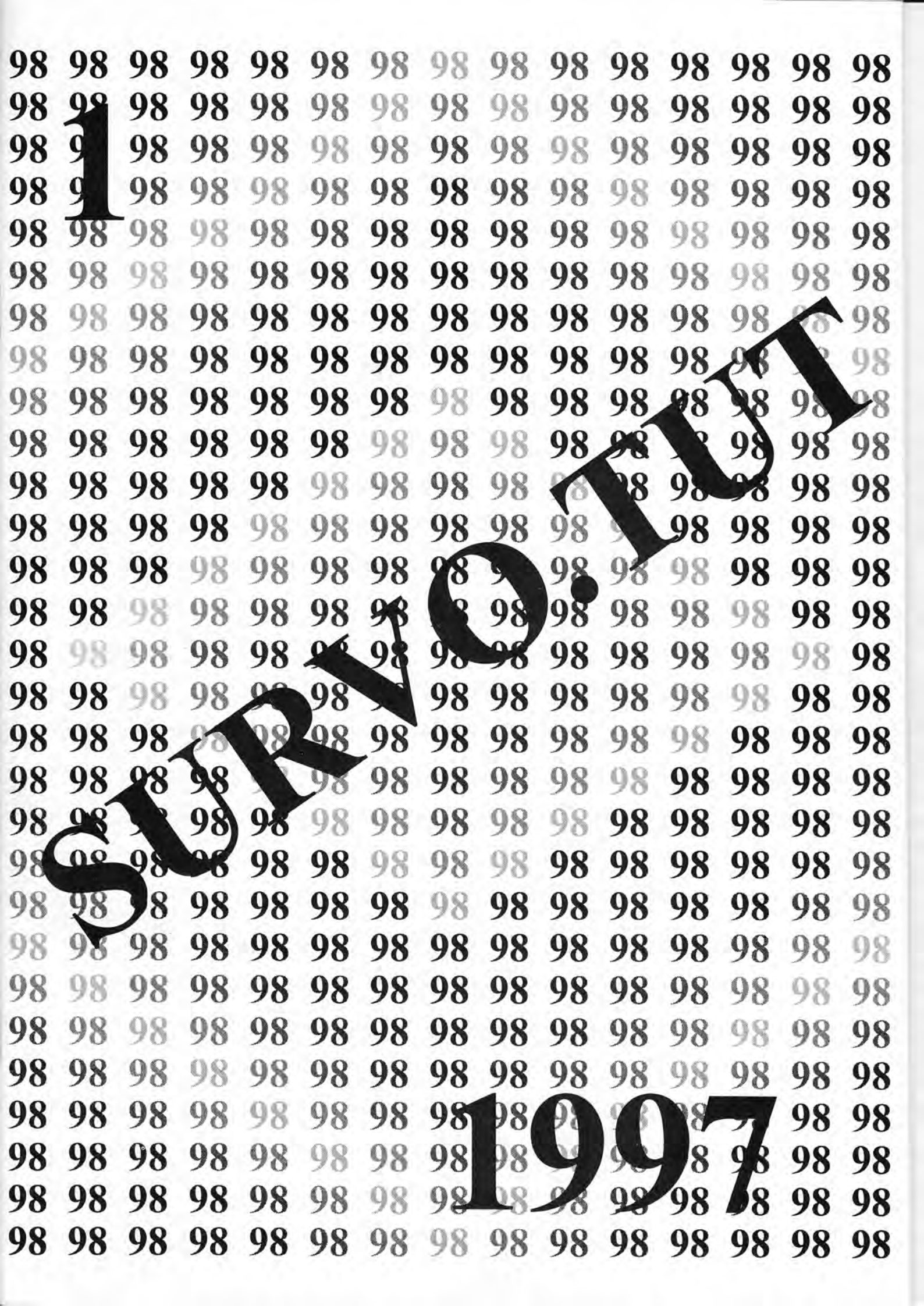




Survo-Käyttäjäyhdistys ry:n
lehti 1/1997

Toimituskunta

Kimmo Vehkalahti	Juha Valtonen
Marjut Schreck	Maija Polus
Jyrki Nummela	Kalevi Kantele

Taitto ja sivunvalmistus

SURVO & Marjut

SISÄLTÖ

Uusi Survo tulossa	2
Survo-seminaarit ja -kurssi syksyllä 97	6
Toimiston tarpeisiin.	7
Taulukoiden tunnistus tekstitiedostoista	15
Desktop	19
Survoilua opettajana ja "oloneuvoksena"	21
Survo ja WWW	26
"X all that apply"	29
Taulukoita ristiin	30
Kuva ja sana	32
SURVOTUT 4 - levyke	33

Survo-Käyttäjäyhdistys ry
Sihteeri: Marjut Schreck
Mäenrinne 11
02160 ESPOO

Survo ja me sekä SM-liiga 98

Tämän lehden numeron 1/96 liitteenä oli Seppo Mustosen Survo ja minä -kirjaa koskeva tiedote. Kirjaanhan oli alunperin tarkoitus tulla omina lukui-
naan myös muiden survoilijoiden kirjoituksia heidän suhteestaan Survoon.
Suunnitelmat muuttuivat kirjaprojektin edetessä, ja lopulta lisäluvut päätet-
tiin jättää pois. Sepon toive oli tällöin, että vuoden sisällä syntyisi uusi teos
teemasta "Survo ja me".

Käyttäjähdistyksessä virisi idea julkaista ylijääneitä kirjoituksia seuraavissa
SURVO.TUT-lehdissä. Lopulta päätimme tehdä käsillä olevasta numerosta
erityisen Survo ja me -teemanumeron.

Lehdykäisestä muodostuikin varsinainen lukupaketti. Komeuden kruunaa
tietenkin SURVOTUT 4 -levyke, joka on täynnä monenmoista mukavaa.

Marjut Schreck kyseli viime numeron pääkirvoituksen otsikossaan: "Survo ja
minä, entäpä sinä ?!". Tämä numero puolestaan uhkuu me-henkeä, eikä
syyttä! Kuitenkin jatkuvasti *minä*, *sinä*, *hän* ja *me* ihmettelemme *teidän*
kanssanne, miksi *he* eivät tunnu oivaltavan Survon ideaa. Se tuntuu käsittä-
mättömältä.

Windows willitsee, ja Survokin haluttaisiin kehystää ja varustaa valikoin ja
painikkein. Mahdollistahan se olisi, mutta kovin pinnallista, pidemmän pääl-
le myös piinallista.

Kyllä windowsmainen työskentely joskus on aivan omiaan; tyypillisesti kerta-
luontoisissa tehtävissä. ATK:ssa tehtävillä on kuitenkin tapana toistua.
Useimmat uudet työt ovat muunnelmia vanhoista. Harvoja asioita tehdään
vain kerran.

Windowsin käyttötavat eivät rohkaise eivätkä anna tarpeeksi mahdollisuuksia
kehittyä: "point-and-click"-työskentelymalli latistaa ihmiset helposti ikui-
siksi aloittelijoiksi. Ja jos edistyneen käyttäjän ominaispiirre on, ettei hän
tarvitse enää hiirtä, koska osaa rykelmän näppäinyhdistelmiä, jotka no-
peammin hoitavat hiiren tehtävät, niin voivoi - otan osaa.

Onneksi Survossa asiat ovat toisin. Sen todistaa jälleen tämä lehti monipuol-
lisine juttuineen ja uutuuttaankiiltelevä **SURVO 98**. Kaukalossa eivät laidat
tule heti vastaan, ja kattokin on korkeammalla. 1.000.000,00 rivin lisäkatso-
mot on jo testattu, eikä pahempia ruuhkia ole päässyt syntymään. Tuttu si-
nivalkoinen tulostaulukin kentän yläpuolella on saanut uutta väriä kylkeen-
sä ja pelaaminen on entistä vikkelmämpää.

Virallisen SM-liiga 98:n alkuun on enää muutamia päiviä, joten kannattaa
tarkistaa varusteet ja päivittää ne kuntoon jo alkupelejä varten. Liigan peli-
sääntöjä löytyy lisää tästä lehdestä, mutta kuten kaikki aiemmissa Survo-
sarjoissa pelanneet tietävät, pelisäännöt laatii pelaaja itse, ei ympäristö. Ete-
neminen kohti sarjataulukon kärkeä ei myöskään kulje pitkin ennalta sovit-
tuja reittejä, vaan vapaan oivalluksen kautta. Peli- ja ottelukieltota ei tässä
sarjassa tunneta ja erätaukokin voi joskus tuntua tarpeettomalta.

Let's play - Survo !

Seppo Mustonen:

Uusi Survo tulossa!

Kaikki Survon kehitystä seuranneet tietävät, että keskusteluja Survon murtautumisesta MS DOS -käyttöjärjestelmän ahtaista rajoista on keskusteltu vuosikaupalla. Viimeksi käyttäjäyhdistyksen vuosiristeilyllä huhtikuussa Jarmo Ahonen esitteli laatimaansa Survon koeversiota, joka toimi OS/2-ympäristössä. Olin sitä ennen koostanut pienen osan Survon C-kielisestä lähdekoodista eräänlaiseksi "minieditoriiksi" ja luovuttanut sen Ahoselle 32-bittisen koeversion tekoa varten. Vuosikokousristeilyllä vaihdettiin innokkaasti mielipiteitä siitä, lähteäkö eteenpäin tuolta pohjalta vai ei.

Jatkokeskustelut osoittivat, että tässä vaiheessa on tärkeintä todella vapautua 16-bittisen maailman kahleista ja muuntaa Survon ohjelmakoodi 32-bittiseksi sopivan C-kääntäjän ja DOS-laajennusohjelman avulla.

Nykyinen SURVO 84C on 16-bittisten ohjelmamodulien muodostama kokonaisuus. Sillä ei päästä ohi 1 megatavun perusmuistin ja siinäkin tilassa yli 64 kilotavun kokoisten taulukkojen käsittely on hankalaa. Kimmo Vehkalahti oli kokeillut jo vuonna 1993 Watcomin C-kääntäjää, joka sallii kätevästi ohjelmien teon kaikkiin yleisiin PC-käyttöjärjestelmiin. Tuo kääntäjä tarjoaa DOS-ohjelmien 32-bittiseen kehittelyyn DOS/4GW-laajentimen (extender).

Vaikka olin päättänyt ottaa tämän kesän rauhallisemmin kuin edellisen (jolloin urakoin "Survo ja minä" -kirjan), innostuin Kimmon kannustamana ja myötävaikutuksella tarttumaan tuohon vaihtoehtoon.

Aluksi oli tavoitteena tehdä 32-bittiset laajennukset pelkästään niistä Survon toiminnoista, joissa rajat pikimmin tulevat vastaan. Huomasin nimittäin ilokseni, että tiettyjen temppujen avulla oli mahdollista kutsua SURVO 84C:n toimittimesta myös 32-bittisiä ohjelmamoduleja aivan samaan tyyliin kuin 16-bittisiä kutsutaan. Niinpä muutamassa viikossa olin saanut aikaan 32-bittiset versiot mm. matriisitulkista, ESTIMATE-operaatiosta ja eräistä muista vaativista tilastollisista toiminnoista. Jo tällöin saatettiin esim. laskea ominaisarvot ja vektorit jopa 1000×1000 -matriisille (90 $\times$ 90 oli aiemmin ehdoton yläraja) ja tehdä faktorianalyysia esim. 700 muuttujan ja mielivaltaisen monen havainnon aineistoilla.

DOS-laajennin antaa oivat keinot hallita koko muistia lineaarisesti teoriassa 2^{32} tavuun eli 4 gigatavuun (4000 megatavuun) asti. Kun koneessa on esim. 32 megatavun muisti, on luonnollista esim. raskaissa iteratiivisissa tilastollisissa toimituksissa poimia laajatin aineistot suoraan muistiin, jolloin toiminta ripeytyy olennaisesti (tyypillisesti 10-kertaiseksi). Myös sisäisessä laskennassa saavutetaan tyypillisesti ainakin 30 prosentin, joskus jopa 100 prosentin nopeuden kasvu. Tämä perustuu siihen, että kääntäjä käyttää 386-prosessorin käskykantaan. Uusi Survo ei siis enää tule toimimaan XT- eikä AT-sarjan koneilla, joista aika lienee jo muutenkin jättänyt.

Tuo parannus ei kuitenkaan sallinut toimituskenttien koon kasvatusta nykyisestä 64 kilotavusta (siis esim. 600 rivin ja 100 sarakkeen kentistä) ylöspäin, sillä Survon toimitin - koko työskentelyn keskus oli edelleen 16-bittinen ohjelma. Heinäkuun alussa päätin - lupaamatta kellekään mitään - hiljaa kokeilla, mahtaisiko senkin avartaminen onnistua tuon DOS/4GW-laajentimen avulla. Suuremmista vastuksista se sitten tapahtui ja toimituskentän saattoi ulottaa vaikka 100000-riviseksi. Tämä tahtoo sanoa, että melkein kaiken, minkä olen itse kirjoitellut viime vuosina erilaisiin toimituskenttiin, voisi nyt sulloa yhteen. Tuommoisella tietojen niputuksella ei kuitenkaan liene mitään tolkkua - halusin vain antaa käsityksen laajennuksen suomista ulottuvuuksista.

Nyt tuli vastaan myös ongelmia. Survon käyttäjät ovat tottuneet siihen, että esim. pienissä laskutoimituksissa vastaus ilmestyy toimituskenttään silmänräpäyksessä. 32-bittisellä toimittimella vaatimattomatkin editorin ulkopuolella tapah-

tuvat - työt veivät hiukan lisääaikaa (noin 0.2 sek.). Syynä oli DOS/4GW-laajennin, joka on yli 200-kiloisena ohjelmana latautui aina uudelleen muistiin jokaisen ulkoisen 32-bittisen modulin käynnistystä varten. Nykyisen toimittimen vikkelyyteen tottuneena koin pienenkä hidastelun sietämättömäksi. Laajoissa sukroissa, pienten toimenpiteiden kasautuessa sadoittain, jopa tuhansittain, myös näin kerrostuvat turhat viivytykset lisäsivät ärsyttävästi kokonaisaikaa.

Pelastukseksi jouduinkin hankkimaan saman laajentimen "paraativersioon" DOS/4G, jota on mahdollista pitää muistissa vakituisesti. Tällöin tuo 0.2 sekunnin ylimääräinen latausaika häviää ja toimintojen välittömyys palautuu. Koko heinä-elokuun olenkin sitten muuntanut satoja Survon ohjelmamoduleita 32-bittisiksi. Kimmo on tehnyt samoin omien moduliensa osalta. Tätä kirjoitettaessa valtaosa Survosta on jo siirretty uudeksi kokonaisuudeksi, josta tullaan käyttämään nimitystä **SURVO 98**.

32-bittisyys kasvattaa luonnostaan ohjelmamodulien kokoa. Tätä kasvua on hillitty kasaamalla samoihin EXE-tiedostoihin toisiinsa liittyviä toimintoja. SURVO 98:ssa on selvästi vähemmän ohjelmatiedostoja, vaikka se sisältää samat ominaisuudet kuin SURVO 84C. Esim. itse toimittimeen (ohjelmatiedosto _S.EXE) on kerätty takaisin lukuisia editoriaaliseen työskentelyyn kuuluvia perustyökaluja, jotka tilansäästön vuoksi olivat hajautettuna useaksi moduliksi. Jopa editoriaalinen laskenta kuuluu suoraan toimittimeen, jonka koko on noin 225 kilotavua (85 kilotavua SURVO 84C:ssä). Vaikka yhdistelyjä on näin tapahtunut, Survon modulaarisesta rakenteesta, mikä on sen kantavia puolia, ei ole tippaakaan luovuttu. Modulien koot vain ovat 100-300 kilotavua aiemman 20-150 kilotavun asemasta.

DOS/4G-laajennin, josta vastaa Tenberry Software, Inc. Bostonin lähellä, ei valitettavasti ole (toistaiseksi) aivan virheetön ohjelma. Huomasin elokuun puolivälissä saatuani riittävän määrän perusmoduleja SURVO 98:n piiriin, ettei suomenkielinen opetussarja toiminut loppuun asti, vaan laajennin keskeytti toiminnan omaan virheilmoitukseensa. Vika on mm. siinä, että laajennin lisää aiheetta DOS-perusmuistin kulutusta 32 tavun verran joka modulin aktivoinnin yhteydessä. Tällä virheellä ei ole mitään yhteyttä Survoon, sillä saman "tuhlauksen" saa näytetyksi toteen hyvin yksinkertaisilla toisiaan kutsuvilla pikkuohjelmilla. Valmistaja on myöntänyt virheen, mutta ei uskalla luvata sille nopeaa korjausta. Niinpä tekemällä viikon verran lisätöitä olen saanut virheen kierretyksi eikä siitä pitäisi olla enää haittaa.

Tavoitteena on, että SURVO 98 tulisi toimimaan, paitsi puhtaissa DOS-koneissa, myös samoissa yleisissä ympäristöissä kuin SURVO 84C eli OS/2, Windowsit ja mahdollisesti jopa Linux. Windows 95:n suhteen ei ole mitään ongelmia, mutta Windows 3.1 ja Windows NT aiheuttivat hieman erityisjärjestelyjä. Niiden DOS-tuki on heikko erityisesti 32-bittisten DOS-laajennusten osalta, mutta tuonkin ongelman olen Kimmon kanssa saanut ratkaistuksi.

SURVO 98 tulee olemaan saatavilla lokakuun alusta lähtien. Sen toiminta edellyttää myös SURVO 84C:n päivitystä versioon 5.30. SURVO 98 kuluttaa levytilaa suunnilleen saman verran kuin SURVO 84C eli yhteinen vaatimus on noin 30 megatavua levyllä. Koneen tulee olla ainakin 386, mieluiten 486 tai Pentium. Keskusmuistia saisi olla 16 megatavua tai enemmän.

SURVO 98 on pieniä poikkeuksia (jotka luetellaan alla) lukuunottamatta yhteensopiva SURVO 84C:n kanssa. Siis SURVO 84C -työt tehdään SURVO 98:ssa täsmälleen samalla tavalla. Myös tiedostorakenteet ja esim. sukat ovat yhteensopivia. SURVO 98:lla laaditut sovellukset kelpaavat SURVO 84C:lle, ellei ylitetä jälkimmäisen rajoituksia esim. toimituskenttien ja matriisien suhteen.

Toistaiseksi, varmuuden vuoksi, SURVO 98 toimii ikäänkuin SURVO 84C:n jatkeena, vaikka uusi Survo onkin täysin itsenäinen, aito 32-bittinen, suojatussa (protected) moodissa toimiva järjestelmä. SURVO 98 käynnistetään suoraan (esim. DOS-tasolta) komennolla SURVO 98 tai SURVO 84C:stä esim. komennolla _S tai näppäimellä ctrl-D. Samalla näppäimellä voi palata takaisin taustalla olevaan SURVO 84C:hen ja jatkaa tekeillä olevaa työtä. Näin yhdellä napinpainalluksella saattaa liikkua edestakaisin uuden ja vanhan Survon välillä ja vertailla niiden toimintoja. Muita kuin nopeuseroja ei pitäisi ilmetä.

SURVO 98:n edut SURVO 84C:hen verrattuna ovat mm. seuraavat:

Toimituskenttien koot ovat käytännössä rajattomat. Kentän maksimileveys on 996 merkkiä. Suurin mahdollinen rivimäärä riippuu keskusmuistin koosta. Mielettömän suurten kenttien (eli esim. 100000-rivisten) käyttöä on syytä kuitenkin välttää eikä suuria numeroaineistoja kannata muuten kuin poikkeustapauksissa pitää jatkuvasti toimituskentässä. Kenttien avartumisesta on varmasti monelle käyttäjälle hyötyä mm. muualla syntyneiden sotkuisten tekstitiedostojen muokkauksessa ennen kuin ne siirretään esim. Survon datatiedostoiksi.

