



# UNSCII uudistaa

## Merkkigrafiikan päivitys Unicode-aikaan

*Unicodessa on monia palikkagrafiikan kannalta herkullisia merkkejä, mutta juuri kukaan ei käytä niitä. Unscii-fontti yrittää ratkaista ongelman.*

Teksti ja kuvat: Ville-Matias Heikkilä

**P**ari vuotta sitten halusin selvittää, onko Unicode löytänyt tiensä perinteiseen merkkigrafiikkaan. Osoitautui, että hyvin vähäisesti. Kyse ei selvästikään ole kehitysvastarinnasta: esimerkiksi Ecmán päätteenohjausstandardin mukaiset laajemmat väripaletit on omaksuttu niin MUDeihin kuin muutamiin suosittuihin merkkigrafiikkaeditoreihin. Palikkavalikoimien osalta on kuitenkin pitäyditty perinnemerkeistöissä – miksi ihmeessä?

Tasalevyisten fonttien vertaileminen vastasi kysymykseen: mikäli fonttissa on grafiikkamerkkejä, yleensä vain PC:stä tutut merkit on toteutettu jokseenkin kunnollisesti. Kyseessä oli selvä muna-kana-ongelma: ennen kuin Unicode-palikkagrafiikka pääsi kehittymään, pitäisi olla kunnollinen referenssifontti, jollaista ei kuitenkaan synny olemassa olevan grafiikkaperinteen puuttuessa.

### Merkkejä tutkimaan

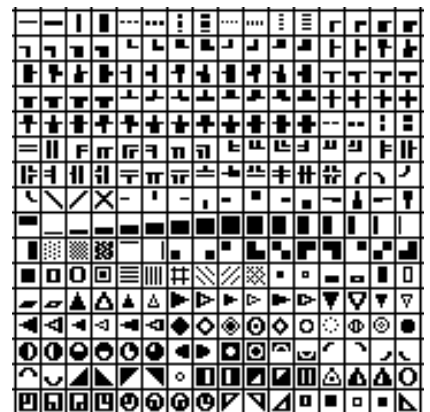
Unscii-fontti syntyi halusta luoda 8×8 pikselin tasalevyinen bittikarttafontti, joka tukisi niin perinteistä kuin uuttakin merkkigrafiikkaa mahdollisimman hyvin. Nimen voi ajatella yhdistelmäksi termeistä Unicode, ANSI ja ASCII. Myöhemmin fontista alkoi syntyä tavallista pääte- ja ohjelmointikäyttöä varten myös 8×16-versio, josta olen mielenyt vaihtoehtoa iänikuiselle PC:n VGA-fontille.

Unicode-konsortio ei ota kantaa siihen, kuinka grafiikkamerkit pitäisi toteuttaa: niille on tarjottu paikat vain alaspäinyhteensopivuuden vuoksi. ”Oikean” toteutustavan löytämiseksi piti siis kaivaa esiin ne vanhat järjestelmäfontit, joilla näitä merkkejä on alun perin käytetty. Koska myös tavallisia kirjaimia, numeroita ja välimerkkejä on käytetty merkkigrafiikkaan, piti myös ne ottaa syyniin.

8×8 pikselin järjestelmäfontteja on karkeasti kahdenlaisia: ohutviivaisia, joissa viiva on yhden pikselin levyinen, ja paksuviivaisia, joissa paksuus on vaakasuunnassa kaksi pikseliä. Jäl-

kimmäisiin kuuluvat niin Commodore 64:n, Amigan, Atari ST:n kuin PC:nkin oletusfontit, joten valinta oli tältä osin selvä. Otin vertailujoukkoon mukaan myös Atari XL:n ja Amstrad CPC:n ROM-fontit ja BBC Micron grafiikkatilafontin.

Paksuviivaiset 8×8-fontit ovat yllättävän samankaltaisia Ascii-alueensa osalta. Monet kirjaimet, numerot ja välimerkit ovat identtisiä liki koko



Unicoden grafiikkamerkit keskittyvät alueelle 2500-25FF, joka näyttää Unsciiin 8×8-versiossa tältä.

vertailuryhmässä, joten niissä pääsi etenemään mainiosti enemmistöperiaatteella. Suuri osa viivapohjaisesta Ascii-grafiikasta on piirretty Amigan Topaz-fontille, joten joitakin sen piirteitä on suosittu. Esimerkiksi X:n ja Y:n viivat yhdistyvät luontevasti merkkeihin \, / ja |.

