Vorlesung am 05.05.06

4.V

IC SS-06

1.5: Das Codierungstheorem von Shannon (Fort.)

2.) Es ist immer möglich, einen Code zu finden, wobei

$\bar{N} - H(X)$ beliebig klein ist

Beweis

M^2 - Elemente

Betrachtet man $X^{(2)} = \{(x_1, x_1), \dots (x_M, x_M)\}$

$$p^{(2)} = \{p_1 \cdot p_1, p_1 \cdot p_2, \dots p_M \cdot p_M\}$$

gedächtnislos angenommen

$$I^{(2)} = \{2I_1, (I_1 + I_2), \dots 2I_M\}$$

$$H(X^{(2)}) = 2H(X)$$

Nach der bereits erklärten Strategie gilt:

$$I_i \leq N_i < I_i + 1$$

$$\Longrightarrow \quad p_i I_i \leq p_i N_i < p_i I_i + p_i$$

$$\underbrace{\sum_{i=1}^{M} p_i I_i}_{H(X)} \leq \underbrace{\sum_{i=1}^{M} p_i N_i}_{\bar{N}} < 1 + \sum_{i=1}^{M} p_i I_i$$

$$\Longrightarrow \quad H(X) \le \bar{N} < H(X) + 1$$

Es gilt auch $\quad H(X^{(2)}) \le \bar{N}^{(2)} < H(X^{(2)}) + 1$

Mit $\quad H(X^{(2)}) = 2\,H(X) \;;\; \bar{N}^{(2)} = 2\,\bar{N}$

erhält man $\quad H(X) \le \bar{N} < H(X) + \frac{1}{2}$

Im Allgemeinen gilt:

$$X^{(n)} = \Big\{ \overbrace{(x_1, x_1, \dots, x_1)}^{n\text{-Elemente}}, \dots\dots, (x_M, x_M, \dots, x_M) \Big\}$$

M^n-Elemente

$$H(X^{(n)}) = n \cdot H(X) \qquad \text{(gedächtnisfrei)}$$

$$\bar{N}^{(n)} = n \cdot \bar{N}$$

$$H(X^{(n)}) \le N^{(n)} < H(X^{(n)}) + 1$$

$$\Longrightarrow \quad H(X) \le \bar{N} < H(X) + \frac{1}{n}$$

1.6: Ausgewählte Codierungsverfahren

1.6.1: Das Shannon-Fano-Verfahren

1.) Die Symbole werden ihrer Wahrscheinlichkeit nach abnehmend geordnet.

2.) Zwei Gruppen mit je "mehr oder weniger" 50 % Auftrittswahrscheinlichkeit werden gebildet; eine bekommt "0" als ersten (z.B. von Links) Bit, die zweite bekommt "1".

3.) Der o.g. Schritt wird mit jeder Gruppe wieder halt, bis die Untergruppen nur noch ein Symbol enthalten

Beispiel $X = \{x_1, x_2, \ldots, x_9\}$

$$P = \{0{,}36, 0{,}2, 0{,}18, 0{,}12, 0{,}08, 0{,}02, 0{,}02, 0{,}01, 0{,}01\}$$

x	p										ΣN	$\sum P_i$	$\sum P_i$
x_1	0,36	0	0									0,36	0,56
x_2	0,20	0	1									0,20	
x_3	0,18	1	0									0,18	
x_4	0,12	1	1	0							0,12		
x_5	0,08	1	1	1	0					0,08			
x_6	0,02	1	1	1	1	0			0,2		0,14	0,26	0,44
x_7	0,02	1	1	1	1	1	0		0,4	0,06			
x_8	0,01	1	1	1	1	1	1	0					
x_9	0,01	1	1	1	1	1	1	1					

$x_1 \longleftrightarrow 00$ ——— $N_1 = 2$

$x_2 \longleftrightarrow 01$ ——— $N_2 = 2$

$x_3 \longleftrightarrow 10$ ——— $N_3 = 2$

$x_4 \longleftrightarrow 110$ ——— $N_4 = 3$

$x_5 \longleftrightarrow 1110$ ——— $N_5 = 4$

$x_6 \longleftrightarrow 11110$ ——— $N_6 = 5$

$x_7 \longleftrightarrow 111110$ ——— $N_7 = 6$

$x_8 \longleftrightarrow 1111110$ ——— $N_8 = 7$

$x_9 \longleftrightarrow 1111111$ ——— $N_9 = 7$

1.6.2: Das Huffmanische Verfahren

1-) Die Symbole werden ihrer Wahrscheinlichkeiten nach abnehmend geordnet

2-) Die letzten zwei Symbole (mit den kleinsten Auftrittswahrscheinlichkeiten) werden in ein Symbol zusammengefasst, dessen Wahrscheinlichkeit die Summe der einzelnen Wahrscheinlichkeiten ist. Nach der Codierung dieses Zusammengefassten Symbols, werden die einzelnen Symbole den gleichen Code nehmen, allerdings mit einer zusätzlichen "0" und "1" am Ende.