Kookkaat (SURVO 84C:n kapasiteetin ylittävät) toimituskentät talletetaan tiivistettyinä niin, etteivät esim. tyhjät rivit eivätkä rivien tyhjät loppuosat vie lainkaan tilaa.

Suurissakin (usean tuhannen muuttujan) datatiedostoissa voidaan käyttää pitkiä muuttujanimiä. Datatiedostot ovat muuten rakenteeltaan täsmälleen samoja kuin SURVO 84C:ssä.

Tilastollisissa toiminnoissa muuttujien lukumäärät voivat olla useita satoja ja jopa tuhannen luokkaa. SURVO 84C:ssä ehdoton yläraja oli 90 ja joissakin tilanteissa vain muutamia kymmeniä.

Matriisien aikaisempi maksimikoko (8100 alkiota) on nyt käytännössä vain keskusmuistin koosta riippuva. On siis mahdollista laskea luokkaa 1000*1000 olevilla matriiseilla (aikaisemman 90*90 asemasta).

Useimmat toiminnot ovat ainakin hieman nopeampia kuin aikaisemmin. Tyyppillinen parannus on 30% joskus jopa 100%, kun puhutaan puhtaasta laskennasta. Joissakin suuren aineiston hallinta- ja analysointitehtävissä nopeus voi olla 10-kertainen.

Myös kuvaruutugrafiikka on nopeampaa ja SURVO 98 tukee sekä SVGA- (800*600) että XRES- (1024*768) grafiikkaa.

Jo tähän mennessä - tilanpuutteen poistuessa - on ollut mahdollista lisätä Survon toimittimen "älykkyyttä". Esim. enää ei tekstiä kirjoitettaessa törmätä ilmoitukseen varjorivien loppumisesta vaan toimitin lisää niille tilaa automaattisesti.

SURVO 98:n rajoitukset (toistaiseksi):

Kuvaruutugrafiikassa kuvien talletustavat eivät ole yhteensopivia. Siis SURVO 84C:llä tehdyt SPX-tiedostot (OUTFILE) eivät kelpaa SURVO 98:lle. Tämä johtuu, siitä, että C-grafiikkakirjastofunktiot `_getimage` ja `_putimage` eivät ole yhteensopivia Microsoftin ja Watcomin C-kääntäjillä.

Kuvien kerrostaminen suoraan graafisessa moodissa (OVERLAY-täsmennyksellä) ei ainakaan toistaiseksi ole mahdollista. INFILE- ja OUTFILE-täsmennykset kuitenkin toimivat.

CHILD-komento (esim. toisen Survon kutsumiseksi) ei ole käytettävissä. On parempi toimia moniajojärjestelmässä ja käyttää siinä tarvittaessa useampaa Survoa.

Suomenkielinen oikoluku (OIO-komento) puuttuu.

Supermatriisikomennot on jätetty pois (tarpeettomina).

SURVO 98

SURVO 98 on hankittavissa lokakuun alusta lähtien. Sen toiminta edellyttää myös SURVO 84C:n päivitystä versioon 5.30.

Ne jotka hankkivat SURVO 98:n tämän vuoden aikana, tulevat saamaan ilmaisen päivityksen helmikuussa 98.

Vanhat Survo-käyttäjät hankkiessaan SURVO 98:n tämän vuoden aikana saavat lisäksi 30 %:n alennuksen alla olevista listahinnoista.

SURVO 98 -hinnasto 1.10.1997 - 31.12.1997

Käyttöoikeus - 1.2.1999 sisältäen päivityksen helmikuussa 1998:

Normaalihinta	4.800 mk
Tutkijahinta	3.500 mk

Lehden teemaan liittyen käyttäjäyhdistyksen jäsenille tarjotaan nyt myös mahdollisuus hankkia Seppo Mustosen kirja "Survo ja minä" erikoishintaan 120 mk.

Tiedustelut:

ComPetit Consulting Oy

✉ Mäenrinne 11, 02160 ESPOO

☎ +358-9-452 1951

Marjut Schreck

Helsingin yliopiston tilastotieteen laitos tulee vastaamaan Survo 98:n jakelusta Helsingin yliopistossa erikseen sovittavilla ehdoilla.

Tiedustelut:

Seppo.Mustonen@helsinki.fi
Kimmo.Vehkalahti@helsinki.fi

Tilastollisen tietojenkäsittelyn seminaari

Tilastotieteen laitos, Helsingin yliopisto

Seminaari kokoontuu tiistaisin klo 16-18 tilastotieteen laitoksen mikroluokassa (Unioninkatu 37, käynti Snellmaninkadun puolelta).

Syksyn 1997 ohjelma on seuraavanlainen:

- 23.9 SURVO 98 tulossa
- 30.9 Jouko Manninen: Hiiri PostScript-kuvan teossa
- 7.10 Kimmo Vehkalahti: KuulumISia Istanbulista
- 14.10 Survo ja me I (uusimman SURVO.TUT-lehden aineiston esittelyä)
- 21.10 SURVO 98:n käyttötavat vs. SURVO 84C
- 4.11 Survo ja me II (uusimman SURVO.TUT-lehden aineiston esittelyä)
- 18.11 klo 13-18: Uuden SURVO 98:n käyttökokemuksia

Survo tutkijan työvälineenä

Tilastotieteen laitos, Helsingin yliopisto

Kurssi pidetään torstaisin klo 13-17 tilastotieteen laitoksen mikroluokassa (Unioninkatu 37, käynti Snellmaninkadun puolelta).

- 25.9 SURVO 98:n ohjelmointi
- 2.10 Ideoita Survon kehittämiseksi
- 9.10 Sukrokielen perusteet I
- 16.10 Sukrokielen perusteet II
- 23.10 Tutkimussovellusten rakentaminen sukrokielellä
- 30.10 (ei kokoontumista)
- 6.11 Tietokantojen hallinta
- 13.11 Simulointikokeet ja niiden suunnittelu
- 20.11 Epälineaarinen regressioanalyysi ja suurimman uskottavuuden estimointi
- 27.11 Tutkimustulosten esittäminen ja julkaiseminen
- 4.12 Kvalitatiivinen tiedonhallinta ja analyysi (LIST-operaatiot)

22.9.1997

Seppo Mustonen

Marjut Schreck:

Toimiston tarpeisiin

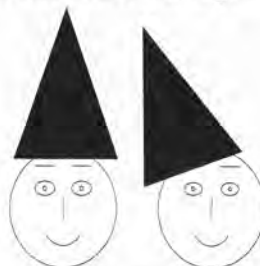
Microsoft mainostaa suurena ja mullistavana uutuutena Office-pakettiaan. Käsittääkseni toimistoruutiineja on kuitenkin ollut pakko hoitaa niin kauan kuin toimistoja on ollut olemassa. Alussa riittävinä välineinä olivat kynä, kumi, kirjoituskone ja helmitaulu, joka myöhemmin korvautui laskukoneella. Näillä välineillä pystyttäisiin yhä hoitamaan monen toimiston juoksevia asioita. Kirjoituskone alkaa kuitenkin olla historiaa, ja tuskinpa yhdeltäkään työpaikalta löytyy pöytää, jota tietokone ei olisi valannut. Kun tutkii hieman tarkemmin, usein sen tietokoneen viereltä kuitenkin vielä löytyy laskukone. Sen sijaan, että tietokone olisi se väline, jolla myös lasketaan, tehdään budjetoiteja, pidetään yllä asiakasrekisteriä, hoidetaan kirjanpitoa ja seurataan yrityksen kehitystä, se valitettavasti monelle on todellakin vain kirjoituskoneen korvike. Ehkä tämä kaikki nyt sitten muuttuu, kun hankitaan koneeseen office-paketti.

Koneiden ja ohjelmien vaatiman kapasiteetin kasvusta löytyy kuvaava tarina eräästä paperitehtaan laadunvalvontaan osallistuvasta henkilöstä, joka 386-mikrollaan ja Survolla vastaanottaa paperikoneen toimintaan liittyviä mittauksia useita tuhansia vuorokaudessa. Ensi töikseen aamuisin hän tuottaa raportin koneen toiminnasta tehtaan johdolle. Tulosten muokkaamiseen ja raportin tuottamiseen kului aikaa reilu tunti, mikä merkitsi aikaista päivänaloitusta, jotta tulokset olisivat johtoryhmällä ajoissa. Sain kuitenkin jokin aika sitten soiton, jossa ko. henkilö tyytyväisenä kertoi, että tehtaan sihteeristö oli vaihtanut tekstinkäsittelyohjelmaan, joka vaati koneen oltavan Pentium-tasoa. Tästä johtuen oli mahdollista, että hän sai vaihdettua koneensa erältä sihteeriltä jääneeseen 486:seen ja nyt "aamuraportin" tekoon kuluu Survolla enää vain vaivaiset 15 minuuttia. Mihinkä joutuisimme ilman tekstinkäsittelyohjelmia?

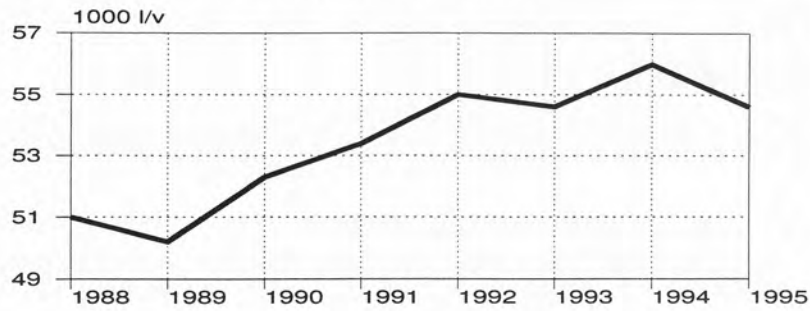
Aluksi talon tavoille

Oma MS-Office-pakettini (lue Marjut Schreck-office) on kehittynyt pikkuhiljaa, mutta ollut olemassa niin kauan kuin olen Survoa käyttänyt (n. 14 vuotta). Näin luottamuksellisesti voin kerota, että Survoon törmättyäni en ole muita ohjelmia tarvinnutkaan. Ensin koneeni oli Wangin PC ja siinä Basic-Survo vaihtuen sitten IBM AT:hen ja C-kieliseen Survoon päätyen nyttemmin 133 MHz Pentiumiin. Konekehitys ei ole ollut Survon vaatimaa vaan koitunut lähinnä omaksi iloksi töitä nopeuttamaan. Survo kulkee tietenkin mukani myös käsilaukussa 300 grammaa painavassa kooltaan taskulaskinta vastaavassa HP:n 200 LX:ssä, jossa on 20 MB kiintolevy.

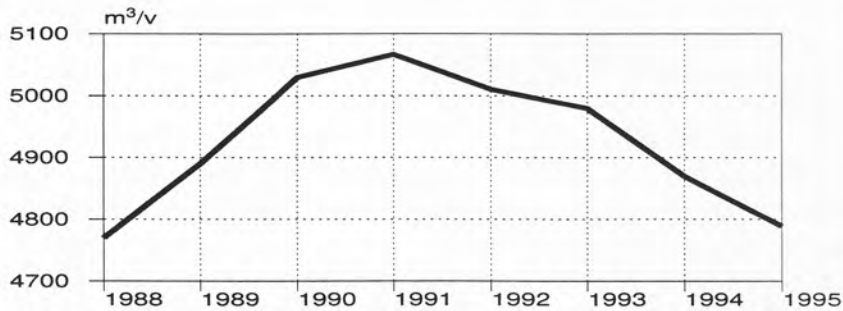
Alunperin kehittelin MS-Officea taloyhtiön ja isännöitsijän tarpeisiin. Asukasrekisteristä tuotettiin kaikki tarvittavat asiakirjat ja lomakkeet osoitetarroista kokouskutsuihin ja isännöitsijäntodistuksiin. Erilaisia teknisiä tietoja, kuten sähkö ja vesi, varten oli omat tiedostonsa, joiden avulla seurattiin ja raportoitiin kehitystä. Kaikkia tulosteita varten oli omat toimituskenttänsä, joten niiden tuottaminen käytännössä tarkoitti ainoastaan muutaman ESC-napin painallusta. Usein osakas tuli kysymään kuinka nopeasti saisi isännöitsijäntodistuksen, jolloin hain valmiin pohjan, lisäsin ko. asunnon numeron, listasin tiedot asukasrekisteristä ja tulostin lomakkeen, keskimäärin siis kymmenkunta sekuntia. Kiitosta tuli ansaitusti lomakkeiden ulkoasusta ja "vakavia" tiedotteita pystyi hauskuuttamaan Survolle ominaisin keinoin.



ÖLJYN KULUTUKSEN SEURANTA 1988-1995



VEDEN KULUTUKSEN SEURANTA 1988-1995



ISÄNNÖITSIJÄNTODISTUS

27.1.1997

Asunto Oy Westendin Mäenrinne 11 c/o Schreck Mäenrinne 11 F 02160 ESPOO	Marjut Schreck, isännöitsijä Mäenrinne 11 F 29 02160 ESPOO p. 4521951 työ / 422617 koti
---	--

KIINTEISTÖTUNNUS

Kunta Espoo	Kaupunginosa 13	Kortteli 13037	Tontti 11
----------------	--------------------	-------------------	--------------

TIETOJA HUONEISTOSTA JA OSAKKEISTA

Osakkeen numerot ja lukumäärä 28861-29640	Porras ja huoneisto F29	Kerros 1-2	Huoneisto- tyyppi	Yht. järj. muk. pa 78	Mahd. turk. pa
Hoitovastike mik 858,-	Rahoitusvastike mik --	Vastike- peruste m²	Vesi yms. maksut mik 35,-	Maksun peruste henkilö/kk	
Osuus yhtiön laineista (bainaus maksuttavissa)		Maksuttomat vastikkeet korkoineen		Ajalta	
Osakeluottoon merkitty omistaja Schreck Lauri ja Marjut			Merkitty osakeluottoon 12.8.1983		Huoneiston käyttötarkoitus asuinhuoneisto

TIETOJA YHTIÖSTÄ

Kaupparekisteri- merkinnän pvm 10.10.1981		Kiinnitykset mik 300.000,-		Yhtiön lainat pvm / mik ---		
Asunnot kpl 29		Pinta-ala yhti. m ² 2964		Osakkeita 29640	Autopaikat 31 katos / 20 muut	
Vakuutusyhtiö Sampo		Vakuutuslaji Täysarvo		Säilytystilat kellarikomero, pyöräkellari		
Tontti oma	Pinta-ala 8372	Rakennus- ten määrä 6	Kerrosluku 1-2	Tilavuus 10315	Valmistus- vuosi 1981	Talo- tyyppi rivitalo
Raken- nuskine bet.elem	Katto- tyyppi harja	Kate tiili	Lämmitys oma öljy	Ilmanvaihto koneellinen	Antenni- jäjest. minikaap.	TV- kanavat 14+2
Lämmönlukitus l/v 1995 54630		Vedenkulutus m3/v 1995 4789		Sähkönkulutus kwh/v 1995 33781		

Liitteet Tuloslaskelma ja tase vuodelta 1995 Toimintakertomus vuodelta 1995, yhtiöjärj.	Isännöitsijän allekirjoitus
---	-----------------------------

Pankkipalveluiden pariin

Olen toiminut Survo-Käyttäjähdistyksen sihteerinä sen perustamisesta lähtien. Yhdistyksen myötä tarpeet hieman muuttuivat ja sen seurauksena MS-Officeen tuli mukaan pankkisiirtojen tuottaminen ja viitenumeroiden laskenta jäsenmaksuja varten.

Kun kyselin pankista ohjeita viitenumeroiden laskentaan, vastattiin, että he tuottavat valmiin listan viitenumeroista, jotka sitten voi naputella esipainettuihin laskulomakkeisiin. Ei niitä millään muulla kuin pankin omalla systeemillä tai sitä varten tehdyllä erillisellä ohjelmalla kykene laskemaan. Sain kuitenkin laskukaavan kuulun itsepäisyyteni ansiosta ja kun vielä ilmoitin, että aion tehdä itse myös pankkisiirtolomakkeet, sain lisäksi myös aika omalaatuisen maineen paikalliskonttorissa.

Pankkisiirtopohjan laadin yksinkertaisesti Survon PLOT /FRAME-kaaviolla, jolla saa piirrettyä erilaisia "laatikoita" ja raameja sekä sijoitettua niihin haluttuja tekstejä ja varjostuksia. Viitenumeroiden laskentaa varten tein oman toimituskenttänsä, jossa ne on helppo laskea uudelleen jäsenmäärän muuttuessa.

Viitenumero lasketaan siten, että haluttu numerosarja, tässä tapauksessa jäsennumero, kerrotaan siten että ensimmäinen luku kerrotaan 1:llä, seuraava luku 3:lla ja sitä seuraava 7:llä. Tätä toistetaan kaikkien lukujen osalta ja kerrotuista luvuista lasketaan summa, joka vähennetään seuraavasta tasakymmenestä. Tämä erotus on pankin vaatima ns. tarkistenumero. Viitenumeroksi tulee siis alkuperäinen numerosarja, jossa liitteenä tarkistenumero. Yksinkertaista vai mitä, mutta miten määrittely onnistuisi "jokapojan" taulukkolaskentaohjelmilla?

Tein tähän lisäksi vielä yksinkertaisen suikon. Se laskee viitteet ja sen jälkeen hakee rekisteristä tarvittavat tiedot, siirtää ne pankkisiirtopohjaan sekä jäsenkirjeeseen, jonka lopulta tulostaa. Minun rooliksi jäi enää lisätä paperia kirjoittimeen. Kun vielä vaihdoin PostScript-laserini Lexmarkin Optra R+:saan, jonka tarkkuus on 1200 dpi ja nopeus 16 sivua minuutissa, ei muutaman sadan jäsenkirjeen tekemiseen kulu aikaa kuin kymmenisen minuuttia.

