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

UTF-8 encoded sample plain-text file

â□¾â

Markus Kuhn [Ë□maÊ³kÊ□s kuË□n] <mkuhn@acm.org> â□□ 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 s

Mathematics and Sciences:

  â□® Eâ□□da = Q, n â□□ â□□, â□□ f(i) = â□□ g(i), â□□xâ□□â□□: â□□xâ□□ =

  â□□ â□□ â□□â□□ â□□ â□¤ â□□ â□□ â□□ â□□ â□□ â□□, â□¥ < a â□ b â□¡ c â□¤

  2Hâ□□ + Oâ□□ â□□ 2Hâ□□O, R = 4.7 kÎ©, â□□ 200 mm

Linguistics and dictionaries:

  Ã°i Ä±ntÉ□Ë□nÃ¦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□]

APL:

```
  ((Vâ³V)=â³â´V)/Vâ,V    â·â³â´â¾â
```

Nicer typography in plain text files:

```
  â
  â                                                          â
  â   â¢ âsingleâ and âdoubleâ quotes                        â
  â                                                          â
  â   â¢ Curly apostrophes: âWeâve been hereâ â
  â                                                          â
  â   â¢ Latin-1 apostrophe and accents: 'Â´`  â
  â                                                          â
  â   â¢ âdeutscheâ âAnfÃ¼hrungszeichenâ       â
  â                                                          â
  â   â¢ â , âi, â°, â¢, 3â4, â, â5/+5, â¢, â¦      â
  â                                                          â
  â   â¢ ASCII safety test: 1lI|, 0OD, 8B     â
  â                        â-âââââââ®         â
  â   â¢ the euro symbol: â 14.95 â¬ â         â
  â                        â°âââââââ¯         â
  â
```

Greek (in Polytonic):

The Greek anthem:

Î£á½² Î³Î½Ï□Ï□á½·Î¶Ï□ á¼□Ï□á½¸ Ï□á½´Î½ ÎºÎ½¹Ï□Î·
Ï□Î¿á¿¦ Ï□Ï□Î±Î¸Î¹Î¿á¿¦ Ï□á½´Î½ Ï□Ï□Î¿Î¼ÎµÏ□á½µ,
Ï□á½² Î³Î½Ï□Ï□á½·Î¶Ï□ á¼□Ï□á½¸ Ï□á½´Î½ á½□Ï□Î·
Ï□Î¿á½º Î¼á½² Î²á½·Î± Î¼ÎµÏ□Ï□á½±ÎµÎ¹ Ï□á½´ Î³á¿□.

á¾¿Î□Ï□á¾¿ Ï□á½° ÎºÎ½¹ÎºÎºÎ±Î»Î± Î²Î³Î±Î»Î¼á½³Î½Î·
Ï□á¿¶Î½ á¿¾Î□Î»Î»á½µÎ½Ï□Î½ Ï□á½° á¼±ÎµÏ□á½±
ÎºÎ±á½¶ Ï□á½°Î½ Ï□Ï□á¿¶Ï□Î± á¼□Î½Î´Ï□ÎµÎ¹Ï□Î¼á½³Ï□Ï□ Î¿
Î»Î½¹Î³Î¿Ï□Ï□ Î¿á½□Ï□ á¼□ÎºÎ¿á½»Ï□Î□ Ï□Î¿á½ºÏ□ Î¼á½²Î½ Î³á½°Ï□ Î»Î½¹
Ï□Î¹Î¼Ï□Ï□á½µÏ□Î±Ï□Î¸Î±Î¹ Î¦á½·Î»Î¹Ï□Ï□Î¿Î½ á½□Ï□á¿¶ Î³Î¹Î³Î½Î¿Î¼á½³
Î¼á¼°Ï□ Ï□Î¿á¿¦Ï□Î¿ Ï□Ï□Î¿á½µÎºÎ¿Î½Ï□Î±, á½¥Ï□Î¸á¾¿ á½□Ï□Ï□Ï□ Î¼á½´
Ï□Ï□á½¹Ï□ÎµÏ□Î¿Î½ ÎºÎ±ÎºÎ¿á¶Ï□ Ï□Îºá½³Ï□Î±Ï□Î¸Î±Î¹ Î´á½³Î¿Î½. Î¿á½□Î´
Î¿á¼± Ï□á½° Ï□Î¿Î¹Î±Î¿á¿¦Ï□Î± Î»á½³Î³Î¿Î½Ï□ÎµÏ□ á¼¢ Ï□á½´Î½ á½□Ï□á½¹Î¸
Î¿á½□Ï□á½¶ Ï□á½´Î½ Î¿á½□Ï□Î±Î½ Ï□Î±Ï□Î¹Ï□Ï□á½±Î½Ï□ÎµÏ□ á½□Î¼á¿□Î½ á¼□
Ï□Î¿Ï□á¾¿ á¼□Î¾á¿□Î½ Ï□á¿□ Ï□á½¹Î»ÎµÎ¹ ÎºÎ±á½¶ Ï□á½° Î±á½□Ï□á¿□Ï□ á¼□
Ï□Î¹Î¼Ï□Ï□á½µÏ□Î±Ï□Î¸Î±Î¹, ÎºÎ±á½¶ Î¼á½±Î»á¾¿ á¼□ÎºÏ□Î¹Î²á¿¶Ï□ Î¿á¼¶¶
Î³á½³Î³Î¿Î½ÎµÎ½ Ï□Î±á¿¦Ï□á¾¿ á¼□Î¼Ï□á½¹Ï□ÎµÏ□Î□ Î½á¿¦Î½ Î¼á½³Î½Ï□Î
Ï□Ï□Î¿Î»Î±Î²Îµá¿□Î½ á¼¡Î¼á¿□Î½ Îµá¼¶Î½Î±Î¹ Ï□á½´Î½ Ï□Ï□á½½Ï□Î·Î½, á½½
Ï□á½½Ï□Î¿Î¼ÎµÎ½. á¼□á½°Î½ Î³á½°Ï□ Ï□Î¿á¿¦Ï□ Î²ÎµÎ²Î±á½·Ï□Ï□ á½□Ï□á½
Ï□á½·Î½Î± Ï□Î¹Î¼Ï□Ï□á½µÏ□ÎµÏ□Î±á½· Ï□Î¹Ï□ ÎºÎ±á½¶ á½□Î½ Ï□Ï□á½¹Ï□Î¿Î½
Ï□á½´Î½ á¼□Ï□Ï□á½´Î½ á½□Ï□Î¸á¿¶Ï□ á½□Ï□Î¿Î¸á½³Ï□Î¸Î±Î¹, Î¼á½±Ï□Î±Î¹Î¿
Ï□ÎµÎ»ÎµÏ□Ï□á¿□Ï□ á½□Î½Ï□Î¹Î½Î¿á¿¦Î½ Ï□Î¿Î¹Îµá¿□Ï□Î¸Î±Î¹ Î»á½¹Î³Î¿Î½