Varsinaisia grafiikkamerkkejä vertailllessani osoittautui, ettei monille niistä ole lainkaan Unicode-paikkaa! Fonttiprojektista tuli siis myös kartoitussuunnitelma, jossa kävin läpi kaikki mahdolliset grafiikkamerkkejä sisältävät perinnemerkistöt ja järjestelmäfontit.

Suosituista merkistöistä esimerkiksi Petscii, Teksti-TV, Videotex ja jopa VT100 sisältävät useita käyttökelpoisia grafiikkamerkkejä, joita ei ole Unicodeissa – tämä siitä huolimatta, että näitä tukevia laitteita on ollut käytössä miljoonittain. Uusia merkkejä löytyi lisäksi CPC:n, Atarin kotimikrojen, EACA Colour Genien, Mattel Aquariuksen, Sharp MZ:n ja TRS-80:n ROM-fonteista sekä Applen Sabine10-merkistöstä. Päätin sijoittaa nämä merkit Unsciiissa PUA-alueelle (Private Use Area), eli koodipaikasta U+E000 alkaen.

## Kaksi eri leveyttä

Monet Unicoden grafiikkamerkkeistä ovat peräisin kiinalais-japanilais-korealaisista merkistöistä (CJK), joten otin tarkasteluun myös japanilaisen NEC PC:n ja muutaman itäisen MSX-koneen ROM-fontit. Näissä koneissa merkeillä on kaksi mahdollista leveyttä: puolilevä (8 pikseliä) ja täyslevä (16 pikseliä). Esimerkiksi kiinalaisperäiset sanamerkit ovat täysleveitä.

Länsimaiset pääteohjelmat käyttävät merkin leveyden määrittämiseen yleensä `wcwidth()`-funktiota. Funktio ilmoittaa itsenäisen merkin leveydeksi aina yhden solun, ellei sitä ole erikseen määritelty täysleväksi itäaasialaismerkiksi, jolloin leveys on kaksi solua. Idässä käytäntö on kuitenkin erilainen: merkin leveys riippuu siitä, käytetäänkö merkin siirtokoodauksessa yhtä vai kahta tavua. Suurin osa grafiikkamerkeistä on kaksitavuisia, joten täydellisen fontin pitäisi pystyä esittämään ne sekä yhden että kahden solun kokoisina.

Monet täysleveät grafiikkamerkit pystyy esittämään tyydyttävästi kah-

```
Atari 8-bit [1004:7] <= EI! X/Y, tai joskus yöllä Q.
Atari ST [1004:7] <= EI! X/Y, tai joskus yöllä Q.
Commodore 64 [1004:7] <= EI! X/Y, tai joskus yöllä Q.
Amiga Topaz [1004:7] <= EI! X/Y, tai joskus yöllä Q.
IBM PC 8x8 [1004:7] <= EI! X/Y, tai joskus yöllä Q.
Amstrad CPC [1004:7] <= EI! X/Y, tai joskus yöllä Q.
Unscii 8x8 [1004:7] <= EI! X/Y, tai joskus yöllä Q.
```

```
X11 8x13B [1004:7] <= EI! X/Y, tai joskus yöllä Q.
```

```
IBM PC 8x16 [1004:7] <= EI! X/Y, tai joskus yöllä Q.
```

```
MS Fixedsys [1004:7] <= EI! X/Y, tai joskus yöllä Q.
```

```
Unscii 8x16 [1004:7] <= EI! X/Y, tai joskus yöllä Q.
```

Unscii muistuttaa tekstimerkkiensä osalta mikrotietokonea maailman järjestelmäfontteja.

den puolilevään Unicode-merkin yhdistelmänä, mutta esimerkiksi taiwanilaisessa BBS-grafiikassa suosittuja kulmakolmioita ei. Niinpä tärkeimmistä grafiikkamerkeistä on Unsciiissa tarjolla myös vasen ja oikea puolisko erikseen. Puolitetut merkit helpottavat myös mittasuhteiden määrittämisessä, esimerkiksi 40x25 merkin ruudulle piirretyn kuvan venyttämiseksi samanmuotoiseen 80x25 merkin ikkunaan.