Vein erään tällaisen jäsenmaksukirjeen pankkisiirtoineen testattavaksi konttoriin, josta olin aiemmin ohjeita pyydellyt, ja ilmeet olivat katsomisen arvoiset. Millä ihmeen ohjelmalla olin moisen tehnyt, oli monen kysymys. Ainoa mihin Survolla en pystynyt, oli repäisyohenne pankkisiirron irrottamiseksi kirjeestä. Sen teitin tutussa painossa. Todennäköisesti siihenkin löytyy jatkossa keinonsa.

```
- - SURVO 84C EDITOR Wed Sep 10 11:18:06 1997      D:\LEHTI197\ 120 120 0
1 *save psiirto / Pankkisiirtolomake
2 *
3 *PLOT /FRAME / DEVICE=PS,PANKKI1
4 *PEN=[SWISS(6)] FRAME=0
5 *FRAMES=A,B,C,D,E,F,G,H,I,J,K
6 *A=[line_type(0)][line_width(1)],0,0,2100,1010,0
7 *B=[line_type(0)][line_width(1)],1525,0,575,1010,0
8 *C=[line_type(0)][line_width(1)],1525,145,575,125,0
.
.
20 *TEXT=A1,B1,C1,D1,D2,D3,D4,D5,D6,D7,...
21 *A1=Tililtä_n:o_Från_konto_nr,1570,115
22 *B1=Tilille_n:o_Till_konto_nr,1570,240
23 *C1=mk,1570,330
24 *D1=Viitenumero_Referensnummer,1570,420
25 *D2=Saaja_ja_maksaja_Mottagare_och_betalare,1570,930
26 *D3=[SWISSB(10)],ComPetit,1570,860
27 *D4=[SWISSB(10)],Consulting_Oy,1570,820
28 *
```



```

- - SURVO 84C EDITOR Tue Jul 30 17:27:38 1996 D:\KYHD\ 100 100 0
1 *SAVE VIITTEET / viitenumeroiden laskenta + rekisteri
2 *
3 *FILE UPDATE JÄSENET
4 *
5 *Survo-Käyttäjähdistyksen jäsenrekisteri / ms
6 *
7 *FIELDS:
8 *12 S 1 J1
9 *13 S 1 J2
10 *14 S 1 J3
11 *15 S 2 JLASKU
12 *16 S 4 VIITENRO
13 *END
14 *
15 *
16 *FILE EDIT JÄSENET+
17 *
18 *VAR str(J1)=str(JÄSENNRO,1,1)
19 *VAR str(J2)=str(JÄSENNRO,2,2)
20 *VAR str(J3)=str(JÄSENNRO,3,3)
21 *
22 *VAR VIITE
23 *VIITE=(1*J1)+(3*J2)+(7*J3)
24 *.....
25 *
26 *CLASSIFY JÄSENET,SÄÄNTÖ,VIITE,EROTUS
27 *CLASSIFICATION SÄÄNTÖ
28 *0-10: 10
29 *11-20: 20
30 *21-30: 30
31 *31-40: 40
32 *41-50: 50
33 *51-60: 60
34 *61-70: 70
35 *71-80: 80
36 *81-90: 90
37 *91-100: 100
38 *101-110: 110
39 *END
40 *.....
41 *
42 *VAR JLASKU=EROTUS-VIITE
43 *VAR str(VIITENRO)=str(JÄSENNRO)&str(JLASKU)
44 *
45 *

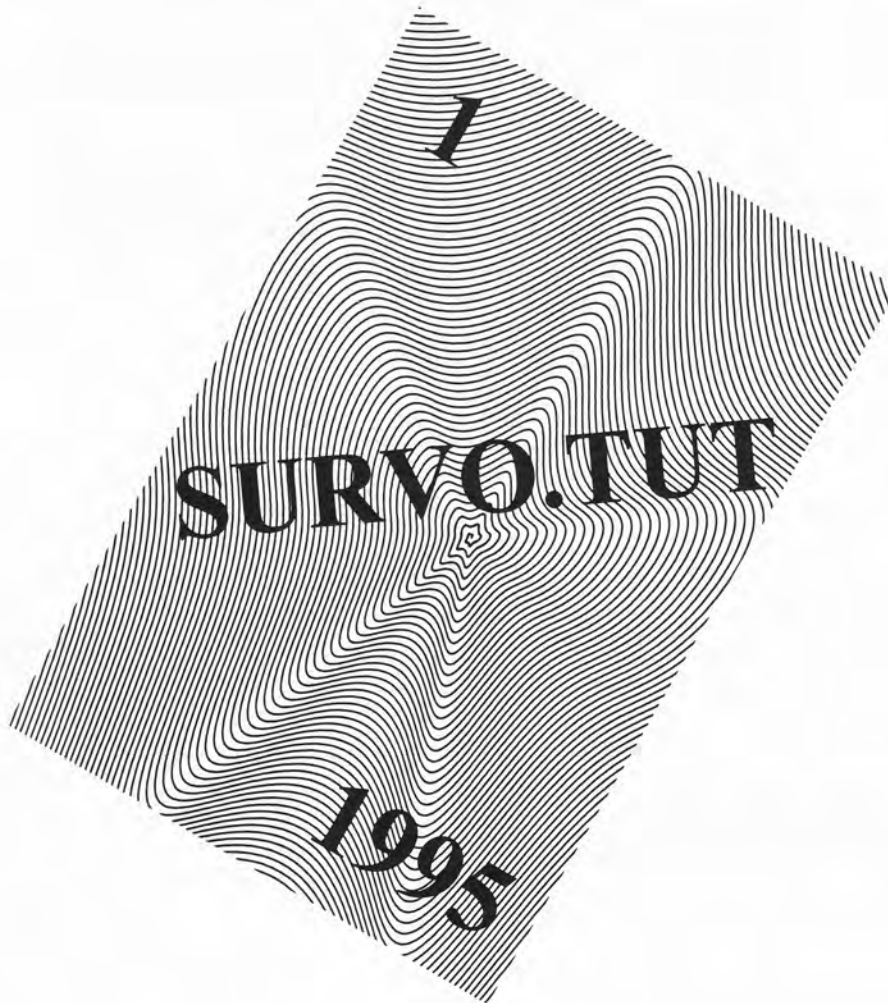
```

Saajan tilinumero, saaja Mottagarens kontonummer, mottagare SHOP Espoo Tapiola 572442-211072 Survo-Käyttäjähdistys ry	TILISIIRTO GIRERING		Saaja ja maksaja Mottagare och betalare Survo-Käyttäjähdistys ry ComPetit Consulting Oy Marjut Schreck Henkilöjäsen 75 mk Yhteisöjäsen 750 mk Jäsenmaksu 1995
Maksaja Betalare ComPetit Consulting Oy Marjut Schreck Allekirjoitus Underskrift	Viitenumero Referensnummer 059 2		Viitenumero Referensnummer 059 2
Tililtä n:o Från konto nr	Eräpv Förf.d. 14 pv	mk	mk
			Tilille n:o Tilli konto nr 572442-211072 Tililtä n:o Från konto nr

Taitaa taittaakin

Erikoisuutena MS-Officessa on toki myös omat pohjansa erilaisten julkaisujen taittoa varten. Esimerkkinä mainittakoon Survo-Käyttäjähdistyksen lehti SURVO.TUT. Survon valmiista esimerkkikuvista pienillä muutoksilla on taiteiltu kannet ja muita Survon operaatioita yhdistelemällä on lehteä tehty jo monien vuosien ajan.

```
- - SURVO 84C EDITOR Wed Jul 31 15:37:38 1996      D:\LEHTI195\ 100 100 0
1 *SAVE KANSI1
2 *
3 *HEADER=
4 *XDIV=0,1,0 YDIV=0,1,0
5 *SIZE=2000,2900
6 *XSCALE=-11(1)12 YSCALE=-8(1)9
7 *LINETYPE=[line_type(0)] [LINE_COLOR(0)] [LINE_WIDTH(1.6)]
8 *FRAME=0 DEVICE=PS,K1
9 *Mutkia tavalliseen spiraaliin sinifunktion avulla:
10 *PLOT X(T)=0.02*T*(1+0.2*sin(3*T))*cos(T),
11 *      Y(T)=0.02*T*(1+0.2*sin(5*T))*sin(T)  T=0,600,0.01
12 *TEXT=T1,T2,T3
13 *T1=[TIMESB(110)],1,400,2500 T2=[TIMESB(110)],1995,1700,200
13 *T3=[TIMESB(110)] [ROT(35)],SURVO.TUT,400,1000
```



Pitää kirjaa

Siirtyessäni päätoimisesti "oman" yrityksen pyörittäjäksi tuli vuoroon kirjanpito-ohjelman teko. Varsinaista C-kielen ohjelmointia en hallitse, joten laadin ohjelman surotekniikkaa hyväksikäyttäen yhdistelemällä Survon olemassa olevia operaatioita. Siitä tuli valikkopohjainen kyselyin etenevä koelma.

ComPetit Consulting Oy / Marjut Schreck puh. 90-4521951

***** KIRJANPITO *****

1. Tapahtumien syöttö
2. Tulosteet
3. Tapahtuman etsintä
4. Tilivuoden vaihto
5. Budjettitietojen syöttö

10. Tilitietojen syöttö
11. Yritystietojen syöttö
12. Tilikausitietojen syöttö
13. Tuloslaskelman ja taseen määrittely

X. LOPETUS



Tilikartat ovat vapaasti käyttäjän määriteltävissä ja muutettavissa. Näistä systeemi tuottaa automaattisesti pää- ja päiväkirjat, tuloslaskelmat sekä taseet käyttäjän haluamalta ajanjaksolta. Alussa annetaan tilitietojen lisäksi yrityksen perustiedot sekä tilikauden alku- ja loppuajankohdat, joita käytetään esimerkiksi eri tulosteiden otsakkeina. Kun syötetään tapahtumia, otetaan pohjaksi tilikartta, josta pyydetään käyttäjää valitsemaan debet- ja kredit-tili, antamaan viennin päiväys, markkamäärä ja vientiselite. Tätä jatketaan kunnes kaikki tapahtumat on syötetty. Ne tallentuvat väliaikaiseen tiedostoon, jolloin käyttäjä voi selailla ja korjata mahdollisia syöttövirheitään.

Kirjanpito-ohjelmani on käytössä muutamissa pienyrityksissä. Erityistä kiitosta on tullut systeemin vapaasta muutettavuudesta ja siitä, että sukro-ohjelmointia osaavan on ollut myös helppoa muuttaa itse ohjelman rakennetta tai vaikkapa lisätä siihen uusia osia tarvittaessa.

Kaiken muunkin yrityksen juokseviin asioihin liittyvän materiaalin teen tietenkin myös Survolla. Valmiita toimituskenttiä toimintakertomuksista, pöytäkirjoihin, tarjouksiin ja laskutukseen löytyy riittävästi. Kaikissa tulosteissa toistuu PostScriptilla tehty yrityksen tunnus. Laskutusta en ole omalta osaltani juurikaan automatisoinut. Se sinänsä tuottaa riittävästi mielihyvää.

ComPetit Consulting Oy

✉ Mäenrinne 11, 02160 ESPOO

☎ +358-9-452 1951

Marjut Schreck

Toimistosta koulunpenkille

Joskus toimistooni, joka sijaitsee kodin yhteydessä, kuuluu muminaa "koulu"-toimiston puolelta. Viime keväänä, jolloin aiheena matematiikassa oli eri suuruusluokkaa olevat luvut, näytti pahasti siltä, ettei asia ollut alkuunkaan hallinnassa. Kun harjoiteltiin yhdessä, huomasin, että tehtävät, joissa luvut oli kirjoitettu kirjaimin ja siitä muutettava numeroiksi, valaisivat asiaa parhaiten. Siispä teke-
mään kirjaimin numeroita kuitenkin niitä kirjoittamatta. Kumpaa silloin käytät, tekstinkäsittelyä vai taulukkolaskentaa?

Niinkuin olet jo varmaan tullut huomanneeksi, Survossa eivät eväät lopu kesken, vaan jo muutamassa minuutissa oli tulostettu sivutolkulla harjoitustehtäviä. Nähtyään ne opettaja oli todennut ihmetel-
len, että kylläpä teidän äiti on nähnyt paljon vaivaa. Salaisuus on edelleen paljastamatta, mutta sinä näet sen oheisesta toimituskentästä.

```
- - SURVO 84C EDITOR Tue Aug 13 10:53:02 1996 D:\SJAMINÄ\ 450 140 0
1 *SAVE KOULU
2 *
3 *DATA LUVUT,A,B,C,D
4 *VAR LUKU=1000+INT(1000000*RND(4521951)) TO LUVUT
5 *
6 *PUTEND A,B,(10:sanoin).=
7 D11111111
8 C LUKU
9 A 224847(10:sanoin).=kaksi sataa kaksi kymmentä neljä tuhatta kahdeksan
10 * 1397(10:sanoin).=tuhat kolme sataa yhdeksän kymmentä seitsemän
11 * 469140(10:sanoin).=neljä sataa kuusi kymmentä yhdeksän tuhatta sata ne
12 * 32464(10:sanoin).=kolme kymmentä kaksi tuhatta neljä sataa kuusi kymm
13 * 866235(10:sanoin).=kahdeksan sataa kuusi kymmentä kuusi tuhatta kaksi
14 * 23797(10:sanoin).=kaksi kymmentä kolme tuhatta seitsemän sataa yhdeks
15 * 384484(10:sanoin).=kolme sataa kahdeksan kymmentä neljä tuhatta neljä
16 * 872003(10:sanoin).=kahdeksan sataa seitsemän kymmentä kaksi tuhatta ko
17 * 991326(10:sanoin).=yhdeksän sataa yhdeksän kymmentä yksituhatta kolme
18 * 516229(10:sanoin).=viisi sataa kuusitoista tuhatta kaksi sataa kaksi k
19 * 907495(10:sanoin).=yhdeksän sataa seitsemän tuhatta neljä sataa yhdeks
20 * 491601(10:sanoin).=neljä sataa yhdeksän kymmentä yksituhatta kuusi sat
21 * 998223(10:sanoin).=yhdeksän sataa yhdeksän kymmentä kahdeksan tuhatta
22 * 603326(10:sanoin).=kuusi sataa kolme tuhatta kolme sataa kaksi kymment
23 * .
24 * .
25 B .
```

Väitkö vastaan ?

Viime aikoina olen Survon kanssa "erikoistunut" väitöskirjoihin. Kaikki alkoi siitä kun kolmisen vuotta sitten menin mukaan Kansanterveyslaitoksen Mielenterveyden osastolle mammografiaseulon-
toihin liittyvään tutkimusprojektiin. Toimenkuvani oli tutkijoiden avustaminen tekemällä tarvittavat laskennat ja analyysit ko. projektiin liittyvillä aineistoilla. Kun parin vuoden uurastuksen jälkeen Ar-
ja Aron väitöskirja oli loppusuoralla, alkoi hän mietiskellä miten sen saisi parhaiten (=mahdollisim-
man nopeasti ja edullisesti) painokuntoon. Ilmoittauduin vapaaehtoiseksi taittajaksi ja ihmettelyä
riitti siitä, mitä kaikkea sillä Survolla voikaan tehdä. Tuli siis konkreettisesti näytettyä monille, mitä
tarkoittaa Lauri Tarkkosen toteama Survosta; ainoa julkaisuohjelma, jolla voi tehdä myös faktori-
analyysin. Kun kirja sitten valmiina kiersi talossa, sain moneen kertaan kehuja sen ulkoasusta ja tyy-
tyväisyys onkin poikunut jo uusia väitöskirjoja työstettäväkseni. Valmiina hyllyssä niitä on jo neljä.
Jokainen projekti on alkanut kommentilla; hei, sullahan on käytössä se tosi hyvä taitto-ohjelma.

Suuri Unelma Ratkaista Vaikeita Ongelmia

Toimenkuvani tällä hetkellä on juoksuhenkilön, kirjanpitäjän, keskuksen ja Survon käyttäjien neuvonnan ja koulutuksen ohella erilaisten asiakassovellusten rakentaminen Survon "päälle". Aiheet vaihtelevat automatisoiduista raporteista, laboratoriotulosten käsittelyyn, johdon informaatiojärjestelmistä asiakastytyväisyystutkimuksiin ja henkilöstöhallinnon seurantasysteemeistä automaattilaskutukseen.

Näistä sovelluksista yhdistelemällä on koottavissa sellainen office-paketti, jollaista ei mistään muusta ympäristöstä löydy. Koneelta vaaditaan ainoastaan kiintolevy ja käyttäjältä luku- ja kirjoitustaito. Näissä kuvioissa ei hiirilavantauti eikä ikonisokeus pääse yllättämään.

Yhteenvetona siis todettakoon, että niin oman yritykseni, Survo-Käyttäjäyhdistyksen, isännöimäni taloyhtiön, lasten koulutehtävien, perheen emännöinnin sekä edellämainittujen asiakassovellusten hallintaan käytetään ainoastaan yhtä ohjelmaa. **Suurta Unelmaa Ratkaista Vaikeita Ongelmia.**

**S uuri
U nelma
R atkaista
V aikeita
O ngelmia**

Kimmo Vehkalahti:

Taulukoiden tunnistus tekstitiedostoista

Tilastoaineistoja voi saada mitä erilaisimmista lähteistä, esim. tutkijoilta, tietopankeista, Tilastokeskuksesta tai vaikkapa Internetistä. Tiedot on tyypillisesti esitetty erilaisina taulukoina. Kun niitä siirrellään paikasta toiseen, on paras siirtomuoto tavallinen tekstitiedosto. Siitä tiedot tulisi saada ohjelmiston omaan muotoon mahdollisimman kätevästi, jotta pääsisi heti aloittamaan varsinaisen työskentelyn aineiston parissa.

Ohjelmilla on myös kykyä suoraan ymmärtää toistensa tiedostomuotoja, mutta näihin kykyihin ei pidä juuri luottaa. Joissakin ohjelmistoissa on tapana muuttaa version vaihtuessa myös tiedostorakenteita, jolloin eri versioilla tehdyt tiedostot eivät olekaan yhteensopivia - eivät aina edes saman ohjelman vanhempien versioiden kanssa! Survoon tottunut voi olla rauhassa, koska Survossa taataan yhteensopivuus taaksepäin.

Tekstitiedostoa ei mikään hetkauta; mitä nyt eri käyttöjärjestelmissä on pieniä eroja rivinvaihtomerkkien ja merkkivalikoimien suhteen. Niistä selviää sopivilla muunnoksilla, joita Survossa tarjoaa esim. sukro /TXT-CONV.

Pitäydytään nyt aluksi tavallisessa ja varsin yksinkertaisessa tekstitiedostossa TEOLL1.TXT. Siihen on listattu tietoja Ahvenanmaan ja Lapin läänin kunnista, joissa teollisuudessa toimivien osuus työvoimasta oli joskus kymmenisen prosenttia - veroäyrestä päätellen vuosia sitten.

Kunta	Lääni	Väestö	Synt.	Maamet	Palvelu	Äyri	Tulotaso
Geta	AHV	468	3	4	3	14.50	11178
Föglö	AHV	590	4	4	4	14.50	13273
Eckerö	AHV	705	5	2	5	14.50	13753
Lemland	AHV	902	11	2	5	14.00	13179
Maarianhamina	AHV	9581	129	0	8	16.00	19178
Pelkosenniemi	LAP	1569	19	4	4	17.50	9813
Savukoski	LAP	1954	27	5	3	17.00	13440
Enontekiö	LAP	2259	29	2	5	17.00	10401
Kittilä	LAP	6490	86	4	4	17.50	10725
Ranua	LAP	5637	106	5	3	18.00	8532
Salla	LAP	7414	121	4	4	18.00	9802
Sodankylä	LAP	10119	140	3	5	17.00	12439

Kuntataulukko pitää siis saada tekstitiedostosta havaintoaineistoksi. Tämä pikkutehtävä toimii johdatuksena luvun varsinaiseen aiheeseen eli taulukoiden tunnistamiseen. Katsotaan aluksi, miten muutamat tilasto-ohjelmat ja Survo selviävät tehtävästä.

Ei pelkkää plussaa

S-Plus:an Windows-versiolla (3.3) työ käy kohtalaisen yksinkertaisesti: valitaan Import, jonka takaa löytyvään valintaikkunaan annetaan tiedoston nimi ja kerrotaan, että muuttujien nimet ovat ensimmäisellä rivillä. Näin kuitenkin kuntien nimistä tulee eräänlaisia rivitunnusiksi. Muuttujien nimissä "ääkköset" eli skandinaaviset kirjaimet muuttuvat pisteiksi. Lienee parempi käyttää komentotasolta funktiota read.table. Luodaan siis näin data teoll1.

```
> teoll1 <- read.table("teoll1.txt", header=T, row.names=NULL)
> teoll1
```

	Kunta	L..ni	V.est.	Synt.	Maamet	Palvelu	.yri	Tulotaso
1	Geta	AHV	468	3	4	3	14.5	11178
2	F\224gl\224	AHV	590	4	4	4	14.5	13273
3	Ecker\224	AHV	705	5	2	5	14.5	13753
4	Lemland	AHV	902	11	2	5	14.0	13179
5	Maarianhamina	AHV	9581	129	0	8	16.0	19178
6	Pelkosenniemi	LAP	1569	19	4	4	17.5	9813
7	Savukoski	LAP	1954	27	5	3	17.0	13440
8	Enonteki\224	LAP	2259	29	2	5	17.0	10401
9	Kittil\204	LAP	6490	86	4	4	17.5	10725
10	Ranua	LAP	5637	106	5	3	18.0	8532
11	Salla	LAP	7414	121	4	4	18.0	9802
12	Sodankyl\204	LAP	10119	140	3	5	17.0	12439

Sarakkeet huojuvat, koska S-Plus ymmärtää vain englantia. Sille oudot merkit se esittää 8-järjestelmän koodeina. Lienee siinä föglöläisillä ihmettelemistä.