3-) Die Symbole werden neu geordnet und der o.g. Schritt wird wiederholt, bis es nur zwei Symbole übrig bleiben.

Beispiel: Das gleich wie bei Shannon-Fano

A	B	C	D	E	F	G	H	I
36	20	18	12	8	2	2	1 $\boxed{0}$	1 $\boxed{1}$

$H^*(2)$

A	B	C	D	E	F	G	H^*
36	20	18	12	8	2	2 $\boxed{0}$	2 $\boxed{1}$

$G^*(4)$

A	B	C	D	E	G^*	F
36	20	18	12	8	4 $\boxed{0}$	2 $\boxed{1}$

$F^*(6)$

A	B	C	D	E	F^*
36	20	18	12	8 $\boxed{0}$	6 $\boxed{1}$

$E^*(14)$

A	B	C	E^*	D
36	20	18	14 $\boxed{0}$	12 $\boxed{1}$

$D^*(26)$

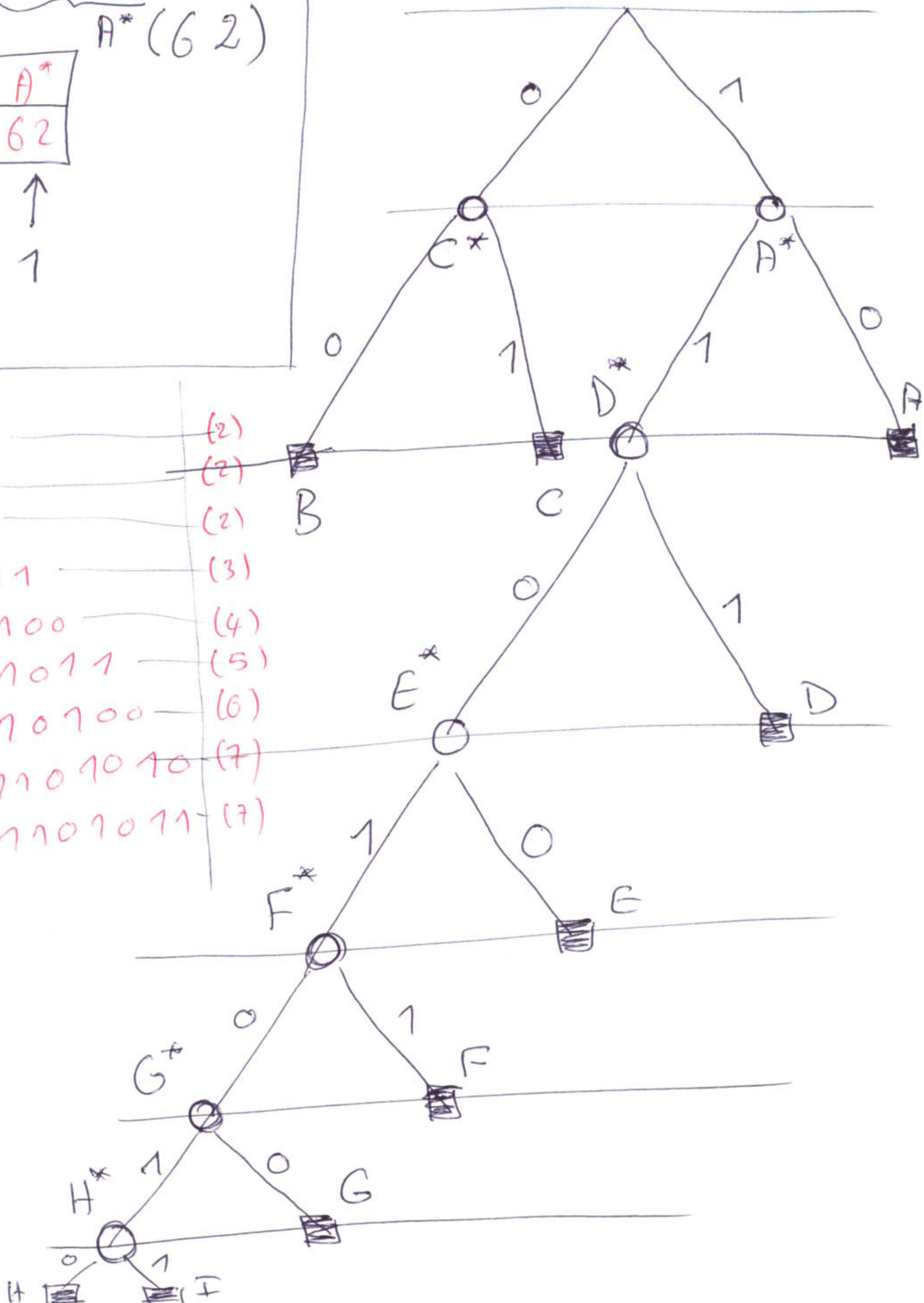
A	D^*	B	C
36	26	20 $\boxed{0}$	18 $\boxed{1}$

