Do. 18.10:

- Informationsgehalt (nach Shannon 1948)

Auftrittswahrscheinlichkeit „p"

Wahrscheinlichkeit für den Buchstaben z.B. R groß $\longleftrightarrow$ Informationsgehalt klein

$$h = f\left(\frac{1}{p}\right)$$

h … Informationsgehalt

p … Auftrittswahrscheinlichkeit eines Zeichens

mehrere voneinander unabhängige Zeichen:

Informationsgehalt = Σ Informationsgehalt der einzelnen Zeichen

$$f\{p(x) + p(y)\}$$

Korrelation … haben Zeichen eine Abhängigkeit voneinander?

z.B. sch, st, sp, …

gedächtnislose Quelle (Unabhängigkeit angenommen):

$$f\{p(x) \cdot p(y)\}$$

(gültig für ein Verbundereignis)

$$f\{p(x) \cdot p(y)\} = f\{p(x)\} + f\{p(y)\}$$

Bsp:

$p(x_i) = 1$

$\rightarrow h(x_i) = 0$

$$\log\left\{\frac{1}{p(x)} \cdot \frac{1}{p(y)}\right\} = \log\left\{\frac{1}{p(x)}\right\} + \log\left\{\frac{1}{p(y)}\right\}$$

Annahme: Alphabet 2 Zeichen : 0 , 1

$$p(0) = p(1) = \frac{1}{2} \text{ mit } p(0) + p(1) = 1$$

1 bit = 1 binary digit ist der Informationsgehalt für Ereignis „0", „1"

Wahl der Basis der Logarithmusfunktion „2"

Informationsgehalt:

$$h = ld\left(\frac{1}{p}\right) = -ld(p)$$ ld ... Logarithmus Dualis

ld:
$$a = 2^b$$
$$\left.\begin{aligned} ld\,a &= b\,ld\,2 \\ ln\,a &= b\,ln\,2 \end{aligned}\right\} \frac{ld\,a}{ln\,a} = \frac{1}{ln\,2}$$

$$\rightarrow \boxed{ld\,a = \frac{ln\,a}{ln\,2}} \approx 1{,}443\,ln\,a$$

$\rightarrow p(x_1) = p(x_2) = \ldots = p(x_n) = \frac{1}{n}$

$\rightarrow$ Informationsgehalt:

$$h = ld\,\frac{1}{p(x)} = ld(n)$$

Informationsgehalt eines Zeichen x in einer Binärquelle

$$h = ld\,\frac{1}{p(x_i)}$$ in bit

Bsp.: eine Dezimalziffer binär codieren

(0 .... 9) Ziffern gleichwahrscheinlich

$$\Rightarrow p(x) = \frac{1}{10}$$

$$h = ld\,\frac{1}{p(x)} = ld\,10 \approx 3{,}32 \text{ Bit}$$

Grundsatz: Binäre Entscheidungen müssen immer ganzzahlig sein!

Man benötigt 4 Bits zur Darstellung von den Ziffern 0 bis 9.

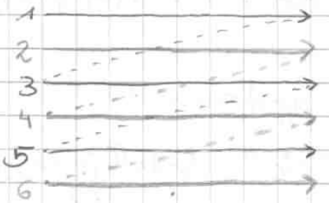
→ keine optimale Codierung mehr!

- Informationsfluss vom Sender S zum Empfänger E

$$C = \frac{h}{t} \quad \text{(in bps oder Bd)}$$

$$1\ Bd = 1\ \frac{bit}{s}$$

Bsp.: Farbfernsehbild (digital)
625 Zeilen
800 Bildpunkte je Zeile
25 Vollbilder pro s
(50 Halbbilder / s)



Halbbilder

16 Helligkeitsstufen für die Bildpunkte
256 Farbstufen

$$C = \underbrace{625 \cdot 800}_{1\ Bild} \cdot 25\ \frac{Bilder}{s} \cdot (\underset{Helligkeit}{4} + \underset{Farbstufen}{8})\ bit$$

$$C = 15\ M\ bit/s$$

Halbbilder: gerade Zahlen und ungerade Zahlen verschiedene Farben!

$h(x_i) = ld \frac{1}{p(x_i)} = - ld(p(x_i))$

Erwartungswert = Mittelwert des Informationsgehalt aller Zeichen als gewichtete Summe

$H = \sum_i p(x_i) \cdot h_i = \sum_i p(x_i) \cdot ld \frac{1}{p(x_i)} = - \sum_i p(x_i) \cdot ld(p(x_i))$

Entropie oder mittlerer Informationsgehalt
(average information contends)

Die Entropie ist ein Maß für die Ungewissheit der Quelle.