Sotkujen selvennykseksi käytän Survon mainioita muunnoksia, joilla nähdään, että $224(8:10)=148$ ja IBM:n PC-merkkivalikoimassa $148(10:ascii)=ö$.

Kaksi ässää ja dollari

SAS:in unix-versiolla (6.09) tehtävä ei oikein ota sujuakseen ellei auta ohjelmistoa muuttujien määrittelyjen kanssa. Tekstitiedostoon ei tarvitse tehdä muutoksia, mutta muuttujien nimiä ja tyyppejä ei sieltä suoraan saa. Muuttujien nimissä ei kelpuuteta ääkkösiä eikä pisteitä. Merkkijonomuuttujat on ilmoitettava \$-merkein ja niiden pituuksia on ilmoitettava erikseen unohtamatta tällöin kaksoispistettä - muussa tapauksessa kuntien nimet pätkitään kahdeksan merkin mittaisiksi. Muuttujien nimet SAS haluaa välttämättä esittää isoin kirjaimin.

```
DATA TEOLL1;
INFILE 'teoll1.txt' FIRSTOBS=2;
INPUT Kunta : $13. Laani $ Vaesto Synt Maamet Palvelu Ayri Tulotaso;
```

Oheinen SAS-ohjelma suoritutetaan SAS:illa vanhaan eräajotyyliin. Tulostetaan näin saatu havaintoaineisto PROC PRINT:illä.

OBS	KUNTA	LAANI	VAESTO	SYNT	MAAMET	PALVELU	AYRI	TULOTASO
1	Geta	AHV	468	3	4	3	14.5	11178
2	Föglö	AHV	590	4	4	4	14.5	13273
3	Eckerö	AHV	705	5	2	5	14.5	13753
4	Lemland	AHV	902	11	2	5	14.0	13179
5	Maarianhamina	AHV	9581	129	0	8	16.0	19178
6	Pelkosenniemi	LAP	1569	19	4	4	17.5	9813
7	Savukoski	LAP	1954	27	5	3	17.0	13440
8	Enontekiö	LAP	2259	29	2	5	17.0	10401
9	Kittilä	LAP	6490	86	4	4	17.5	10725
10	Ranua	LAP	5637	106	5	3	18.0	8532
11	Salla	LAP	7414	121	4	4	18.0	9802
12	Sodankylä	LAP	10119	140	3	5	17.0	12439

SAS:in Windows-versio (6.10) ei tarjoa mitään tämän kummempaa.

Kolme ässää ja tähdenlento

SPSS:n unix-versio (5.0) muistuttaa SAS:ia. Muuttujien nimissä ei kelpuuteta pisteitä, ei etenkään niitä tiettyjen vokaalien päällä käytettäviä. Myös SPSS suosii suuria kirjaimia. Sanalliset muuttujat täytyy "ajovirrassa" erikseen ilmoittaa pituuksineen. Lisäksi tulee erikseen "tähdentää" numeeristen muuttujien roolia. Muuttujien nimet tiedoston alussa häiritsivät SPSS:ää, kunnes poistin ne (!).

```
DATA LIST FILE 'teoll1b.txt' FREE / Kunta (A13)
Laani (A3) Vaesto * Synt Maamet Palvelu Ayri Tulotaso .
```

Listaus sisältää 60 ylimääräistä nollaa - vaikuttaako tarkemmalta?

KUNTA	LAANI	VAESTO	SYNT	MAAMET	PALVELU	AYRI	TULOTASO
Geta	AHV	468.00	3.00	4.00	3.00	14.50	11178.00
Föglö	AHV	590.00	4.00	4.00	4.00	14.50	13273.00
Eckerö	AHV	705.00	5.00	2.00	5.00	14.50	13753.00
Lemland	AHV	902.00	11.00	2.00	5.00	14.00	13179.00
Maarianhamina	AHV	9581.00	129.00	.00	8.00	16.00	19178.00
Pelkosenniemi	LAP	1569.00	19.00	4.00	4.00	17.50	9813.00
Savukoski	LAP	1954.00	27.00	5.00	3.00	17.00	13440.00
Enontekiö	LAP	2259.00	29.00	2.00	5.00	17.00	10401.00
Kittilä	LAP	6490.00	86.00	4.00	4.00	17.50	10725.00
Ranua	LAP	5637.00	106.00	5.00	3.00	18.00	8532.00
Salla	LAP	7414.00	121.00	4.00	4.00	18.00	9802.00
Sodankylä	LAP	10119.00	140.00	3.00	5.00	17.00	12439.00

SPSS:n Windows-versiolla (6.1) osoitan asioita hiirellä. Yritys klikkaa muuttujien nimiin, jotka pitää kuitenkin kirjoittaa uudestaan ja kertoa niiden tyypit ja pituudet. Tuottamassaan syntaksissa ohjelma esittää kaikki numeeriset muuttujat tähdellisinä. Muuten tuotos vastaa edellä itse kirjoittamaani; sen tähden en esitä sitä tässä.

Yhden komennon peliä

Survossa tehtävän hoitaa komento FILE SAVE, joka tallettaa tekstitiedoston sisällön Survon datatiedostoksi. Poimitaan aineisto näytteeksi kenttään.

45	1	SURVO 84C EDITOR Sat Jul 27 00:23:23 1996							D:\A\SMKIRJA\		50	72	0
1	*												
2	*	FILE SAVE TEOLL1.TXT TO TEOLL1 / tiedot sisään											
3	*	FILE LOAD TEOLL1 / mitäs saatiin											
4	*	DATA TEOLL1*,A,B,C											
5	C	Kunta	Lää	Väestö	Synt.	Maamet	Palvelu	Äyri	Tulotaso				
6	A	Geta	AHV	468	3	4	3	14.50	11178				
7	*	Föglö	AHV	590	4	4	4	14.50	13273				
8	*	Eckerö	AHV	705	5	2	5	14.50	13753				
9	*	Lemland	AHV	902	11	2	5	14.00	13179				
10	*	Maarianhamina	AHV	9581	129	0	8	16.00	19178				
11	*	Pelkosenniemi	LAP	1569	19	4	4	17.50	9813				
12	*	Savukoski	LAP	1954	27	5	3	17.00	13440				
13	*	Enontekiö	LAP	2259	29	2	5	17.00	10401				
14	*	Kittilä	LAP	6490	86	4	4	17.50	10725				
15	*	Ranua	LAP	5637	106	5	3	18.00	8532				
16	*	Salla	LAP	7414	121	4	4	18.00	9802				
17	B	Sodankylä	LAP	10119	140	3	5	17.00	12439				
18	*												

Mitä tähän nyt toteaisi?

Datan sisäänluku em. tilasto-ohjelmien tyyliin aiheuttaa liikaa vaivaa; tuntuu turhauttavalta neuvoa ohjelmia pikkutarkasti, kun jälki on silti puutteellista. En ihmettele, että Survoon tottuneelle mm. näiden ohjelmien käyttö vaikuttaa hankalalta - useinhan työ alkaa aineiston sisäänlukemisesta. Onneksi useimpien survoilijoiden ei yleensä tarvitsekaan muihin ohjelmiin kajota.

Kosolti kiusaa aiheuttavat myös kaikenkokoiset ja -näköiset kirjaimet. Joillakin ohjelmilla tuntuu olevan ns. kansainvälisyydestään huolimatta vaikeuksia kansallisten merkkivalikoimien kanssa. Areaksi vetää tammaalaisten olon, kun nakee hameen, laanin, vailla aita ja oita, eika millaan lailla erota noita.

Tunnistus toiseen

Entäpä jos tekstitiedosto sisältääkin useita taulukoita? Kuinka ne saisi poimittua sieltä mahdollisimman vähäisellä työllä? Tämän kysymyksen esitti Seppo Mustonen minulle vuonna 1993. Hänellä oli jo vuosia ollut mielessään visio siitä, kuinka paperilla esitetyistä taulukoista saisi kätevästi graafisen yhteenvedon. Kuvanlukijat ja tekstintunnistusohjelmat ovatkin sittemmin yleistyneet, joten taulukot voidaan saada periltäkin tekstitiedostoksi. Vision toteutumiseen tarvittiin enää *taulukontunnistusohjelma*.

Seppo totesi silloin, ettei ollut saanut aloitettua sellaisen ohjelman toteuttamista arvellen työtä verrattain laajaksi ja aikaavieväksi. Niinpä näimme siinä otollisen aiheen opinnäytetyökseksi. Keskustelimme aiheesta useasti, mutta toteuttamisen aloitin vasta vuoden 1995 alussa. Työ keskeytyi kesällä, kun päätimme vaihtaa pro gradu -työni aiheeksi reliabiliteettimittojen tilastollisten ominaisuuksien tutkimisen. Graduni peruspilarin muodostivat Survolla toteutetut laajat simulointikokeet, joiden perusteella mm. laajalti käytetty reliabiliteettimitta, ns. Cronbachin alfa, osoittautui virattomaksi. Taulukoiden tunnistustehtävästä muodostui erikoistyö, jonka viimeistelin vuoden 1996 alussa.

Tehtävä saattaa tuntua liiankin yksinkertaiselta: onhan ihmissilmän aivan helppo erottaa taulukot tekstistä. Mutta miten sen tekee "sokea" ohjelma, joka hapuilee tekstimassan seassa? Siinäpä ongelma.

Alkurytmi

Yritin aluksi keksiä tunnistamisen tueksi kaikenlaisia yksinkertaisia mittoja kuten sanojen ja lukujen määrän rivillä. Tein nopeasti Survo-modulin, jolla saatoin laskea tällaisten esiintymien määriä havaintotiedostoihin. Niitä tilastollisesti analysoimalla koetin saada selville, mitä ovat ne kriteerit, joilla ohjelma voisi taulukon tunnistaa. Monetkaan näistä periaatteista eivät tulleet käyttöön lopullisessa toteutuksessa, koska ne eivät osoittautuneet riittäviksi.

Kun esittelin ensimmäisiä aikaansaannoksiani laitoksemme tilastollisen tietojenkäsittelyn seminaarissa, Seppo ehdotti että siirtyisin jokien tutkimiseen. Tällä hän tarkoitti, että pitäisi ehkä tutkia tarkemmin sanojen väleissä olevia tyhjiä alueita, jotka pystysuunnassa muodostavat jokia taulukoiden kohdalla vaan eivät juuri muualla. Lisäksi hän ehdotti, että kiinnittäisin tarkempaa huomiota sanojen alkuihin ja loppuihin peräkkäisillä riveillä. Tästä alkoi ohjelman algoritmin runko hahmottua.

Lopullisen Survo-modulin nimeksi tuli DATAFIND. Sen toimintaperiaatteen voi kuvata lyhyesti: luetaan annettu tekstitiedosto kertaalleen läpi ja suodatetaan se uudeksi tekstitiedostoksi, johon sijoitetaan löydettyt taulukot otsikkorivineen. Otsikot eli muuttujien nimet haetaan taulukon yläpuolelta ja muunnetaan yksikäsitteisiksi. Tarvittaessa tyydytään nimiin X1, X2 jne. Lopuksi tulostetaan toimituskenttään komennot, joilla taulukot voidaan siirtää Survon havaintotiedostoiksi.

DATAFINDin toimintaa ohjaavat tilastolliset ja heuristiset päättelysäännöt. Joskus päättely menee pieleen, ja tuloksena on kummallisiakin "taulukoita". Se ei haittaa, kunhan selvät tapaukset hoituvat kunnolla. Tätä voisi verrata oikolukuohjelmiin: nekään eivät saisi luulla itsestään liikoja. Ratkaisioon käyttäjä epäselvät tapaukset, ei ohjelman sitä tarvitse yrittää tehdä.

Tunnistustesti

Jotta nähtäisiin kuinka tunnistusohjelma kunnostautuu, tehdään koe. Lisään PRINT-kaaviooni ohjauksen - ascii_text kv.txt ja tulostan tähän asti kirjoittamani tekstin kolmesta toimituskentästä, jolloin saan n. 250-rivisen tekstitiedoston juuri lukemastasi materiaalista.

```

22 1 SURVO 84C EDITOR Mon Aug 28 10:35:33 1996 D:\A\SMKIRJA\ 100 80 0
1 *SAVE TESTI / tunnistustesti luvun alkupuolen tekstistä
2 *
3 *DATAFIND KV.TXT,CUR+1
4 *SHOW KV.DAF / Data tables found by DATAFIND
5 *FILE SAVE KV.DAF TO KV1 / NAMES=4 FIRST=5 LAST=16
6 *FILE SAVE KV.DAF TO KV2 / NAMES=19 FIRST=20 LAST=31
7 *FILE SAVE KV.DAF TO KV3 / NAMES=34 FIRST=35 LAST=46
8 *FILE SAVE KV.DAF TO KV4 / NAMES=49 FIRST=50 LAST=60
9 *FILE SAVE KV.DAF TO KV5 / NAMES=63 FIRST=64 LAST=75
10 *INDEX KV*.SVO,CUR+1 / List of data files
11 *FILE SHOW KV1.SVO / Copied from text file KV.DAF 3698 28.08.96 10:35
12 *FILE SHOW KV2.SVO / Copied from text file KV.DAF 3810 28.08.96 10:35
13 *FILE SHOW KV3.SVO / Copied from text file KV.DAF 3810 28.08.96 10:35
14 *FILE SHOW KV4.SVO / Copied from text file KV.DAF 3816 28.08.96 10:35
15 *FILE SHOW KV5.SVO / Copied from text file KV.DAF 3922 28.08.96 10:35
16 *

```

Käynnistän DATAFINDin riviltä 3. Parin sekunnin tunnustelun jälkeen teksti tunnustaa sisältävänsä viisi taulukkoa, ja DATAFIND kirjoittaa rivit 4-10. Selailen hetken aikaa tulostiedostoa KV.DAF rivin 4 SHOW-komennolla, minkä jälkeen täytyy minun jatkuvan aktiivoinnin (F2 ESC) seuraavalta riviltä. Näin luvun alkupuolen viisi taulukkoa kopioituvat Survon datatiedostoiksi KV1-KV5. Lopuksi aktivoituu rivin 10 INDEX-komento. Tähän aktivointi loppuisikin, mutta INDEX tuottaa riveille 11-15 FILE SHOW -komennot, jotka edelleen tulevat aktivoiduiksi. Varsinainen ketjureaktio!

Pieni lisäys

Lisään vuotta myöhemmin, siis syyskuun alussa 1997 tähän juttuun pari seikkaa. Edellä käsittelin vain yksinkertaista tekstitiedostoa, jossa kenttien arvot on erotettu toisistaan yhdellä tai useammalla välilyöllä. Vaikka tämä onkin yleisluontoisin tapa esittää tilastotietoja järjestelmistä riippumatta, niin käytännössä on usein tarjolla muitakin vaihtoehtoja. Eräs kätevä tapa on käyttää kenttien erottajana jotakin erikoismerkkiä, tyypillisesti TAB-merkkiä. Monet tilasto-ohjelmat, mm. SPSS:n Windows-versio ja taulukkolaskentaohjelmat, esim. Excel, osaavat tallettaa dataa tässä "TAB-delimited"-muodossa.

Survoon tällaiset tiedostot saadaan erittäin helposti antamalla FILE SAVE -komennossa täsmennys DELIMITER=char(9), joka tarkoittaa että kenttien oletetaan olevan tiedostossa erotettu merkillä, jonka ASCII-koodi on 9, siis juuri TAB-merkillä. (Erotinmerkki voi siis Survon puolesta olla mikä hyvänsä.) Tässä täytyy huomata se kierous, että mikäli käytössä on suomenkielinen Windows, niin sen ohjelmat käyttävät yleensä desimaalierottimena pilkkua pisteen sijasta. Tällöin täytyy aluksi muuttaa tekstitiedostossa pilkut pisteiksi. Se sujuu TXTCONV-operaatiolla kuten alla olevassa esimerkissä. SURVO 98:ssa homman voi hoitaa myös toimituskentässä esim. REPLACE-komennolla, vaikka tekstitiedosto olisi vähän isompikin.

```

1 1 SURVO 98 Wed Sep 3 00:06:43 1997 D:\A\SURVO&ME\ 100 100 0
1 *
2 *TXTCONV WIN.TXT LOOSE.DAT
3 *CONVERSIONS:
4 *R , .
5 *END
6 *
7 *FILE SAVE LOOSE.DAT TO SURVODAT / DELIMITER=char(9)
8 *

```

```
1 1 SURVO 84C EDITOR Sat Sep 13 21:27:11 1997 D:\A\SURVO&ME\ 300 120 0
1 *DESKTOP?
```

Several modules are available for managing the 'desktop' and organizing jobs, files and directories. These modules are made by K.Vehkalahti.

- 1 = Creating index of files to edit field (INDEX)
- 2 = Navigating and controlling files and directories (DD)
- 3 = Searching occurrences of text string from files (SEARCH)
- 4 = Changing directory (CD)
- 5 = Managing day and week numbers (DATE)
- 6 = Calendar functions (CALENDAR)
- 7 = Making a telephone call (PHONE)
- 8 = Searching files from the disks (WHERE)
- 9 = Matching two directories (DM)
- 0 = Browsing directory trees (TREE)

! = Copyright notice and contact information
C = More information on control operations

Select 1-C or press ENTER!

työpöydältä

Työpöytämoduleita koskevista asioista on vähitellen tullut yksi lehden vakiopalstoista. Viime lehden aikaan uutuuksena tuli mukaan DM (Directory Matcher), jolla vertaillaan kahta hakemistoa ja kopioidaan tiedostoja niiden välillä. Aivan äskettäin olen lisännyt DM:ään pari uutta ominaisuutta, joita oli toivottu. Nyt kohdehakemistosta näytetään kaikki tiedostot, lähtö- ja kohdehakemistoa voi lennossa vaihtaa ja tiedoston sisälle voi kurkistaa SHOW:n avulla. Nämä uutuuksien ovat käytettävissä ainoastaan SURVO 98:ssa.

puuhun

Kevään risteilyllä tulin vilauttaneeksi jälleen uutta tulokasta, joka levittää tiedostojen ja hakemistojen hallinnan lopulta koskemaan kokonaisia hakemistorakenteita. Tulokkaan nimi on TREE, mutta sitä kutsutaan SURVO 84C:ssä komennolla WHERE /TREE tai sukokomennolla /TREE. SURVO 98:ssa myös komentona TREE kelpaa. Lisäksi TREE-toiminto on käytettävissä DD:n sisältä napilla T.

puhelunsiirrot

PHONE-toimintoon olen tehnyt pieniä parannuksia. Numeroissa voi nyt käyttää väliviivan lisäksi sulkuja, esim. (09)-1918861. Numeroiden lisäksi voi käyttää merkkejä * ja #, esim. puhelunsiirron voi aktivoida komennolla PHONE *21*9876543210#.

satelliittiin

Jos annetaan täsmennys NMEA=1, PHONE odottaakin saavansa parametreinaan kaksi rivitunnusta, esim. PHONE CUR+1,END. Tällöin annetuilla riveillä olevat merkit lähetetään valittuun tietoliikenneporttiin (esim. COM1). Toiminnon avulla voidaan mm. siirtää Survossa ylläpidettäviä reittipistetietoja tietoliikennekaapelia pitkin GPS-satelliittinavigointilaitteeseen ns. NMEA-muodossa. Arvelen, että tästä PHONE-modulin uudesta ominaisuudesta saattaisi olla hyötyä muunkinlaisissa sovelluksissa.

neula heinäsuovasta

SEARCH:iin olen taas lisännyt yhden namun: SEARCH=FIRST lopettaa heti ensimmäisen löydettyään. Tarvitsin tätä HTML-muunnoksissa automaattisen linkkilistan teossa ja uskoisin siitäkin olevan hyötyä muisakin sovelluksissa. Aina kun ei kannata hakea, jos on jo löytänyt etsimänsä. Vähän samaan sarjaan kuuluu se, että WHERE-haun voi nykyään keskeyttää pistettä painamalla.

