

TK1100

FORELESNING 0x02 TXT MEDIA

1. Tegnsett og kodetabeller

- a) Hva er en glyf?

En glyf er et tegn, i datasammenheng så er dette tegnet vi ser på skjermen som har blitt bestemt av en kodetabell, hvordan dette tegnet skal se ut bestemmes av font-tabellen.

- b) Hva er et charset/tegnsett?

Det er det datamaskiner bruker før å finne ut hvilket tegn det er som skal vises på skjermen av et gitt heltall, eksempler på charset er UTF-8 og UTF 16.

- c) Skriv litt om kodetabellene ASCII, ISO-8859-1, Windows 1252 og forskjellen mellom de.

I ditt svar burde du minimum vært innom:

- ASCII kom først og er 7-bit, og dermed er hvert tegn alltid 1 byte
- ISO-8859-1 og Windows 1252 er basert på ASCII
- Windows 1252 bygger også på ISO-8859-1 da denne kom sist av disse tre kodetabellene
- De første tegnene er kontrolltegn
- ASCII har kun det «amerikansk» alfabetet
- ISO-8859-1 har «undefined» tegn i tabellen
- Windows 1252 inneholder flere tegn som i dag ikke brukes ofte
- ISO-8859-1 og Windows 1252 inneholder «åæø» og andre nordiske bokstaver

- d) Hva er ulempen med ASCII-kodetabellen?

Det støtter kun «amerikansk» alfabet og trenger dermed alternative kodetabeller og fonter for å fungere med andre språk med spesialtegn.

- e) Slå opp i ASCII-tabellen og fyll inn riktig svar.

0x21 - !

0x44 - D

0x4F - O

0x43 - C

0x54 - T

0x45 - E

0x20 - Space

f) Hvorfor bruker ASCII kun syv bit?

Med 7 bit vil den åttende biten fungere som en paritets bit, se på dette som en sjekk-bit, disse brukes som en error-sjekker og er lagd for å finne ut av om noe har gått galt. Fordi ASCII tabellen kun har 7 bit til å bruke til koding av tegnet/symbolet så finnes det 128 ulike tegn/symboler (0-127) istedenfor 256 som det er i for eksempel Windows 1252.

g) Bruk ASCII-tabellen til å sortere tegnene under i riktig rekkefølge. Det med minst tallverdi i tabellen til venstre, høyest verdi til høyre.

A
2
;
>
R
Y
V
!
\$

Svar: ! \$ 2 ; > A R V Y

h) Parr tegnene til venstre med sine respektive ASCII-koder(Desimal).

A	47
D	65
y	43
+	68
/	121

2. UNICODE

a) Skriv kort om UTF-8 og UTF-16, og forklar forskjellen mellom de.

I ditt svar burde du minimum vært innom:

- Sier hvordan kodepunktene skal kodes
- Er de mest brukte char-settene/tegnsettene i dag
- UTF-8 = minimum 1 byte
- UTF-16 = minimum 2 byte
- UTF-8 koder på samme måte som ASCII med 8 bit for U+0000 – U+007F, over U+007F bruker UTF-8 to eller flere byte for å kode et tegn
- Primært bruker UTF-16 to byte for å kode et tegn, her kommer UTF-8 å bruke tre eller mer byte for tegn der UTF-16 bruker kun to
- Å bruke UTF-8 når ASCII tegn er en stor del av de som skal kodes så er det en fordel, hvis ikke så er UTF-16 passende å bruke

b) Hva er prefix-en til Unicode?

Svar: U+

c) Hva er spesielt med kodene U+0000 til U+007F i Unicode?

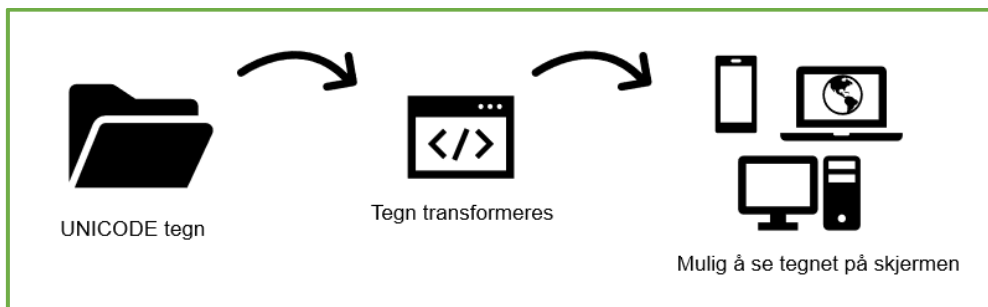
Det er tegnene som er i ASCII tabellen, med andre ord så er de første tegnene i Unicode samme som i ASCII.

d) Se videoen: The Unicode Consortium Overview (5 min) og skriv kort om hva Unicode er og hva det brukes til.

