TSKS10 Signaler, information & Kommunikation

Föreläsning 6

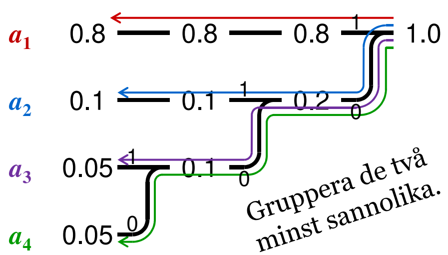
Datakompression, explicita koder

Introduktion kanalkodning

Mikael Olofsson
Institutionen för Systemteknik (ISY)
Ämnesområdet Kommunikationssystem

LINKÖPINGS UNIVERSITET

Huffmankodning (förenklad algoritm)



$c^{(1)} = 1$

$c^{(2)} = 01$

$c^{(3)} = 001$

$c^{(4)} = 000$

p_i	$c^{(i)}$	l_i	$p_i l_i$
0.8	1	1	0.8
0.1	01	2	0.2
0.05	001	3	0.15
0.05	000	3	0.15

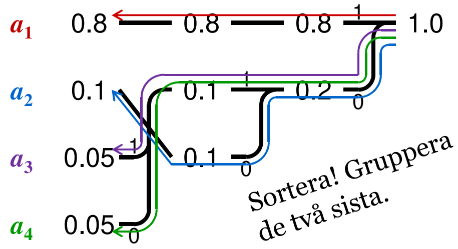
$m_L = 1.3$

Entropi: $H(A) = -\sum_{i=1}^{N} p_i \log_2 p_i \approx 1.022$

Redundans: $m_L - H(A) \approx 0.278$

Kompressionskvot: $\frac{\log_2 N}{m_L} \approx 1.54$

Huffmankodning (fullständig algoritm)



$c^{(1)} = 1$

$c^{(2)} = 00$

$c^{(3)} = 011$

$c^{(4)} = 010$

p_i	$c^{(i)}$	l_i	$p_i l_i$
0.8	1	1	0.8
0.1	00	2	0.2
0.05	011	3	0.15
0.05	010	3	0.15

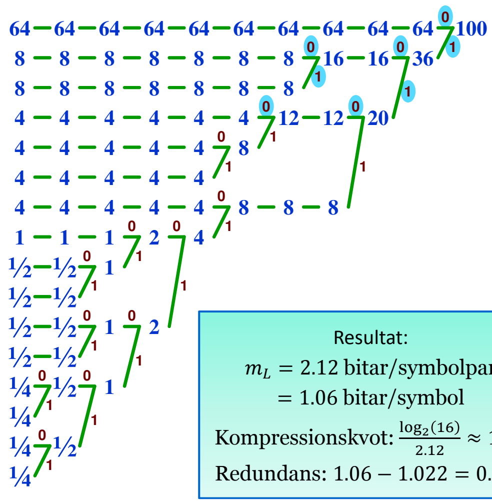
$m_L = 1.3$

Entropi: $H(A) = -\sum_{i=1}^{N} p_i \log_2 p_i \approx 1.022$

Redundans: $m_L - H(A) \approx 0.278$

Kompressionskvot: $\frac{\log_2 N}{m_L} \approx 1.54$

Förenklad huffmankod för en utvidgad källa



a_1a_1	0
a_1a_2	100
a_2a_1	101
a_1a_3	1100
a_1a_4	11010
a_3a_1	11011
a_4a_1	1110
a_2a_2	111100
a_2a_3	1111010
a_2a_4	1111011
a_3a_2	1111100
a_4a_2	1111101
a_3a_3	11111100
a_3a_4	11111101
a_4a_3	11111110
a_4a_4	11111111

Resultat:

$m_L = 2.12$ bitar/symbolpar
$= 1.06$ bitar/symbol

Kompressionskvot: $\frac{\log_2(16)}{2.12} \approx 1.89$

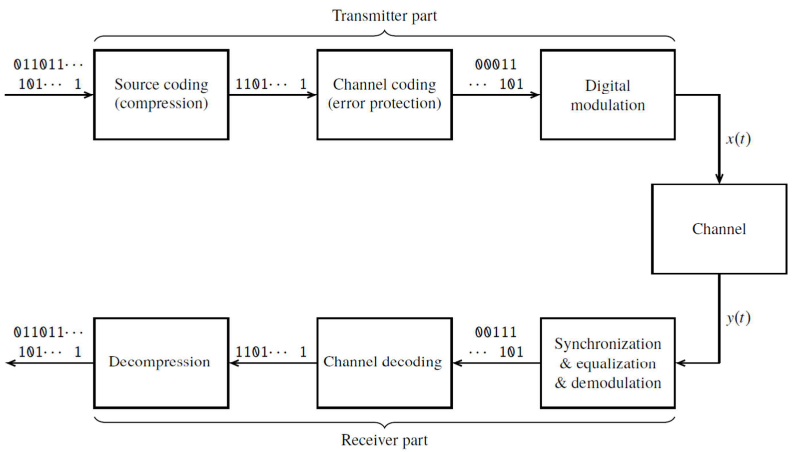
Redundans: $1.06 - 1.022 = 0.038$

Resultat av källkodning

Nästan lika sannolika
nästan okorrelerade bitar.

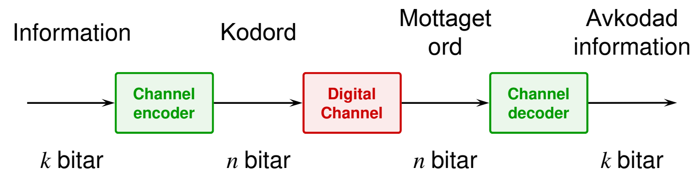
Ju närmare entropin vi kommer, desto närmare kommer vi lika sannolika och okorrelerade bitar.

Ett kommunikationssystem



Figur 1.11 ur kursboken

Block-koder – Grundidé



Beräkna r paritetsbitar från k informationsbitar.

Skicka $n = k + r$ kodordsbitar.

Mottaget: n eventuellt felaktiga bitar.

Avkoda till det mest troliga sända kodordet givet de mottagna bitarna.

Mer generellt:
Avbilda k informationsbitar på $n = k + r$ kodordsbitar.

Om kodorden är välvalda, då kommer vi att kunna korrigera fel.

Block-kod – Mängd av binära vektorer

Kod: $\mathscr{C} = \left\{ \bar{c}_i \in \{0,1\}^n \right\}_{i=1}^{2^k}$

Kodord: $\bar{c}_i = (c_{i,1}, \ldots, c_{i,n})$, $\in \{0,1\}$

Takt: $R = \dfrac{k}{n}$

Exempelkod:

Information	Kodord
00	10101010
01	11010000
10	01100111
11	00011101
$k = 2$	$n = 8$

Principer för avkodning:

Anta att alla fel är lika allvarliga. $\Rightarrow$

Välj det mest troliga kodordet givet den mottagna vektorn.

Avkodning

Stokastiska variabler: Sänt kodord: $\overline{C}$
Mottaget ord: $\overline{X}$

MAP-avkodning (Maximim À-Posteriori):

Avkodningsregel 1: Sätt $\hat{c} = \overline{c}_i$ om $\Pr\{\overline{C} = \overline{c}_k \mid \overline{X} = \overline{x}\}$ maximeras för $k = i$.

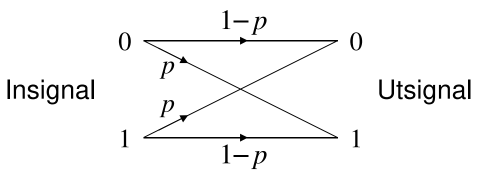
Bayes regel $\Rightarrow$

Avkodningsregel 2: Sätt $\hat{c} = \overline{c}_i$ om $\Pr\{\overline{C} = \overline{c}_k\}\Pr\{\overline{X} = \overline{x} \mid \overline{C} = \overline{c}_k\}$ max för $k = i$.

ML-avkodning $\left(\Pr\{\overline{C} = \overline{c}_k\} = 1/2^k\right)$ (Maximum Likelihood):

Avkodningsregel 3: Sätt $\hat{c} = \overline{c}_i$ om $\Pr\{\overline{X} = \overline{x} \mid \overline{C} = \overline{c}_k\}$ maximeras för $k = i$.

Binärsymmetriska kanaler (BSC)



Konsekutiva användningar av kanalen är oberoende.