vuosituhannen tunnustus

Tunnustan: vanhat työpöytämodulit käyttäytyvät siivottomasti, kun vuosituhat vaihtuu. Olin alun alkaen hieman huolimaton tässä suhteessa. SURVO 98 -remontin yhteydessä ryhdistäydyin ja tein vaadittavat korjaukset. Kävin testaamassa ne 2000-luvulla (terveisii!). SURVO 98:ssa asiat ovat siis kunnossa, 84C:n uusimmissa versioissa (väh. v.5.30) samoin.

suurtilalliset

Työpöytämodulit, erityisesti DD, hyötyvät suuresti SURVO 98:n tarjoamista avarammista tiloista. Esim. tämän lehden numeron 1/95 sivulla 9 esittämiäni DOS-muistin maksimointitemppuja DD:n jouhevan toiminnan takaamiseksi ei enää lainkaan tarvita. Tämän voi hauskaasti todeta WHERE-operaation avulla. Laina edellämäinitsemaani lehden aiempaa numeroa sivulta 8: *"...ei kannata antaa liian laveita hakuehtoja, esim. CDEF:*,* , koska siihen muisti ei varmastikaan riitä. Silloin kone saattaa mennä jumiin, ja se on käynnistettävä uudelleen."* Siispä aktivoitin heti komennon WHERE CDEF:*,* omassa koneessani, jossa on 32 megatavua muistia ja gigatavu levytilaa. Suorittimena on 133 MHz:n Pentium ja käyttöjärjestelmänä Windows 95.

kerskailevat

Tietojen lataaminen levyiltä kestää 2 min 40 s, minkä jälkeen minulla on lista KAIKISTA tiedostoistani: 18233 tiedostoa 636 hakemistosta, yht. 668994158 tavua. Voin katsella tiedostojen sisältöjä Enteriä painamalla, olivat ne sitten toimituskenttiä, datoja, sukroja tai jotain muita. Lajittelu on kivaa: Mikä on suurin tiedostosi? Entä pienin? Onko paljon samannimisiä? Alla minulla oli 200 rivin ja 100 sarakkeen toimituskenttä, jossa oli tilaa 50 varjoriville. Toistin kokeen vielä sellaisellakin kentällä, jossa oli 5555 riviä, 555 saraketta ja tilaa 5555 varjoriville. Ei ongelmia! Pystyin kevyesti katselemaan DD:n alla jopa 40 megatavun kokoista Survon datatiedostoa (1000 muuttujaa ja 10000 havaintoa).

hoikin möhkäle

SURVO 84C:ssä työpöytäohjelmat muodostuvat yhdeksästä tiedostosta Survon päähakemistossa. Ne vievät tilaa yhteensä n. 450000 tavua. SURVO 98:ssä asiat ovat vallan toisin: samat työpöytäohjelmat koostuvat vain kahdesta tiedostosta (&DESKTOP.EXE ja _DATE.EXE). Jälkimmäinen sisältää DATE:n lisäksi CALENDAR- ja PHONE-modulit. &DESKTOP-möhkäle on ahminut kaikki muut lukuunottamatta CD:tä, joka on fiinisti sanottuna integroitu Survon 32-bittiseen editoriin. Näiden kahden yhteenlaskettu koko on vain n. 360000 tavua, eli siirtyminen 16-bittisestä 32-bittiseen ympäristöön ei ole päässyt lihottamaan koodia - päinvastoin!

näytä, poista, lisää

Viime lehdessä esittelin laiskuuksiani yhden kirjaimen komentojen muodossa. Komentoja S, D ja I vastaavat ohjelmat olivat myös mukana SURVOTUT 3-levykkeellä. SURVO 98:ssä nämä SHOW- DELETE- ja INSERT-toimintojen pikakomennot kuuluvat vakiokalustoon. Silti pidän suositeltavana, että sukroihin nämä komennot kirjoitetaan kokonaisina. Toimituskentässä työskentelyssä yksikirjaimiset ovat käteviä. Nythän SURVO 98:ssä voi suttuisia, muualta saatuja tekstitiedostoja panna ojennukseen vaikka ne olisivat vähän kookkaampiakin. Työkaluistahan meillä ei ole puutetta, verstasta on vain kohennettu entisestään.

aave - maista

Ghostscript-ohjelma on tullut monille tutuksi viime vuosien aikana. Jos toimii moniajojärjestelmän (kuten esim. Windows 95) alaisuudessa, voi lisäksi hyötyä GSview-nimisestä Ghostscriptin laajennuksesta. Sitä voi pitää auki omassa ikkunassaan samalla kun survoilee toisessa ikkunassa, kuten parasta aikaa teen. Tosin Survoa käytän aina mieluummin kokoruudussa. Tarvittaessa vaihdan alt-TAB:illa GSview:n puolelle.

GSview:ssä Survolla tehtyjä sivuja voi zoomata, ja keväällä PRINT- ja EPS-operaatioihin tehtyjen lisäysten ansiosta myös katsella missä järjestyksessä tahansa. Itse olen käyttänyt GSview:ta kuten Ghostscriptia aiemmin; erona vain se, että voin samaan aikaan pitää auki lähtökohtana olevia toimituskenttiä ja tehdä niihin korjailuja tarpeen mukaan. Olen myös säätänyt sivun koon siten että leveys vastaa A4-arkkia. Pi-tuussuunnassa sivujen selailu käy kätevästi nuolinapeilla. Minulla GSview on tehnyt /GS-PRINT:in täysin tarpeettomaksi.

3²2³ – eli mitä se "32-bittisyys" käytännössä oikein tarkoittaa?

DOS on 16-bittinen järjestelmä. Siinä voidaan kerralla käsitellä keskusmuistissa korkeintaan 2¹⁶=65536 tavun kokoisia alueita. 65536 on yhtä kuin 64*1024 eli 64 kilotavua, lyhennettynä 64K. Näin ollen esim. matriisin alkioiden maksimimäärä SURVO 84C:ssä on 8192 (65536/8), kun se talletetaan 8 tavun kaksoistarkkuudella. Tästä seuraa, että neliömatriisin rivien/sarakkaiden lkm voi olla korkeintaan int(sqrt(8192))=90.

Nämä rajoitukset kummittelevat matriisitulkin lisäksi yhden jos toisenkin operaation taustalla (CORR, ESTIMATE, REGDIAG, CORRESP, FACTA, RELIAB, DIST, DISTV, GENREG, CANON, COMPARE, LSCAL). 64K asettaa lisäksi rajoituksia mm. TAB- ja STAT-operaatioihin, datatiedostojen muuttujien kuvauksiin sekä toimituskentän kokoon.

Toinen DOS:in tilarajoitus on 640K, DOS-muistin määrä. Tämä raja tulee SURVO 84C:ssä vastaan tyypillisesti verkko- ja Windows-ympäristöissä, joissa on helposti pulaa vapaasta muistista. Mm. DD joutuu tällöin ahtaalle eikä kykene enää kutsumaan SHOW- ym. operaatioita.

SURVO 98:lle nuo kaikki rajoitukset ovat täysin tuntemattomia! Se toimii 32-bittisessä maailmassa, jossa voidaan periaatteessa käsitellä kerralla 2³²=4294967296 tavun kokoisia alueita. Todellisuudessa koneen keskusmuisti asettaa rajat, mutta käytännössä ne eivät taida tulla vastaan.

Jouko Manninen:

SURVOILUA OPETTAJANA JA "OLONEUVOKSENA"

Ihmisellä on ns. varttuneemmalla iällä puolellaan se etu, että hän voi tarkastella maailman menoa hiukan pitemmällä horisontilla. Niinpä ajatus helposti välillä siirtyy muisteluksiin, jotka usein saavat ylimääräistäkin hohtoa.

Omaa työrupeamaani korkeakoulussa ajatellessani voin todeta, että siihen aikaväliin sisältyy lähes koko tietokoneiden kehityshistoria. Muistan kuinka ollessani 1950-luvulla nuorena assistenttina Helsingin kauppakorkeakoulussa tilastotieteen jatkokurssin oppilaat kävivät kurssin tietyssä vaiheessa professorin johdolla ekskursiolla tutustumassa IBM:n reikäkorttikoneisiin, joita pidettiin jo pienoisina ihmeinä. Paljon myöhemmin, joskus 1960-luvun lopulla, hankittiin sitten koulullemme ensimmäiset tietokoneet ja perustettiin laskentakeskus. Tässä vaiheessa tutustuin FORTRAN-ohjelmointikieleen ja haaveilin jo tietokoneiden käytöstä tilastotieteen opetuksen apuna ja pian jopa kokeilinkin sitä eräillä vapaaehtoisilla kursseilla. 1970-luvun lopulla tuli ajankohtaiseksi tietokoneen hankinta Menetelmätieteiden laitokselle nimenomaan tilastotieteen opetuksen ja tutkimuksen tarpeisiin ja siinä vaiheissa päädyttiin Wang 2200 pientietokoneeseen ja SURVO 76 -järjestelmään. Luonnollinen jatke mikrotietokoneiden esiinmarssin jälkeen oli sitten Wangin mikrolla toimiva Survo-versio BASIC-kielineen ja edelleen nykyinen SURVO 84C.

Tutustuttuani SURVO 84C:n entisestä paljon helppokäyttöisemmäksi kehittyneeseen makrokieleen eli 'sukrokoodiin' jäin varsinaisesti "koukkuun" ja rupesin sukrotekniikkaan nojaten kehittämään laajempaa kokonaisuutta tilastotieteen perusopetukseen. Syntyi TILJAT-opetuspaketti, jonka toimivuus testattiin syyskauden 1990 opetuksessa. Palaute oppilaiden taholta oli erittäin myönteistä ja tuo kurssi on sen jälkeen näihin asti pitänyt pintansa opetusohjelmissa. Tiljat-opetusta olen kuvannut aikaisemmin tämän lehden palstoilla numeroissa 1/91 ja 2/93, joten sivuutan sen tässä.

Graafinen liittymä - sittenkin

Eläkkeelle siirryttyäni havaitsin itsessäni eräänlaista "Survo-riippuvuutta", ja totesin samalla ilokseni, että minulla oli nyt vapaus kohdistaa survoiluni mihin suuntaan tahansa. Kotioloissa on ollut vähemmän tarvetta paperitulostusten aikaansaamiseen, joten olen keskittynyt lähinnä vuorovaikutukseen kuvaruudun kanssa ja pyrkinyt kehittämään sukroja, jotka olisivat mahdollisesti muillekin Survon käyttäjille hyödyllisiä. Survon uudemmissa versioissa on sukrokoodi yhä täydentynyt ja sillä tavalla on avautunut uudenlaisia Survon käyttömahdollisuuksia, varsinkin kun samanaikaisesti mikrot ovat tulleet huomattavasti nopeammiksi.

Vaikka Survon parhaat ominaisuudet perustuvat editoriaaliseen käyttöliittymään, toimintaa voidaan myös sukrojen avustuksella ohjailta ikäänkuin käytettäisiin graafista liittymää. Tämä on kätevää eräissä yhteyksissä, esimerkiksi opetukseen liittyvissä demonstraatioissa tai muissa esitystilanteissa. Tällaisissa sukroissa Survon toimituskenttä voidaan tietyllä sukroon sisältyvällä koodilla "sammuttaa" eli muuttaa "piilokentäksi". Vaikkei kenttä näykaan kuvaruudussa, se on silti koneen toiminnan kannalta yhä olemassa ja kohdistin liikkuu samalla tavoin kuin muutenkin. Kun sukrossa palataan jonkin kuva-aktivoinnin jälkeen tekstitilaan ja toimituskenttä on sammutettuna, esitetty kuva jää näyttöön kunnes uusi kuva sen syrjäyttää. Tällä välin sukro voi piilokentässä ohjailta kohdistinta, kirjoittaa kenttään näkymätöntä tekstiä ja aktivoida siellä olevia Survon komentoja. Näin voidaan koneella teettää mitä tahansa Survon operaatioita. Vuorovaikutus käyttäjän kanssa - "graafinen käyttöliittymä" - saadaan aikaan sukrokielen 'on key'-rakenteella, jossa kone pysähtyy odottamaan käyttäjän napinpainallusta ja jatkaa sitten haarautuen painetusta napista riippuen erilaisiin jatkoihin.

Nämä johtavat yleensä uuteen näyttökuvaan, joka syrjäyttää edellisen, sekä uuteen 'on key'-tilanteeseen. Koko ajan on käyttäjä pidettävä selvillä kuhunkin valintatilanteeseen sisältyvistä mahdollisuuksista. Tämä voi tapahtua esimerkiksi siten, että esillä olevaan kuvaan liitetään ala- tai yläreunaan tekstirivejä, jotka kertovat eri nappeja vastaavat jatkovaihtoehdot.

Yllä kuvatulla periaatteella voidaan siis tehdä "graafisen liittymän" sukroja. Vaihtuvat näytöt voivat olla erilaisia tilastollisia graafisia esityksiä, "tekstikalvoja" erilaisin ja erikokoisin fontein jne.. Kuvaruudulle voidaan myös kerrostaa päällekkäisiä kuvia ja siten rakentaa vaikka kuinka komplisoituja uusia kuvia.

Aikasarja-analyysin graafiseksi apuvälineeksi

Tekemistäni graafisen liittymän sukroista tulokoon esimerkkinä mainituksi vaikkapa /TIMEPLOT, jota on kuvattu tämän lehden numerossa 1/95 ja joka sisältyy levykkeeseen SURVOTUT 2. Tällä sukrolla voidaan tarkastella havaintotiedostoksi talletettua aikasarjaa graafisesti ja valita jatkoksi jokin tasoitus- tai analyysimenetelmä, jolloin aikasarjaa esittävään peruskuvaan saadaan esim. halutulle osavälille liukuvan keskiarvon käyrä tahi lineaarinen tai eksponentiaalinen trendi. Saatu trendi voidaan edelleen eliminoida joko jakamalla tai vähentämällä, jolloin uuteen kuvaan samassa kuvaruudussa saadaan esim. trendistä puhdistettu aikasarja. Tätä voidaan sitten tarkastella eri tavoin. Monessa tapauksessa voidaan näin suoraan graafisesti tehdä erilaisia johtopäätöksiä. Tarkemmat numeeriset tulokset tallettuvat samalla levyllä ja ne voidaan tarpeen vaatiessa tulostaa tai niillä voidaan suoraan jatkaa erilaisia analyyseja.

Myös voidaan kahta samaan aikaväliin kohdistuvaa aikasarjaa käsitellä yhdessä. Nämä aikasarjat saadaan aluksi samaan kuvaan, jossa on nyt kaksi pystyasteikkoa. Alapuolelle voidaan saada aikasarjojen summa, erotus, osamäärä jne. tai jokin muu relaatio, johon voidaan myös sisällyttää ajallisia viiveitä.

Esimerkiksi jokin taloudellinen aikasarja voidaan puhdistaa rahan arvon vaihtelusta jakamalla hintaindeksillä. Samoin voidaan tarkastella kahden aikasarjan, vaikkapa myönnettyjen rakennuslupien ja valmistuneiden asuinrakennusten lukumääriä kuvaavien aikasarjojen välisiä ristikorrelaatioita, jolloin nähdään suunnilleen näiden välillä esiintyvä viive ja voidaan samaan kuvaruutuun muodostaa vaikkapa aikasarja kuvaamaan valmistuneiden rakennusten ja vuotta aikaisemmin myönnettyjen rakennuslupien lukumäärien suhdetta.

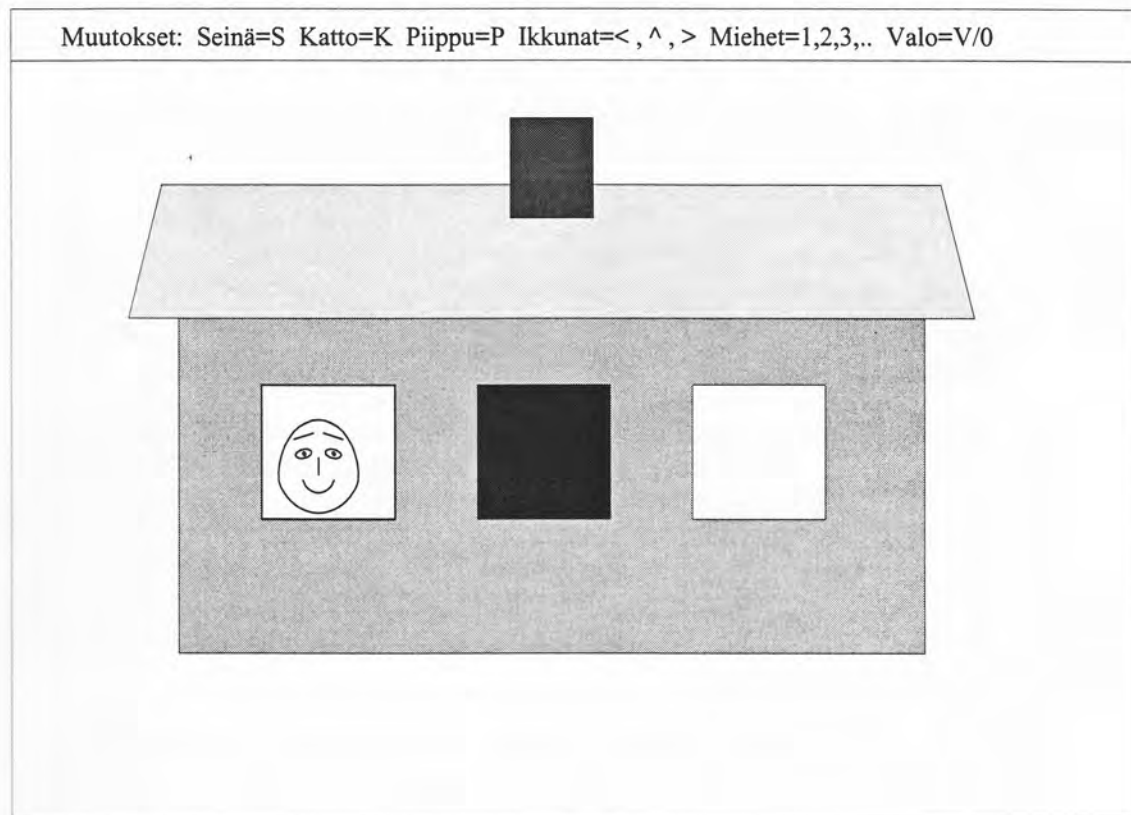
Lapseksi jälleen

Eräs tärkeä kohderyhmä, jolle olen viime aikoina sukroja tehnyt, on alle kouluikäiset lapset. Kotimikrossani on hakemisto C:\NURSERY, jossa tällä hetkellä näyttää olevan talletettuna toistasataa "lapsellista" tiedostoa, pääasiassa sukroja ja kuvia. Juuri graafiset interaktiiviset sukrot ovat lasten käyttöön omiaan.

Eräällä tällaisella sukrolla esimerkiksi voidaan tehdä Survon monimuuttujagrafiikkaan sisältyviä ns. Chernoffin naamoja siten, että ilmeet ja väritykset tulevat käyttäjän mieleisiksi: Kasvoissa on kaikkiaan 18 piirrettä, joita voi muuttaa (esim. silmien ja suun asennot jne..) ja 6 eri värityskohdetta. Muuteltavilla piirteillä on kullakin toisistaan riippumatta 10 eri tasoa ja kullekin värityskohteelle valittavissa 16 erilaista väriä, joten siis itse asiassa voidaan tehdä $10^{18} \cdot 16^6$ eli noin $1.67 \cdot 10^{25}$ erilaista naamaa. Kun halutun näköinen naama on valmis, se voidaan "leikata" ja tallettaa. Koko homma sujuu graafisessa näytössä muutamia nappeja painellen ja lapsikin oppii sen helposti.

