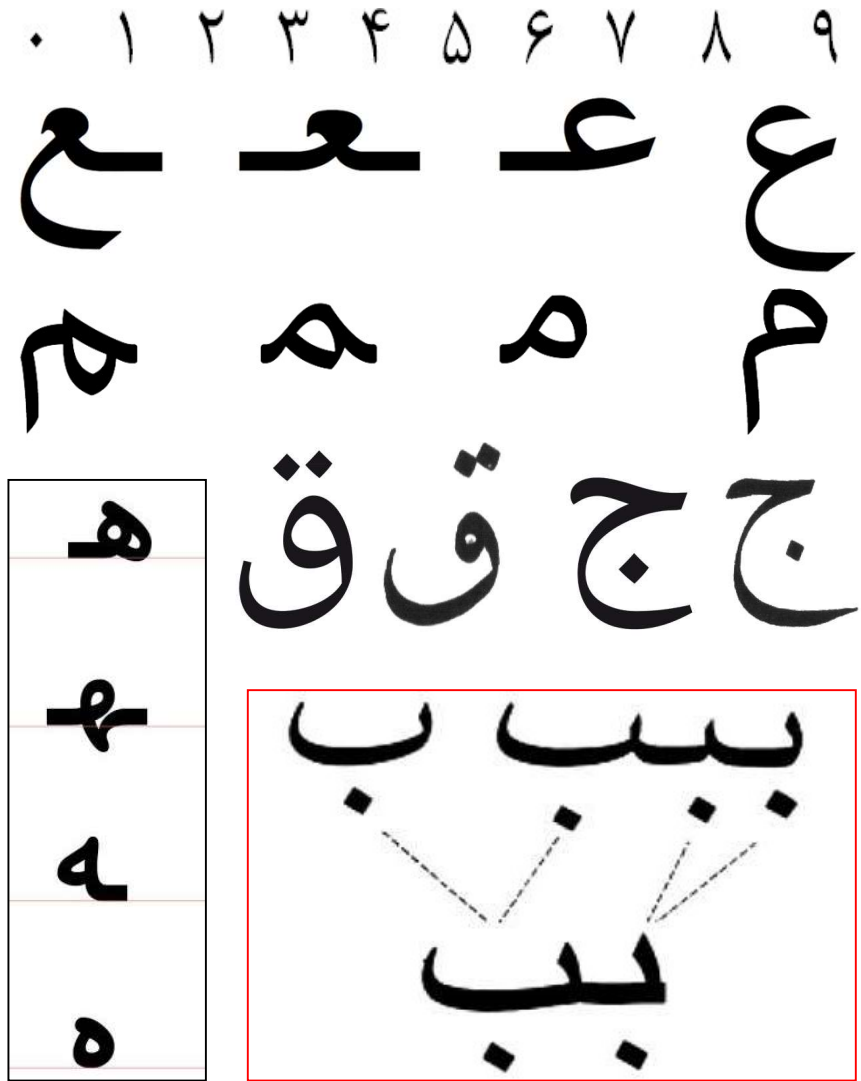
Persian Typewriters: Mechanical/Electric



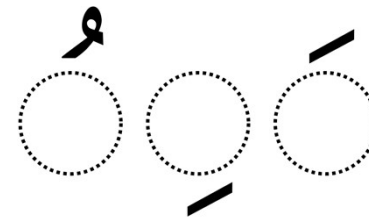
From Print to Typewriter Typefaces

Persian Alphabet / الفبای فارسی

Detached	Initial	Medial	Final	Roman	Name	Detached	Initial	Medial	Final	Roman	Name
ا	ا	ا	ا	á	alef	ص	ص	ص	ص	ş	sád
ب	ب	ب	ب	b	be	ض	ض	ض	ض	đ	zád
پ	پ	پ	پ	p	pe	ط	ط	ط	ط	t	tá
ت	ت	ت	ت	t	te	ظ	ظ	ظ	ظ	z	zá
ث	ث	ث	ث	th	se	ع	ع	ع	ع	'	ayn
ج	ج	ج	ج	j	jim	غ	غ	غ	غ	gh	ghayn
چ	چ	چ	چ	ch	che	ف	ف	ف	ف	f	fe
ح	ح	ح	ح	h	he	ق	ق	ق	ق	q	qáf
خ	خ	خ	خ	kh	khe	ک	ک	ک	ک	k	káf
د	د	د	د	d	dál	گ	گ	گ	گ	g	gáf
ذ	ذ	ذ	ذ	dh	zál	ل	ل	ل	ل	l	lám
ر	ر	ر	ر	r	re	م	م	م	م	m	mím
ز	ز	ز	ز	z	ze	ن	ن	ن	ن	n	nún
ژ	ژ	ژ	ژ	zh	zhe	و	و	و	و	v/ú	váv
س	س	س	س	s	sin	ه	ه	ه	ه	h	he
ش	ش	ش	ش	sh	shin	ی	ی	ی	ی	y/i	ye



Persian Typewriters: Memory and Intelligence



ک
کم
کمت
کمتر

حرف "ع" در چاپ فارسی چهار گونه دارد:

گونه ی اول، یا "ع"

گونه ی وسط، یا "ع"

گونه ی آخر، یا "ع"

گونه ی تنها، یا "ع"

حرف "ع" در چاپ فارسی چهار گونه دارد:

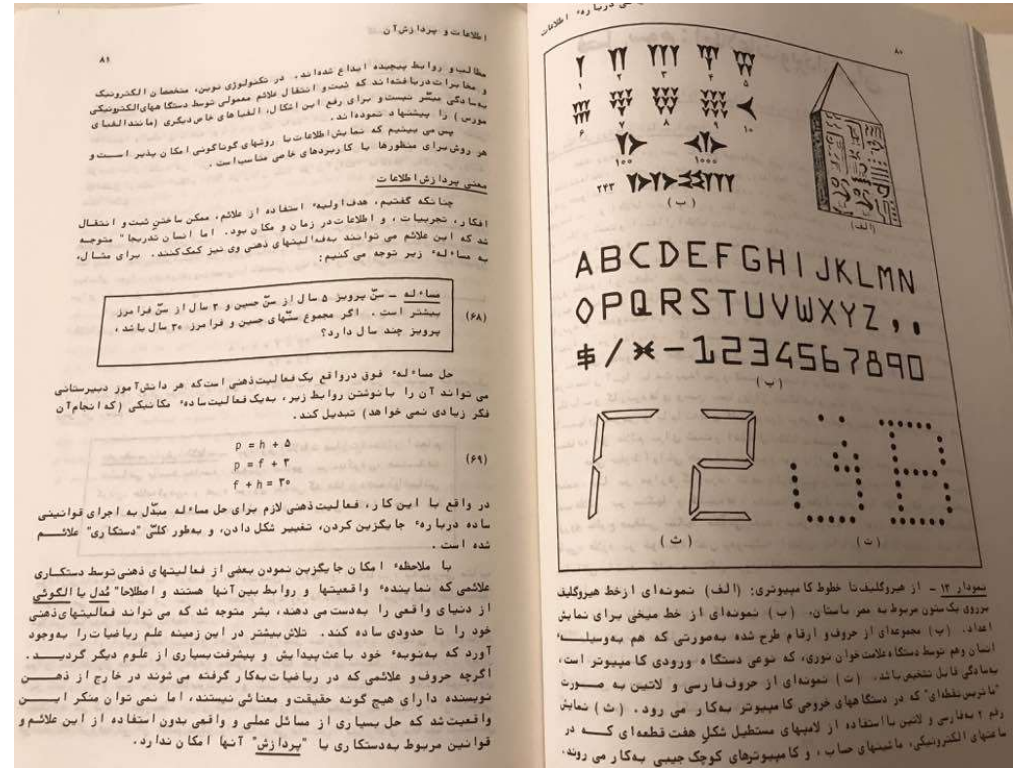
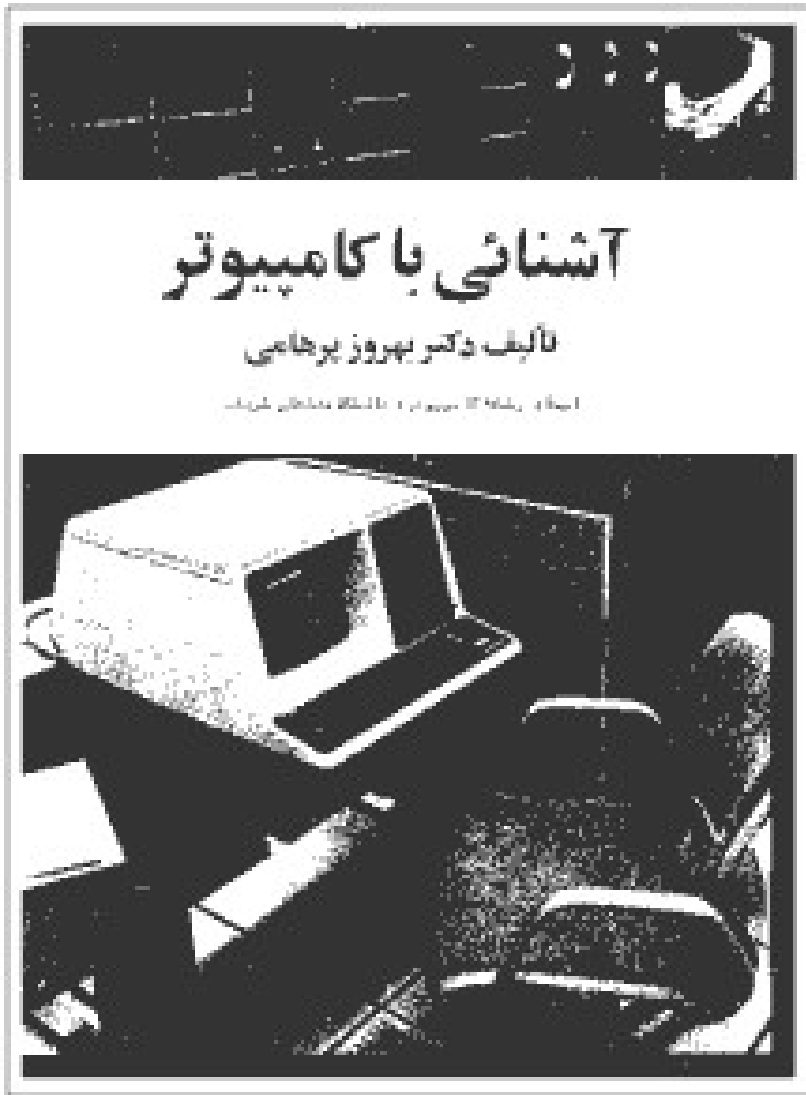
گونه ی اول، یا "ع"