Î□Î·Î¼Î¿Ï□Î¸á½³Î½Î¿Ï□Ï□, Î□á¿½ á¾¿Î□Î»Ï□Î½Î¸Î±Îºá½¸Ï□

Georgian:

From a Unicode conference invitation:

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□ªá□□á□□á□□á□□ á□□á□□á□iá□□á□iá□¬á□ á□□á□□á□□á□□,
á□¥. á□□á□□á□□á□□á□ªá□¨á□□, á□□á□□á□ á□□á□□á□□á□□á□□á□¨á□□. á□□á□□á□□á□¤á
á□□á□¥á□iá□□á□□á□ á□¢á□□á□□á□i á□□á□iá□□á□□ á□□á□□á□ á□□á□□á□□á□¨á□□ á□ á
á□□á□□á□¢á□□á□ á□□á□□á□ªá□□á□□á□□á□□á□□á□□á□□á□□á□ªá□□á□□ á□□á□□ á□□á□□á□□
á□□á□□á□□á□ á□□á□ªá□□á□£á□□ á□iá□□á□iá□¢á□□á□□á□□á□□á□iá□□, á□□á□□ á□□á□□
á□¢á□□á□¥á□iá□¢á□□á□□á□□á□i á□□á□□á□□á□£á□¨á□□á□□á□□á□□á□□á□iá□□ á□□á□□ á

Russian:

From a Unicode conference invitation:

Ð□Ð°Ñ□ÐµÐ³Ð¸Ñ□Ñ□Ð¸Ñ□Ð¹Ñ□ÐµÑ□Ñ□ Ñ□ÐµÐ¹Ñ□Ð°Ñ□ Ð½Ð° Ð□ÐµÑ□Ñ□Ñ□Ñ□Ñ□ Ð□
Unicode, ÐºÐ¾Ñ□Ð¾Ñ□Ð°Ñ□ Ñ□Ð¾Ñ□Ñ□Ð¾Ð¸Ñ□Ñ□Ñ□ 10-12 Ð¼Ð°Ñ□Ñ□Ð° 1997 Ð³Ð¾
Ð□Ð¾Ð½Ñ□ÐµÑ□ÐµÐ½Ñ□Ð¸Ñ□ Ñ□Ð¾Ð±ÐµÑ□ÐµÑ□ Ñ□Ð¸Ñ□Ð¾ÐºÐ¾Ð¹ ÐºÑ□Ñ□Ð³ Ñ□ÐºÑ□Ð
Ð□Ð½Ñ□ÐµÑ□Ð½ÐµÑ□Ð° Ð¸ Unicode, Ð»Ð¾ÐºÐ°Ð»Ð¸Ð·Ð°Ñ□Ð¸Ð¸ Ð¸ Ð¸Ð½Ñ□ÐµÑ□Ð
Ð¿Ñ□Ð¸Ð¼ÐµÐ½ÐµÐ½Ð¸Ñ□ Unicode Ð² Ñ□Ð°Ð·Ð»Ð¸Ñ□Ð½Ñ□Ñ□ Ð¾Ð¿ÐµÑ□Ð°Ñ□Ð¸Ð
Ð¿Ñ□Ð¸Ð»Ð¾Ð¶ÐµÐ½Ð¸ÑÑ□, Ñ□Ñ□Ð¸Ñ□Ñ□Ð°Ñ□, Ð²ÐµÑ□Ñ□Ñ□ÐºÐµ Ð¸ Ð¼Ð½Ð¾Ð³Ð¾Ñ