Toisella sukrolla taas saadaan kuvaan kolmi-ikkunainen talo (ks. kuvaa), jonka seinät, katto ja piippu ovat väritettävissä 16 värillä painelemalla nappeja S, K ja P. Tähän taloon voidaan sijoittaa "asukkaiksi" esimerkiksi yhdeksän edellä kuvatulla sukrolla itse tehtyä erilaista naamaa. Talon kolmessa

huoneessa (valinta nuolinapeilla) voidaan kussakin sytyttää ja sammuttaa valo (napit V ja 0) ja kuhunkin ikkunaan voidaan kutsua mikä tahansa talon yhdeksästä asukkaasta painamalla jotakin napeista 1 - 9. Eräillä napeilla voidaan vielä saada aikaan äänitehosteita (murinaa, kiljahduksia ym.). Olen viettänyt riemastuttavia hetkiä leikkiessäni tällä talolla lastenlasteni kanssa, jotka sepittävät samalla kaikenlaisia tarinoita siitä mitä asukkaille kulloinkin tapahtuu.



Summa summarum: voin siis omakohtaisesti todeta, että Survosta on moneksi. Sillä voi suorittaa tilastollisia analyyseja, kirjoittaa raporteja, sitä voi käyttää opetusvälineenä. Ja kuten edellisestä ilmenee, se toimii viihdykkeenäkin "vauvasta vaariin".

Suunnilleen edellä olevan tekstin kirjoitin Sepon pyynnöstä syksyllä -96 hänen silloin julkaisemaansa kirjaa "Survo ja minä" varten, johon oli alun perin tarkoitus sisällyttää erillisiä lukuja eräiden muiden survoilijoiden henkilökohtaisista kokemuksista "Survo ja me"-hengessä. Lisäksi silloisessa tekstissäni oli vielä väliotsikolla "Hiiri osaaikatöihin" kertomus kokeiluistani hiiren käytössä graafisissa sukroissa. En siinä vaiheessa vielä arvannut, millaisia mahdollisuuksia tähän sisältyy, joten käytän nyt tilaisuutta hyväkseni ja kerron viimeisimmistä kokemuksistani uudella otsikolla.

Survo ja hiiri sekä minä

Kaikki sai alkunsa survivoristeilyllä keväällä -96, jolloin Juha Puranen esitteli mm. hiiren ja Survon yhteistyötä. Hänen hiirensä tuli näyttämölle GPLOT-komennon CALL-täsmennyksellä ja talletti näpäytyksellä senhetkiset koordinaatit tekstitiedostoon, josta ne voitiin sen jälkeen lukea esim. LOADP-komentoa käyttämällä. Huomasin heti pienen kokeilun jälkeen, että saisin hiirestä tehokkaan apuvälineen graafisiin sukroihiin, joiden parissa niihin aikoihin puuhastelin. Juha ystävällisesti modifioi hiirimoduliaan tarkoitukseen sopivaksi ja antoi käyttöön: tämänhetkisen hiirimodulin käyttönimi on JMPEN2.EXE.

Aluksi kokeilin hiiren käyttöä graafisessa liittymässä edellä kuvaamani 'on key'-rakenteen ohessa. Tällöin voidaan tietystä graafisen sukron vaiheessa jollakin näyttöruudussa ilmoitetulla näppäimellä saada kuvaan hiiri, jonka voi siirtää haluttuun paikkaan ennen näpätystä. Näpättämällä talletetut hiiren koordinaatit voidaan lukea sukromuistiin, jolloin jatkossa sukron eteneminen voidaan haaroittaa luettujen koordinaattien mukaan. Aluksi käytin hiirioptiota graafisessa sukrossa nopeuttamaan kohdistimen siirtämistä, mutta huomasin pian, että tällainen "puolihiiri-liittymä" tarjosi myös uusia mahdollisuuksia. Survo-risteilyllä keväällä -97 esittelin muutamia tällaisia kehittelemiäni uusia sukroja.

Tiedoston graafisesta esityksestä voi esim. hiiren koordinaattien avulla päästä käsiksi kuvassa näkyviin havaintoihin, joita vastaavia eri muuttujien arvoja voi kirjoittaa piilokenttään FILE LOAD operaatiolla ja sieltä edelleen graafiseen näyttöön TEXTS-täsmennystä käyttäen. Tällä tavalla toimii mm. sukro /SHOWPLOT. Sukrolla /IOPLOT voidaan taas kuvassa näkyvään korrelaatiodiagrammiin lisätä tai siitä poistaa havaintoja hiirtä näpättämällä ja samanaikaisesti nähdä miten esim. regressiosuora ja korrelaatiokerroin muuttuvat. Sukrolla /CROSSCUT saadaan korrelaatiodiagrammista (2-ulotteisesta jakaumasta) hiirellä osoitetusta kohdasta koordinaattiakselien suuntaiset leikkaukset välittömästi histogrammeina samaan kuvaan.

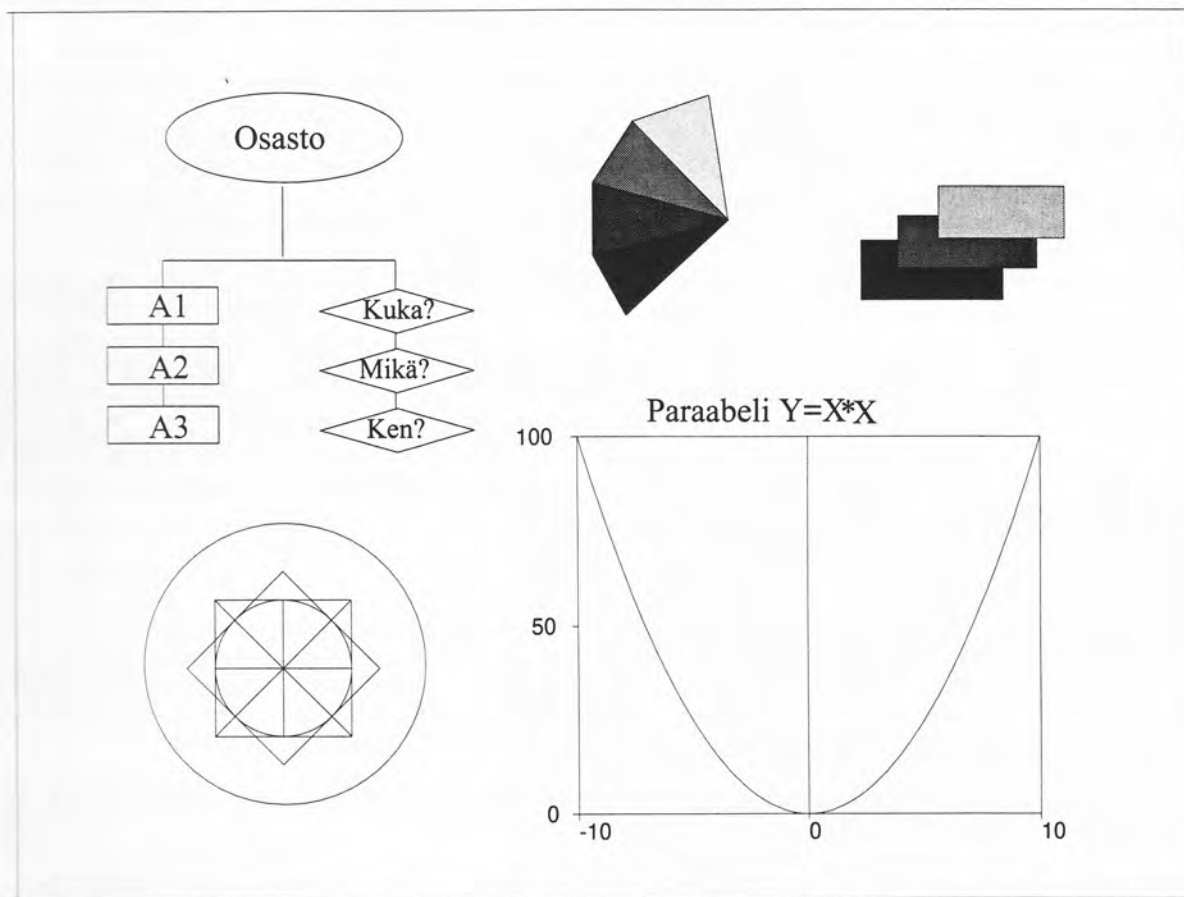
Kuluneen kevään ja kesän aikana olen erityisesti ahertanut hiirenkesyttäjänä tavoitteena "täyshiiri-liittymä", jossa survoilija"noviisikin" voisi helposti tehdä vapaamuotoista grafiikkaa, myös paperille asti. Tuntuu siltä, että SURVOTUT 4 -levykkeeseen sisältyvässä sukrossa /MOUSE apusukroineen tavoite on jopa kohtalaisen hyvin saavutettu.

Lyhyesti kuvattuna tällä sukrolla voi rakentaa kuvaruutuun kuvaa alkaen tyhjästä kuvaruudusta tai jostakin Survolla samalla kertaa tehtävästä peruskuvasta (histogrammista, korrelaatiodiagrammista, aikasarjasta, käyrän kuvaajasta jne.) tai myöskin aikaisemmin talletetusta spx-kuvapohjasta. Sukron käynnistyttyä jälkeen alkukuvan ympärille saadaan tilapäinen kehys, jossa on ikonin tapaisia merkkejä, sekä liikuteltava hiirimerkki. Hiirellä voi valita kehyksen reuna-alueelta halutun toiminnon. Kulloinkin käytössä oleva toiminto voidaan nähdä oikeassa alakulmassa olevasta kirjaimesta. Valinnan jälkeen siirrytään kuvan sisäalueelle, jossa hiirinäpäytyksin suoritetaan haluttu toiminto. Tuloksena saatu uusi kuvio on aluksi tilapäinen ja sitä voidaan korjata vasemmalla hiirinäpäällä, vasta oikeanpuolisen napin painallus kiinnittää kuvion kuvaan. Erilaisia valittavissa olevia toimintoja ovat murtoviivan piirto, vaaka- ja pystysuorien viivojen teko, erilaisten ja erisuuruisten POINT-merkkien teko, suorakaiteiden, vinoneliöiden ja ellipsien piirto, rajattujen alueiden fill-väritys, FRAME-kehysten teko, osasuorakaiteiden irtileikkaaminen ja haluttaessa kopioiminen uuteen paikkaan samassa kuvassa, ennakolta tehtyjen spx-osaruutujen (logojen) sijoittaminen kuvaan, tekstin kirjoittaminen halutulla tekstilajilla haluttuun kohtaan kuvassa jne. Käytettävissä on myös HELP-järjestelmä: tietyn toiminnan ollessa käytössä saadaan toimintaohjeet näpättämällä oikealla hiirellä vasemman alanurkan kysymysmerkkiä.

Erityisen mukava on sukrolla tehdä erilaisia kaavioita, joihin sisältyy suorakaiteita, vinoneliöitä tai ellipseja. Samoin vakiomuotoisiin graafisiin esityksiin on helppo lisätä täydentäviä merkkejä ja tekstejä tai poistaa osia kuvasta.

Tärkeä sukron ominaisuus on, että tehty kuva saadaan myös automaattisesti PostScript-kuvana (siis tavallaan WYSIWYG- tai oikeastaan WYSIAWYG "What You See Is Almost What You Get" -periaatteella). PS-kuvaa voidaan nimittäin vielä tarvittaessa hienosäätää valmista kaaviota apuna käyttäen. FILL-värit jäävät pois PS-kuvasta, viivavärit korvautuvat PS-ajurissa määritellyillä väreillä. FRAME-kuvion värit samoin kuin suorakaiteiden ja vinoneliöiden peittovärit, jotka valitaan kehyksen fill-värikartasta, näkyvät PS-kuvassa tummennuksina,

Alla olevassa kuvassa on esimerkkinä eräitä kuvioita, jotka pystytään tekemään MOUSE-sukrolla kätevästi hiiren avulla "hygieenisesti käsin koskematta".



Edellä kuvatun sukroperheen eri osien hiominen on vienyt melkoisesti aikaa. Aikoinaan sodassa rintaman pysyessä paikallaan sotilaat tekivät ajankulukseen kaikenlaisia puhdetöitä, joita sitten kotirintamalla ihmeteltiin. Minäkin olen välistä sukroa tehdessäni tuntenut rakentavani "purjelaivaa pullon sisälle". Mutta tulipahan tehtyä.

K. Vehkalahti:

Survo ja WWW

Survo tarjoaa mainion alustan myös WWW-sivujen tekoon. Tässä jutussa vertailen PostScript- ja HTML-dokumenttien ominaisuuksia sekä kerron miten WWW-sivuja Survolla luodaan.

PS vs HTML

PostScript on ns. sivunkuvauskieli. Survon PLOT- ja PRINT-komennot tuottavat PostScript-kielisiä tiedostoja. Sivunkuvauskielille on tyypillistä, että pienimmätkin yksityiskohdat voidaan säätää tarkasti kohdalleen. PostScriptille on lisäksi ominaista skaalattavuus, ts. kaikkia objekteja (kuten esim. kuvia) voidaan suurentaa, pienentää, venyttää ja kutistaa mielin määrin ilman että tulosteen laatu kärsii.

Kokonaan toisenlaista suuntausta edustavat niin sanotut rakenteenkuvauskielet. Eräs sellainen on HTML (HyperText Markup Language), jota käytetään internetin dokumenttien määrittelyyn. HTML määrittelee vain dokumentin rakenteen, ei sen täsmällistä ulkoasua. Ulkoasusta vastaa suurimmalta osin kulloinkin selainohjelma (esim. Netscape). Rakenteeseen kuuluvat esimerkiksi eri tasoiset otsikot, tekstikappaleet ja taulukot. Dokumentti voi sisältää linkkejä muihin dokumentteihin (yleisemmin: mihin tahansa verkon resursseihin, mm. kuvaja äänitiedostoihin) ja siihen voidaan viitata muista dokumenteista.

Viittaukset tapahtuvat tavallisesti HTTP- (HyperText Transfer Protocol) tekniikalla, jossa jokaisella dokumentilla on yksikäsitteinen nimi koko internetissä tai vaikkapa yrityksen sisäisessä intranetissä. Esimerkiksi Survon englanninkielisen kotisivun internet-osoite on <http://www.helsinki.fi/survo/eindex.html>. Tällaista osoitetta kutsutaan myös nimellä URL (Uniform Resource Locator). Jos nyt vielä yksi kirjainyhdistelmä tähän lisätään, niin se on tietysti WWW (World Wide Web), eli suoraan käännettynä maailmanlaajuinen verkko. WWW ja internet käsitetään usein samaksi asiaksi. Tarkalleen ottaen internet on hie- man laajempikin käsite, mutta samasta asiasta on joka tapauksessa kyse. En mene tässä enempiin yksityiskohtiin, koska WWW:stä ja siihen liittyvistä asioista on nykyään aivan tolkkottomasti materiaalia saatavilla joka puolella.

Paperi vs selain

Paperi ja verkkoselain ovat kaksi erityyppistä mediaa: se mikä näyttää hyvältä paperilla, ei välttämättä ole lainkaan hyvä selainohjelman ikkunassa. Ei siis kannata edes yrittää orjallisesti jäljitellä paperitulosteiden vakiintuneita ulkoasuja, vaan luoda verkkoa varten oma tyylinsä. Tyylin voi luoda määrittelemällä rakenteen, mutta ulkoasun viilaaminen HTML:llä on turha(uttava)ja. Tärkeintä on, että informaatio saadaan kohtuullisella vaivalla siirrettyä verkkoon ja että sitä voidaan sieltä kätevästi hyödyntää.

Kannattaa muistaa, että selaimia on muitakin kuin Netscape 3, näyttöruudun tarkkuuksia muitakin kuin 1024x768 ja tiedonsiirtonopeuksia muitakin kuin 10 Mbit/s.

Itse surffailen webissä useimmiten Lynx-nimisellä unix-ohjelmalla, joka ei edes osaa näyttää kuvia lainkaan. Tekstitkin se näyttää varsin vaatimattomasti, mutta nopeasti. Enitenhän aikaa suttaantuu kuvien lataamiseen - useimmiten täysin turhaan. Ne sivut, joissa oletetaan selaimeni olevan joko Netscape tai Explorer, jäävät tällä tavoin näkemättä, ja saavat jäädäkin. Tyylikkäintä on, kun sivut on suunniteltu selaisiksi, ettei niiden mahdollinen informaatiolisä ole riippuvainen tietyistä selaimista.

Kuvausta kuvista

En silti sano, ettei kuvia pitäisi laittaa verkkoon lainkaan. Niiden kohdalla vain kannattaa harkita joka kerta, onko se mitenkään välttämätöntä. Kuvien lataamisen odottelu aiheuttaa turhautumista jopa nopeilakin yhteyksillä. WWW:täkin on jo alettu kutsua nimellä *World Wide Wait*. Erityisesti kannattaa miettiä, onko kuvaa pakko tyrkyttää automaattisesti ladattavaksi vai kannattaisiko se pikemminkin sijoittaa linkin taakse, jolloin käyttäjä voi katsoa kuvan niin halutessaan. Jälkimmäinen vaihtoehto on erittäin suositeltava, silloin kuvia voi laittaa enemmänkin, eivätkä ne hidasta varsinaisen sivun lataamista. Linkkeihin kannattaa kirjoittaa hyvä kuvaus ao. kuvasta ja mahdollisesti sen koko, jolloin käyttäjän on helpompi arvioida, kannattaisiko linkkiä seurata.

Survon PostScript-kuvat voidaan muuntaa Ghostscript-ohjelman (version 4) avulla TIFF-muotoon (Tagged Image File Format). Se on hyvä välimuoto, koska se säilyttää kuvan sisältämän informaation erittäin tarkasti. Tämä tapahtuu tietenkin tilankäytön kustannuksella - tyypillinen TIFF-muotoinen kuva vie levytilaa n. puolitoista megatavua. Niitä ei onneksi tarvitse säilyttellä pitempään. TIFF-kuva kannattaa nimittäin muuntaa edelleen GIF-muotoon (Graphics Interchange Format), joka soveltuu verkkoon hyvin eikä vie paljoakaan tilaa (yleensä vähemmän kuin alkuperäinen PS-tiedosto).

Valitettavasti Ghostscript ei osaa GIF-muunnosta tehdä; siksi se on tehtävä jollakin muulla ohjelmistolla. Itse käyttämäni Paint Shop Pro on erittäin kätevä ja varsin edullinen. Tilasin sen verkon kautta ohjelman eurooppalaiselta maahantuojalta Tanskasta, jolloin hinta oli n. 500 mk. Olen nähnyt sitä myös täällä painkaupoissa. MikroPC-lehden numerossa 6-7/97 oli kuvankäsittelyohjelmien vertailu, jossa PsP:tä myös keuhuttiin. Ohjelmasta on saatavilla sekä Win95/WinNT- että Win3-versiot. Molemmista täytyy todeta, että ovat Windows-ohjelmiksi harvinaisen mukavia käyttää: ohjelma mm. "muistaa" aina, mistä viimeksi haettiin mitään kuvatiedostoja ja minkälaisiksi ne viimeksi muunnettiin. Toisesta käyttökerrasta lähtien hommat sujuvat siis joutuisasti. Tyypillistähän Windows-ohjelmille on, että aina saa tehdä samat valinnat uudelleen kuin jokin robotti. Minua sellainen kypsyttää aivan suunnattomasti. Tyypillinen esimerkki on juuri tiedoston haku melkein mihin tahansa Windows-ohjelmaan, vaikka Notepadiin ("Muistio"). Kamalaa hommaa. No, se ei kuulu tähän juttuun enempää kuin sika satelliittiin.