گونه ی وسط، یا "ع"

گونه ی آخر، یا "ع"

گونه ی تنها، یا "ع"

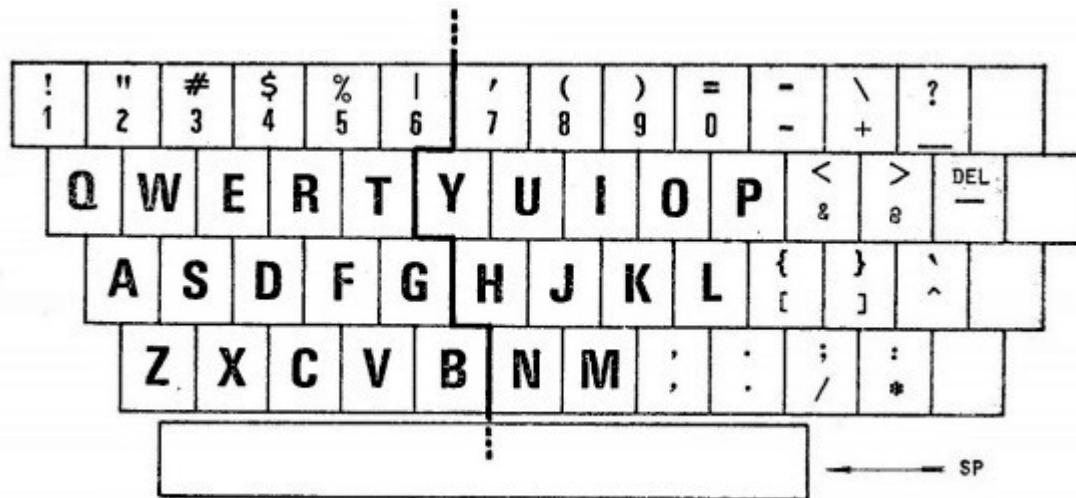
My 1984 Book Typed on IBM Selectric



کامپیوتر در ایران

کاربرد کامپیوتر در ایران از سال ۱۳۴۱ خورشیدی در کوچک "آی بی ام ۱۶۲۰" آغاز گردید. برطبق آمار موجود ایران در سال ۱۳۵۰ به حدود ۱۰۰ دستگاه رسید، در سال ۵۶ به روایتی بالغ بر ۷۰۰ دستگاه گردید، و از آن پس تقریباً

Keyboard Layout and Exchange Code

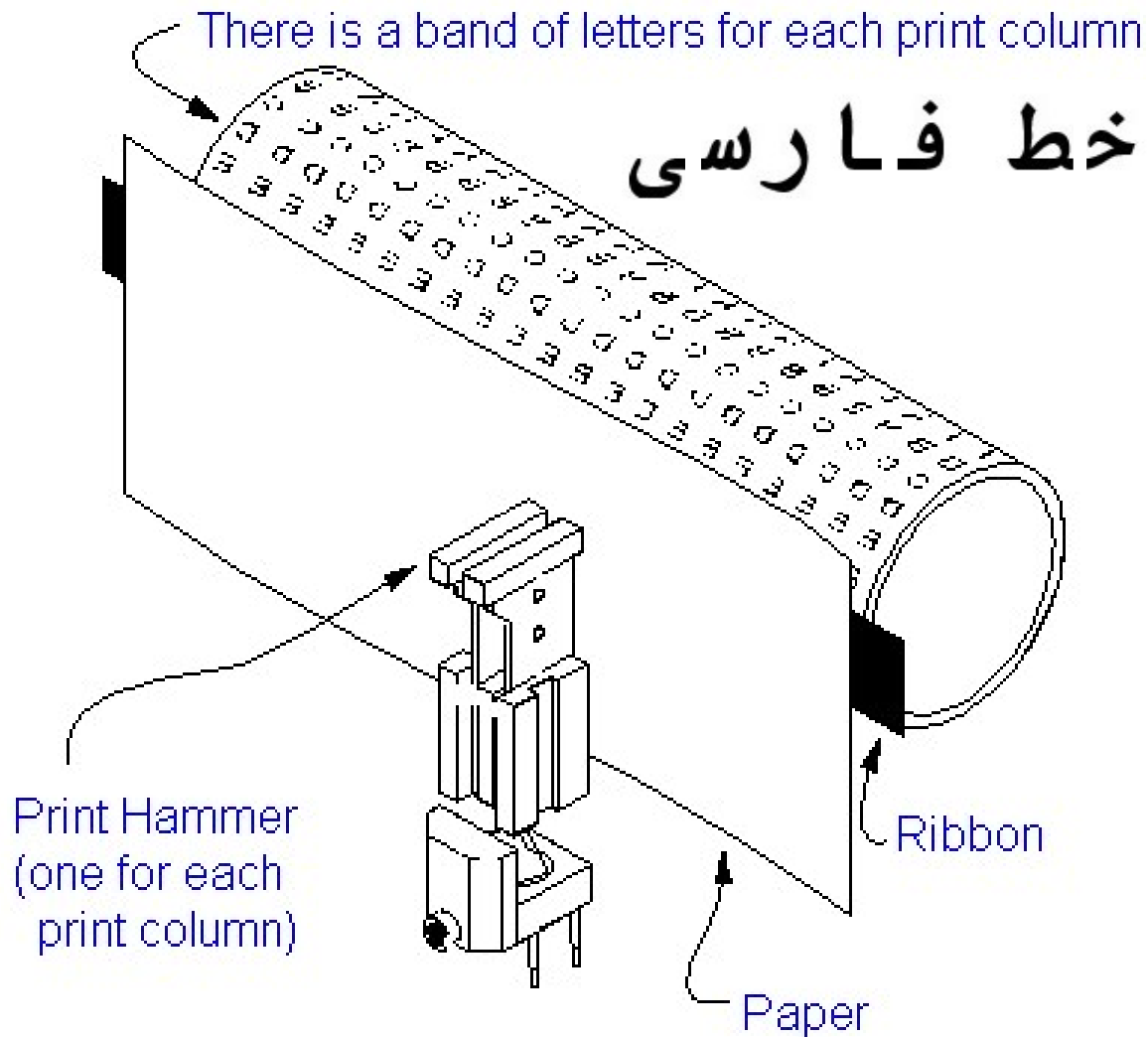


		G1						
		Column No.						
		2	3	4	5	6	7	
Row No.	0	SP	o	x	>	-	ل	
	1	!	ا	ب	خ	ض	ز	
2	"	ب	ط	ظ	م			
3	#	ا	ط	ن	م			
4	٪	ا	ن	ظ	ن			
5	÷	ا	ر	ظ	ن			
6	'	ا	ز	ع	و			
7	'	ا	ز	ع	و			
8	(ا	غ	ه				
9)	ا	غ	ه	ی			
10	*	:	ب	ف	ی			
11	+	:	ج	ق	"			
12	'	»	ق	ا				
13	-	=	ج	ک	ع			
14	.	»	ق	ک	-			
15	/	؟	ح	-	لا	DEL		

		G0							
		Column No.							
		0	1	2	3	4	5	6	7
CONTROL CHARACTERS	0	@	P	`	p				
	1	A	Q	a	q				
	2	B	R	b	r				
	3	C	S	s	s				
	4	D	T	d	t				
	5	E	U	e	u				
	6	F	V	f	v				
	7	G	W	g	w				
	8	H	X	h	x				
	9	I	Y	i	y				
	10	J	Z	j	z				
	11	K	[k	{				
	12	L	\	l					
	13	M]	m	}				
	14	N	^	n	~				
	15	O	_	o	o				

ISO Code
(International Reference Version, ISO-IRV)

Early Persian Drum Printers



Drum line-printers:
Could print an entire
line at once (during
one drum revolution)

Gap between
hammers and letters
on the drum did not
allow good
connectivity of letters

