

Kompression, kap. 7 i komp.

Digital musikk måler typisk lydsignalet 44100 pr. sekund, i stereo, og lagres hver måling som et heltall med to bytes.

På ett sekund: $44100 \cdot 2 \cdot 2 = 176400$ bytes

På ett minutt: ca 10 MB.

Låt på 4 min: ca 40 MB

Konklusjon: Lett å generere store datamengder.

Huffman coding

Grunnprinsippet. Bruk kort kode for tegn som forekommer hyppig og lang kode for tegn som forekommer sjelden.

Ex. Anta at vi har en tekst

$\bar{X} = DBACDBD$, A-alfabett

$f(A) = 1, f(B) = 2, f(C) = 1$ = $\{A, B, C, D\}$

$f(D) = 3$ \rightarrow frekvens

Vi velger coding:

$c(D) = 0, c(B) = 1, c(C) = 01, c(A) = 10$

Kodet tekst $\bar{Z} = 011001010$ - Kan ikke dekodes.

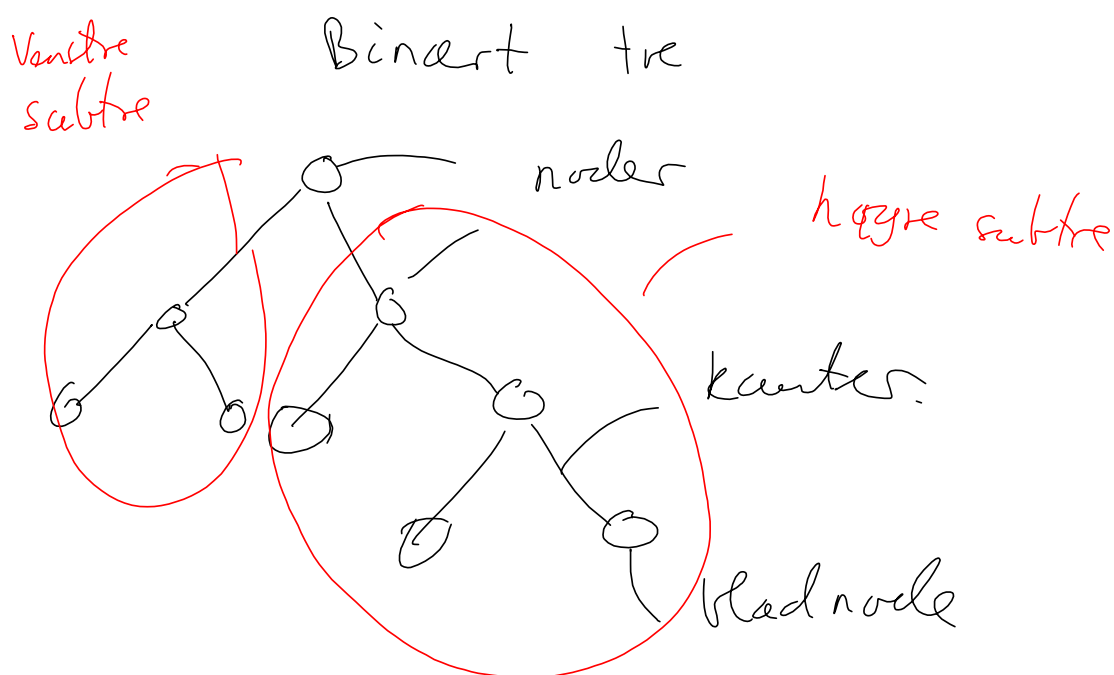
Ex2. Samme tekst $\bar{X} = DBACDBD$

$c(D) = 1, c(B) = 01, c(C) = 001, c(A) = 000$

Kodet tekst:

$\bar{Z} = 1010000011011$

Ingen kode har en annen kode som prefiks (forstavelse) - prefiksegenskaper.

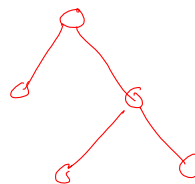


Kernen i Huffman-koding er at konstruere et binært træ som kaldes Huffman-træet.

Huffman-tre

Def 7.6 Et Huffman-tre er et binært tre assosiert med et alfabet med symboler $\{\alpha_i\}_{i=1}^n$ og frekvenser $f(\alpha_i), i=1, \dots, n$.

1. Hver bladnode er assosiert med nøyaktig ett symbol α_i og alle symboler er assosiert med en bladnode.
2. Hver node har en heltallig vekt
 - (a) Vekten til en bladnode er vekten til symbolet
 - (b) Vekten til en node er summen av vektene til røttene av nodens subtrær.
3. Alle noder som ikke er bladnoder har nøyaktig to subtrær ("barn")
4. Huffman-koden til et symbol framkommer ved å følge kantene fra roten ned til symbolets bladnode. Hver kant legger ett bit til koden: 0 for venstre kant og 1 for høyre kant.

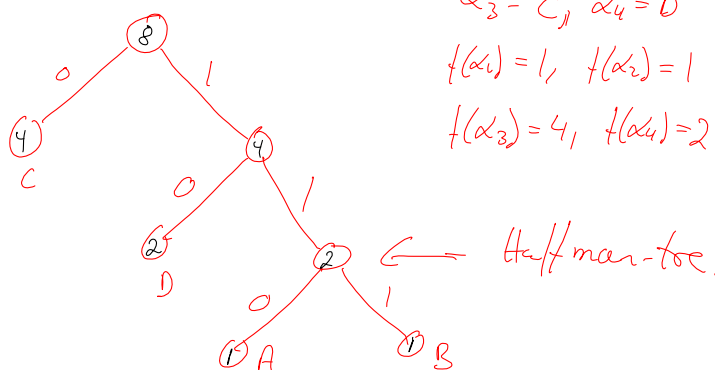


Eks. $x = CCDA CBDC, \alpha_1 = A, \alpha_2 = B$

$\alpha_3 = C, \alpha_4 = D$

$f(\alpha_1) = 1, f(\alpha_2) = 1$

$f(\alpha_3) = 4, f(\alpha_4) = 2$



$c(C) = 0, c(D) = 10, c(A) = 110$

$c(B) = 111$

Prefixegenskap OK.

Huffman algoritmen.

1. Lag en-node træer fra hvert av de n symbolene x_i med frekvens $f(x_i)$
2. Gjenta inntil vi bare har ett tre:
 - a) Velg de to trærne T_0 og T_1 med minst vekt og slå de sammen til et tre der T_0 er venstre subtre og T_1 er høyre subtre.
3. Træet som står igjen til slutt er ett Huffman-træ

Eks. 7.9.