HØGSKOLEN I STAVANGER

Institutt for elektroteknikk og databehandling

Stavanger, 9. mai 1995

4542 Informasjonsteori og feilkorrigerende koding: Øving 4

Oppgave 1

Løpelengde koding. En kilde genererer en følge av uavhengige binære symboler med sannsynligheter $P(0) = 0.9$ og $P(1) = 0.1$. Vi koder kildesymbolene ved en to-trinnsprosess som følger: Først teller en opp antallet direkte etterfølgende 0'ere mellom to 1'ere. Antallet nuller og den påfølgende 1'er representeres ved *foreløpige* kodeord (tall) i samsvar med følgende tabell:

Kildestreng	Foreløpig kodeord (tall)
1	'0'
01	'1'
001	'2'
0001	'3'
00001	'4'
000001	'5'
0000001	'6'
00000001	'7'
00000000	'8'

Sven Ole Aase

Førsteamanuensis, dr.ing.

Høgskolen i Stavanger

Institutt for elektroteknikk og databehandling

Postboks 2557 Ullandhaug

4004 Stavanger

Tel.: 51 83 20 09

Fax.: 51 83 17 50

E-mail: aase@hsr.no

Eksempel:

$$\underbrace{1}_{'0'}\underbrace{001}_{'2'}\underbrace{00000000}_{'8'}\underbrace{001}_{'2'}\underbrace{1}_{'0'}\underbrace{00001}_{'4'} \qquad (1)$$

Siste steg i kodingsprosedyren går ut på å tilordne de foreløpige kodeordene *ett* binært symbol for '8', og fire binærsymboler for hvert av de foreløpige kodeordene '0', '1', '2', '3', '4', '5', '6', '7'.

a) Konstruer en kode som er entydig dekodbar.

b) Finn midlere antall kildesymboler (binære symboler) pr. foreløpig kodeord, $\overline{n_1}$.

c) Finn midlere antall endelige kodesymboler (binære symboler) pr. foreløpig kodeord, $\overline{n_2}$.

d) Vis, – ved å argumentere med utgangspunkt i Tsjebysjeffs ulikhet, at for lange kildestrenger vil forholdet mellom antallet binære symboler i endelig enkodet streng og antallet binære symboler i kildestrengen være $\overline{n_2}/\overline{n_1}$. Sammenlign dette tallet med midlere kodeordslengde som oppnås ved å Huffmankode en utvidet kilde hvor kildesymbolene utgjøres av strenger av fire opprinnelige binære kildesymboler. Sammenlign også med kildens entropi. Er løpelengdekoding en god idé her?

Oppgitt: Tjebysjeffs ulikhet:

$$P(\mid \frac{1}{N}\sum_{i=1}^{N} X_i - \mu_x \mid < \epsilon) \geq 1 - \frac{\sigma_x^2}{N\epsilon^2} \tag{2}$$

Oppgave 2

En diskret hukommelsesfri kilde (en annen måte å si at symbolene genereres uavhengig og med identisk fordeling) har 6 kildesymboler med sannsynligheter gitt av $\{0.3, 0.2, 0.15, 0.15, 0.15, 0.05\}$.

a) Finn Huffman koden til kilden. Hva er midlere kodeord lengde.

b) En kode sies å tillfredsstille *paritet-sjekk-betingelsen* hvis *alle* kodeordene enten har et odde eller et like antall 1'ere. Tilfredsstiller koden som ble konstruert over paritet-sjekk-betingelsen? Hvis så ikke er tilfelle, går det an å konstruere en Huffman kode som tilfredsstiller denne betingelsen?

Oppgave 3

Gitt en binær kilde som genererer symbolene $\{S, M\}$ med like stor sannsynlighet. Informasjonen som kilden genererer skal overføres over en upålitelig kanal (f.eks. kanal beheftet med støy). Sannsynligheten for å motta en S når M ble sendt er $1/6$, mens sannsynligheten for å motta en M når S er sendt er $1/2$.

a) Tegn en modell for kanalen og identifiser "overførings- sannsynligheter" med grener i modellen. I figuren kan du betegne hendelsen at S er sendt med a_1 og at M er sendt med a_2. Tilsvarende kan det være hensiktsmessig å betegne hendelsene at S er mottatt med b_1 og at M er mottatt med b_2.

b) Finn følgende sannsynligheter (for $i, j = 1, 2$):

- $P(a_i, b_j)$.
- $P(b_j)$.
- $P(a_i)$.
- $P(a_i \mid b_j)$.

c) • Finn entropien i utsendt signal, $H(A)$.

- Finn entropien i mottatt signal, $H(B)$.
- Finn midlere gjensidig informasjon, $I(A;B)$.

Oppgave 4

Vi har i en tidligere øving definert simultan entropi ved

$$H(X,Y) = -\sum_i \sum_j P(x_i, y_j) \log_2(P(x_i, y_j)). \tag{3}$$

Vi viste da at hvis X og Y var uavhengige så gjaldt $H(X,Y) = H(X) + H(Y)$. Vis generelt at $H(X,Y) = H(X) + H(Y \mid X) = H(Y) + H(X \mid Y)$.