<https://www.youtube.com/watch?v=-n2nlPHEMG8&feature=youtu.be>

I ditt svar burde du minimum vært innom:

- En struktur med kodepunkter delt opp i 17 plan, er ikke et charset/tegnsett
- Unicode sier IKKE hvordan tegnene skal kodes binært
- Mulig å sende informasjon mellom alle, uavhengig av språk
- Alla typer av tegn og emojis
- Ulike charset/tegnsett, for eksempel UTF-32, UTF-16, UTF-8, Windows-1252, ISO-8859-1
- Charsetet/tegnsettet sier hvordan tegnene skal kodes binært



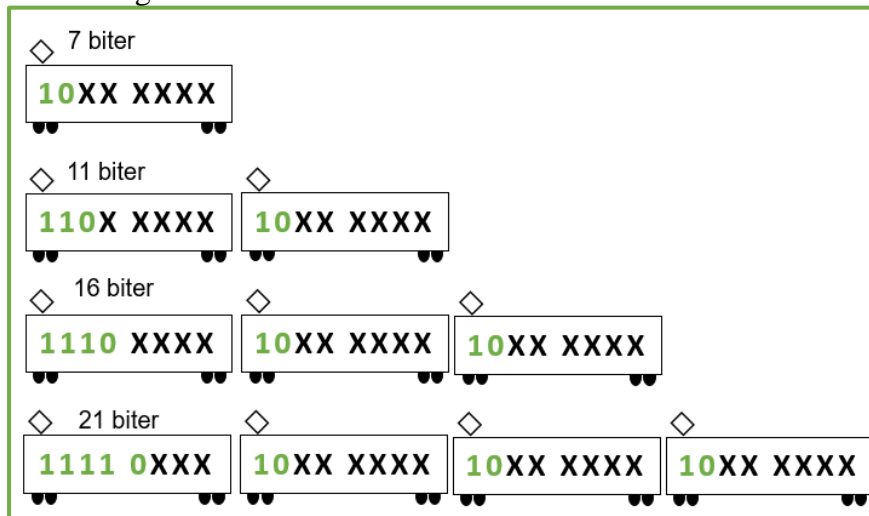
e) Hva er strukturen til Unicode?

17 plan hvor et plan inneholder 65536 kodepunkter, plan 0 (Basic Multilingual Plane) er det planet vi bruker mest, dette planet inneholder tegn for de vanligste språkene. Resterende plan inneholder, emojis, noter, uvanlige tegn, spesial tegn osv.

f) Hva brukes for å transformere Unicode-kodepunkter til binære tall, så datamaskiner kan lese de og vise tegn på skjermen?

Det er charset/tegnsett som brukes for å transformere disse til binære tall, de vi i dag bruker mest er UTF-8 og UTF-16.

- g) Tegn opp «vognene» i Unicode UTF-8 (7, 11, 16 og 21 bit). Hint: se sildesettet fra forelesningen.



- h) Forklar hva padding er i UTF-8 transformering.

Hvis «vognene» i transformeringen ikke blir fylt opp (se svar fra forrige spørsmål) vil plassene i vognene som er tomme fylles opp med padding, padding består alltid av '0'.

- i) Regn disse punktene i UNICODE (se eksempel med karakteren «Å» i slidesettet fra forelesningen):

U+0042 - B

U+0003 - 3

U+00F1 - ñ

U+0067 - g

U+0021 - !

- j) Hvor mange adresselokasjoner kan adresseres dersom start- og stoppadressene er som angitt: 0x3CA8 og 0x3CFF?

Svar: 88 (Desimal).

Hvis man konverterer de heksadesimale tallene til desimal, får man 15 528 og 15 615. Differansen mellom disse er 87, men siden det første tallet er start, og det siste er stopp, teller begge tallene, og dermed blir svaret 1 høyere.

3. Filformater

- a) Hva er data og metadata i filformater?

Data ligger i Body og dette er data som skal brukes til noe, metadata forteller hvordan filen skal behandles av programmet som skal lese filen.

- b) Hva er et «magic number», hvorfor trenger man det?

«Magic number» er starten av en fil fungerer som en ID, det vill si at dette nummert sier til noe om hva det er for typ av filformat.

- c) Hva er «magic-number»-ene til filtypene .java, .exe og .pdf?

.java - CAFE BABE

.exe - MZ

.pdf - %PDF

- d) Nevn noen typer tekst-filer, og noen typer binære filer.

Eksempel på svar:

Tekst-filer: HTML, CSS, SVG...

Binære filer: .class, .exe, .pdf...

4. Bilder, lyd og komprimering

- a) Gi eksempel på når man bruker CMYK og når man bruker RGB. Forklar også kort hva additiv og subtraktiv er og om CMYK og RGB er additive eller subtraktive.

Eksempel på svar:

Subtraktiv: "Blekker" reflekterer den fargen vi vil se og absorberer de andre

Additiv: Skjermen stråler ut den fargen vi vil se

CMYK: Printere (Subtraktiv)

RGB: Skjerm (Additiv)

- b) Hva er bruksområdene for vektorgrafikk mot Bitmap?