One solution:
Print vertically

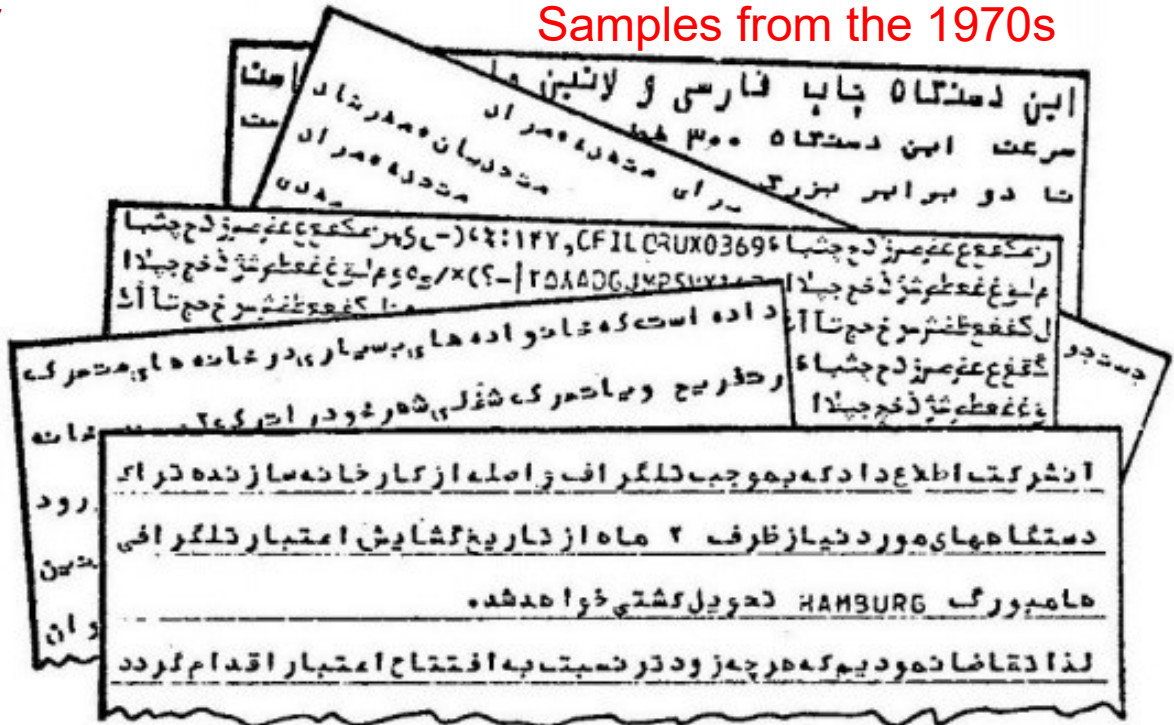
Samples from Early Persian Printers

Smudge problem of hammer-type printers was actually a plus, as it created the illusion of connectivity

However, it wasn't so good in other respects, because small features such as the loop in letters such as قاف and ميم tended to disappear



Samples from the 1970s

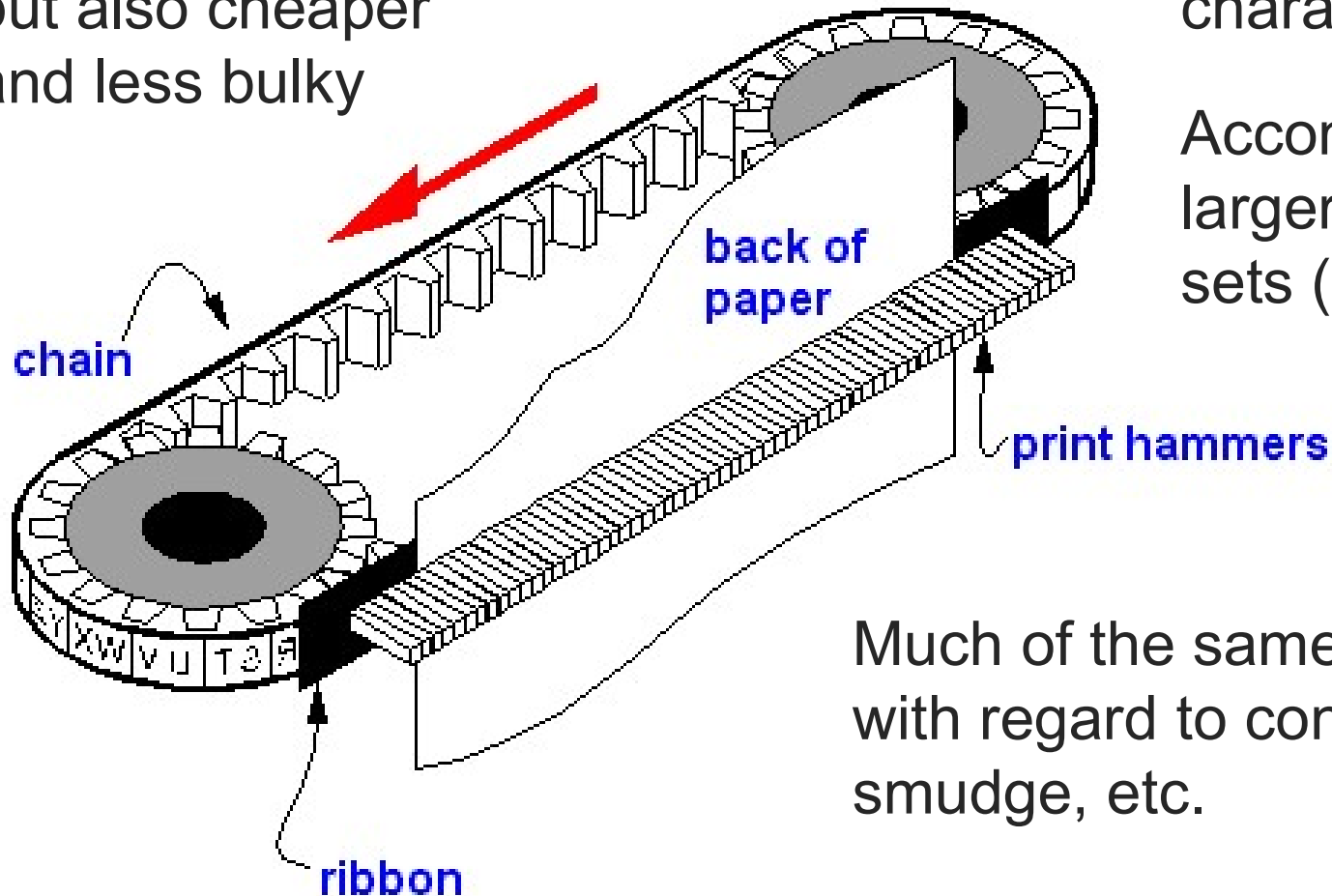


Early Persian Chain Printers

Slower than drum printers,
but also cheaper
and less bulky

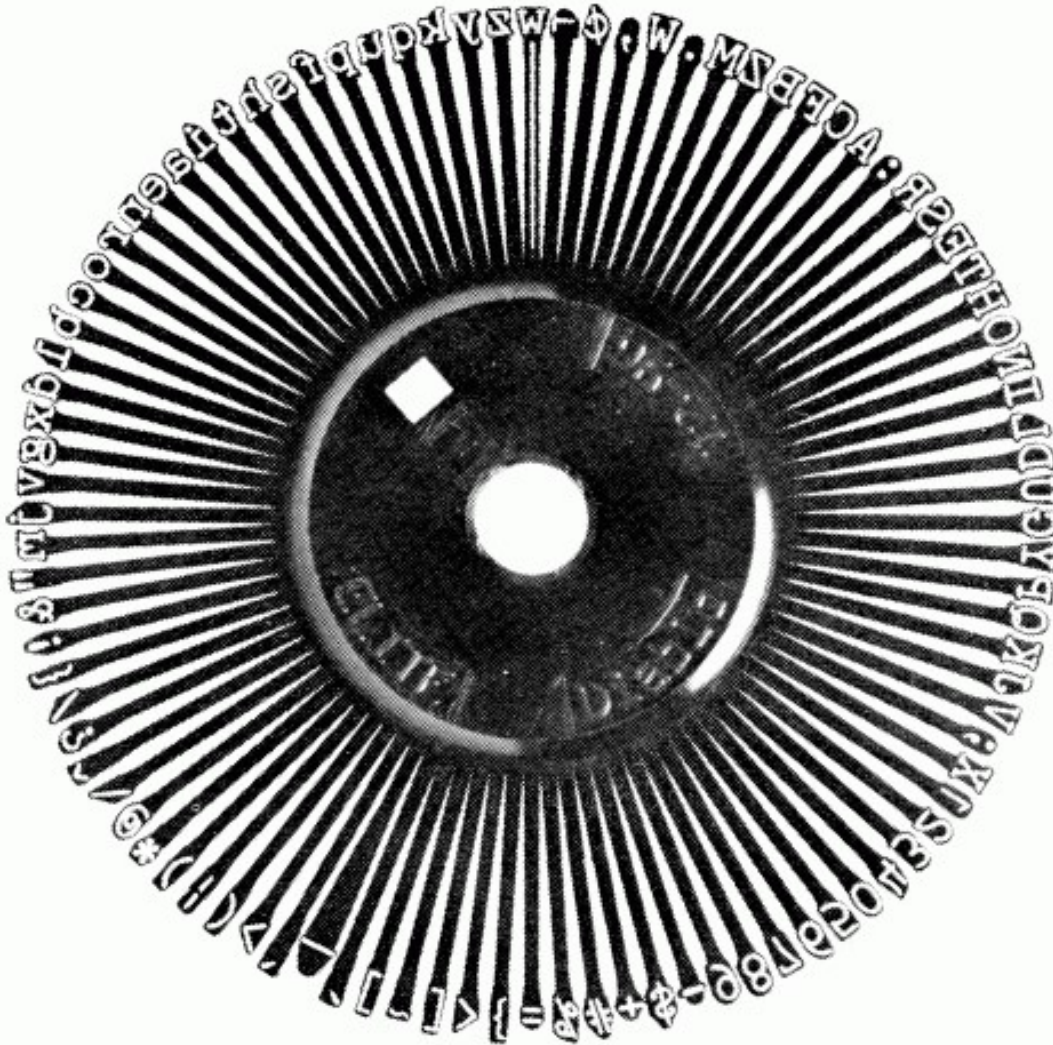
Fixed-width
characters

Accommodates
larger character
sets (bilingual?)



Much of the same problems
with regard to connectivity,
smudge, etc.

Early Persian Daisy-Wheel Printers

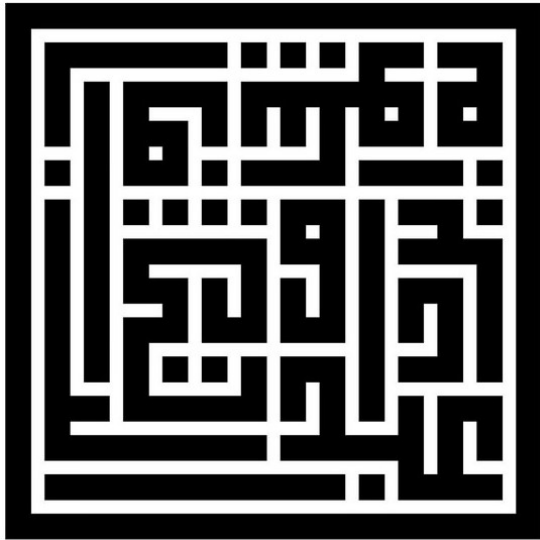


Even slower, but also cheaper and more compact

Fixed-width characters

Does not support multilingual printing or multiple font sizes

Dot-Matrix Printers to the Rescue



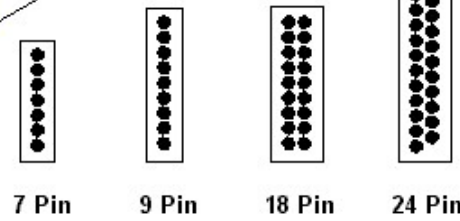
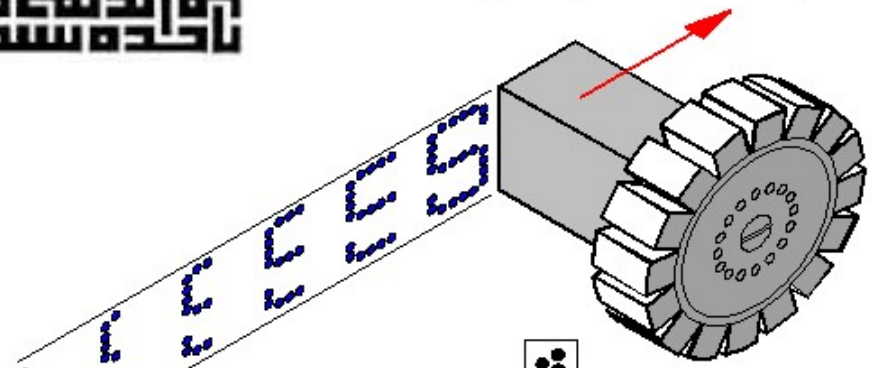
Kufi Script



PIXEL FONT

A B C D E F G H I
 J K L M N O P Q R
 S T U V W X Y Z
 1 2 3 4 5 6 7 8 9 0
 ? ! @ # \$ % & * + = - _ #
 ^ / \ | . : ; ' () \$

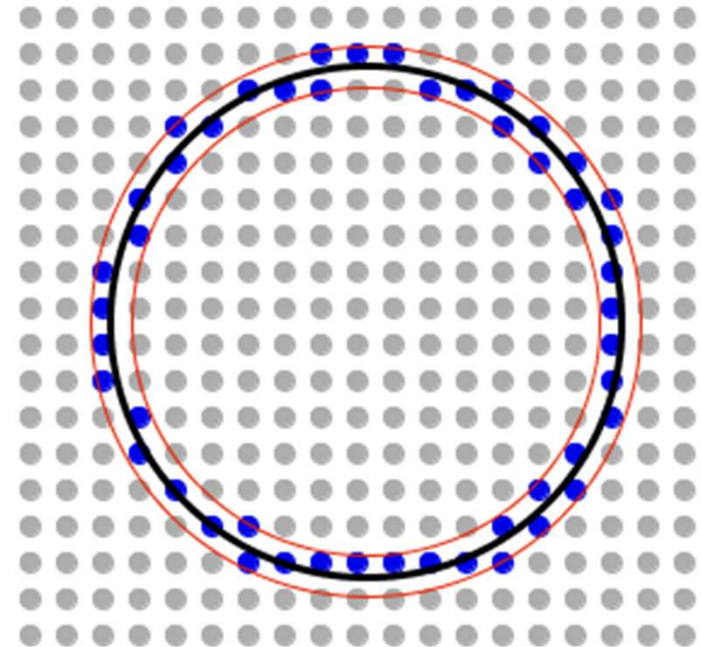
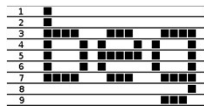
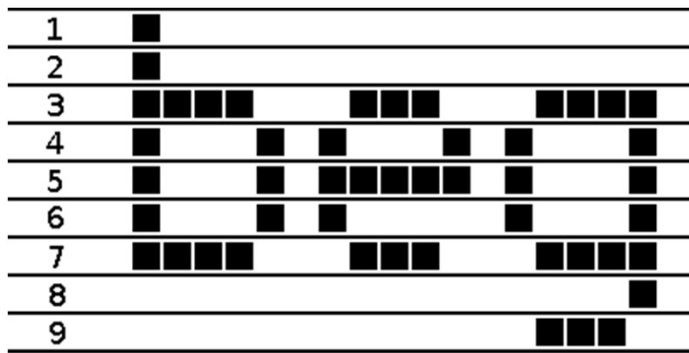
ystem where a
 ld allow us t
 mercial supplier.



More Flexible, But Not a Cure-All

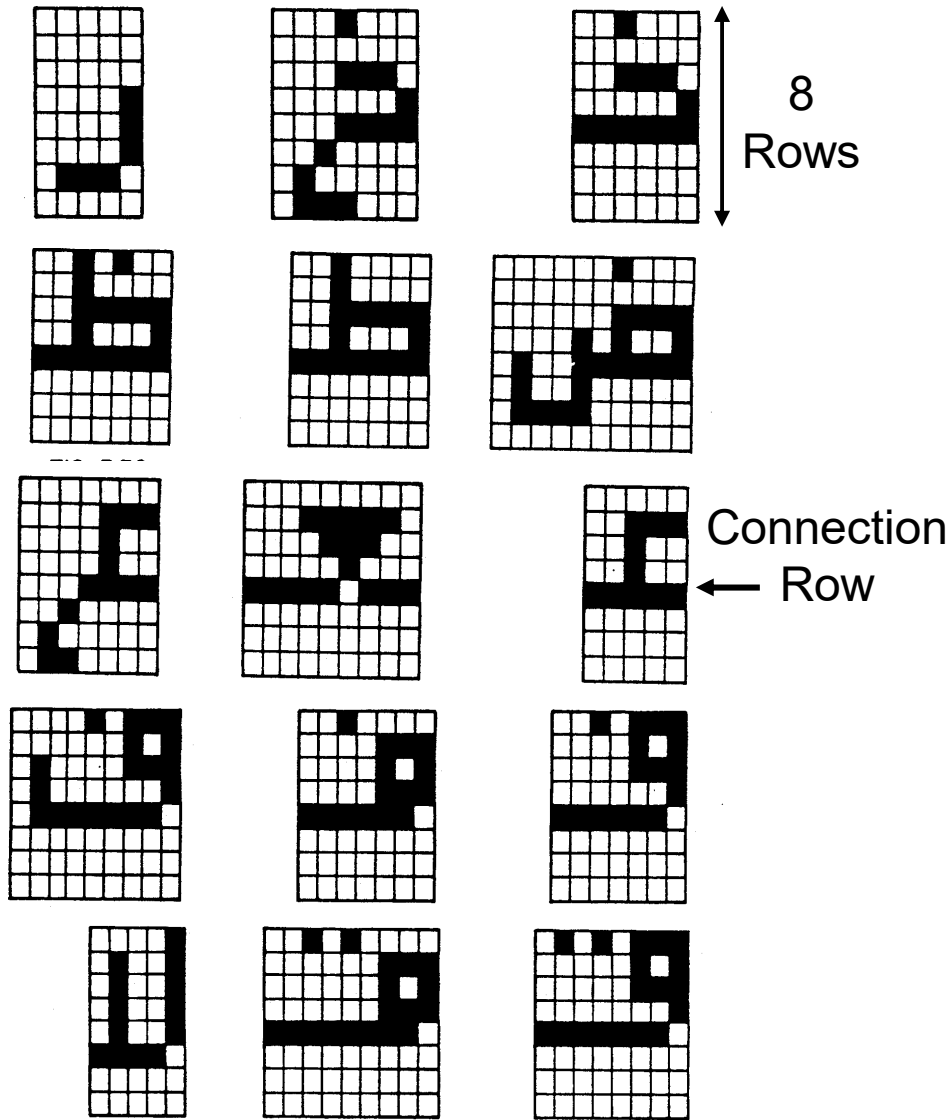


The same print mechanism usable for any language (software fonts) or for rendering graphic images



