

TSKS10 Signaler, information & Kommunikation

Föreläsning 6

Datakompression, explicita koder

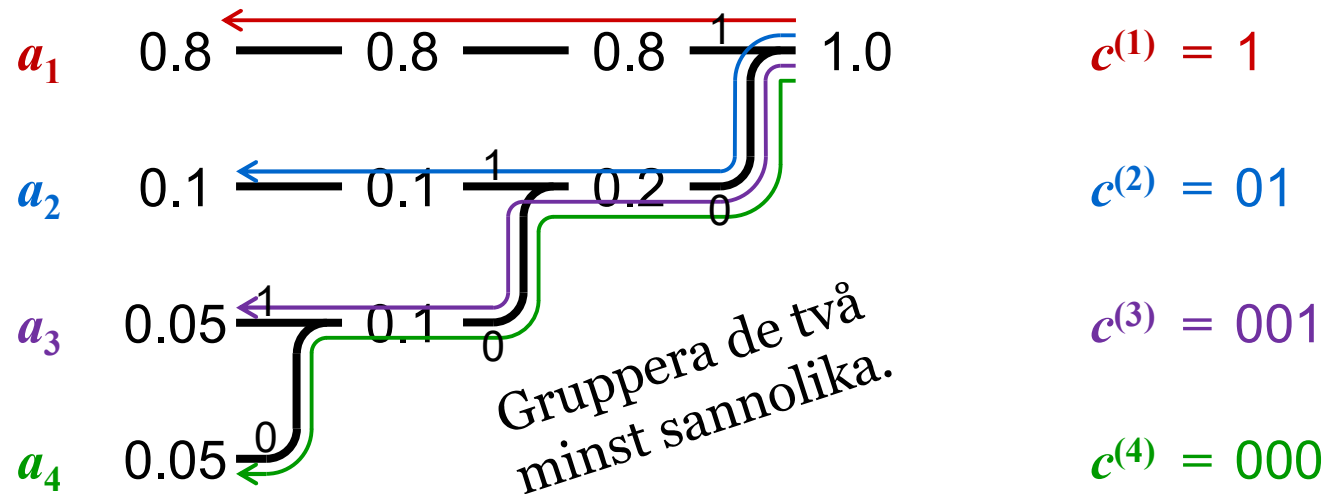
Introduktion kanalkodning

Mikael Olofsson

Institutionen för Systemteknik (ISY)

Ämnesområdet Kommunikationssystem

Huffmankodning (förenklad algoritm)



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

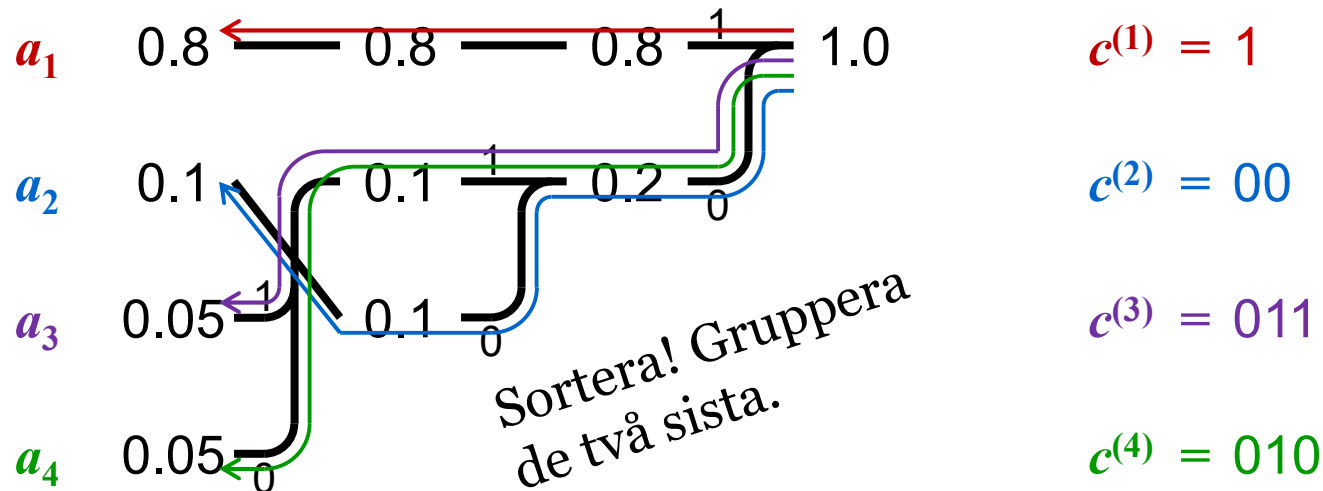
$$m_L = 1.3$$

Entropi: $H(A) = - \sum_{i=1}^M p_i \log_2(p_i) \approx 1,022$

Redundans: $m_L - H(A) \approx 0,278$

Kompressionskvot: $\frac{\log_2(M)}{m_L} \approx 1,54$

Huffmankodning (fullständig algoritm)



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: $H(A) = - \sum_{i=1}^M p_i \log_2(p_i) \approx 1,022$

Redundans: $m_L - H(A) \approx 0,278$

Kompressionskvot: $\frac{\log_2(M)}{m_L} \approx 1,54$

Undre och övre gräns

Källstatistik: $p_i = P(A = a_i)$

Kodordslängd: $L: A = a_i \Rightarrow L = l_i$

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

”Ideala kodordslängder”: $l_i = -\log_2(p_i)$

Problem: Är kanske inte heltal.

Möjligt val: $l_i = \lceil -\log_2(p_i) \rceil$

Resultat: Krafts olikhet uppfylld.

Gränser: $-\log_2(p_i) \leq l_i < 1 - \log_2(p_i)$

$$- \sum_i p_i \log_2(p_i) \leq \sum_i p_i l_i < \sum_i p_i (1 - \log_2(p_i))$$

$$H(A) \leq m_L < 1 + H(A)$$

Källutvidgning

Koda en symbol i taget: $H(A) \leq m_L < 1 + H(A)$ [bitar/symbol]

Oberoende symboler: $H(A_1, \dots, A_K) = K \cdot H(A)$ [bitar/ K symboler]

Koda två ober. symboler: $H(A_1, A_2) \leq m_{L_2} < 1 + H(A_1, A_2)$ [bitar/symbolpar]

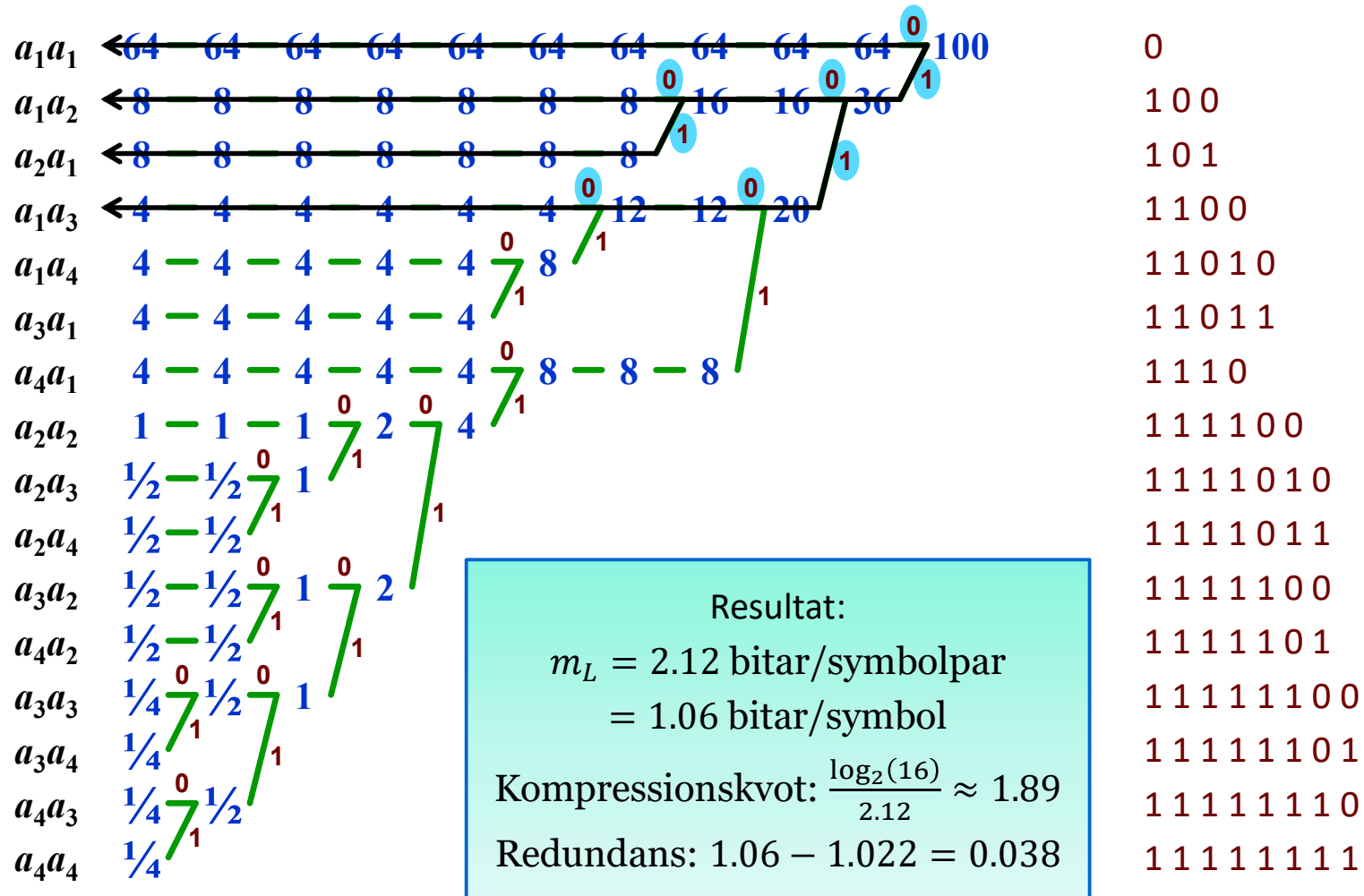
$2H(A) \leq m_{L_2} < 1 + 2H(A)$ [bitar/symbolpar]

$H(A) \leq \frac{m_{L_2}}{2} < \frac{1}{2} + H(A)$ [bitar/symbol]

Koda K ober. symboler: $H(A_1, \dots, A_K) \leq m_{L_K} < 1 + H(A_1, \dots, A_K)$ [bitar/ K symboler]

$H(A) \leq \frac{m_{L_K}}{K} < \frac{1}{K} + H(A)$ [bitar/symbol]

Förenklad huffmankod för en utvidgad källa



Skurlängdskodning

– för källor med kraftigt minne

1 1 1 1 0 0 0 1 1 0 1 1 1 1 1 0 0 0 0 0 0 0 1 1 1 0 0 1 1 1 1 1 1 0 ...

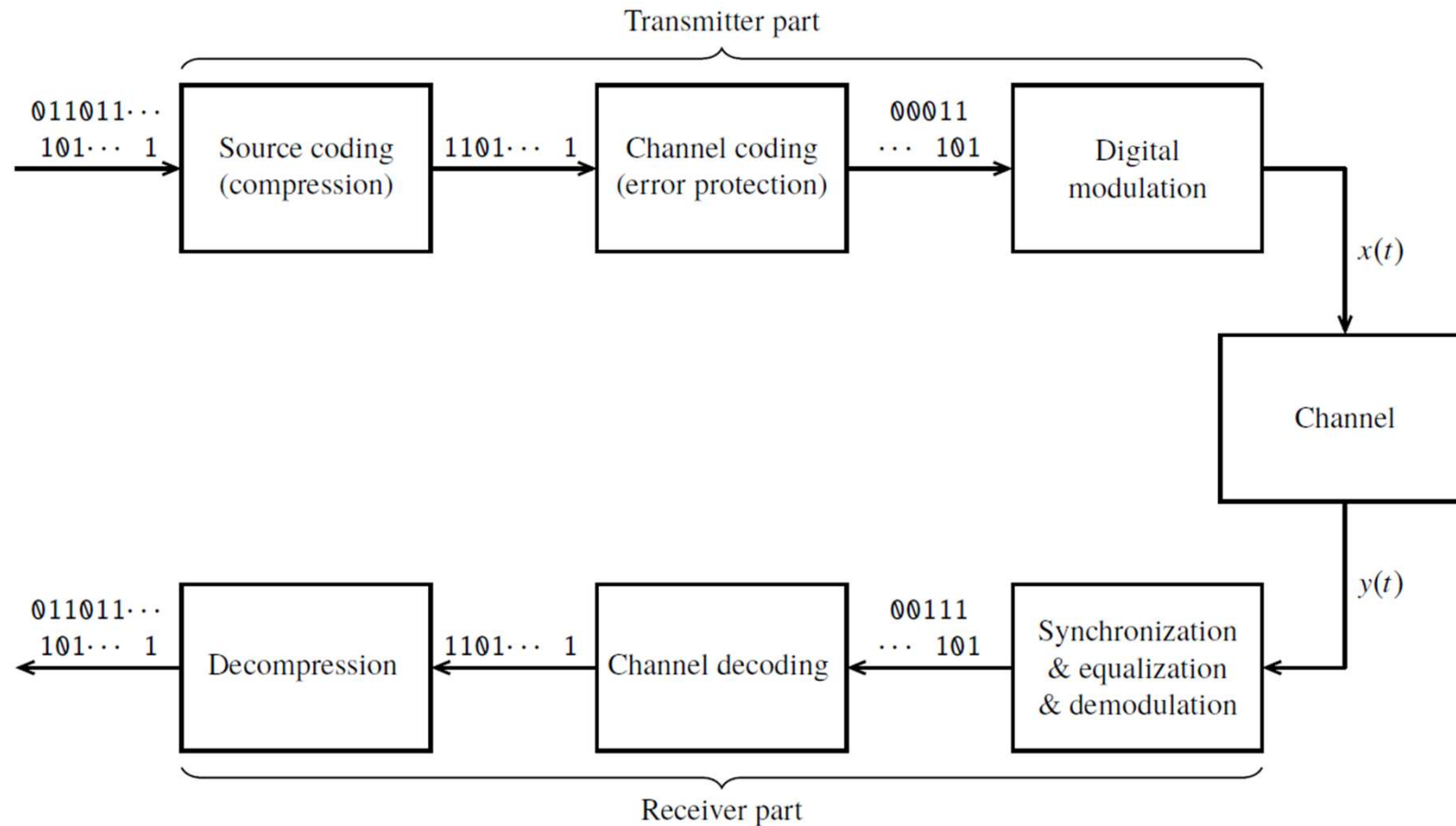
4 3 2 1 5 7 3 2 6 ...

Resultat av källkodning

Nästan lika sannolika
nästan okorrelerade bitar.

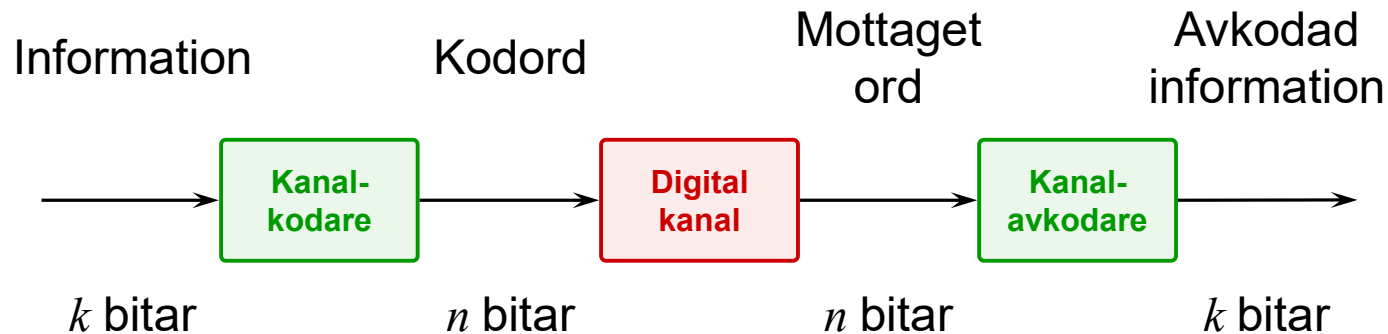
Ju närmare entropin vi kommer, desto
närmare kommer vi lika sannolika och
okorrelerade bitar.

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

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

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

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.

Bayes regel

Betingad sannolikhet:

$$P(A|B) = \frac{P(A, B)}{P(B)}$$

$$P(B|A) = \frac{P(A, B)}{P(A)}$$

$\Rightarrow$

$$P(A, B) = P(A|B)P(B) = P(B|A)P(A)$$

$\Rightarrow$

$$P(A|B) = \frac{P(B|A)}{P(B)} P(A)$$

Avkodning

Stokastiska variabler: Sönt kodord: $\bar{C}$
Mottaget ord: $\bar{X}$

MAP-avkodning (Maximim À-Posteriori):

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

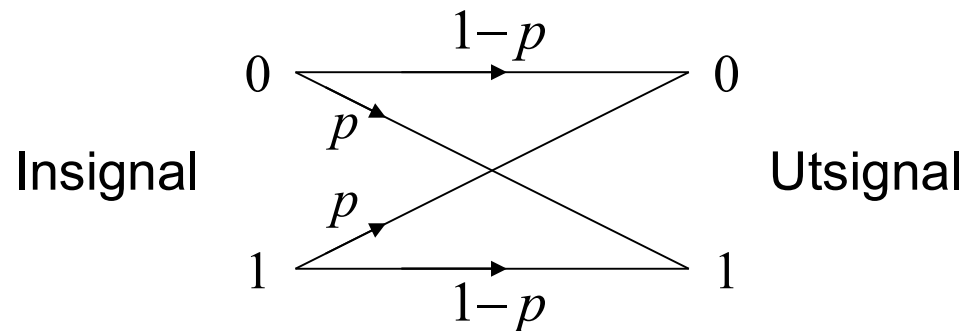
Bayes regel $\Rightarrow$

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

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

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

Binärsymmetriska kanaler (BSC)



Konsekutiva
användningar av
kanalen är oberoende.

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

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

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

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

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

Exempel på avkodning

Information	Kodord		
00	10101010		$d_H(10101010, 11010000) = 5$ $d_H(10101010, 01100111) = 5$ $d_H(10101010, 00011101) = 6$ $d_H(11010000, 01100111) = 6$ $d_H(11010000, 00011101) = 5$ $d_H(01100111, 00011101) = 5$
01	11010000		
10	01100111		
11	00011101		
$k = 2$	$n = 8$		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

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

Feldetekteringsförmåga

$$v = d-1 \quad \text{Här } v=4$$

Mikael Olofsson
ISY/CommSys

www.liu.se