Thai (UCS Level 2):

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

[----------|----------]
à¹□ à¹□à¸□à¹□à¸□à¸□à¸´à¸□à¸®à¸±à¹□à¸□à¹□à¸ªà¸·à¹□à¸-à¸ià¹□à¸□à¸£à¸i
à¸ªà¸´à¸□à¸ªà¸-à¸□à¸□à¸©à¸±à¸□à¸£à¸´à¸¢à¹□à¸□à¹□à¸-à¸□à¸«à¸□à¹□à¸²à¹□
à¸□à¸£à¸□à¸□à¸±à¸□à¸□à¸·à¸-à¸□à¸±à¸□à¸□à¸µà¹□à¸□à¹□à¸□à¸□à¸µà¹□à¸□à
à¹□à¸®à¸□à¸´à¹□à¸□à¹□à¸£à¸µà¸¢à¸□à¸□à¸±à¸□à¸□à¸±à¹□à¸§à¸«à¸±à¸§à¹□à¸
à¹□à¸«à¸ià¸·à¸-à¸□à¸□à¸±à¸□à¹□à¸ªà¹□à¸¥à¹□à¹□à¸ªà¸·à¸-à¸□à¸²à¸□à¹□à
à¸□à¹□à¸²à¸¢à¸-à¹□à¸-à¸□à¸-à¸¸à¹□à¸□à¸¢à¸¸à¹□à¸¢à¸□à¹□à¸«à¹□à¹□à¸□à¸
à¸□à¸¥à¸±à¸□à¸¥à¸´à¸□à¸¸à¸¢à¸□à¸¸à¸¢à¸□à¸µà¸□à¸¥à¸±à¸□à¹□à¸-à¹□à
à¸□à¹□à¸-à¸□à¸£à¸□à¸£à¸²à¸□à¹□à¸²à¸□à¸±à¸□à¸□à¸□à¸£à¸£à¸¥à¸¢

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

Ethiopian:

Proverbs in the Amharic language:

á□°á□□á□- á□ á□-á□³á□¨á□µ 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□µ á□ á□□á□°á□□á□□á□¢
á□¨á□ á□□ á□□á□□á□□á□³ á□ á□□á□¤ á□ á□-á□³á□½á□□á□¢
á□ á□-á□¥ á□ á□ á□□ á□³á□□ á□°á□□á□³á□¢
á□²á□°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□³á□□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□³á□°á
á□¥á□□á□□á□- á□¨á□¨á□□á□°á□□á□□ á□□á□®á□® á□³á□-á□□á□□á□□ á□ á□-á□µá□-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□³á□µ á□□á□□á□□ á□□á□□á□³á□µá□¢
á□□á□£á□- á□□á□°á□ªá□« á□¨á□□á□□á□¥ á□□á□□á□µ 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□¢
á□□á□³á□□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□½á□□ á□□á□- á□□á□-á□□á□¢

Runes:

á□»á□□ á□³á□¹á□«á□¦ á□¦á□«á□□ á□»á□□ á□□á□¢á□□á□□ á□©á□¾ á□¦á□«á□□ á□□á□

(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:

âi□â □â §â □ â ¼â □â □ âi□â □â □â □â ¹â °â □ âi£â □â □

âi□â □â □â □â ¹ â ºâ □â □ â □â □â □â □â □ â □â □ â □â □â □â □ â ºâ □â
â ±â □â □â □â §â » â □â □â ³â □ â ¹â □â □â ² âi¹â □ â □â □â □â □â □â
â □â □â □â □â « â □â ¹ â ¹â □ â □â □â »â □â ¹â □â □â □ â ¹â □ â □â
â □â □â □ â ¹â □ â ¡â □â □â □ â □â ³â □â □â »â ² âi□â □â □â □â □â □

âi□â □â □â □â □â □â °â □ â □â □â □â □ â ºâ □â □ â □â □â □â □ â ¥â □
â ¡â □â □â □ â □â □ â □â ¥â □ â □â □â □ â □â □â □â □ â □â □â ²