Pixel or digital geometry

Dot-Matrix Font Design



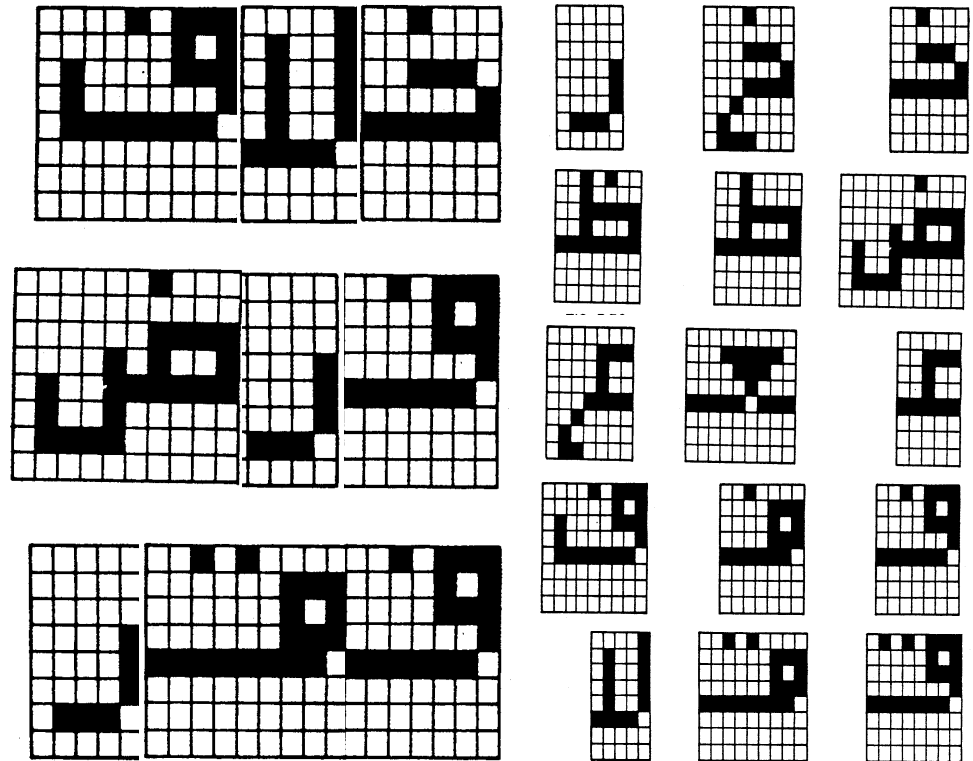
ABCDEFGHIJKLMNOP
 QRSTUVWXYZÀÁÂËÜÖ
 abcdefghijklmnopqr
 stuvwxyzÀÁÂËÜÖ&12
 34567890(\$£€.,!?)

 abcdefghijklm
 nopqrstuvwxyz
 abcdefghijklm
 nopqrstuvwxyz

Persian Script Needs Larger Dot-Matrix

For a given dot-matrix size, legibility and pleasantness of the script is up to the font designer's skill and artistry

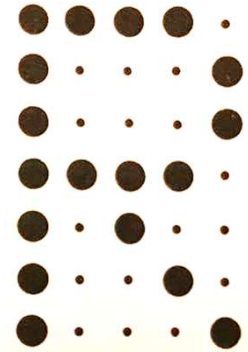
Characters must be viewed not in isolation, but in various combinations (with each other and with common punctuations)



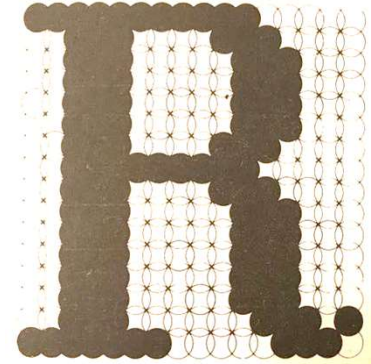
Large Matrix + Small Dots = High Quality

Overlapped dots also help

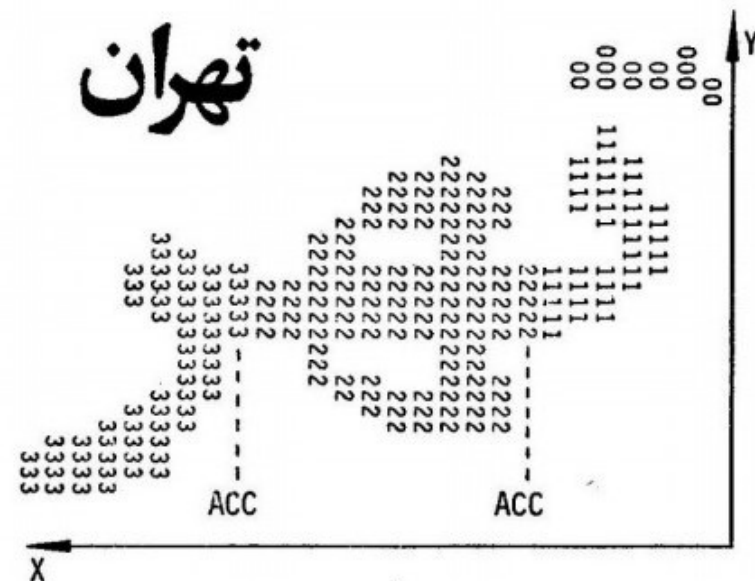
Variable widths is helpful



7 × 5 matrix,
disjoint dots



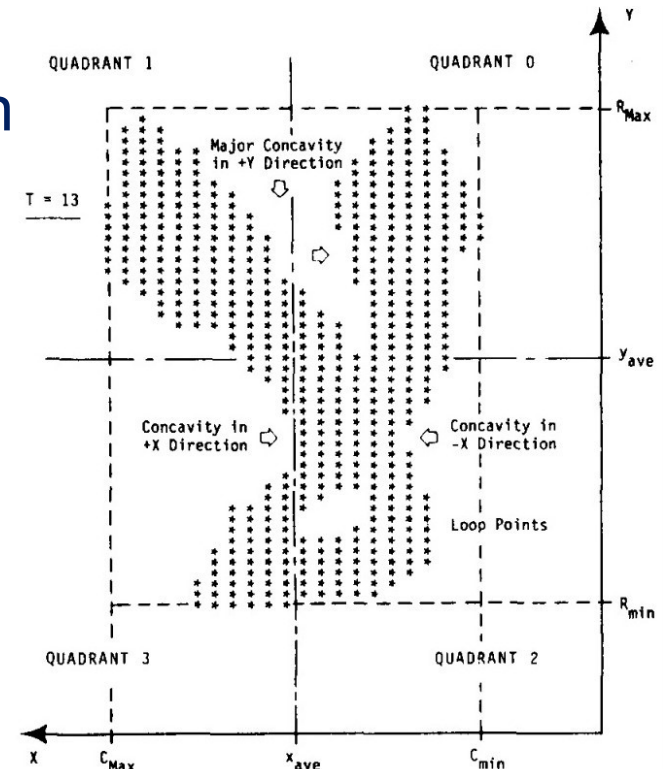
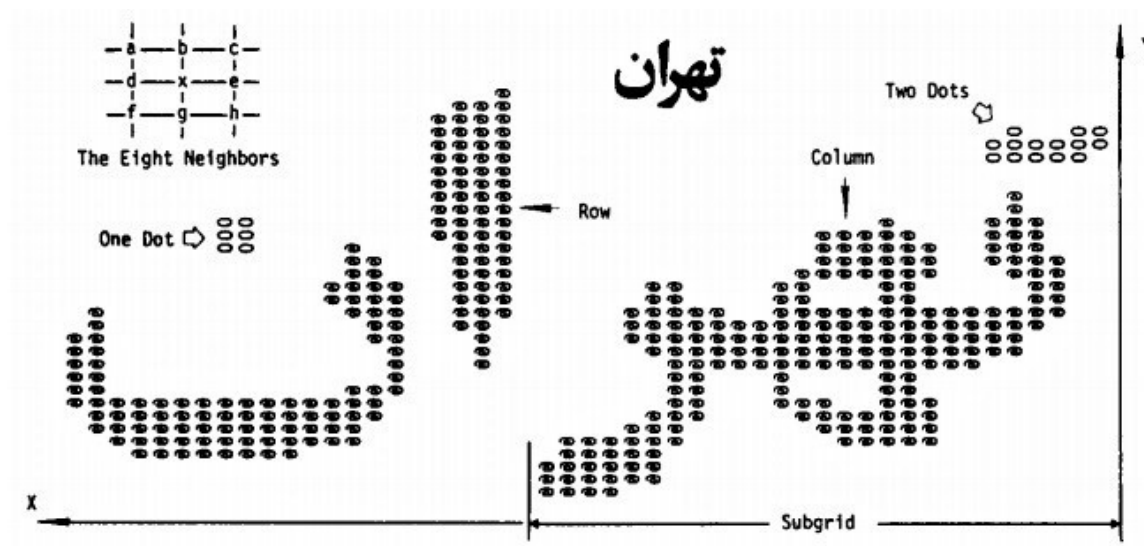
16 × 16 matrix,
overlapped dots



Persian Print Script Recognition

Examples of Features Used for Recognition

- Distribution of “weight” in the four quadrants
- Concavities (major, minor, in various directions)
- “Islands” (e.g., dots) and “lakes” (loops)
- Connectivity to the right or left symbol
- Row/column segment-length distribution



Lower Bound for Dot-Matrix Size

$$7 \times 5$$

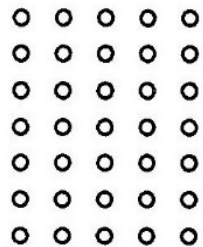
این جمله و ترکیبات زیر
نمایند: ی نوع خط حاصلند:
ابجد هوز طی کلیم سعفص
قرشت شخض طغخ آنهم شش
چنگ حاذق دلائل عطش گنج
 $11293 + 7527 = 13020 \bar{A}$

Departures		13:52	
FLIGHT	DESTINATION	BOARDING TIME	GATE STATUS
NZ119	SYDNEY	12:20	07 DEPARTED
FJ410	NADI	13:10	01 GATE CLOSED
QF190	SYDNEY	13:10	04 GATE CLOSED
NZ175	PERTH	13:25	10 FINAL CALL
MH130	KUALA LUMPUR	13:30	03 FINAL CALL
CX100	HONG KONG	13:40	08 BOARDING
QF126	BRISBANE	13:50	07 BOARDING IN 5 MINS
NZ725	MELBOURNE	14:50	BOARDING IN 60 MIN
NZ739	BRISBANE	14:50	BOARDING IN 60 MIN
TG390	BANGKOK	14:50	BOARDING IN 60 MIN
NZ105	SYDNEY	16:05	PLEASE WAIT
	MELBOURNE	15:45	PLEASE WAIT