Hammingavstånd: $d_H(\overline{a}, \overline{b})$ # positioner där $\overline{a}$ och $\overline{b}$ är olika.

Egenskaper: $d_H(\overline{a}, \overline{a}) = 0$ $d_H(\overline{a}, \overline{b}) \geq 0$ $d_H(\overline{a}, \overline{c}) \leq d_H(\overline{a}, \overline{b}) + d_H(\overline{b}, \overline{c})$

Då får vi: $\Pr\{\overline{X} = \overline{x} \mid \overline{C} = \overline{c}_k\} = p^{d_H(\overline{x}, \overline{c}_k)}(1-p)^{n - d_H(\overline{x}, \overline{c}_k)}$

ML-avkodning för BSC med felsannolikhet p (vi antar $0 < p < 0.5$):

Avkodningsregel 4: Sätt $\hat{c} = \overline{c}_i$ om $d_H(\overline{x}, \overline{c}_k)$ minimeras för $k = i$.

Exempel på avkodning

Information	Kodord
00	10101010
01	11010000
10	01100111
11	00011101
$k = 2$	$n = 8$

$d_H(10101010, 11010000) = 5$
$d_H(10101010, 01100111) = 5$
$d_H(10101010, 00011101) = 6$
$d_H(11010000, 01100111) = 6$
$d_H(11010000, 00011101) = 5$
$d_H(01100111, 00011101) = 5$

Minavstånd $d = 5$

Avkodning: Välj närmaste kodord

(10111010) avkodas till (10101010)
(11110111) avkodas till (01100111)

(00111000) är på avstånd 3 från både (10101010) & (00011101)

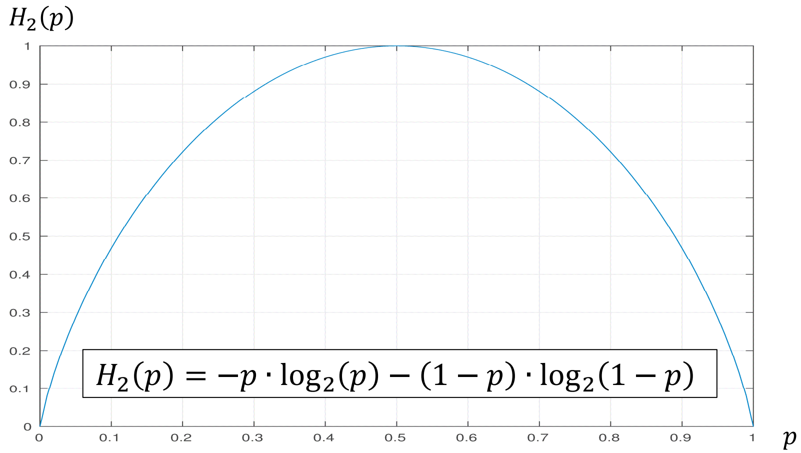
Felrättningsförmåga

$t = \left\lfloor \frac{d-1}{2} \right\rfloor$ Här $t = 2$

Feldetekteringsförmåga

$v = d - 1$ Här $v = 4$

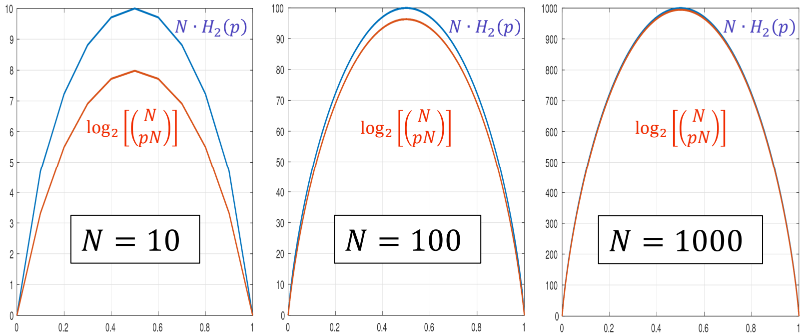
Binära entropifunktionen (Fö 7, slide 2)



Motsvarar figur 6.2 i kursboken

Användbar approximation (Fö 7, slide 3)

$$\log_2\left[\binom{N}{pN}\right] \approx N \cdot H_2(p)$$



Mikael Olofsson
ISY/CommSys

www.liu.se