

INF1040 obligatorisk oppgave 2: Tekst og tall

Innlevering: Fredag 10. oktober 2008.

Formaliteter

- Besvarelsen skal være et tekst-dokument.
- Besvarelsen skal konverteres til en pdf-fil.
- Pdf-filen skal ha navn på formen `oblig2-dittbrukernavn.pdf`, for eksempel `oblig2-ragnhilk.pdf`
- Besvarelsen sendes til gruppelærer som et vedlegg til en e-post innen fristen.
- E-posten skal ha “subject” INF1040: Innlevering Oblig 2.

For å lage en pdf-fil kan du først lage en postscript-fil med kommandoen `print` (men send output til en fil istedenfor til en printer), og så kjøre programmet `ps2pdf` med postscript-filen som input. Her er et eksempel der filen `oblig2-ragnhilk.txt` konverteres til filen `oblig2-ragnhilk.pdf`:

```
print -o oblig2-ragnhilk.ps oblig2-ragnhilk.txt
ps2pdf oblig2-ragnhilk.ps
```

Hvis du ønsker å formattere besvarelsen som noe mer enn ren tekst, kan du f.eks. bruke Word, OpenOffice (alternativ til Microsoft Office, bruk kommandoen `ooffice` i Unix/Linux), eller \LaTeX (mye brukt i akademiske miljøer, se introduksjon på <http://heim.ifi.uio.no/~dag/LaTeX-intro.pdf> og/eller <http://heim.ifi.uio.no/~ifidrift/doc/LaTeX-for-nybegynnere.pdf>). Uansett hva du velger skal besvarelsen konverteres til pdf før levering.

Oppgaven skal løses individuelt, og må være godkjent for å kunne gå opp til eksamen.

Forsiden på besvarelsen skal inneholde navn, brukernavn, e-postadresse på UiO, samt formuleringen

Jeg har lest og forstått reglene som er gitt i dokumentet “Krav til innleverte oppgaver ved Institutt for Informatikk” på <http://www.ifi.uio.no/studinf/skjemaer/erklaring.pdf>.

Spør gruppelæreren om hjelp hvis du har problemer!

Oppgave 1: Tekst

Ta utgangspunkt i hjemmesiden du laget i oppgave 1 i Oblig1. Se bort fra alle bilder, formattering og lignende, og se på selve teksten som står på hjemmesiden.

Oppgave a

Gjør et overslag over hvor stor plass denne teksten ville tatt lagret som “ren tekst” ved hjelp av hver av Unicode-transformasjonene UTF-32, UTF-16 og UTF-8. Overslaget skal bygge på regnestykker ut fra hvordan tegnene representeres. (Hvis du ønsker det, kan du selvsagt kontrollere resultatet av regnestykkene ved å lagre teksten i de ulike transformasjonene og se hvor langt det blir, men dette skal ikke være den eneste fremgangsmåten!)

Oppgave b

Anta at du har en tekst med like mange tegn som teksten (hjemmesiden) over, men der du ikke vet noe mer om teksten enn at den bare bruker tegn fra ISO 8859-1.

1. Hva er den minste plassen denne teksten kan ta hvis den lagres som UTF-8?
2. Hva er den største plassen denne teksten kan ta hvis den lagres som UTF-8?

Oppgave 2: Tall

I denne oppgaven skal du ta utgangspunkt i de fire siste bokstavene i brukernavnet ditt,¹ der det første av disse skal gjøres om til stor bokstav. Vi vil i det følgende referere til disse som X_1 , X_2 , X_3 og X_4 . Med brukernavn ragnhilk, blir altså $X_1=H$, $X_2=i$, $X_3=l$ og $X_4=k$.

Oppgave a

Skriv opp UTF-8 representasjonen til hver av de fire bokstavene. La Y_1 stå for UTF-8 representasjonen av X_1 , Y_2 stå for representasjonen av X_2 , osv.

¹Hvis brukernavnet ditt er kortere enn fire bokstaver, kan du finne på de manglende bokstavene selv.

Oppgave b

I denne oppgaven skal vi tolke Y_1 , Y_2 , Y_3 og Y_4 som binære tall.

1. Konverter de binære tallene Y_1 , Y_2 , Y_3 og Y_4 til titallsystemet.
2. Regn ut i det binære tallsystemet summen $Y_1 + Y_2$ og produktet $Y_3 * Y_4$.
3. Regn ut i det binære tallsystemet $Y_3 + (-Y_1)$ og $Y_1 + (-Y_4)$, der negative tall representeres med bruk av toerkomplement.

For alle oppgavene gjelder at besvarelsen skal inneholde selve regnestykkene, ikke bare svarene.

(Hint: Sjekk at regnestykkene dine stemmer ved å gjøre de samme regnestykkene i titallsystemet!)

Oppgave c

Anta at du ved overføring har mottatt byte-sekvensen $Y_1Y_2Y_3Y_4$.

1. Hvis vi følger “big endian” prinsippet, vil dette da være et lovlig “single precision” flyttall i henhold til IEEE 754? I tilfelle, hvilket?
2. Hvis vi følger “little endian” prinsippet, vil dette da være et lovlig “single precision” flyttall i henhold til IEEE 754? I tilfelle, hvilket?

Det er her tilstrekkelig å angi svarene på formen $m \cdot 2^e$, der m er et binært “desimaltall” og e er et heltall i titallsystemet, som vist midt på side 93 i læreboken.