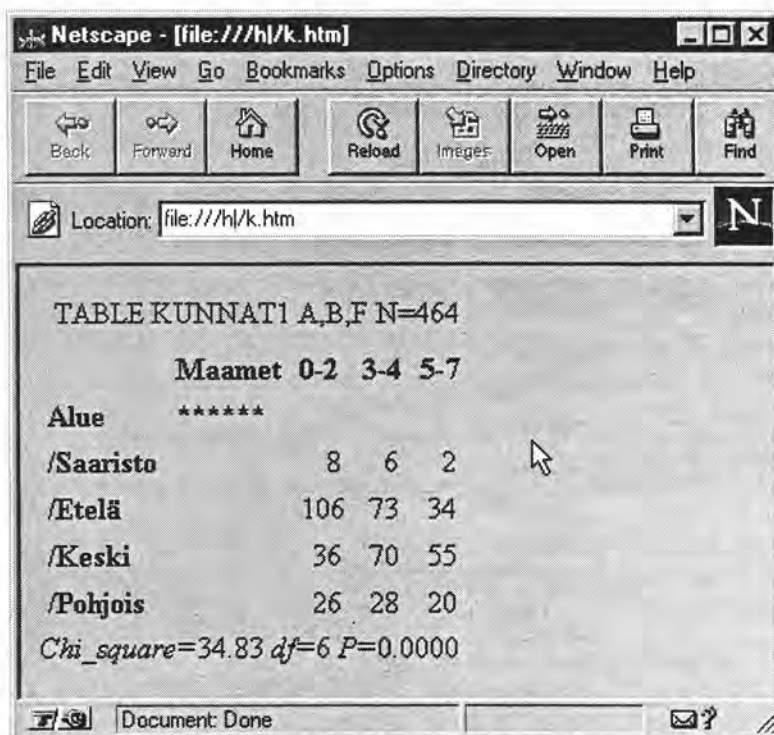
Asiaan

Keväällä -95 Lauri Tarkkonen usutti minua tekemään Survoon jonkinlaisen HTML-kytkennän. Ideoinkin silloin alustavan version tekemällä HTML.DEV-nimisen kirjoitinajurin PRINT-operaation yhteyteen. Huomasin näet ilokseni, että PRINT:issa käytettävä kuvauskieli tarjoaa mahtavat työkalut myös HTML:n koodaamiseen sivistyneemmin kuin yleisesti suosituilla tavoilla. (Eli että HTML-koodia kirjoitetaan enemmän tai vähemmän käsipöydällä, mikä on tietenkin julmaa ja ihmisarvoa alentavaa.) HTML:ään pitää teknisesti suhtautua kuin PostScriptiin: Survo-käyttäjän ei tarvitse käyttännössä koskaan liata käsiään kumpaankaan vaan likaisesta työstä huolehtii PRINT saatuaan ohjeet sopivan kirjoitinajurin muodossa. Käyttäjän tehtäväksi jää toimituskenttien valojen ja varjojen säätäminen (ks. *Survo ja minä*, s. 83-84).

Päivitin HTML.DEV-ajuria vielä syksyllä -96, jolloin tein Survon kotisivut (<http://www.helsinki.fi/survo/>). Vaatimusten kasvaessa määrittelin keväällä -97 kokonaan uuden ajurin, HTML32.DEV. 32-bittisyyden kanssa sillä ei ole mitään tekemistä vaan sen määrittelyt perustuvat HTML 3.2 -standardiin, joka on tällä hetkellä vallitseva ja useimpien selaimien tuntema.

Kieroisimpia HTML 3.2 -standardin määrittelemiä asioita ovat taulukot. Alla oleva koodirykelmä on monien HTML-oppaiden suositusten mukainen esitys pienelle KUNNAT-aineistosta tehdyille frekvenssitaulukolle. Koodi koostuu kulmasuluissa olevista avainsanapareista, esim. `<TABLE>` ja `</TABLE>`. Tällaisia suosittelaa siis tehtäväksi käsipöydällä, ts. jollakin tekstieditorilla.

```
<TABLE>
<CAPTION>TABLE KUNNAT1 A,B,F N=464</CAPTION>
<TR ALIGN=RIGHT>
<TH><BR></TH>
<TH ALIGN=LEFT>Maamet</TH>
<TH>0-2</TH>
<TH>3-4</TH>
<TH>5-7</TH>
</TR>
<TR ALIGN=RIGHT>
<TH ALIGN=LEFT>Alue</TH>
<TH ALIGN=LEFT>*****</TH>
<TH><BR></TH>
<TH><BR></TH>
<TH><BR></TH>
</TR>
<TR ALIGN=RIGHT>
<TH ALIGN=LEFT>Saaristo</TH>
<TH><BR></TH>
<TD> 8</TD>
<TD> 6</TD>
<TD> 2</TD>
</TR>
<TR ALIGN=RIGHT>
<TH ALIGN=LEFT>Etelä</TH>
<TH><BR></TH>
<TD>106</TD>
<TD> 73</TD>
<TD> 34</TD>
</TR>
<TR ALIGN=RIGHT>
<TH ALIGN=LEFT>Keski</TH>
<TH><BR></TH>
<TD> 36</TD>
<TD> 70</TD>
<TD> 55</TD>
</TR>
<TR ALIGN=RIGHT>
<TH ALIGN=LEFT>Pohjois</TH>
<TH><BR></TH>
<TD> 26</TD>
<TD> 28</TD>
<TD> 20</TD>
</TR>
</TABLE>
<EM>Chi_square=34.83 <EM>df=6 <EM>P=0.0000
```



Enpä menisi kehuaan koodia luettavuudella. Tai saa olla ainakin melkomoinen liero, jos tuossa muodossa pystyy taulukoitaan käsittelemään. Netscape on toki tähän esitykseen aivan tyytyväinen kuten näkyy; koodi on HTML-syntaksin mukaista. Mutta kyllä virheitä tulee liian helposti, jos tuollaista koodia itse kirjoittelee. No, onhan sitten hitusen kehittyneempiä ohjelmia, ns. HTML-editoreita. Niiden ainoa ero on kuitenkin yleensä se, että eri painikkeita klikkailemalla saa noita kulmasulkujen välissä esitettyjä koodeja liisäiltyä tekstinsä sekaan. Ei siis paljoa hurraamista. Sitten on vielä uudempia välineitä, joissa sivuja määritellään graafisesti, mutta niihin en mene tässä, kun en liiasta grafiikasta perusta muutenkaan. Siirrytään siis Survon maisemaan.


```

1 1 SURVO 84C EDITOR Wed Sep 10 20:12:19 1997 D:\A\SURVO&ME\ 100 100 0
1 *SAVE HTML / Hauska Taulukko Menneistä Lääneistä
2 *
3 *CLASSIFY KUNNAT,Aluejako,Lääni,Alue
4 *CLASSIFICATION Aluejako
5 *AHV: Saariisto
6 *HÄM,KYM,TUR,UUS: Etelä
7 *KAR,KES,KUO,MIK,VAA: Keski
8 *LAP,OUL: Pohjois
9 *END
10 *
11 */HTML-TAB KUNNAT,END+4 / VARIABLES=Maamet,Alue
12 *Maamet=0,2(0-2),4(3-4),7(5-7) Alue=/Saariisto,/Etelä,/Keski,/Pohjois
13 *
14 *PRINT CUR+1,END TO ALUEET.HTM
15 - include HTML32.DEV
16 - [TABLE]
17 *TABLE KUNNAT1 A,B,F N=464
18 r Maamet 0-2 3-4 5-7
19 r Alue *****
20 r /Saariisto 8 6 2
21 r /Etelä 106 73 34
22 r /Keski 36 70 55
23 r /Pohjois 26 28 20
24 - [/TABLE]
25 *Chi_square=34.83 df=6 P=0.0000
26 - [END]
27 *

```

Aluksi on muodostettu alueet luokittelemalla aineisto Lääni-muuttujan perusteella. (Esitän luokittelunkin tässä vaikkei se tähän juttuun nyt mitenkään sen kummemmin liity; halusin oikeastaan vain palstatilaa vanhalle kunnon CLASSIFY-operaatiolle!)

Taulukko on tehty sukrolla /HTML-TAB, joka toimii kuten TAB, mutta muuntaa tulostaulukon Survon HTML-koodaustavan mukaiseksi. Tulos näkyy riveillä 16-25. Taulukko on edelleen täysin luettavassa muodossa toisin kuin edellä esitetty HTML-tilausta, joka on aivan käsittämätöntä mössöä. Taulukkoa on vain TAB-operaation jälkeen hieman ehostettu erilaisilla väreillä ja muilla koristeilla.

HTML-tiedoston luontia varten on sitten lisätty PRINT-komento riville 14, HTML-ajurin kutsu riville 15 ja HTML-koodauksen lopetusmerkki [END] riville 26. Kun PRINT-komento aktivoidaan, saadaan tulokseksi itse asiassa vielä hiukan enemmän mössöä kuin mitä edellä nähtiin, mutta oleellisesti samat asiat. Kun mössö on syötetty Netscape:lle, saadaan sama näkymä kuin edellä. Korostan vielä, että HTML-koodia ei survoilijan tarvitse edes nähdä, saatikka sitten editoida. Kaikki tehdään niin kuin PRINT-operaation kanssa yleensäkin, koodit vain eroavat PostScript-tulostuksen vastaavista.

HTML32.DEV-ajuri sisältää määrittymiset 127 hakasulkukoodille, 41 varjomerille ja 17 kontrollisarakerille. (Vastaavat lukemat PS.DEV-ajurista ovat muuten 94, 20 ja 4, mutta sen yhteydessä käytetään vastaavasti usein muitakin ajureita). Sukro /HTML-CODES kertoo, mitä kaikkia koodeja ja merkkejä on käytettävissä. Itse HTML-koodeista löytyy selostuksia lukuisista kirjoista sekä tietysti WWW:stä.

/HTML-sukroperheessä on vielä pari muutakin jäsentä: /HTML-PRINT toimii kuten PRINT, mutta muuntaa tulostettavat tekstit, taulukot ja kuvaviittaukset uudeksi listaksi (joukoksi toimituskenttiä). Se tekee myös automaattisen linkkilistan muodossa olevan sisällysluettelon, jos tekstiin on merkitty avainsanoja. Tämän sukron kehittäminen on jatkuvasti kesken, sillä aina tulee vastaan uusia tilanteita, joissa se ei toimi kovin fiksumasti. Toisaalta, täydellistä siitä ei saakaan, mutta tyydyttävästi se on toiminut jo viime kesästä lähtien. Tärkeintä on, että saa työläimmät muunnokset tehtyä automaattisesti; /HTML-PRINT:in tuottamaa listaa voi sen jälkeen hienosäätää manuaalisesti.

Täysin automaattisesti muunnokset paperiversiosta verkossa esitettävään versioon sujuvat, jos paperiversio on mahdollista määrittellä sopivasti. Jos sen muoto on jo kiinnitetty, on joko tehtävä jonkinasteisia kompromisseja tai tehtävä verkkoraportille omat määrittämisensä. Kummastakin tavasta on jo saatu hyviä kokemuksia.

Sokerina pohjalla on sukro /HTML-7BIT, joka yksinkertaisesti muuntaa saadun HTML-tiedoston muotoon, jossa skandinaaviset merkit on esitetty niiden HTML-koodeina. Ilman tätä muunnostakin skandeille tehdään konversio. Sen tekee PRINT-komento HTML.BIN-koodimuunnostiedoston välityksellä. Siinä ääkköset muunnetaan ISO-Latin1-merkistön mukaisiksi. Tämä merkitseminen on WWW:ssä yleisesti käytetty. Kuitenkin, jos haluaa olla pedantti ja varmistaa, että ääkköset näkyvät oikein missä tahansa, niin voi käyttää /HTML-7BIT:iä, jonka jälkeen **hyvää yötä** onkin **hyvä yä yätä**.

"X all that apply"

Tulipa kerran männä kesänä vastaan kiero ASCII-tiedosto, suuren kansainvälisen projektin amerikkalaisen aikaansaannos. Siitä saatiin toki aikaan Survon datatiedosto - n. tuhat muuttujaa ja kolmisentuhatta havaintoa - eräiden muuttujien jakaumat vain alkoivat hieman ihmetyttää. Tyypiesimerkki tällaisesta oli kysymys: *Mitä harrastatte? Rastittakaa seuraavista kaikki, jotka sopivat teihin.* Ja perässä luettelo 59 harrastuksesta. Eihän tässä mitään - järkevää olisi olettaa, että aineistossa olisi 59 dikotomista muuttujaa, jotka saavat esim. arvoja 0 ja 1 sen mukaan, onko rastitettu vai ei. No, aineistossa **oli** siinä kohdassa 59 muuttujaa, *so what's the problem?*

Tässä vaiheessa täytyy selvyiden vuoksi siirtyä pieneen esimerkkiin, joka kuvaa tilannetta pienoiskoossa. Alla olevassa HOBBY-datassa on vain 10 muuttujaa ja 9 havaintoa, mutta se riittää tähän. Logiikka alla on seuraava: Jos harrastat golfia, Golf-muuttujalla on arvo yksi (tämä on vielä järkeenkäypää). Sen sijaan, jos harrastat esim. TV:n katselua, ei WatchTV-muuttujalla välttämättä ole mitään arvoa! Se johtuu siitä, että luvut kertovat rastitettujen harrastusten järjestysnumeron lomakkeella. WatchTV-muuttujalla on siis arvo 10 vain jos olet rastittanut kaikki 10 harrastusta. Toisaalta, jos harrastat TV:n katselua, jollakin näistä muuttujista on arvo 10, riippuen siitä mitä muuta harrastat. Selkeä koodaus!

1 1 SURVO 84C EDITOR Thu Sep 04 00:00:01 1997 D:\A\SURVO&ME\ 100 100 0														
1	*	11	11	11	11	11	11	11	11	11	11	11	11	11
2	*	DATA HOBBY A	A+8	A-1	A-3									
3	*	Golf	Fitness	Books	Skiing	Hiking	Hunting	Sailing	Sewing	Crafts	WatchTV			
4	*	1	1	1	1	1	1	1	1	1	1	1	1	1
5	*	2	2	2	2	2	2	2	2	2	2	2	2	2
6	*	3	3	3	3	3	3	3	3	3	3	3	3	3
7	*	4	4	4	4	4	4	4	4	4	4	4	4	4
8	*	5	5	5	5	5	5	5	5	5	5	5	5	5
9	*	6	6	6	6	6	6	6	6	6	6	6	6	6
10	*	7	7	7	7	7	7	7	7	7	7	7	7	7
11	*	8	8	8	8	8	8	8	8	8	8	8	8	8
12	*	9	9	9	9	9	9	9	9	9	9	9	9	9
13	*	10	10	10	10	10	10	10	10	10	10	10	10	10
14	*													

Voitte uskoa, että projektityöntekijä oli hämmästynyt tällaisten muuttujien keskiarvoista dataa tarkastaessaan. Hän tuli tässä vaiheessa kysymään minulta neuvoa. Olin aivan yhtä ihmeissäni. Totesin pian, ettei muunnoksia kannata alkaa tehdä millään VAR-operaatioilla; siinä menisi ikä ja terveys meiltä molemmita. Kiirekin oli (tottakai).

15 1 SURVO 84C EDITOR Thu Sep 04 00:00:02 1997 D:\A\SURVO&ME\ 100 100 0														
1	*	11	11	11	11	11	11	11	11	11	11	11	11	11
2	*	DATA HOBBY A	A+8	A-1	A-3									
3	*	Golf	Fitness	Books	Skiing	Hiking	Hunting	Sailing	Sewing	Crafts	WatchTV			
4	*	1	1	1	1	1	1	1	1	1	1	1	1	1
5	*	2	2	2	2	2	2	2	2	2	2	2	2	2
6	*	3	3	3	3	3	3	3	3	3	3	3	3	3
7	*	4	4	4	4	4	4	4	4	4	4	4	4	4
8	*	5	5	5	5	5	5	5	5	5	5	5	5	5
9	*	6	6	6	6	6	6	6	6	6	6	6	6	6
10	*	7	7	7	7	7	7	7	7	7	7	7	7	7
11	*	8	8	8	8	8	8	8	8	8	8	8	8	8
12	*	9	9	9	9	9	9	9	9	9	9	9	9	9
13	*	10	10	10	10	10	10	10	10	10	10	10	10	10
14	*													
15	*	XALL HOBBY Golf	10											
16	*													
17	*													

Tämä oli sen verran kiero probleemi, ettei auttanut muu kuin kääriä hihansa ja kaivaa Survon C-ohjelmointioppaan esiin. Vaikeinta oli modulin nimen keksiminen - kuten aina. Lomakkeen täyttöohjeen perusteella päädyin lopulta nimeen XALL. Parametreina annetaan aineiston nimi, ensimmäinen koodattu muuttuja ja käsiteltävien muuttujien lukumäärä.

47 1 SURVO 84C EDITOR Thu Sep 04 00:00:03 1997 D:\A\SURVO&ME\ 100 100 0														
1	*	11	11	11	11	11	11	11	11	11	11	11	11	11
2	*	DATA HOBBY A	A+8	A-1	A-3									
3	*	Golf	Fitness	Books	Skiing	Hiking	Hunting	Sailing	Sewing	Crafts	WatchTV			
4	*	1	1	1	1	1	1	1	1	1	1	1	1	1
5	*	2	2	2	2	2	2	2	2	2	2	2	2	2
6	*	3	3	3	3	3	3	3	3	3	3	3	3	3
7	*	4	4	4	4	4	4	4	4	4	4	4	4	4
8	*	5	5	5	5	5	5	5	5	5	5	5	5	5
9	*	6	6	6	6	6	6	6	6	6	6	6	6	6
10	*	7	7	7	7	7	7	7	7	7	7	7	7	7
11	*	8	8	8	8	8	8	8	8	8	8	8	8	8
12	*	9	9	9	9	9	9	9	9	9	9	9	9	9
13	*	10	10	10	10	10	10	10	10	10	10	10	10	10
14	*													
15	*	XALL HOBBY Golf	10											
16	*	TRANSFORM HOBBY BY	if (X=MISSING) then (0) else (1)											
17	*													

XALL:in tuottamasta muodosta on jo helppo viimeistellä tilanne TRANSFORM-operaation avulla, jolloin aineisto alkaa näyttää jollakin tavoin järjelliseltä. Tästä muodosta on jo hieman helpompi laskea esim. purjehdusta harrastavien lukumäärä.

1 1 SURVO 84C EDITOR Thu Sep 04 00:00:04 1997 D:\A\SURVO&ME\ 100 100 0														
1	*	11	11	11	11	11	11	11	11	11	11	11	11	11
2	*	DATA HOBBY A	A+8	A-1	A-3									
3	*	Golf	Fitness	Books	Skiing	Hiking	Hunting	Sailing	Sewing	Crafts	WatchTV			
4	*	1	1	1	1	1	1	1	1	1	1	1	1	1
5	*	2	2	2	2	2	2	2	2	2	2	2	2	2
6	*	3	3	3	3	3	3	3	3	3	3	3	3	3
7	*	4	4	4	4	4	4	4	4	4	4	4	4	4
8	*	5	5	5	5	5	5	5	5	5	5	5	5	5
9	*	6	6	6	6	6	6	6	6	6	6	6	6	6
10	*	7	7	7	7	7	7	7	7	7	7	7	7	7
11	*	8	8	8	8	8	8	8	8	8	8	8	8	8
12	*	9	9	9	9	9	9	9	9	9	9	9	9	9
13	*	10	10	10	10	10	10	10	10	10	10	10	10	10
14	*													
15	*	XALL HOBBY Golf	10											
16	*	TRANSFORM HOBBY BY	if (X=MISSING) then (0) else (1)											
17	*													

Ei tämä tähän loppunut. Seuraavaksi tuli vastaan muuten samanlaisia tapauksia, mutta jokaiseen kohtaan oli liitetty lisäkysymys, esim. kuinka tärkeänä (asteikolla 1-7) pitää ao. harrastusta.