## Esikuvia ja ryöstöretkiä

Unsciiin 8x16-versio on muodostettu 8x8-fontin pohjalta siten, että grafiikkamerkkien mitat on pidetty solun suhteen mahdollisimman ennallaan mutta tekstiglyyfejä on mataloitettu ja kaunistettu. Esikuvana ovat olleet etenkin PC:n 8x16-kokoinen VGA-fontti ja Windowsin 8x15-kokoinen Fixedsys, jotka ovat läheistä sukua paksuviivaisille 8x8-fonteille. Muita esikuvia ovat olleet X-ikkunointijärjestelmän lihavoidut 8x13- ja 9x15-fontit sekä VT420-päätteen 10x16-fontti. Tarvittaessa mallia on otettu myös suosituista tasalevyisistä vektorifonteista, joita ovat DejaVu Sans Mono, Lucida Console, Monaco ja Inconsolata.

Unscii on toteutettu piirtämällä sen glyyfit tekstitiedostoihin pistettä ja risuaitaa käyttäen. Tekstitiedostoissa on lisäksi fonttien Unicode-paikat sekä erilaisia ohjauksikäskyjä, joita on projektin edetessä määritelty lisää tarpeen mukaan. Tekstitiedostot käännetään BDF-muotoon Perl-skripteillä ja BDF-muodosta tuotetaan PCF- ja TTF-versiot erillisillä työkaluilla.

Unsciiin Unicode-kattavuus ei ole sellaisenaan kovinkaan hyvä, minkä huomaa esimerkiksi jonkun kopioidessa IRC:hen kiinankielistä tekstiä tai foneettisia ääntämisohteja. Tämän vuoksi 8x16-Unsciiista on myös versio, johon on haettu puuttuvat glyyfit skriptin avulla GNU Unifontista.

Unifont pyrkii olemaan mahdollisimman kattava bittikarttafontti. Sen glyyfit ovat oikeankokoiset – 8x16 ja 16x16 pikseliä – mutta ne eivät ohutviivaisina istu Unsciiin tyyliin. Skripti asenoi glyyfit uusiksi ja skaalaa ne tarvittaessa `wcwidthin` mukaisiksi, mutta viivanleveyteen ei kosketa. Fixedsysin harrastelijavetoinen Unicode-versio Fixedsys Excelsior olisi tyyliään yhteensopivampi, mutta sen lisenssi on turhan epämääräinen.

## Koeajo

Unsciiä on kehitysvaiheessa testattu niin grafiikassa kuin normaalissa päätöksenteossa: ohjelmoinnissa, tekstinmuokkauksessa ja irkkauksessa. Muutamia typografisia ongelmia on korjattu: toisiinsa helposti sekoittuvien merkkien erottuvuutta on parannettu etenkin ohjelmointikäyttöä ajatellen ja tekstijäljestä on tehty tasaisempaa pidentämällä kapeiden I-, i- ja l-kirjainten päätteitä.

Olemassa olevat merkkigrafiikkaeditorit eivät taivu kovinkaan helposti Unicodeen eivätkä tekstieditorit merkkigrafiikkaan. Tein palikkagrafiikkakokeiluja varten aluksi joitakin makroja tekstieditoriin ja myöhemmin puhtaalta pöydältä oman grafiikkaeditorin. Grafiikan piirtelyyn sain innoketta myös Simulaatio-demotahtumaan työstämäni merkkigrafiikkademon kautta. Demoa varten on tosin koottu merkeistä suppeampi, 256 merkin valikoima.

Tätä kirjoitettaessa Unscii on vaiheessa, jossa sen kehtaa jo laittaa julkisesti ladattavaksi. Käy siis lataamassa fontti vaikkapa Skrollin sivulta <http://skrolli.fi/2015.2> ja kerro kokemuksistasi! Koska seuraava etappi tulee olemaan jonkinlaisen grafiikkakilpailun järjestäminen, ovat myös merkkigrafiikkapiirroksia hyvin tervetulleita – niin värittömät kuin värillisetkin! 🐣