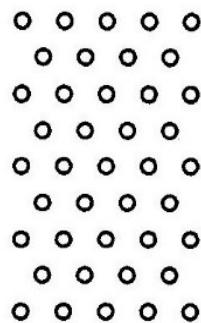
$$7 \times 9/2$$

این جمله و ترکیبات زیر
نمایند: ی نوع خط حاصلند:
ابجد هوز طی کلیم سعفص
قرشت شخض طغخ آنهم شش
چنگ حاذق دلائل عطش گنج
 $11243 + 7527 = 13048 \bar{A}$

$$7 \times 5$$



$$9 \times 9/2$$



ملاحظه کاری، طبخ غذا، خلاصی، گل آلاله

ملاحظه کاری، طبخ غذا، خلاصی، گل آلاله

ملاحظه کاری، طبخ غذا، خلاصی، گل آلاله

ملاحظه کاری، طبخ غذا، خلاصی، گل آلاله

Tall letters with 7×5 , $7 \times 9/2$,
 $9 \times 9/2$, and $9 \times w$ matrices

$$9 \times 9/2$$

این جمله و ترکیبات زیر
نمایند: ی نوع خط حاصلند:
ابجد هوز طی کلیم سعفص
قرشت شخض طغخ آنهم شش
چنگ حاذق دلائل عطش گنج
 $11293 + 7527 = 13020 \bar{A}$

Improved Script with Dot-Matrix Printers

Samples from the 1980s

Variable widths
is helpful

Overlapped dots
also help

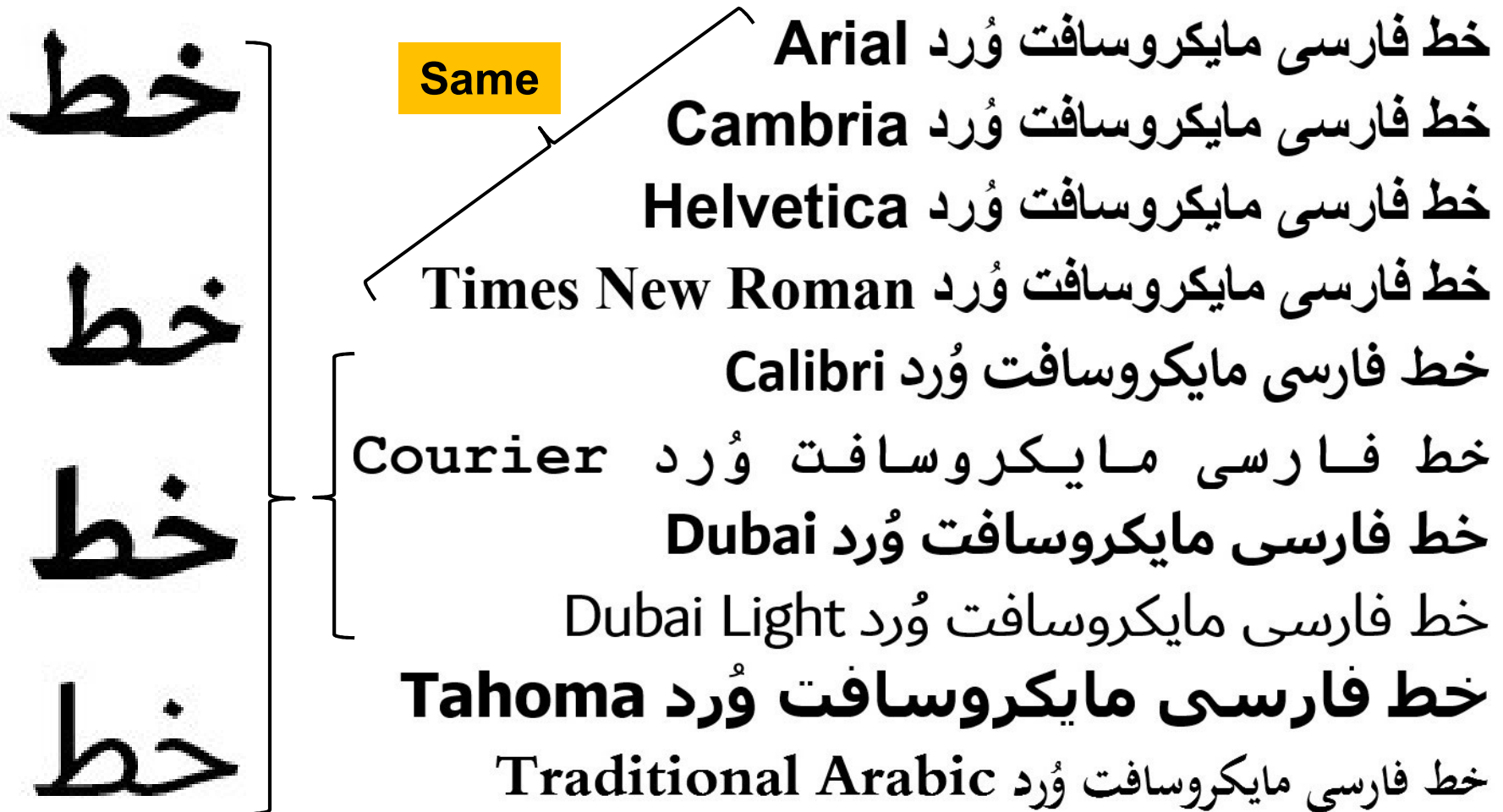
Highly flexible,
software-defined
font

Multiple fonts &
font weights,
with no added
hardware cost



Microsoft Word Persian Fonts

All fonts 32, converted to JPG from MS Word doc



Scalability of Persian Fonts and Script

The best response to fools is silence **Arial 28**

جواب ابلهان خاموشی است **Arial 28**

جواب ابلهان خاموشی است **32**

جواب ابلهان خاموشی است **36**

جواب ابلهان خاموشی است **40**

The best response to fools is silence **Arial 14**

جواب ابلهان خاموشی است **Arial 14**

جواب ابلهان خاموشی است **16**

جواب ابلهان خاموشی است **18**

جواب ابلهان خاموشی است **20**

The best response to fools is silence **Arial 14**

جواب ابلهان خاموشی است **Arial 9**

جواب ابلهان خاموشی است **10**

جواب ابلهان خاموشی است **11**

جواب ابلهان خاموشی است **12**

**Font-size
adjustment**

The best response to fools is silence **Arial 28**

جواب ابلهان خاموشی است **Arial 28**

جواب ابلهان خاموشی است **32**

جواب ابلهان خاموشی است **36**

جواب ابلهان خاموشی است **40**

The best response to fools is silence **Arial 28**

جواب ابلهان خاموشی است **Arial 28**

جواب ابلهان خاموشی است **32**

جواب ابلهان خاموشی است **36**

جواب ابلهان خاموشی است **40**

**Image
resizing**

Modern Persian Display Typefaces



Behrooz Parhami

1 hr · 🌐 ▼

کار پاکان را قیاس از خود مگیر
 گر چه باشد در نوشتن شیر، شیر
 آن یکی شیر است اندر بادیه
 آن دگر شیر است اندر بادیه
 آن یکی شیر است کادم می خورد
 و آن دگر شیر است کادم می خورد

کار پاکان را قیاس از خود مگیر
 آن یکی شیر است اندر بادیه
 و آن دگر شیر است کادم می خورد

کار پاکان را قیاس از خود مگیر
 آن یکی شیر است اندر بادیه
 و آن دگر شیر است کادم می خورد

**Computer Persian
 script samples**

کار پاکان را قیاس از خود مگیر
 آن یکی شیر است اندر بادیه
 و آن دگر شیر است کادم می خورد

کار پاکان را قیاس از خود مگیر
 آن یکی شیر است اندر بادیه
 و آن دگر شیر است کادم می خورد

**Which one is
 most legible?**

کار پاکان را قیاس از خود مگیر
 آن یکی شیر است اندر بادیه
 و آن دگر شیر است کادم می خورد

کار پاکان را قیاس از خود مگیر
 آن یکی شیر است اندر بادیه
 و آن دگر شیر است کادم می خورد



Behrooz Parhami

September 10, 2011 · 🌐 ▼

Selected verses from a poem by Hooshang Ebtehaj

عمری ست تا از جان و دل ، ای جان و دل می خوانمت
 تو نیز خواهان منی ، می دانمت ، می دانمت
 ای خنده ی نیلوفر ی ، در گریه ام می آوری
 بر گریه می خندی و من ، در گریه می خندانمت
 ای زاده ی پندار من ، پوشیده از دینار من
 چو کودک نادانسته ، گپواره می جندانمت
 ای من تو بی من کیستی ، چون سایه بی من نیستی
 همراه من می ایستی ، همپای خود می رانمت

Samples from Facebook and other Internet sites

دانش و فن آوری

حذف نرم افزارهای ایرانی از فروشگاه مجازی گوگل

- 📄 ویدئو: سلول های بنیادی
- 📄 محدودیت اینترنت: آزادی های آنلاین پس از انتخابات ۹۶
- 📄 چهل سال پس از وویجرها: دو نامه از یک تمدن
- 📄 محمدجواد آذری جهرمی و همزیستی شغلی با فیلترینگ و شنود

گوگل نیز پس از اپل
 نرم افزارهای ایرانی را از
 فروشگاه مجازی خود حذف
 کرد. دلیل آن تبعیت از قانون
 تحریم های ۲۲ ساله آمریکا
 علیه ایران است.



Aug. 2019



Evolutionary Changes in Persian & Arabic Scripts



Slide 32

Persian Script on Large & Small Displays



Jam-e Jam Online, news site of the government-run IRI Broadcasting:
Laptop screen image from July 16, 2019

BBC Persian news site on iPhone 7 screen:
Captured on July 16, 2019



Digikala e-commerce site on iPhone 7 screen:
Captured on July 16, 2019



Where We Are, Where We Go from Here

Technology to the rescue

- Improved resolution/contrast in displays
- HD (1920 x 1080) and UHD display units
- Obsolescence of line-segment displays
- High-res printers (600 dpi or more)
- Accumulated know-how about font design



Path to further improvements

- Trade-offs between legibility and aesthetics
- Learning from large data sets
- Crowd-sourcing the evaluation of aesthetics
- Formatting, especially in bilingual output
- Expect more to get more!