Vektorgrafikk som kan «blåses opp» i hvilken som helst fysisk størrelse, eksempel .svg filer, her bruker man seg av en matematisk formel for å lage formen du vil ha og da lagrer den formelen. Bitmap koder og lagrer hver enkelt piksel slik de skal vises på skjermen.

- c) Hvorfor komprimerer man filer, og hvilke typer filer komprimerer man ofte?

Gjennom å komprimere filer vill man ha mulighet til å eksempel sende store filer over internett forttere, typiske filer som blir komprimert er media og bilder.

- d) Gitt setningen: «Min a♣en år gamle ka♣ var beta♣ av en ha♣ så han må♣e på aku♣sykehus»

Finn ut hva ♣ er.

Svar: ♣ = tt

5. Repetisjon addisjon binært

Regn alle oppgaver med penn og papir, viktig at du skriver ned hele utregningen for hver oppgave.

- a) $1001\ 0001 + 1111\ 0000 = ?$ (8 bit presisjon)

Svar: 1000 0001

- b) $1111\ 0111 + 0111\ 1111 = ?$ (8 bit presisjon)
Svar: 0111 0110
- c) $7 + 123 = ?$ (Hvis utregning i binært og gi svar i 16 bit presisjon)
Svar: 0000 0000 1000 0010
- d) $0000\ 0100 + 0000\ 1111 = ?$ (16 bit presisjon)
Svar: 0000 0000 0001 0011
- e) $37 + 23 = ?$ (Hvis utregning i binært og gi svar i 8 bit presisjon)
Svar: 0011 1100
- f) $1000\ 1111\ 0011\ 0001 + 1010\ 0001 = ?$ (16 bit presisjon)
Svar: 1000 1111 1101 0010
- g) $0001\ 0110\ 0001\ 1010 + 0111\ 1111\ 0001\ 0101 = ?$ (16 bit presisjon)
Svar: 1001 0101 0010 1111

6. Repetisjon Toerkomplement

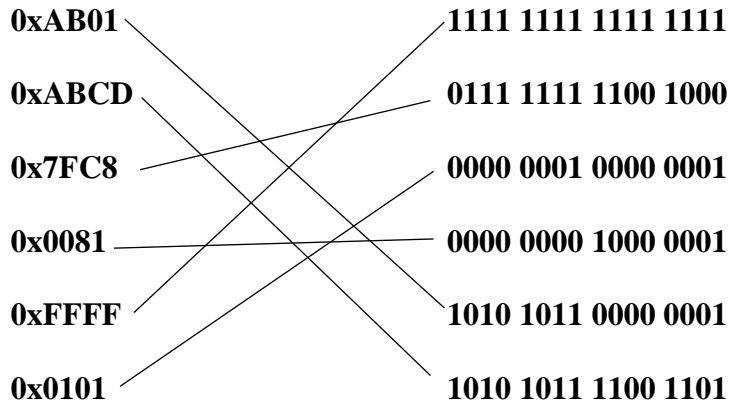
Regn alle oppgaver med penn og papir, viktig at du skriver ned hele utregningen for hver oppgave.

- a) $1001\ 1101 - 1111\ 0100 = ?$ (Gi svar i desimal)
Svar: -87 (1010 1001 i binært, MSB indikerer på at dette er et negativt tall)
- b) $15 - 121 = ?$ (Hvis utregning i binært og gi svar i 8 bit presisjon)
Svar: 1001 0110
- c) $45 - 14 = ?$ (Hvis utregning i binært og gi svar i 8 bit presisjon)
Svar: 0001 1111
- d) $1110\ 1101\ 0000\ 0001 - 1000\ 0001 = ?$ (16 bit presisjon)
Svar: 1110 1100 1000 0000
- e) $0011\ 0010\ 1101\ 1110 - 0101\ 1101\ 1001\ 1111 = ?$ (16 bit presisjon)
Svar: 1110 1100 1000 0000

7. Repetisjon Heksadesimal

Regn alle oppgaver med penn og papir, viktig at du skriver ned hele utregningen for hver oppgave.

- a) Par i hop heksadesimal-tall med binær-tall:



- b) $0x007B + 0x7B06 = ?$ (Gi svar i heksadesimal, 16 bit presisjon binært og desimal)
Svar: $0x7B81$, $0111\ 1011\ 1000\ 0001$ og $31\ 617$
- c) $0x000F + 0x08C5 = ?$ (Gi svar i heksadesimal, 8 bit presisjon binært og desimal)
Svar: $0x8D4$, $1101\ 0100$ og $2\ 260$
- d) $0x001A - 0x0023 = ?$ (Gi svar i 8 bit presisjon binært, desimal og bruk toerkskomplement)
Svar: $1111\ 0111$ og -9
- e) $0011\ 0010\ 1101\ 0110 = ?$ (Gi svar i heksadesimal og desimal)
Svar: $0x32D6$ og $13\ 014$
- f) $0x0006 - 0x000E = ?$ (Bruk toerkskomplement og gi svar i desimal)
Svar: -8