

Original by Markus Kuhn, adapted for HTML by Martin Dürst.

UTF-8 encoded sample plain-text file

Markus Kuhn [$\ddot{E}\ddot{m}\dot{a}\dot{\mathring{E}}^3\dot{k}\ddot{E}\ddot{u}\ddot{s}$ $\dot{k}\ddot{u}\ddot{\mathring{E}}\ddot{u}\ddot{n}$] <mkuhn@acm.org> $\dot{a}\ddot{u}\ddot{\mathring{u}}$ 1999-08-20

The ASCII compatible UTF-8 encoding of ISO 10646 and Unicode plain-text files is defined in RFC 2279 and in ISO 10646-1 Annex R.

Using Unicode/UTF-8, you can write in emails and source code things such as

Mathematics and Sciences:

$2H_{\text{eff}} + 0.1 H_{\text{eff}}^2$, $R = 4.7 \text{ k}^{\circ}\text{C}$, 200 mm

Linguistics and dictionaries:

Ä° i Ä±ntÉ¶Ë¶nÄ¡ Ê¶É¶nÉ¶l fÉ¶Ë¶nÉ¶tÄ±k É¶soÊ¶siË¶eÄ±Ê¶n
Y [Ë¶Ê¶psilÉ¶n], Yen [jÉ¶n], Yoga [Ë¶joË¶gÉ¶]

API :

$$((V \sqcap^3 V) = \neg \sqcap^3 \neg \sqcap' V) / V \sqcap \neg V$$

Nicer typography in plain text files:

Greek (in Polytonic):

The Greek anthem:

Îfá $\frac{1}{2}$ ² î $\frac{3}{2}$ î $\frac{1}{2}$ î $\frac{1}{2}$ á $\frac{1}{2}$ · î $\frac{1}{4}$ î $\frac{1}{2}$ á $\frac{1}{2}$, î $\frac{1}{2}$ á $\frac{1}{2}$ · î $\frac{1}{2}$ î $\frac{1}{2}$ á $\frac{1}{2}$ ·
 î $\frac{1}{2}$ î $\frac{1}{2}$ á $\frac{1}{2}$; î $\frac{1}{2}$ î $\frac{1}{2}$ î $\frac{1}{2}$ · î $\frac{1}{2}$, î $\frac{1}{2}$ î $\frac{1}{2}$ á $\frac{1}{2}$; î $\frac{1}{2}$ á $\frac{1}{2}$ · î $\frac{1}{2}$ î $\frac{1}{2}$ î $\frac{1}{2}$ á $\frac{1}{2}$ μ,
 î $\frac{1}{2}$ á $\frac{1}{2}$ ² î $\frac{3}{2}$ î $\frac{1}{2}$ î $\frac{1}{2}$ á $\frac{1}{2}$ · î $\frac{1}{4}$ î $\frac{1}{2}$ á $\frac{1}{2}$, î $\frac{1}{2}$ á $\frac{1}{2}$ · î $\frac{1}{2}$ á $\frac{1}{2}$ î $\frac{1}{2}$ á $\frac{1}{2}$ ·
 î $\frac{1}{2}$ î $\frac{1}{2}$ á $\frac{1}{2}$ ⁰ î $\frac{1}{4}$ á $\frac{1}{2}$ ² î $\frac{1}{2}$ á $\frac{1}{2}$ · î \pm î $\frac{1}{4}$ î $\frac{1}{2}$ á $\frac{1}{2}$ ± î $\frac{1}{4}$ î $\frac{1}{2}$ á $\frac{1}{2}$ · î $\frac{1}{2}$ á $\frac{1}{2}$ · î $\frac{1}{2}$ á $\frac{1}{2}$.

From a speech of Demosthenes in the 4th century BC:

$\hat{I}_1 \hat{I} \cdot \hat{I}^{\frac{1}{4}} \hat{I} \hat{I} \hat{I} \hat{I}$, $a^{\frac{1}{2}} {}^3 \hat{I}^{\frac{1}{2}} \hat{I} \hat{I} \hat{I} \hat{I}$, $\hat{I} \hat{I} a \hat{I}^{\frac{1}{2}}$ $a^{\frac{3}{4}} \hat{I} \hat{I} \hat{I} \hat{I} \hat{I} \hat{I} \hat{I}$, $\hat{I}^1 \hat{I} + \hat{I}^0 a^{\frac{1}{2}}$, $\hat{I} \hat{I}$

Georgian:

From a Unicode conference invitation:

Russian:

From a Unicode conference invitation:

Điện °ÑđμĐ³Đ, ÑđÑñÑđĐ, ÑđÑđĐ¹ÑđμÑñÑđ ÑđμĐ¹Ñđ°Ñđ Đ½Đ° ĐđđμÑñÑñÑñÑđ ÐđμĐ¶Đ 'ÑđĐ½Đ° Unicode, Đ°Đ¾Ñđ¾Ñđ°Ñđ Ñđ¾ÑđÑđ¾Đ, ÑđÑñÑđ 10-12 Đ¼Đ°ÑđÑđ° 1997 Đ³Đ¾Đ'Đ° Đ² ÐđĐ
Ðđ¾Đ½ÑđμÑđμĐ¹Ñđ, Ñđ Ñđ¾Đ±ĐμÑđμÑđ Ñđ, Ñđ¾Đ°Đ, Đ¹ Đ°ÑđÑđ³ Ñđ°Ñđ¿ĐμÑñÑđ¾Đ²
Ðđ½ÑđμÑđ½ĐμÑđ° Đ, Unicode, Đ»Đ¾Đ°Đ»Đ, Đ·Đ°Ñđ, Đ, Đ, Đ½ÑđμÑđ½Đ°Ñđ, Đ¾Đ½Đ
Đ¿Ñđ, Đ½ĐμĐ½ĐμĐ½Đ, Ñđ Unicode Đ² Ñđ°Đ·Đ»Đ, Ñđ½ÑđÑđ Ð¾Đ¿ĐμÑđ°Ñđ, Đ¾Đ½Đ½ÑđÑđ Ñđ,
Đ¿Ñđ, Đ»Đ¾Đ°ĐμĐ½Đ, ÑđÑđ, ÑđÑđ°Ñđ, Đ²ĐμÑñÑđÑđ°Đμ Đ, Đ½Đ½Đ¾Đ³Đ¾Ñđ·ÑđÑđ½Ñđ

Thai (UCS Level 2):

Excerpt from a poetry on The Romance of The Three Kingdoms (a Chinese classic 'San Guo'):

(The above is a two-column text. If combining characters are handled correctly, the lines of the second column should be aligned with the | character above.)

Ethiopian:

Proverbs in the Amharic language:

á»° á»á»- á» á»-á»³ á».. á»µ á»á»á»á»¥ á» á»-á».. á»° á»µá»¢
á»¥á» á»«á»á» á»¥á»á»° á» á»£á»' á» á»á»á» á»á»¢
á»á»¥ á»«á»á»¤á»± á»á»á»¥á» á»á»á»¢
á»° á» á» á»á»á» á» á»á»¤ á»£á»-á» á»£ á»á»£á»µ á» á»á»° á»á»á»¢
á».. á» á» á»á»á»á»³ á» á»á»¤ á» á»-á»³ á»¤á»¢
á» á»-á»¥ á» á» á» á»³ á» á»° á»á»³ á»¢
á»² á»° á».. á»á» á»-á»° á».. á»á»á»¢
á»á»µ á» á»á»µá»¥ á»á»á»á»á»á» á» á»¥á»á»© á»-á»á»³ á»á»¢
á»µá»- á»¤á»«á»¥á»- á» á»á» á»³ á»«á»µá»-á»¢
á»° á» á»¥á»á»° á»¤á»± á»¥á»á» á»¥á»á»° á»¤á».. á»¤á»± á» á»-á»° á»³ á»° á»-á»á»¢
á»¥á»á»á»- á».. á».. á»á»° á»á» á»¤á»¤á»® á»³ á»-á»á»á» á» á»-á»µá»-á»á»¢
á».. á»¤á».. á»¤á»µ á»á»£á»¥ á»¤á»«á»©á»µ á»-á»µá» á»£á»«á»©á»µ á»«á» á»á»á»¢
á»¥á»« á».. á»¤á»á»³ á»µ á»á»á» á»¤á»³ á»µá»¢
á»¤á»- á»¤á»° á»¤á»« á».. á»¤á»á»¥ á»¤á»á»µ á»-á» á» á»-á»¤á»«á»¤á»¢
á».. á»¥á»µá»á» á» á»¤á»© á»¤á»« á».. á»¤á»« á» á»¤á»© á»¤á»-á»¤á»¢
á»° á»¤á»á» á»¤á»° á» á»¤á»¤á» á»£á»¤á»¢
á»¤á»³ á»¤á» á»¤á»- á»¤á»¤á» á».. á»-á»µá» á» á»µá»¤á»° á»¤á»¢
á»¥á»¤á»-á»¤á» á» á»¤á»«á»¤á» á»¤á»- á»¤á»-á»¤á»¢

Runes:

(Old English, which transcribed into Latin reads 'He cwaeth that he bude thaem lande northweardum with tha Westsae.' and means 'He said that he lived in the northern land near the Western Sea.')

Braille:

â i | â | â | §â | | â | $\frac{1}{4}$ â | | â | | â i | â | | â | | â | | â | | $\frac{1}{4}$ â | | â i f â | | â | |

(The first couple of paragraphs of "A Christmas Carol" by Dickens)

Compact font selection example text:

ABCDEFGHIJKLMNOPQRSTUVWXYZ /0123456789

abcdefghijklmnopqrstuvwxyz ÁÉÁ©ÁµÁ¶Á¶Á¶Á¶Á¶Á©Á¶Á¿

Greetings in various languages:

Hello world, $\hat{I} \square \hat{I} + \hat{I} \square \hat{I} \cdot \hat{I}^{\frac{1}{4}} \hat{A}^{\frac{1}{2}} {}^3 \hat{I} \square \hat{I} + \hat{I}^0 \hat{A}^{\frac{1}{2}} {}^1 \hat{I} \square \hat{I}^{\frac{1}{4}} \hat{I} \mu$, $\hat{a}^0 {}^3 \hat{a}^1 {}^3 \hat{a}^2 \square \hat{a}^3 \square \hat{a}^4 \square$

Box drawing alignment tests:

â

â