Questions or Comments?

parhami@ece.ucsb.edu

<http://www.ece.ucsb.edu/~parhami/>



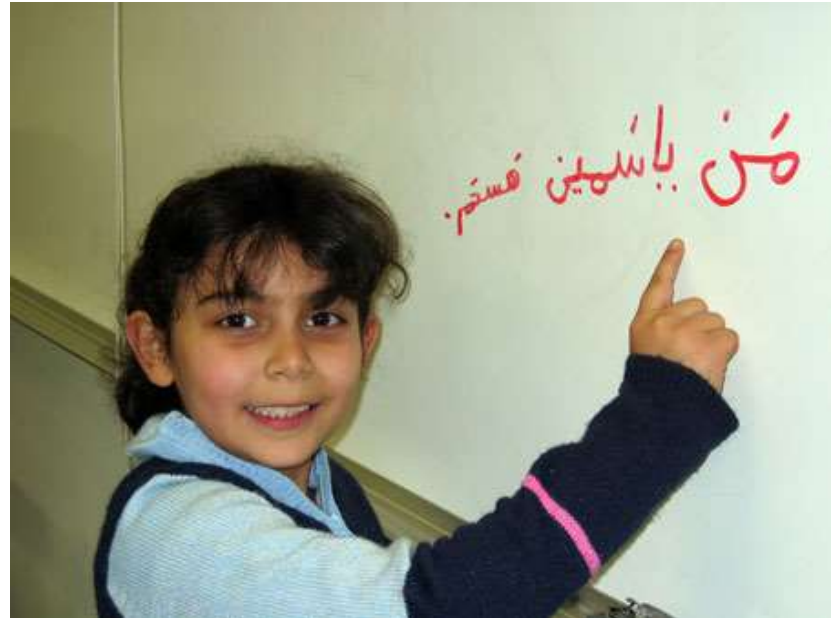
Persian Handwriting

Examples of beautiful handwriting

پیشتر
 به عشق سینه از دیده همبسته ای عشق
 کفایت از زبان دران نیز جور و عادت به دلایر از نوزاد و خرد گشته شد، بخت قدر از دست استار گاه و با تجربه از رخ گشته و در یکجهت و در دماغه شریک
 بر طاعت از چشم سیر که همه بر آرزو داشته باشم و در دهن و در دهان و در کف دست و در کف پا و در کف کمر و در کف کمر و در کف کمر و در کف کمر
 مردم با او نیز جوان نه تنها عارفان از او سپاس می گفتند بلکه شیوه آنها خطی نازیب و دلنواز و در نوشته گشته که از زبان طاعت است
 و در جبهه از خوشتر از آن است که با کمال سیرت با بهر بنامه و با عارفان و با خطایه در پیشه و در کف دست و در کف پا و در کف کمر و در کف کمر
 مستور صرف نان بر دست، کجف عارفان که کف دست و در کف پا و در کف کمر و در کف کمر و در کف کمر و در کف کمر و در کف کمر
 گشته و در سیر از این گشته و در کف دست و در کف پا و در کف کمر و در کف کمر و در کف کمر و در کف کمر و در کف کمر
 بلکه کف دست و در کف پا و در کف کمر و در کف کمر و در کف کمر و در کف کمر و در کف کمر و در کف کمر
 باشد و در کف دست و در کف پا و در کف کمر و در کف کمر و در کف کمر و در کف کمر و در کف کمر
 با تو خیر که کاش در کف دست و در کف پا و در کف کمر و در کف کمر و در کف کمر و در کف کمر و در کف کمر
 از قدر هر که کف دست و در کف پا و در کف کمر و در کف کمر و در کف کمر و در کف کمر و در کف کمر
 رضاشاهی کف دست و در کف پا و در کف کمر و در کف کمر و در کف کمر و در کف کمر و در کف کمر

عشق
 به سینه از دیده همبسته ای عشق
 ۱۳۶۱

A schoolchild's handwriting



Ali-Akbar Dekhoda's handwriting

سینه از دیده همبسته ای عشق
 به سینه از دیده همبسته ای عشق

Computer-Generated Nasta'liq

<http://www.NastaliqOnline.ir>



سریع و آسان
بدون دردسر

نستعلیق آنلاین

کامپیوتر و مشکلات نگارش فارسی
کامپیوتر و مشکلات نگارش فارسی

On-Line Access to My Pertinent Publications

Three-digit numbers in square brackets refer to entries in B. Parhami's Publications list:

<http://www.ece.ucsb.edu/~parhami/publications.htm>

- [111] B. Parhami, "On Lower Bounds for the Dimensions of Dot-Matrix Characters to Represent Farsi and Arabic Scripts," *Proc. 1st Annual CSI Computer Conf.*, Tehran, Iran, December 1995, pp. 125-130.
- [045] B. Parhami, "Standard Farsi Information Interchange Code and Keyboard Layout: A Unified Proposal," *J. Institution of Electrical and Telecommunications Engineers*, Vol. 30, pp. 179-183, 1984.
- [041] B. Parhami and M. Taraghi, "Automatic Recognition of Printed Farsi Texts," *Pattern Recognition*, Vol. 14, Nos. 1-6, pp. 395-403, 1981.
- [038] B. Parhami, "Language-Dependent Considerations for Computer Applications in Farsi and Arabic Speaking Countries," *System Approach for Development (Proc. IFAC Conf.)*, 1981, pp. 507-513.
- [037] B. Parhami and M. Taraghi, "Automatic Recognition of Printed Farsi Texts" (summary), *Proc. Conf. Pattern Recognition*, Oxford, England, January 1980.
- [033] B. Parhami, "On the Use of Farsi and Arabic Languages in Computer-Based Information Systems," *Proc. Symp. Linguistic Implications of Computer-Based Information Systems*, November 1978.
- [032] B. Parhami, "Optically Weighted Dot-Matrix Farsi and Arabic Numerals," *Information Technology 78 (Proc. 3rd Jerusalem Conf.)*, August 1978, pp. 207-210.
- [023] B. Parhami and F. Mavaddat, "Computers and the Farsi Language: A Survey of Problem Areas," *Information Processing 77 (Proc. IFIP World Congress)*, 1977, pp. 673-676.
- [. . .] Parhami, B., "Impact of Farsi Language on Computing in Iran," *Mideast Computer*, Vol. 1, pp. 6-7, September 1978. [Not part of my academic publications; link to be supplied here]

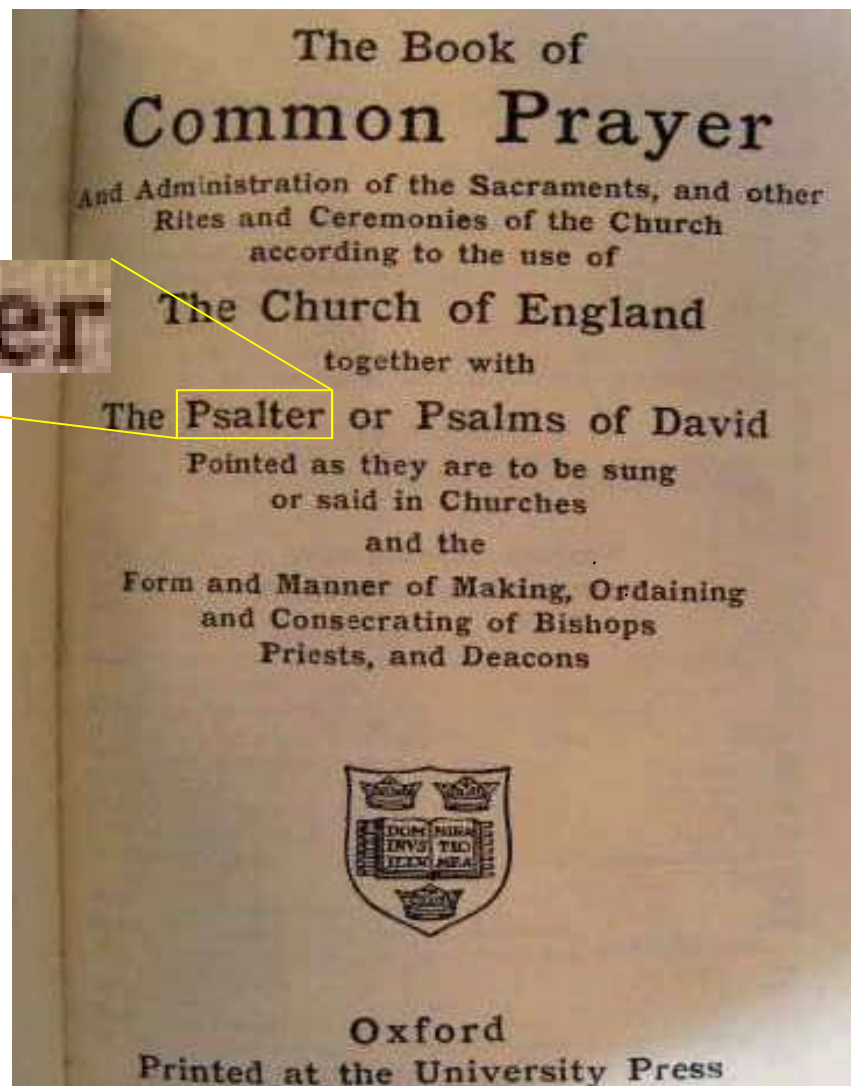
Early Computer Typesetting Challenges

Manually typeset text had:

- Alignment problems
- Spacing irregularities
- Broken fonts
- Ink imperfections
- Nonuniform shapes

Early computer-typeset
Texts were disliked because
they were just too perfect

So, a little randomness was
introduced in font shapes
and alignment



Persian vs. Arabic Printing: Problems are Quite Similar

Arial Bold:

ليبيا: الامم المتحدة تدعو لوقف
الصراع تخفيفا لمعاناة الناس

Arabic Transparent Bold:

ليبيا: الامم المتحدة تدعو لوقف
الصراع تخفيفا لمعاناة الناس

Tahoma Bold:

ليبيا: الامم المتحدة تدعو لوقف
الصراع تخفيفا لمعاناة الناس

Times New Roman Bold:

ليبيا: الامم المتحدة تدعو لوقف
الصراع تخفيفا لمعاناة الناس

Simplified Arabic Bold:

ليبيا: الامم المتحدة تدعو لوقف
الصراع تخفيفا لمعاناة الناس

العربية فارسی

بلادیدن balā-dīdan [ع. - ف.]
(مصل. ٠) ١ - رنج دیدن. ٢ - بمصیبت
دچار شدن.

بلادیده b.-dīda(-e) (ص.مف.) ١ -
رنجیدیده. ٢ - دچار مصیبت شده.

بلاذر balāzor [= بلادر = بلادور]
(١.) (گیا.) ١ - بلادر (م.ه.)

بلار balār [= بلال] (١.) (گیا. ٠)
آذر بویه (م.ه.)، اشنان (م.ه.)

بلار ballār [= بلور] (١.) ← بلور.
بلارج balāraj (١.) (جان.) لکک

← ٢ - لکک

الله ل + ا = لا
نہا نہا نہا

Some Properties of the Persian Script: Greater Variety in Letter Shapes and Sizes

The Latin alphabet

a b c d e f g h i j k l m n o p q r s t u v w x y z + / # % \$ & (
A B C D E F G H I J K L M N O P Q R S T U V W X Y Z

The quick brown fox jumped over the fence.

Arial 24

The quick brown fox jumped over the fence.

Times N R 24

The quick brown fox jumped over the fence.

Courier 20

The Persian alphabet



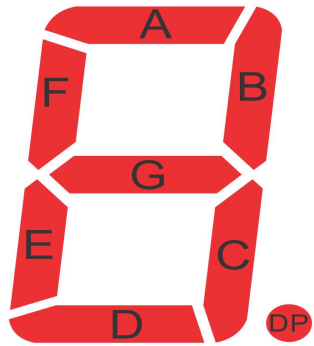
Fixed-width characters

گر چه باشد در نوشتن شیر، شیر	کارِ پاکان را قیاس از خود مگیر
آن دگر شیر است اندر بادیه	آن یکی شیر است اندر بادیه
و آن دگر شیر است کادم می‌خورد	آن یکی شیر است کادم می‌خورد

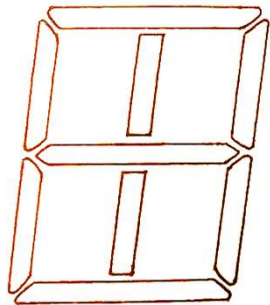
Variable-width characters

گر چه باشد در نوشتن شیر، شیر	کارِ پاکان را قیاس از خود مگیر
آن دگر شیر است اندر بادیه	آن یکی شیر است اندر بادیه
و آن دگر شیر است کادم می‌خورد	آن یکی شیر است کادم می‌خورد

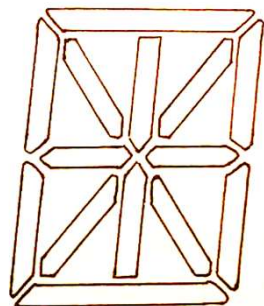
Line-Segment Displays: Latin Script



7-segment display for calculators and other numerical apps



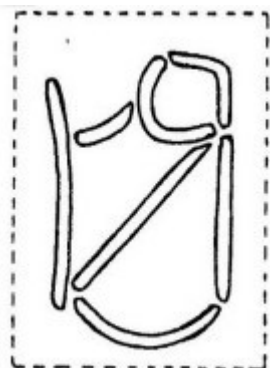
Two kinds of line-segment alphanumeric displays



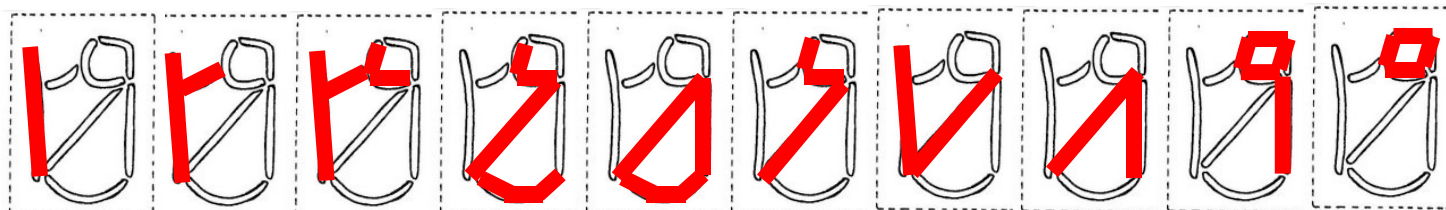
Okay for uppercase letters, but problematic for all else



Line-Segment Displays: Persian Script

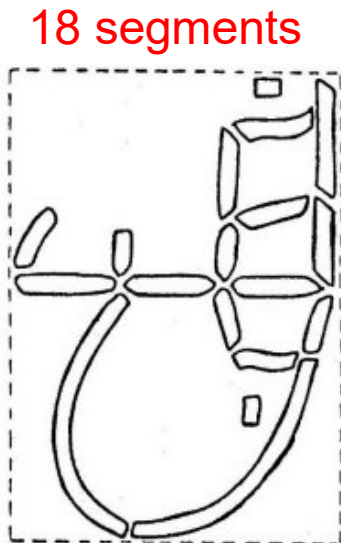


7 segments

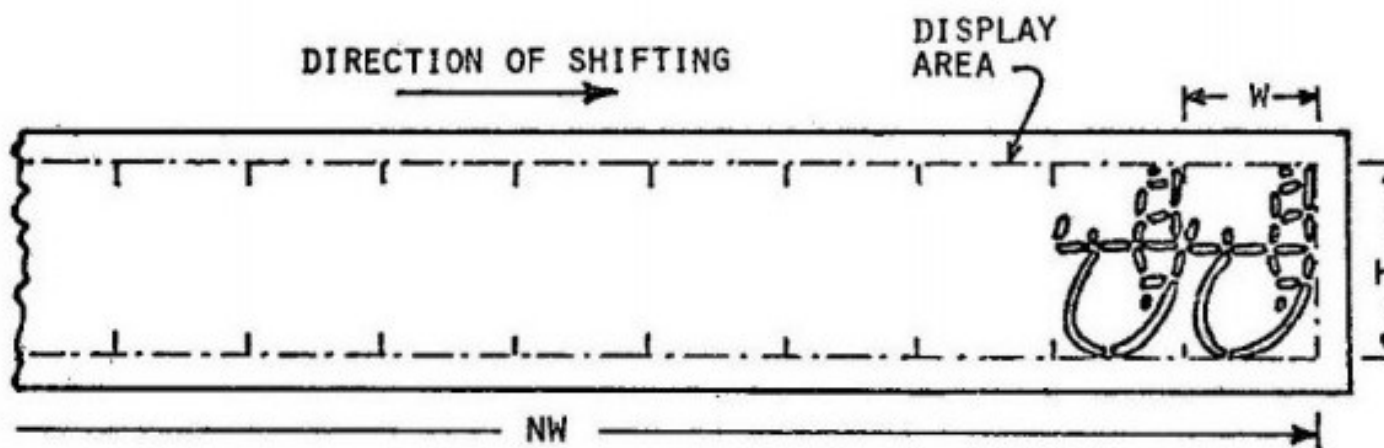


Seven-segment display for Persian digits feasible, but ...

**At least 18 segments are needed for displaying letters
Even then, alphanumeric symbols cannot be combined
Special characters also present problems**



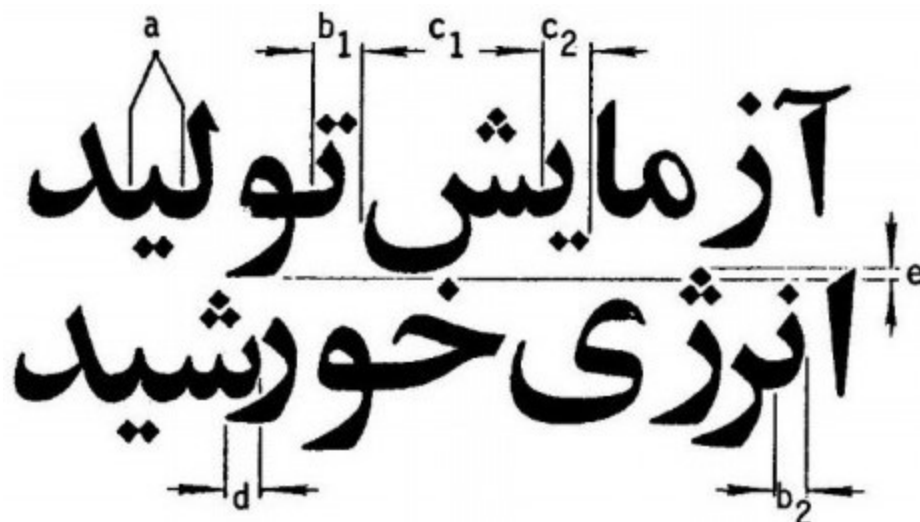
18 segments



Challenges in Persian Print Recognition

Technical Challenges

- Connectivity of letters
- Wide variations in letter widths/heights
- Context-dependent letter shapes
- Presence of 1-3 dots
- Small loops and other fine features
- Absence of short vowels (usually)



Five-Step Algorithm

- Decompose into lines
- Identify connection axis
- Decompose into letters
- Evaluate features
- Match extracted features