24 1 SURVO 84C EDITOR Thu Sep 04 00:00:05 1997 D:\A\SURVO&ME\ 100 100 0														
1	*	11	11	11	11	11	11	11	11	11	11	11	11	11
2	*	DATA HOBBY A	A+8	A-1	A-3									
3	*	Golf	Fitness	Books	Skiing	Hiking	Hunting	Sailing	Sewing	Crafts	WatchTV			
4	*	1	1	1	1	1	1	1	1	1	1	1	1	1
5	*	2	2	2	2	2	2	2	2	2	2	2	2	2
6	*	3	3	3	3	3	3	3	3	3	3	3	3	3
7	*	4	4	4	4	4	4	4	4	4	4	4	4	4
8	*	5	5	5	5	5	5	5	5	5	5	5	5	5
9	*	6	6	6	6	6	6	6	6	6	6	6	6	6
10	*	7	7	7	7	7	7	7	7	7	7	7	7	7
11	*	8	8	8	8	8	8	8	8	8	8	8	8	8
12	*	9	9	9	9	9	9	9	9	9	9	9	9	9
13	*	10	10	10	10	10	10	10	10	10	10	10	10	10
14	*													
15	*	XALL HOBBY Golf	10											
16	*	TRANSFORM HOBBY BY	if (X=MISSING) then (0) else (1)											
17	*													

Näihin XALL ei tietenkään purrut. Olin jo ohjelmoida uuden modulin, mutta onneksi huomasi, että XALL:ia voi yleistää lisäämällä yhden parametrin, jolla kerrotaan, monenko muuttujan arvoja siirretään kerralla. Välissä olevat muuttujat (kuten tässä Q1, Q2, ...) vain kuljetetaan varsinaisten mukana vastaaviin paikkoihin.

- ku

Taulukoita ristiin

Tulipa taannoin tehtäväksi tuottaa tuhoton tukku tavallisia taulukoita tilaajan toivomuksesta tietenkin ta-
juttoman toimeliaasti. Toive: tulevaksi tiistaiksi. Tiedossa: tavattoman tuskainen työ.

Oli ryhmitelty aineistoa ristiin ja rastiin, etsitty mahdollisia uusia kuluttajasegmenttejä. Tyypillistä mark-
kinatutkimusta siis. Ryhmien profileista oltiin kiinnostuneita, haluttiin taulukoida erilaiset ryhmittelyt
kaikkien muuttujien kanssa vastakkain, luokittelumuuttujista frekvenssi-, muista keskiarvotaulukot. No
hyvä. Eikä kuin TAB aineisto, END+2 / VARIABLES=G1,mja G1=1(1)9 ja mja=1(1)5 . PRINT CUR+1,END
jne. jne. Heti kävi selväksi, että edessä on v-mäinen työ - enkä nyt viittaa V-sukeroon. Mainittakoon, että
taulukoitavia muuttujia per yksi ryhmittelymuuttuja oli luokkaa tuhat kappaletta, ryhmittelymuuttujia
toistakymmentä ja havaintoja kolmisentuhatta. Tulisiko joku vähän jeesaamaan TAB-komentojen kanssa?

Homma alkoi sujua varsin mallikkaasti, kun otettiin käyttöön mainio /TABULATE-sukro ja annettiin sen
hoidella likainen työ. Tulostus hoituikin jo sitten siististi, ja isokuriiri sai viedäkseen laatikollisen pikkuku-
riirilla prääntätyä paperia.

Unohtuikohari tästä jutusta nyt jotain oleellista... ainiin, sen jälkeen kun tajuttiin käyttää /TABULATE-
sukroa, oli lyhyt vaihe jossa havaittiin, ettei sellaista ole. Ja ennen kuin sitä päästiin käyttämään, oli vä-
hän pidempi vaihe, jonka jälkeen sellainen oli.

Se /TABULATE-sukron syntyhistoriasta. Seuraavassa esimerkki käytöstä:

```
33 1 SURVO 84C EDITOR Fri Sep 05 12:39:36 1997 C:\E\D\ 100 100 0
1 *SAVE KUN / /TABULATE-esimerkki perinteikkäällä kunta-aineistolla
2 *
3 *FILE STATUS KUNNAT / taulukoitavat muuttujat valittu FILE ACTIVATE:lla
4 * Suomen kunnat aakkosjärjestyksessä
5 * Tiedot ovat pääosin vuosilta 1978-80. 5.2.84/SM
6 * SORT:Kunta
7 *FIELDS: (active)
8 * 2 Sr- 3 Lääni [UUS,TUR,AHV,HÄM,KYM,MIK,KAR,KUO,KES,VAA,OUL,LAP]
9 * 8 Sm- 1 Palvelu Ammatissa ja palveluelinkeinoissa toim.osuus (10%)
10 * 11 NX- 2 Tulotaso Veroäyryjä asukasta kohti v.1979 (####)
11 * 12 Nm- 4 SYNT 1000*Synt./Väestö (##.###)
12 *END
13 *SURVO 84C data file KUNNAT: record=128 bytes, M1=30 L=64 M=14 N=464
14 *.....
15 *
16 */TABULATE KUNNAT BY Maamet TO KUN_
17 *Ryhmittelymuuttujan luokitus, tässä Maamet=1(1)7 , on annettava itse.
18 *Muiden muuttujien luokituksia voi antaa tarvittaessa, kuten
19 *esim. Tulotaso=0,13999 (Alle_14000),17999 (14000-18000),30000 (Yli_18000)
20 *
```

Muuttujat on aktivoitu erilaisilla kirjaimilla: Lääni-muuttujan **r** tarkoittaa frekvenssi- ja riviprosenttitaulu-
koita, **m** Palvelu- ja SYNT-muuttujissa keskiarvotaulukkoa. Iso **X** Tulotaso-muuttujan kohdalla saa aikaan
taulukot sekä havaituista että odotetuista frekvensseistä ja ns. χ^2 -kontribuutioista (χ^2 -testisuureen solu-
kohtaisista komponenteista).

Tuloksena syntyy KUN-niminen lista, ts. joukko toimituskenttiä, joiden nimien alkuosa on KUN ja lopussa
on juokseva numero. Kentät on valmiiksi sulutettu ja pilkottu kappaleiksi, mikä helpottaa niiden tulosta-
mista. Koko listan voi itse asiassa nykyään tulostaa varsin ytimekkäällä PRINT-kaaviolla:

```
18 1 SURVO 84C EDITOR Fri Sep 05 12:40:17 1997 C:\E\D\ 100 100 0
20 *
21 *PRINT CUR+1,CUR+2_
22 - [Courier(10)] [line_spacing(11)] [margin(300)]
23 - list KUN.LST
24 *
```

Edellisen sivun pienessä kuntaesimerkissä /TABULATE-sukron tulostusta koostuu vain yhdestä kentästä:

1 1 SURVO 84C EDITOR Fri Sep 05 12:40:27 1997 C:\E\D\ 350 150 0									
1	*SAVE KUN1 / LIST MAKE KUN OF TABULATE.OUT								
2	*LOAD KUN2								
3	*								
4	*DEF A,5,END								
5	(TABLE KUNNAT1 N=464								
6	*	Maamet	1	2	3	4	5	6	7 sum
7	*Lääni *****								
8	*/AHV		1	7	1	5	1	1	0 16
9	*/HÄM		17	8	10	12	2	0	0 49
10	*/KAR		2	3	2	5	3	4	0 19
11	*/KES		8	2	2	10	7	3	0 32
12	*/KUO		4	1	1	7	8	3	0 24
13	*/KYM		8	3	5	3	3	6	0 28
14	*/LAP		7	6	3	4	2	0	0 22
15	*/MIK		4	1	6	4	14	0	0 29
16	*/OUL		9	4	10	11	17	1	0 52
17	*/TUR		26	15	17	17	18	2	1 96
18	*/UUS		23	6	5	4	1	1	0 40
19	*/VAA		6	5	17	16	10	2	1 57
20	*	sum	115	61	79	98	86	23	2 464
21	)Chi_square=156.3 df=66 P=0.0000								
22	*								
23	(TABLE KUNNAT2 N=464								
24	*	Maamet	1	2	3	4	5	6	7 sum
25	*Lääni *****								
26	*/AHV		6.3	43.8	6.3	31.3	6.3	6.3	0.0 100.0
27	*/HÄM		34.7	16.3	20.4	24.5	4.1	0.0	0.0 100.0
28	*/KAR		10.5	15.8	10.5	26.3	15.8	21.1	0.0 100.0
29	*/KES		25.0	6.3	6.3	31.3	21.9	9.4	0.0 100.0
30	*/KUO		16.7	4.2	4.2	29.2	33.3	12.5	0.0 100.0
31	*/KYM		28.6	10.7	17.9	10.7	10.7	21.4	0.0 100.0
32	*/LAP		31.8	27.3	13.6	18.2	9.1	0.0	0.0 100.0
33	*/MIK		13.8	3.4	20.7	13.8	48.3	0.0	0.0 100.0
34	*/OUL		17.3	7.7	19.2	21.2	32.7	1.9	0.0 100.0
35	*/TUR		27.1	15.6	17.7	17.7	18.8	2.1	1.0 100.0
36	*/UUS		57.5	15.0	12.5	10.0	2.5	2.5	0.0 100.0
37	*/VAA		10.5	8.8	29.8	28.1	17.5	3.5	1.8 100.0
38	*	sum	24.8	13.1	17.0	21.1	18.5	5.0	0.4 100.0
39	*								
40	(TABLE KUNNAT1 A,B,FMS N=464 Mean and SD of Palvelu								
41	*	Maamet	1	2	3	4	5	6	7 Total
42	*		115	61	79	98	86	23	2 464
43	*	Mean	4.30	3.38	2.86	2.78	2.19	1.96	1.50 3.09
44	*	SD	1.33	1.02	0.69	0.68	0.54	0.37	0.71 1.20
45	*								
46	(TABLE KUNNAT1 A,B,FCR N=464								
47	*	Maamet	1	2	3	4	5	6	7 sum
48	*Tulotaso *****								
49	*Alle_14000		18	46	74	96	86	23	2 345
50	*14000-18000		69	14	5	2	0	0	0 90
51	*Yli_18000		28	1	0	0	0	0	0 29
52	*	sum	115	61	79	98	86	23	2 464
53	)Chi_square=296.6 df=12 P=0.0000								
54	*								
55	(TABLE KUNNAT1S C,D,Expected_frequencies								
56	*	Maamet	1	2	3	4	5	6	7 sum
57	*Tulotaso *****								
58	*Alle_14000		85.5	45.4	58.7	72.9	63.9	17.1	1.5 345.0
59	*14000-18000		22.3	11.8	15.3	19.0	16.7	4.5	0.4 90.0
60	*Yli_18000		7.2	3.8	4.9	6.1	5.4	1.4	0.1 29.0
61	*	sum	115.0	61.0	79.0	98.0	86.0	23.0	2.0 464.0
62	*								
63	(TABLE KUNNAT1S E,F,Contributions_to_Chi-square								
64	*	Maamet	1	2	3	4	5	6	7 sum
65	*Tulotaso *****								
66	*Alle_14000		53.3	0.0	4.0	7.3	7.6	2.0	0.2 74.4
67	*14000-18000		97.7	0.4	7.0	15.2	16.7	4.5	0.4 141.8
68	*Yli_18000		60.3	2.1	4.9	6.1	5.4	1.4	0.1 80.3
69	*	sum	211.3	2.5	15.9	28.7	29.7	7.9	0.7 296.6
70	*								
71	(TABLE KUNNAT1 A,B,FMS N=464 Mean and SD of SYNT								
72	*	Maamet	1	2	3	4	5	6	7 Total
73	*		115	61	79	98	86	23	2 464
74	*	Mean	14.34	12.49	11.98	10.86	10.74	9.31	9.83 12.02
75	*	SD	2.67	3.42	2.75	3.09	3.73	3.26	2.76 3.46
76	*								

KUVA JA SANA

Viime kevään Survo-risteilyllä käytiin ajoittain kiivastakin keskustelua Survon tulevasta suunnasta. Lähinnä siitä onko syytä pyrkiä muuttamaan Survon käyttöjärjestelmä muuksi kuin dos-pohjaiseksi. Ymmärsin, että monen toiveena olisi täysverinen graafinen moniajoympäristö. Koska ymmärrykseni asiasta on vähäinen, en malta olla ottamatta kantaa siihen, pitäisikö Survon hukkaa persoonattomasti rannattomaan Windows-mereen.

Lisäksi haluaisin ottaa esille pari seikkaa, joilla on mielestäni merkitystä yleisemminkin, kun ajatellaan tietotekniikan kehitystä ja sitä mihin käytetään ja tullaan käyttämään. Vieläkin suurin osa keskustelusta, käytiin se sitten alan lehdissä tai muualla, koskee itse välinettä eikä sen sisältöä. Asia on melkein sama, jos televisiosta käytävä keskustelu koskisi sitä, kuinka suuri TV:n kuvakoko on tai mikä on sen herzimäärä eikä siitä onko Kummeli hyvä vai huono ohjelma.

On kysyttävä mitä tietokoneella tehdään ja kuinka vaivattomasti se onnistuu. Samoin on kysyttävä, mikä Internetissä on hyvää ja mitä lisäarvoa se antaa. On kysyttävä myös, mitä tietoyhteiskunta on ja mitä se ei ole. En aio pohtia tässä näitä kysymyksiä, vaan sitä kuinka kuvien käyttäminen vaikuttaa ajatteluamme yleisesti ja tietokoneen käyttämiseen erityisesti.

Ymmärtääkseni kielitieteilijät hyväksyvät yleisesti ajatuksen, että muutos kuvista (hieroglyfeistä) foneettiseen kirjoitukseen on ollut kehitystä. Vasta foneettinen kirjoitus on antanut mahdollisuuden loogiseen ja rationaaliseen ajatteluun. Tietysti jo paljon ennen kirjapainon keksimistä on ollut loistavia ajattelijoita ja tiedemiehiä, mutta vasta kirjapaino teki rationaalisesta ajattelusta "massatuotteen".

Rationaalisuuden valta-asema kesti vahvana viime vuosisadan loppupuolelle asti, jolloin sen asemaa alkoi heikentää kaksi lähes yhtäaikaista keksintöä, lennätin ja valokuva. Näiden seuraajat, erityisesti televisio, on muuttanut ajattelutapojamme uskomattoman paljon.

Voidaan tietysti kysyä, mitä tällä kaikella on tekemistä tietokoneiden kanssa. Mielestäni hyvinkin paljon. Neil Postman kirjoitti jo vuonna 1982, siis silloin kun henkilökohtaiset tietokoneet olivat vasta tulossa markkinoille, kirjassaan "Lyhenevä lapsuus": "On olemassa taloudellisia ja poliittisia intressejä, joiden kannalta on parempi, että suuri puolittain lukutaitoinen väestönosa viihdyttää itseään visuaalisilla tietokonepeleillä, että se käyttää tietokoneita ja tietokoneet sitä ilman syvempää ymmärrystä asiasta."

Usean viestintätutkijan mukaan väline sellaisenaan vaikuttaa ajattelutapaamme. Kirja edellyttää loogista ja rationaalista ajattelua. Kielellisesti voimme esittää yleistyksiä, abstratioita ja koherentteja propositioita. Kuvallisesti emme. Emme voi esittää kuvalla metsää, puuta tai kukkaa yleisesti, voimme esittää vain tietyn metsän, puun tai kukan tietyllä hetkellä tietyssä paikassa.

Sama koskee kuvallista tietotekniikkaa. Vaikka ikoni toimisi vain ohjausyksikkönä, on sillä kuitenkin samanlainen vaikutus ajatteluamme kuin kuvalla. Tämä koskee niin Internetiä kuin muitakin tietoverkkoja.

Vaatimukset enemmän kuvasta, äänestä ja ennen kaikkea liikkuvasta kuvasta ovat viemässä verkkoja enemmän ja enemmän televisiomaaiseen suuntaan. Se, että tietokoneesta on tulossa kuvallinen viestin, ei ole oleellista, vaan se, että kuvallinen väline muuttaa ajattelutapojamme radikaalisti. Vaarana on irrationaalinen ja lyhytjännitteinen ajattelu, aivan samoin kuin on käynyt television kanssa.

Kaikki tietoyhteiskunnasta puhuvat intoilijat selittävät meille kuinka vallankumouksellinen tietokone, multimediat ja Internet ovat. Niitä on verrattu milloin tulen, pyörän tai kirjapainon keksimiseen. Ylittämättömän vertauksen esitti taannoin Erkki Kureniemi, joka väitti, että Internet on ihmisen yhdeksäs evoluutiohyppy. Sen ylittäminen edellyttäisi Internetin vertaamista erään raamatullisen henkilön toiseen tulemiseen.

Olen vakuuttunut, että Negroponten, Gatesin ja Windowsin viitoittamalla tiellä tietokoneet, uudet mediat ja Internet ovat muutaman vuoden päästä päättyneet siihen, mistä voi sanoa hiukan Irving Berliniä muuntaen: "There's No Business But Show Business".

Haluankin kysyä: mikä on Survon paikka tietotekniikan kehityksessä. Oma vastaukseni on selvä. Tiedän hyvin, että Survosta ei tule koskaan massatuotetta, mutta mielestäni on hyvä, että on olemassa ohjelma niillekin, jotka ennemmin tai myöhemmin huomaavat, että Windows ei ole ruma ankanpoikanen vaan keisari maankuuluissa vaatteissaan.

Tästä tässä on kysymys. Nykyinen vaatimus helppoudesta ja vaivattomuudesta on vienyt siihen, että tietokoneita käytetään ohjelmansuunnittelijoiden antamien mahdollisuuksien puitteissa. Ikonit mahdollistavat juuri ne ja vain ne asiat, joihin suunnittelijan mielikuviutus on riittänyt.

Kalevi Kantele



SURVOTUT 4 - levyke

Lehden mukana jaetaan kaikille SURVOTUT 4 - levyke, joka sisältää Survo-Käyttäjäyhdistyksen jäsenten tekemiä ohjelmia, sukroja ja sovellusesimerkkejä. Seuraavaa levyä varten odottelemmekin jo, että mahdollisimman moni käyttäjä osallistuisi sen kokoamiseen lähettämällä tekemiään lisäohjelmia tai sukroja tai muita omia oivalluksia kuvauksineen ja mahdollisine esimerkkeineen alla olevaan osoitteeseen. Ohjelmien ja sukrojen ei tarvitse olla mitään laajoja kokonaisuuksia, vaan ne voivat olla hyvinkin pieniä Survo-työskentelyyn liittyviä "apuvälineitä". Oivallusten suhteen ei kannata käyttää minäänlaista sensuuria.

Kentässä INDEX kerrotaan levykkeen sisältö. Useimmat toiminnot eivät edellytä mitään tiettyä Survon versiota, mutta erityisesti sukroissa on saatettu käyttää ominaisuuksia, jotka edellyttävät suhteellisen uutta versiota. Asentamisessa on syytä huomioida myös se, että SURVOS-ohjelmistoon voi lisätä sukrot, mutta ei C-kielisiä ohjelmia.

Survo-Käyttäjäyhdistys ry

Survo-Käyttäjäyhdistys on toiminut vuodesta 1989 lähtien ohjelmiston käyttäjien, kehittäjien ja markkinoijien yhdyssiteenä.

Yhdistys järjestää vuosittain erilaisia koulutus- ja kongressitilaisuuksia, joissa käyttäjät voivat tutustua Survon eri käyttömuotoihin ja ohjelmiston uusiin piirteisiin. Yhdistys toimii myös erilaisten Survoon liittyvien sovellusten ja muun materiaalin välittäjänä jäsenistön piirissä ja kädessäsi oleva lehti on esimerkki yhdistyksen omasta julkaisusta.

Yhdistyksen vuosittainen jäsenmaksu on 75 mk henkilöjäseneltä ja 750 mk yhteisöjäseneltä.

Voit liittyä yhdistyksen jäseneksi tekemällä ilmoituksen yhdistyksen sihteerille tai maksamalla suoraan jäsenmaksun.

Pankkiyhteys:

OKO Pankki Tapiola 572442-211072

Survo-Käyttäjäyhdistys ry / Marjut Schreck
Mäenrinne 11
02160 ESPOO

[illegible]