âi□â □â □ âi□â □â □â □â ¹ â ºâ □â □ â □â □ â □â □â □â □ â □â □ â □ â □

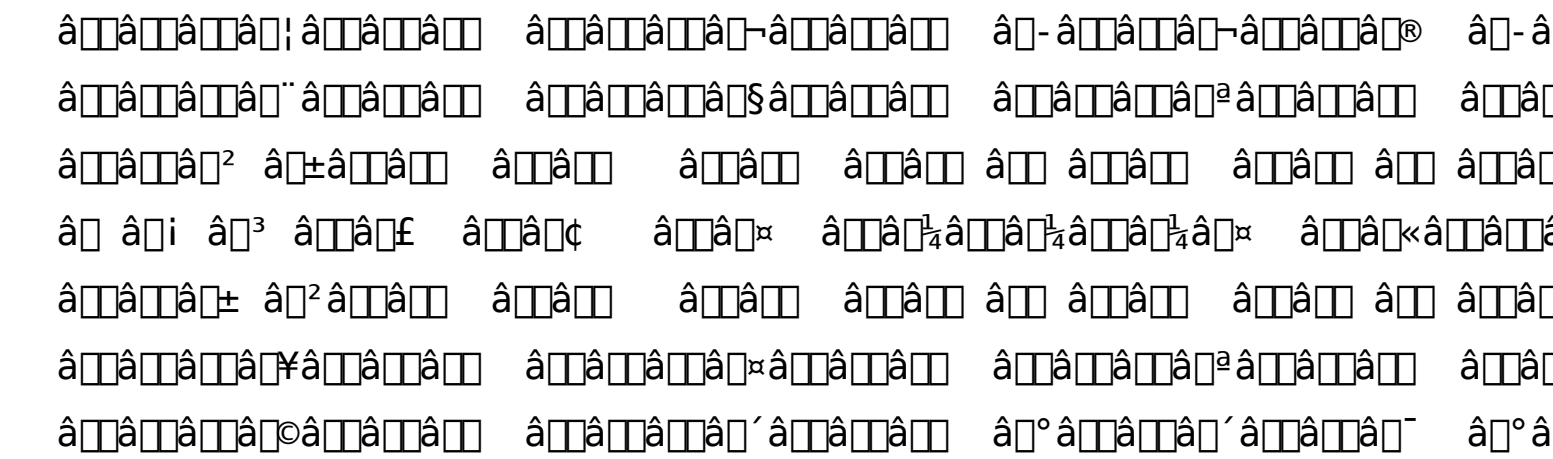
âi□â □â □â □ âi□ â □â □â □â °â □ â □â □â □â □ â □â □ â □â □â ¹ â ¹â □â
â ªâ □ â □â □â ªâ □â «â □â □â □ â ±â □â □ â ¹â »â □ â □â □ â □â □â □â
â □ â □â □â □â □â ¤â □â □â □â □â ² âi□ â □â □â £â □ â □â □â §â □ â □â
â □â □â □â □â □ â □ â □â □â □â □â □â ¤â □â □â □â □ â □â □ â ¹â □ â □â
â □ â ¹â □ â □â □â □â □â □â ² âi□â ¥â □ â ¹â □ â ºâ □â □â □â □â □ â □â
â □â □ â □ â ¹â □ â □â □â □â □â □â □â □ â □â □â □ â □â ¹ â ¥â □â □â □â
â ©â □â □â □ â □â □â □ â □â □â □â ¥â □â □ â □â □â □ â □â □ â ¹â □ âi□â
â ºâ □â □â □ â ¹â »â □â □â □â □â □ â □â »â □â □â □ â □â □ â □â □ â □â
âi□â □â □â □â ¹ â ºâ □â □ â □â □ â □â □â □â □ â □â □ â □ â □â □â □â □â

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

Compact font selection example text:

ABCDEFGHIJKLMNOPQRSTUVWXYZ /0123456789
abcdefghijklmnopqrstuvwxyz Â£Â©ÂµÃ□Ã□Ã□Ã□Ã□Ã©Ã¶Ã¿
â□□â□□â□□â□□â□□â□□â□ â□¢â□¦â□°â□¢Å□Å Å¸Å¾â□¬ Î□Î□Î□Î□Î©Î±Î²Î³Î´Ï□ Ð□Ð□Ð
â□□â□□â□□â□□â□§â□ªâ□¡â□□ â□□â□□â□¨â□»â□£ â□□â□¼â□□â□□â□□â□ºâ□ºâ□□ ï¬□ï¿½â

Greetings in various languages:

Hello world, Î□Î±Î»Î·Î¼á½³Ï□Î± Îºá½¹Ï□Î¼Îµ, ã□³ã□³ã□□ã□□ã□□

Box drawing alignment tests: