

# المواصفات والمقاييس

حسني المحتسب

بِسْمِ اللّٰهِ الرَّحْمٰنِ الرَّحِیْمِ

## الحواسب العربية (تعريف الحاسبات) - 8

حسني عبدالغني المحتسب

تَقْبِلُ اللهُ مِنَّا وَمِنْكُمْ  
وَكُلُّ عَامٍ وَأَنْتُمْ بِخَيْرٍ

# المواصفات والمقاييس

# المسار

- تعريف المواصفة وأهميتها
- تقييم صفحة المحارف أو الشفرة
- تقييم لوحة المفاتيح
- أشكال الحروف حسب السياق

# المواصفة

- تعريف: أفضل حل يلبي متطلبات المعنيين
  - الصانعين
  - المستخدمين
  - الحكومات
- المطورون: خبراء في مجال تخصص المواصفة
- المستويات:
  - وطني
  - اقليمي
  - دولي

# المواصفة

- أهمية المواصفة
  - تضمن حدا أدنى للجودة
  - تحمي المصلحة الوطنية
  - التوافق بين الصانعين
- من المنظمات الإقليمية للتقييس
  - اللجنة الأوروبية للتقييس (ECS)
  - المنظمة الإفريقية للتقييس (ARSO)
  - المنظمة العربية للمواصفات والمقاييس (ASMO)
- عالميا
  - المنظمة الدولية للتقييس International Organization for Standardization (ISO) 1946

# المقاييس العربية في مجال الحاسب

- صفحة المحارف أو الشفرة
- لوحة المفاتيح
- أشكال الحروف حسب السياق
- تحويل الكتابة العربية إلى كتابة صوتية



# جهود التقّيس

## • أهميته

• تسهيل تبادل المعلومات

• كسر الاحتكار

• سهولة الاستخدام، وسرعة نقل التقنية

## • الهموم المختلفة

• الطباعون: تبسيط أشكال الحروف وتسهيل صفها

• الحاسوبيون: شفافية البرامج

• اللغويون: الكفاية اللغوية، تغطية الشفرة لجميع الحروف

• المنتجون والموردون: حصر التعديلات والجهد المبذول من طرفهم

# الصعوبات التي تواجه التقييس

- ضرورة اعتبار ما هو مستقر ومقبول، مع مراعاة ما سيستجد والسرعة في ذلك
- تلبية احتياجات المستخدمين والمنتجين في مجالات وأماكن شتى
- ضرورة التوفيق بين الاعتبارات التقنية والتسويقية
- غياب منظمات التقييس العربية

# المعايير الواجب توفرها في الشفرة

- الكفاية
  - تغطية جميع الحروف اللازمة لتبادل المعلومات
- الأداء
  - شفرة لكل حرف
  - تقليل عدد الشفرات
- طول الشفرة يساوي عدد أحرف الكلمة
- ترتيب الشفرات يجب أن يوافق ترتيب الحروف
- يفضل أن تكون الشفرات متتالية

# تاريخ الشفرات العربية

- 1977 الشفرة المغاربية
  - CUDAR-U
  - طورت في المغرب
  - معهد الدراسات للأبحاث والتعريب في الرباط
  - 1980 أقرت في المملكة
- 1981 لجنة الحرف العربي في المعلوماتية
  - المغرب، تونس، الأردن، السعودية، العراق، سوريا
  - الشفرات، لوحات المفاتيح، أشكال الحروف
- 1982 الشفرة القياسية الأولى ASMO 449
  - The Arab Standardization and Metrology Organization
  - أقرتها المنظمة العربية للمواصفات والمقاييس عام 1982م
  - أصبحت دولية عام ISO 9036: 1986

# الشفرة القياسية ٤٤٩

- ASMO 449
- تستخدم سبع خانات 7-bit
- رمز لكل حرف وليس لكل شكل من أشكال الحروف
- ترمز ٤٤ حرفا
- اسقطت رموز لأ، لإ، لا، لآ لأن كل منها حرفين
- وحدت شفرات الرموز المشتركة مثل <math>\langle \rangle</math> “ ()
- سمحت لعدة أشكال للهمزة (ء، أ، إ، ء، ئ، آ) وموقعها مع الألف
- شكلين للتاء (ت، ة) – فرق بين التاء المربوطة والتاء

## ISO 9036

7-bit Coded Arabic Character Set for Information Interchange  
(a.k.a. ECMA?, ASMO 449)

	0	1	2	3	4	5	6	7
0	NUL <sub>0</sub>	DLE <sub>16</sub>	ASP <sub>32</sub>	• <sub>48</sub>	@ <sub>64</sub>	ن <sub>80</sub>	- <sub>96</sub>	، / 112
1	SOH <sub>1</sub>	DC1 <sub>17</sub>	! <sub>33</sub>	١ <sub>49</sub>	ء <sub>65</sub>	و <sub>81</sub>	ف <sub>97</sub>	، - 113
2	STX <sub>2</sub>	DC2 <sub>18</sub>	" <sub>34</sub>	٢ <sub>50</sub>	آ <sub>66</sub>	ز <sub>82</sub>	ق <sub>98</sub>	° - 114
3	ETX <sub>3</sub>	DC3 <sub>19</sub>	# <sub>35</sub>	٣ <sub>51</sub>	أ <sub>67</sub>	س <sub>83</sub>	ك <sub>99</sub>	115
4	EOT <sub>4</sub>	DC4 <sub>20</sub>	£ <sub>36</sub>	٤ <sub>52</sub>	ؤ <sub>68</sub>	ش <sub>84</sub>	ل <sub>100</sub>	116
5	EQ <sub>5</sub>	NAK <sub>21</sub>	‰ <sub>37</sub>	٥ <sub>53</sub>	إ <sub>69</sub>	ص <sub>85</sub>	م <sub>101</sub>	117
6	ACK <sub>6</sub>	SYN <sub>22</sub>	& <sub>38</sub>	٦ <sub>54</sub>	ى <sub>70</sub>	ض <sub>86</sub>	ن <sub>102</sub>	118
7	BEL <sub>7</sub>	ETB <sub>23</sub>	' <sub>39</sub>	٧ <sub>55</sub>	ا <sub>71</sub>	ط <sub>87</sub>	ه <sub>103</sub>	119
8	BS <sub>8</sub>	CAN <sub>24</sub>	) <sub>40</sub>	٨ <sub>56</sub>	ب <sub>72</sub>	ظ <sub>88</sub>	و <sub>104</sub>	120
9	HT <sub>9</sub>	EM <sub>25</sub>	( <sub>41</sub>	٩ <sub>57</sub>	ة <sub>73</sub>	ع <sub>89</sub>	ى <sub>105</sub>	121
A	LF <sub>10</sub>	SUB <sub>26</sub>	* <sub>42</sub>	: <sub>58</sub>	ت <sub>74</sub>	غ <sub>90</sub>	ي <sub>106</sub>	122
B	VT <sub>11</sub>	ESC <sub>27</sub>	+ <sub>43</sub>	; <sub>59</sub>	ث <sub>75</sub>	] 91	- <sub>107</sub>	{ 123
C	FF <sub>12</sub>	FS <sub>28</sub>	• <sub>44</sub>	> <sub>60</sub>	ج <sub>76</sub>	\ 92	- <sub>108</sub>	124
D	CR <sub>13</sub>	GS <sub>29</sub>	- <sub>45</sub>	= <sub>61</sub>	ح <sub>77</sub>	[ 93	- <sub>109</sub>	} 125
E	SO <sub>14</sub>	RS <sub>30</sub>	• <sub>46</sub>	< <sub>62</sub>	خ <sub>78</sub>	^ 94	- <sub>110</sub>	- 126
F	SI <sub>15</sub>	US <sub>31</sub>	/ <sub>47</sub>	§ <sub>63</sub>	د <sub>79</sub>	- 95	- <sub>111</sub>	DEL <sub>127</sub>

## الشفرة القياسية ٧٠٨

- شفرة عربية لاتينية ذات ثمان خانوات
- ASMO 708
- نفس ترتيب ٤٤٩
- ذات شفافية مع ASCII
- أقرت كمواصفة دولية ISO 8859-6
- تستخدم من قبل Macintosh

## ISO 8859-6

Right-Hand Part of Latin/Arabic Alphabet  
(a.k.a. ECMA-114, ASMO-708, ISO-IR-127)

	8	9	A	B	C	D	E	F
0	PAD <sub>128</sub>	DCS <sub>144</sub>	NBSP <sub>160</sub>	• <sub>176</sub>	@ <sub>192</sub>	ن <sub>208</sub>	- <sub>224</sub>	- <sub>240</sub>
1	HOP <sub>129</sub>	PU1 <sub>145</sub>	! <sub>161</sub>	١ <sub>177</sub>	ء <sub>193</sub>	و <sub>209</sub>	ف <sub>225</sub>	- <sub>241</sub>
2	BPH <sub>130</sub>	PU2 <sub>146</sub>	" <sub>162</sub>	٢ <sub>178</sub>	آ <sub>194</sub>	ز <sub>210</sub>	ق <sub>226</sub>	- <sub>242</sub>
3	NBH <sub>131</sub>	STS <sub>147</sub>	# <sub>163</sub>	٣ <sub>179</sub>	أ <sub>195</sub>	س <sub>211</sub>	ك <sub>227</sub>	243
4	IND <sub>132</sub>	CCH <sub>148</sub>	£ <sub>164</sub>	٤ <sub>180</sub>	ؤ <sub>196</sub>	ش <sub>212</sub>	ل <sub>228</sub>	244
5	NEL <sub>133</sub>	MW <sub>149</sub>	‰ <sub>165</sub>	٥ <sub>181</sub>	إ <sub>197</sub>	ص <sub>213</sub>	م <sub>229</sub>	245
6	SSA <sub>134</sub>	SPA <sub>150</sub>	& <sub>166</sub>	٦ <sub>182</sub>	ى <sub>198</sub>	ض <sub>214</sub>	ن <sub>230</sub>	246
7	ESA <sub>135</sub>	EPA <sub>151</sub>	' <sub>167</sub>	٧ <sub>183</sub>	ا <sub>199</sub>	ط <sub>215</sub>	ه <sub>231</sub>	247
8	HTS <sub>136</sub>	SOS <sub>152</sub>	) <sub>168</sub>	٨ <sub>184</sub>	ب <sub>200</sub>	ظ <sub>216</sub>	و <sub>232</sub>	248
9	HTJ <sub>137</sub>	SGCI <sub>153</sub>	( <sub>169</sub>	٩ <sub>185</sub>	ة <sub>201</sub>	ع <sub>217</sub>	ى <sub>233</sub>	249
A	VTS <sub>138</sub>	SCI <sub>154</sub>	* <sub>170</sub>	: <sub>186</sub>	ت <sub>202</sub>	غ <sub>218</sub>	ي <sub>234</sub>	250
B	PLD <sub>139</sub>	CSI <sub>155</sub>	+ <sub>171</sub>	: <sub>187</sub>	ث <sub>203</sub>	] <sub>219</sub>	- <sub>235</sub>	251
C	PLU <sub>140</sub>	ST <sub>156</sub>	• <sub>172</sub>	> <sub>188</sub>	ج <sub>204</sub>	\ <sub>220</sub>	- <sub>236</sub>	252
D	RI <sub>141</sub>	OSC <sub>157</sub>	- <sub>173</sub>	= <sub>189</sub>	ح <sub>205</sub>	[ <sub>221</sub>	- <sub>237</sub>	253
E	SS2 <sub>142</sub>	PM <sub>158</sub>	• <sub>174</sub>	< <sub>190</sub>	خ <sub>206</sub>	^ <sub>222</sub>	- <sub>238</sub>	254
F	SS3 <sub>143</sub>	APC <sub>159</sub>	/ <sub>175</sub>	؟ <sub>191</sub>	ـ <sub>207</sub>	_ <sub>223</sub>	- <sub>239</sub>	255



# الشفرة العالمية Unicode

Universal Character Set UCS / Unicode / ISO 10646 •

• تستخدم ٣٢ خانة لترميز الحروف (جزئية ١٦)

• ترمز جميع لغات العالم في شفرة واحدة

• من الممكن استخدام أكثر من لغتين في نفس الوقت

• تستخدم حالياً في Linux, Win NT, XP, 95, 98, ME

• تؤدي إلى مضاعفة أحجام الملفات

• تعرف أشكال الحروف الخاصة مثل: ل+ا = لا

# عندما يريد العالم أن يتكلم ، فهو يتحدث بلغة يونكود

- Arabic: عندما يريد العالم أن يتكلم ، فهو يتحدث بلغة يونكود
- Arabic x : دوکینوی ةغلب تُدحتي وهف ،مَلکتیناً ملاعلا ديري امدنع
- Catalan : Quan el món vol conversar, parla Unicode
- Chinese (Simplified): 当世界需要沟通时, 请用 Unicode!
- Chinese (Traditional): 當世界需要溝通時, 請用統一碼 (Unicode)
- Danish : Når verden vil tale, taler den Unicode
- Dutch : Als de wereld wil praten, spreekt hij Unicode
- English : When the world wants to talk, it speaks Unicode

# عندما يريد العالم أن يتكلم ، فهو يتحدث بلغة يونكود

- Esperanto : Kiam la mondo volas paroli, gi parolas Unicode
- Finnish : Kun maailma haluaa puhua, se puhuu Unicodea
- French : Quand le monde veut communiquer, il parle en Unicode
- German : Wenn die Welt miteinander spricht, spricht sie Unicode
- Hebrew : Unicodeברבדמאווה ,רבדלהצורםלועהרשאכ
- Hungarian : Ha a világ beszélni akar, azt Unicode-ul mondja
- Irish Gaelic : Nuair a bhualann fonn cainte an domhan, is as Unicode a labhrann sé
- Italian : Quando il mondo vuole comunicare, parla Unicode
- Japanese : 世界的に話すなら、Unicode です。

# عندما يريد العالم أن يتكلم ، فهو يتحدث بلغة يونكود

- Korean : , .
- Norwegian (Bokmål å ål): Når verden vil snakke, snakker den Unicode
- Norwegian (Nynorsk): Når verda ønskjer å snakke, talar ho Unicode
- Occitan : Quan lo mond vòl conversar, parla en Unicode
- Portuguese : Quando o mundo quer falar, fala Unicode
- Romanian : Când lumea vrea să comunice, vorbește Unicode
- Russian : Когда мир желает общаться, он общается на Unicode
- Slovenian : Ko se želi svet pogovarjati, govori Unicode
- Spanish : Cuando el mundo quiere conversar, habla Unicode
- Swedish : När världen vill tala, så talar den Unicode
- Yiddish : דאָקײַנאָ יזט דער ײִדערליטלע ידִאַ

## • صفحة المحارف

- أعطي لكل كل محرف شفرة فريدة
- كل شفرة تشير إلى محرف
- مثال

• مرحبا

• U+0645 U+0631 U+ 062D U+0628 U+ 0627  
(Hexadecimal)

• (Decimal) &#1605;&#1585;&#1581;&#1576;&#1575

• ملاحظة: تستخدم U للتعبير ولا تخزن وكذلك &#

• الإصدار الرابع (٤.٠) من يونكود يشمل ٩٦٤٤٧ محرف

• الإصدار (٤.١) تم إضافة 1273 محرف

	060	061	062	063	064	065	066	067	068	069	06A	06B	06C	06D	06E	06F
0	◻ 0600	◌ 0610	◻	ذ 0630	- 0640	◌ 0650	◌ 0660	◌ 0670	پ 0680	ڈ 0690	غ 06A0	گ 06B0	ة 06C0	ې 06D0	◌ 06E0	◌ 06F0
1	◻ 0601	◌ 0611	ء 0621	ر 0631	ف 0641	◌ 0651	◌ 0661	آ 0671	خ 0681	ڑ 0691	ف 06A1	گ 06B1	ہ 06C1	ې 06D1	◌ 06E1	◌ 06F1
2	◻ 0602	◌ 0612	آ 0622	ز 0632	ق 0642	◌ 0652	◌ 0662	أ 0672	خ 0682	ز 0692	ب 06A2	گ 06B2	ء 06C2	ے 06D2	◌ 06E2	◌ 06F2
3	◻ 0603	◌ 0613	أ 0623	س 0633	ك 0643	◌ 0653	◌ 0663	إ 0673	ج 0683	ر 0693	ب 06A3	گ 06B3	ة 06C3	ے 06D3	◌ 06E3	◌ 06F3
4	◻	◌ 0614	ؤ 0624	ش 0634	ل 0644	◌ 0654	◌ 0664	ء 0674	ج 0684	ر 0694	ف 06A4	گ 06B4	و 06C4	- 06D4	◌ 06E4	◌ 06F4
5	◻	◌ 0615	إ 0625	ص 0635	م 0645	◌ 0655	◌ 0665	أ 0675	خ 0685	ر 0695	پ 06A5	ل 06B5	و 06C5	ہ 06D5	ر 06E5	◌ 06F5
6	◻	◻	ئ 0626	ض 0636	ن 0646	◌ 0656	◌ 0666	و 0676	ج 0686	ر 0696	ق 06A6	ل 06B6	و 06C6	◌ 06D6	◌ 06E6	◌ 06F6
7	◻	◻	ا 0627	ط 0637	ھ 0647	◌ 0657	◌ 0667	و 0677	ج 0687	ت 0697	ف 06A7	ث 06B7	و 06C7	◌ 06D7	◌ 06E7	◌ 06F7
8	◻	◻	ب 0628	ظ 0638	و 0648	◌ 0658	◌ 0668	ء 0678	ئ 0688	ڈ 0698	ق 06A8	پ 06B8	و 06C8	◌ 06D8	◌ 06E8	◌ 06F8
9	◻	◻	ة 0629	ع 0639	ی 0649	◻	◌ 0669	ٹ 0679	د 0689	ت 0699	ک 06A9	ن 06B9	و 06C9	◌ 06D9	◌ 06E9	◌ 06F9
A	◻	◻	ت 062A	غ 063A	ی 064A	◻	◌ 066A	ن 067A	د 068A	ب 069A	ک 06AA	ن 06BA	ق 06CA	◌ 06DA	◌ 06EA	بش 06FA
B	◻	◌ 061B	ث 062B	◻	◌ 064B	◻	◌ 066B	ر 067B	ب 068B	ڈ 069B	ک 06AB	ن 06BB	و 06CB	◌ 06DB	◌ 06EB	ض 06FB
C	◌ 060C	◻	ج 062C	◻	◌ 064C	◻	◌ 066C	ت 067C	ت 068C	ڈ 069C	ک 06AC	ن 06BC	ی 06CC	◌ 06DC	◌ 06EC	نج 06FC
D	◌ 060D	◻	ح 062D	◻	◌ 064D	◻	◌ 066D	ت 067D	د 068D	ص 069D	ک 06AD	ن 06BD	ی 06CD	◻	◌ 06ED	ء 06FD
E	م 060E	◻	خ 062E	◻	◌ 064E	◻	◌ 066E	ب 067E	پ 068E	ڈ 069E	ض 06AE	پ 06BE	ی 06CE	◻	ذ 06EE	م 06FE
F	ع 060F	؟ 061F	د 062F	◻	◌ 064F	◻	◌ 066F	ت 067F	ت 068F	ڈ 069F	ظ 06AF	گ 06BF	ف 06CF	◌ 06DF	ر 06EF	ھ 06FF

# المواصفات والمقاييس

- الشفرة الموحدة
  - وصف الهيكل العام لمحارف الشفرة
  - تحديد المسطح الأساسي متعدد اللغات
  - تحديد الشفرة القائمة على ٢ ، ٤ بايت لمحارف المجموعة
  - تحديد الشفرات الممثلة لوظائف التحكم
  - تحديد أسلوب الإضافات المستقبلية
- مواصفة قياسية ١٠٦٤٦
  - جمعت بين كل المحارف في اللغات الأساسية في ٤ بايت

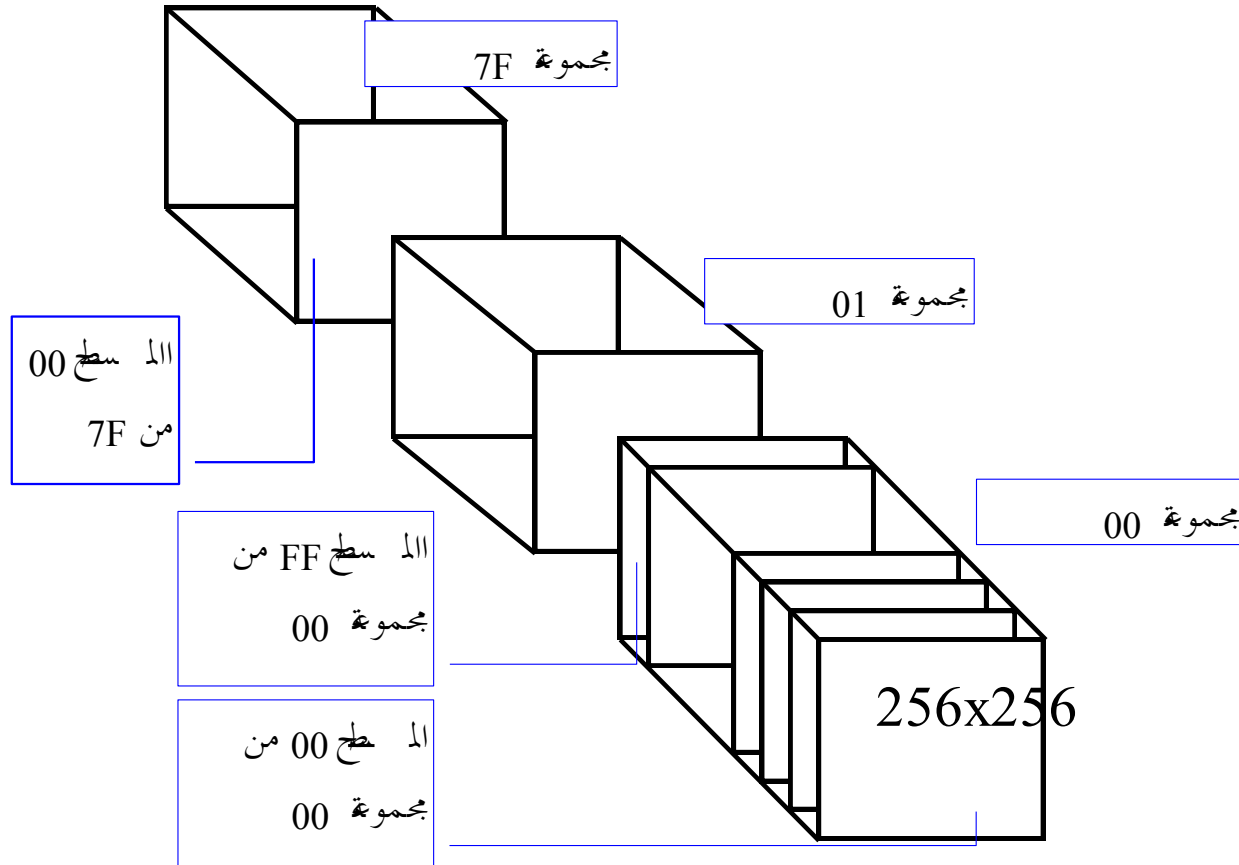
# الشفرة العالمية

- تتكون من مجموعات ومسطحات وسطور وخلايا
- المسطح الأول هو المسطح الأساسي متعدد اللغات
- يتطابق سطرها العربي ٦ مع ASMO 449

المجموعة	المسطح	السطر	الخلية
أعلى قيمة	0000	0041	A
	0000	062C	ج
			أدنى قيمة



# الشفرة الموحدة (Unicode)



# مرجع <http://www.unicode.org/standard/translations/arabic.html>

- ما هي الشفرة الموحدة "يونكود" ؟
- أساسًا، تتعامل الحواسيب فقط مع الأرقام، وتقوم بتخزين الأحرف والمحارف الأخرى بعد أن تُعطي رقما معينًا لكل واحد منها. وقبل اختراع "يونكود"، كان هناك مئات الأنظمة للتشفير وتخصيص هذه الأرقام للمحارف، ولم يوجد نظام تشفير واحد يحتوي على جميع المحارف الضرورية. وعلى سبيل المثال، فإن الاتحاد الأوروبي لوحدته، احتوى العديد من الشفرات المختلفة ليغطي جميع اللغات المستخدمة في الاتحاد. وحتى لو اعتبرنا لغة واحدة، كاللغة الإنجليزية، فإن جدول شفرة واحد لم يكف لاستيعاب جميع الأحرف وعلامات الترقيم والرموز الفنية والعلمية الشائعة الاستعمال.
- وتجدر الملاحظة أن أنظمة التشفير المختلفة تتعارض مع بعضها البعض. وبعبارة أخرى، يمكن أن يستخدم جدول شفرة نفس الرقم لتمثيل حرفين مختلفين، أو رقمين مختلفين لتمثيل نفس الحرف. ولو أخذنا أي جهاز حاسوب، وبخاصة جهاز النادل (server)، فيجب أن تكون لديه القدرة على التعامل مع عدد كبير من الشفرات المختلفة، ويتم تصميمه على هذا الأساس. ومع ذلك، فعندما تمر البيانات عبر أنظمة مختلفة، توجد هناك خطورة لضياع أو تحريف بعض هذه البيانات.
- "يونكود" تغير هذا كليا !
- تخصص الشفرة الموحدة "يونكود" رقما وحيدا لكل محرف في جميع اللغات العالمية، وذلك بغض النظر عن نوع الحاسوب أو البرامج المستخدمة. وقد تم تبني مواصفة "يونكود" من قبل قادة الصانعين لأنظمة الحواسيب في العالم، مثل شركات آي.بي.إم. (IBM)، أبل (APPLE)، هينولت باكرد (Hewlett-Packard)، مايكروسوفت (Microsoft)، أوراكل (Oracle)، صن (Sun) وغيرها. كما أن المواصفات والمقاييس الحديثة (مثل لغة البرمجة "جافا" "JAVA" ولغة "إكس إم إل" "XML" التي تستخدم لبرمجة الانترنت) تتطلب استخدام "يونكود". علاوة على ذلك، فإن "يونكود" هي الطريقة الرسمية لتطبيق المقياس العالمي إيزو ١٠٦٤٦ (ISO 10646).
- إن بزوغ مواصفة "يونكود" وتوفر الأنظمة التي تستخدمه وتدعمه، يعتبر من أهم الاختراعات الحديثة في عولمة البرمجيات لجميع اللغات في العالم. وإن استخدام "يونكود" في عالم الانترنت سيؤدي إلى توفير كبير مقارنة مع استخدام المجموعات التقليدية للمحارف المشفرة. كما أن استخدام "يونكود" سيمكن المبرمج من كتابة البرنامج مرة واحدة، واستخدامه على أي نوع من الأجهزة أو الأنظمة، ولاي لغة أو دولة في العالم أينما كانت، دون الحاجة لإعادة البرمجة أو إجراء أي تعديل. وأخيرا، فإن استخدام "يونكود" سيمكن البيانات من الانتقال عبر الأنظمة والأجهزة المختلفة دون أي خطورة لتحريفها، مهما تعددت الشركات الصانعة للأنظمة واللغات، والدول التي تمر من خلالها هذه البيانات.
- Arabic translation by Mike Ksar

# UTF-8

- **Unicode Transformation Format-8**
- ترميز المحارف بعدد متغير من الثمانيات (بايت) حسب الرقم الذي يمثل شفرة المحرف
- كل محارف الانجليزية US-ASCII تمثل بثمانية واحدة حيث أن شفراتها ممثلة بين U+0000 و U+007F
- UTF-8 هو التمثيل الافتراضي لصيغ XML

# UTF-8

- خصائص هذا الترميز
- المحارف من U+0000 إلى U+007F تمثل كل منها ببساطة كثمانية واحدة (بايت) من 0x00 إلى 0x7F (محارف ASCII)
- بالنظام الثنائي من 00000000 وحتى 01111111
- هذا يعني أن الملفات التي تحوي محارف ASCII سباعية لها نفس ترميز UTF-8
- كل المحارف التي هي أكبر من U+007F تمثل بثمانيتين أو أكثر وتكون قيمة الخانة العليا في كل منها واحد
- هذا يعني أن ثمانية الأسكي السباعية لن تظهر في أي من المحارف الأخرى

# UTF-8

- تمثل الثمانية الأولى في السلسلة التي تمثل حرفا خارج الأسكي عدد الثمانيات التي تتبعها لهذا المحرف
- يمكن تمثيل  $2^{31}$  محرف (كل محارف اليونكود)
- نظريا، قد يصل طول السلسلة الواحدة الممثلة لمحرف واحد إلى ٦ ثمانيات

# مدى محارف الـ Unicode والـ UTF-8

**Unicode Range**

**UTF-8 Representation**

**(hexadecimal)**

**(binary)**

0000 0000-0000 007F

0xxxxxxx

0000 0080-0000 07FF

110xxxxx 10xxxxxx

0000 0800-0000 FFFF

1110xxxx 10xxxxxx 10xxxxxx

0001 0000-0010 FFFF 11110xxx 10xxxxxx 10xxxxxx 10xxxxxx

# التحويل من Unicode إلى UTF-8

- حدد عدد الثمانيات اللازمة للمحرف من الجدول السابق
- حضرّ الخانات العليا لكل ثمانية حسب الجدول السابق
- حدد الخانات المعلمة بـ  $x$ . ابدأ بتحديد أدنى الخانات في الثمانية الدنيا بوضع أدنى خانات شفرة المحرف ثم الخانة التالية وهكذا.
- عند تحديد آخر  $x$  في الثمانية الدنيا انتقل إلى أدنى خانة في الثمانية التالية لها، وهكذا.
- لاحظ أن ترميز UTF-8 لا يشمل المحارف التي في المدى U+D800 وحتى U+DFFF لأن هذه المحارف محفوظة لـ UTF-16

# نشاط صفي : التحويل من Unicode إلى UTF-8

- مرت معنا خطوات التحويل من Unicode إلى UTF-8 والملخصة في ما يلي:
- حدد عدد الثمانيات اللازمة للمحرف من الجدول المعطى
- حضرّ الخانات العليا لكل ثمانية حسب الجدول
- حدد الخانات المعلمة بـ  $x$ . ابدأ بتحديد أدنى الخانات في الثمانية الدنيا بوضع أدنى خانات شفرة المحرف ثم الخانة التالية وهكذا.
- عند تحديد آخر  $x$  في الثمانية الدنيا انتقل إلى أدنى خانة في الثمانية التالية لها، وهكذا.
- إذا علمت أن كلمة "مرحبا" تمثل باليونكود كما يلي
- U+0645 U+0631 U+ 062D U+0628 U+ 0627
- ما هو تمثيل هذه الكلمة باستخدام UTF-8؟

Unicode Range (hexadecimal)	UTF-8 Representation (binary)
0000 0000 - 0000 007F	0xxxxxxx
0000 0080 - 0000 07FF	110xxxxx 10xxxxxx
0000 0800 - 0000 FFFF	1110xxxx 10xxxxxx 10xxxxxx
0001 0000 - 0010 FFFF	11110xxx 10xxxxxx 10xxxxxx 10xxxxxx



# التحويل من UTF-8 إلى Unocode

- اقترح طريقة لذلك

# المواصفات والمقاييس

- تقييس الحروف العربية دوليا
- الحاسبات تستخدم شفرات مختلفة



# معايير تصميم لوحة المفاتيح

• تسهيل وتسريع الطباعة

• لوحة QWERTY مقابل Dvorak

• المواصفة ASMO 663

• تقليل استخدام SHIFT

• مشابهة للوحة الآلة الكاتبة

• تستخدمها الآن Mac

- Pronounced *kwer-tee*, refers to the arrangement of keys on a standard English computer keyboard or typewriter. The name derives from the first six characters on the top alphabetic line of the keyboard. The arrangement of characters on a QWERTY keyboard was designed in 1868 by Christopher Sholes, the inventor of the typewriter
- The Dvorak keyboard layout is a control panel option on almost every current computer. It is a vastly more comfortable and efficient alternative to the standard "QWERTY" pattern, which was designed in the 1800s with no effective attempt at typing comfort

# QWERTY Keyboard layout

~ '	! 1	@ 2	# 3	\$ 4	% 5	^ 6	& 7	* 8	( 9	) 0	= -	+ =	BS
Tab	Q	W	E	R	T	Y	U	I	O	P	{ [	} ]	
Caps Lock	A	S	D	F	G	H	J	K	L	: ;	" '	Return	
Shift	Z	X	C	V	B	N	M	< ,	> .	? /	Shift		
Ctrl .	Alt	Space									Alt	 ;	Ctrl

# Dvorak keyboard layout



Dvorak Keyboard Layout

# ASMO 663

F1 PF1	F2 PF2	F3 PF3	F4 PF4	F5 F5*	F6 F6	F7 F7	F8 F8	F9 F9	F10 F10	F11 F11	F12 F12
--------	--------	--------	--------	--------	-------	-------	-------	-------	---------	---------	---------

~	!	@	#	\$	%	^	&	*	(	)	-	+	BS
1	١	2	٢	3	٣	4	٤	5	٥	6	٦	7	٧
8	٨	9	٩	0	.	-	=						
Tab	Q	W	E	R	T	Y	U	I	O	P	{	}	
	ق	س	ث	ق	ج	خ	ع	د	ح	ط	[	]	\
Ctrl	A	S	D	F	G	H	J	K	L	:	"	'	Return
	ش	س	ي	ب	ل	ا	ت	ن	م	;	'	'	
Shift	Z	X	C	V	B	N	M	<	>	.	?	/	Shift
	ظ	خ	ذ	د	ز	ر	و	<	>	.	/	/	

Ins	Home	PgUp
Ins	Find	Next
Del	End	PgDn
Rem	Sel	Prev
	up	
left	down	right



<http://www.langbox.com/arabic/asm0663.html>

# معايير تصميم لوحة المفاتيح

- تصميمات أخرى بناء على دراسات إحصائية
  - معدلات استخدام الحروف
  - احتمالات الانتقال من حرف إلى آخر
  - توزيع الحروف بهدف تقليل حركة الأصابع
- التوفيق بين متطلبات المستخدمين والمصنعين