Exkurs:

Tätigkeitsfelder des Informatikers

USW	User Software
OS	Operating systems
QAS	Quality Assurance (Software Qualitätssicherung)
SYSADMIN	Systemadministration
XPS, AI	Expert Systems, Artificial Intelligence
SW-Tools	Entwicklung von SW-Tools
SW-Management	Entwicklung / Einsatz von Management-Tools

Ab Tätigkeitsfeldern eines Informatikers: Fr. 19.10:

- Informationsgehalt:

$$h\,[p(x_i)] = ld\,\frac{1}{p(x_i)} = -\,ld\,(p(x_i))$$

- Informationsfluss:

$$c = \frac{h}{t} \quad \text{(als Mittelwert)}$$

- mittlere Informationsgehalt = Erwartungswert als gewichtete Summe

$$\underbrace{H = \sum_i p(x_i) \cdot h(x_i)}_{\text{Entropie}} = \sum_i p(x_i) \cdot ld\,\frac{1}{p(x_i)}$$

- mittlere Wortlänge

l_i = Länge des i-ten Codewortes

$$L = \sum_i p(x_i) \cdot l_i$$

Bsp: Nachrichtenquelle mit 3 Zeichen x, y, z

Voraussetzungen:
Auftrittswahrscheinlichkeit und Codierung bekannt

Codierung		p(...)	h	p · h	l	p · l
1	x	0,7	0,515	0,360	1	0,7
01	y	0,2	2,322	0,464	2	0,4
00	z	0,1	3,322	0,332	2	0,2

$$H = \sum_i p(x_i) \cdot h(x_i) \;\rightarrow\; 1{,}156 \text{ Bit} \qquad L = \sum_i p(x_i) \cdot l_i \;\rightarrow\; 1{,}300 \text{ Bit}$$

$$H < L$$

Die Differenz zwischen Wortlänge und Entropie ist die Redundanz (in bit)

$$R = L - H$$

R ... gibt an, um wieviel Bit ein Wort eines Binärcodes im Mittel länger ist, als im optimalen Fall nötig wäre

Relative Redundanz:

$$r = \frac{R}{L}$$

Optimalcode ... Redundanz minimal!

◦) Die Shannonsche Funktion

Entropie $= H = \sum_i p_i \cdot ld \frac{1}{p_i}$

Spezialisierung: Binärquelle (log 0 und log 1)

Zeichen	$p(\,)$
0	p
1	$1-p$
	$\sum p_i = 1$

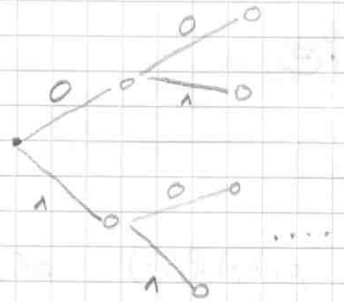
$$\rightarrow H = p \cdot \overset{\text{„}0\text{"}}{ld \frac{1}{p}} + (1-p) \cdot \overset{\text{„}1\text{"}}{ld \frac{1}{(1-p)}}$$

$$\downarrow$$

$$S = -p\, ld\, p - (1-p)\, ld\, (1-p)$$

$$\lim_{p \to 0} \left(p \cdot ld \frac{1}{p}\right)$$

$$\lim_{p \to \infty} \frac{ld\left(\frac{1}{p}\right)}{\frac{1}{p}} = \frac{-\infty}{0} \quad \text{De l'Hospital}$$



i	x_i	$p(x_i)$	l_i	Codewort
1	x_1	$\frac{1}{4}$	2	00
2	x_2	$\frac{1}{4}$	2	01
3	x_3	$\frac{1}{8}$	3	100
4	x_4	$\frac{1}{8}$	3	101
5	x_5	$\frac{1}{8}$	3	110
6	x_6	$\frac{1}{16}$	4	1110
7	x_7	$\frac{1}{16}$	4	1111

$$H = \sum_{i=1}^{7} p(x_i) \cdot \mathrm{ld} \frac{1}{p(x_i)} = 2{,}625 \text{ Bit}$$

$$L = \sum_{i=7}^{7} p(x_i) \cdot l_i = 2{,}625 \text{ Bit}$$

$$R = L - H = 0$$

$\Rightarrow$ Optimalcode!

R kann null werden, wenn die Auftritts=wahrscheinlichkeiten der Zeichen durch $\frac{1}{2^{l_i}}$ dargestellt werden können.

3 Größen:

- H = mittlerer Informationsgehalt
- L = Mittlere Codewort länge
- H_0 = Entscheidungsgehalt

dabei:

$R = L - H$ Redundanz

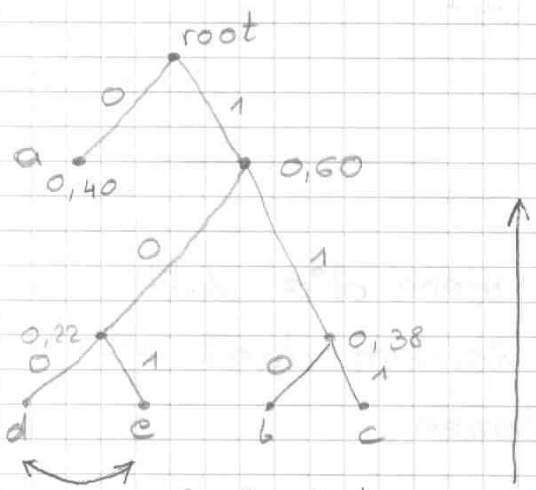
Codierungstheorie

- Huffman-Code

x_i	$p(x_i)$			$\Sigma = 1$	Code	$p(x_i)\,h(x_i)$	$p(x_i)\,l_i$
a	0,40	0,40	0,40	0,60	0	0,53	0,40
b	0,20	0,22	0,38	0,40	110	0,46	0,50
c	0,18	0,20	0,22		111	0,45	0,54
d	0,11	0,18			100	0,35	0,33
e	0,11				101	0,35	0,33
						2,14 Bit	2,20 Bit
						H	L

muss bekannt sein

$R = L - H = 0{,}06$ Bit



lexikographische Ordnung

bottom-up