Unix ja shell-ohjelmointi (TIE 205)

Demo 4, 20.10.1999, ratkaisut

Tehtävä 1

```
# Demo 4
# Tehtävä 1
#
s/&auml;/ä/g
s/&Auml;/Ä/g
s/&ouml;/ö/g
s/&Ouml;/Ö/g
# jne.
```

Tehtävä 2

```
# Demo 4
# Tehtävä 2
#
# Toiminta: Ensimmäinen korvaa-komento poistaa kaikki samalla rivillä
# kokonaan sijaitsevat tagit. Ensimmäinen tOPEN on vain "lipun tyhjennys",
# jotta ehdollista hyppyä voidaan käyttää myöhemmin. Tämän jälkeen silmukassa
# yhdistetään eri riveille jakautuneet tagit ja poistetaan ne, kunnes avoimia
# tageja ei enää ole.
#
s/<[^>]*>/ /g
tOPEN
:OPEN
/</ {
s/<[^>]*>/ /g
tOPEN
N
s/\n/ /g
tOPEN
}
```

Tehtävä 3

```
# Demo 4
# Tehtävä 3
#

s/^[^:]*:[^:]*:[^:]*:[^:]*://
s/:.*//
```

tai

```
# Demo 4
# Tehtävä 3
#
# huomattavasti hitaampi ratkaisu kuin edellinen

s/\(^.*:\)\(.*\)\(:.*:.*$\)/\2/
```

Tehtävä 4

```
# Demo 4
# Tehtävä 4
#
# Korvataan regexp, out_file in_file omilla. Huomaa d-komennon
# paikka.

/regexp/ {
  w out_file
  r in_file
  d
}
```

tai

```
#!/bin/sh
#
# Demo 4
# Tehtava 4
#
# Esim. shell-toteutuksesta.
```

```
echo "regexp>\c"
read regexp

echo "input file>\c"
read ifile

echo "output file>\c"
read ofile

sed '
/'$regexp'/ {
  w '$ofile'
  r '$ifile'
  d
}' $1 > $2
```

Tehtävä 5

```
#n
#
# Demo 4
# Tehtävä 5
#
# Kommentti: GNU sed toteutus (\| -käyttö), voitaisiin korvata käyttämällä
# kahta erillistä korvaus-komentoa.
#
# Toiminta: Kopioidaan työmuisti apumuistiin. Otetaan ensimmäinen pieni sanan
# aloittava merkki talteen ja muunnetaan se isoksi. Vaihdetaan apumuisti ja
# työmuisti keskenään. Korvataan talteen otettu / muunnettu merkki jollakin
# merkkijonolla (jonka ei uskota esiintyvän tiedostossa, tässä ğğğ).
# Vaihdetaan taas apumuisti ja työmuisti keskenään ja kopioidaan apumuisti
# työmuistin perään (muutettu kirjain nyt työmuistin alussa). Järjestetään
# työmuisti uudestaan (poistetaan ğğğ).
#
:ALKU
h
s/\(^\|.*[^a-zA-ZåäöÅÄÖ]\)\([a-zåäö]\)\([a-zA-ZåäöÅÄÖ.]*.*\)/\2/
```

```
tKORV
bLOPPU
:KORV
y/abcdefghijklmnopqrstuvwxyzåäö/ABCDEFGHIJKLMNOPQRSTUVWXYZÅÄÖ/
x
s/\(^\|.*[^a-zA-ZåäöÅÄÖ]\)\([a-zåäö]\)\([a-zA-ZåäöÅÄÖ.]*.*\)/\1ğğğ\3/
x
G
s/\(^[A-ZÅÄÖ]\)\(.*\)\(ğğğ\)\(.*\)/\2\1\4/
s/\n//g
tALKU
:LOPPU
p
```

tai

```
# Demo 4
# Tehtävä 5
#
# GNU-sed toteutus. Huomattavasti nopeampi kuin edellinen. Skandit
# ei tässä toimi. Voitaisiin käsitellä erikseen. \< voi olla myös \b.
#
s/\<a/A/g
s/\<b/B/g
s/\<c/C/g
s/\<d/D/g
s/\<e/E/g
# jne.
```