$C^*(38)$

C^*	A	D^*
38	36 $\boxed{0}$	26 $\boxed{1}$

A^* (62)

C^*	A^*
38	62
↑ 0	↑ 1



A ↔ 10 (2)
B ↔ 00 (2)
C ↔ 01 (2)
D ↔ 111 (3)
E ↔ 1100 (4)
F ↔ 11011 (5)
G ↔ 110100 (6)
H ↔ 1101010 (7)
I ↔ 1101011 (7)

1.7: Diskrete Informationsquellen mit Gedächtnis

1.7.1: Die bedingte Entropie und die Entropie einer Quelle mit Gedächtnis

Gegeben sei die Quelle $X = \{x_1, x_2, \dots, x_M\}$

Wir definieren: $x_i^{(m)}$ = das ausgewählte Symbol zum Zeitpunkt mT_0

$x_i^{(m-1)}$ = das ausgewählte Symbol zum Zeitpunkt $(m-1)T_0$

$\vdots$

Definitionen:

Gedächtnis 1. Ordnung

$$P(x_i^{(m)} \mid x_j^{(m-1)}) \neq P(x_i)$$

Gedächtnis 2. Ordnung.

$$P\left(x_i^{(m)} \mid \left(x_j^{(m-1)}, x_h^{(m-2)}\right)\right) \neq P\left(x_i^{(m)} \mid x_j^{(m-1)}\right)$$

⋮

etc.

Es ist zu erwarten, daß $P\left(x_i^{(m)} \mid x_j^{(m-1)}\right)$ hetrogener als $P(x_i)$ ist, falls Gedächtnis vorhanden ist. Das gleiche gilt für $P\left(x_i^{(m)} \mid \left(x_j^{(m-1)}; x_h^{(m-2)}\right)\right)$, die hetrogener als $P\left(x_i^{(m)} \mid x_j^{(m-1)}\right)$ ist und so weiter

Wir definieren nun die <u>bedingten</u> Entropien

$$H\left(X \mid x_j^{(m-1)}\right) = \sum_{i=1}^{M} P\left(x_i^{(m)} \mid x_j^{(m-1)}\right) \log_2 \frac{1}{P\left(x_i^{(m)} \mid x_j^{(m-1)}\right)}$$

Wegen der Tatsache, daß $P\left(x_i^{(m)} \mid x_j^{(m-1)}\right)$ hetrogener als $P(x_i)$ ist, haben wir:

$$H(X|x_j^{(m-1)}) \leq H(X)$$

Die Entropie der Quelle mit Gedächtnis wird als der Mittelwert der bedingten Entropien definiert:

$$H^{(n)}(X) = \sum_{j=1}^{M} P(x_j) \cdot H(X|x_j^{(m-1)})$$

$$= \sum_{j=1}^{M} \sum_{i=1}^{M} \overbrace{P(x_j)\, P(x_i^{(m)}|x_j^{(m-1)})}^{= P(x_i^{(m)}; x_j^{(m-1)})} \cdot \log_2 \frac{1}{P(x_i^{(m)}|x_j^{(m-1)})}$$

$$H^{(n)}(X) = \sum_{i=1}^{M} \sum_{j=1}^{M} P(x_i^{(m)}; x_j^{(m-1)}) \log_2 \frac{1}{P(x_i^{(m)}|x_j^{(m-1)})}$$

Da $H(X|x_j^{(m-1)}) \leq H(X)$

$$\Rightarrow H^{(n)}(X) \leq \sum_{j=1}^{M} p(x_j) \cdot H(X)$$

$$= H(X) \underbrace{\sum_{j=1}^{M} p(x_j)}_{=1}$$

$$= H(X)$$

$$\Downarrow$$

$$\boxed{H^{(n)}(X) \leq H(X)}$$

<u>Verallgemeinerung</u>

$$H^{(m)}(X) \leq H^{(m-1)}(X) \leq \cdots \leq H^{(1)}(X) \leq H(X)$$

Wegen der Tatsache, daß $H^{(n)}(X) \leq H(X)$, ist die Kompression der Quelle mit Gedächtnis effizienter als die der Quelle ohne Gedächtnis. Dies bedeutet, daß die mittlere Codelänge weiter minimiert werden kann.