

Felix Pistorius

pistorius@kit.edu

Karlsruher Institut für Technologie (KIT)

Übung - Digitaltechnik

2. Übung

www.kit.edu

Shannons Informationstheorie



Definition Der Informationsgehalt eines Zeichens hängt logarithmisch von dessen Auftrittswahrscheinlichkeit ab:

$$H_e = \log_2\left(\frac{1}{p}\right)$$

Definition Mittlere Codewortlänge:

$$\overline{m} = \sum_{i=1}^{n} p(x_i) \cdot m(x_i)$$

Definition Entropie der Sendequelle:

$$H = \sum_{i=1}^{n} p(x_i) \cdot H_e(x_i) = \sum_{i=1}^{n} p(x_i) \cdot ld \frac{1}{p(x_i)}$$



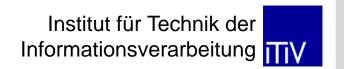
Shannons Informationstheorie



- Postulat Es gibt kein Codierungsverfahren mit
 - $\overline{m} < H$
 - → H ist das theoretische Minimum der mittleren Codewortlänge
- Codierungsverfahren, die \overline{m} möglichst dicht an ihren möglichen (Minimal-)Wert H heranführen, heißen optimale Codes
- Vereinfacht gesagt gilt für optimale Codes:
 - Häufig auftretende Symbole
 - → möglichst kurze Codewörter
 - seltene Symbole
 - → längere Codewörter







Felix Pistorius

pistorius@kit.edu

Karlsruher Institut für Technologie (KIT)





- Einige Studenten wollen ihre Verabredungen über das Internet koordinieren.
- Zu diesem Zweck legen sie den Zeitpunkt fest, nur über den Treffpunkt kommt keine Einigkeit zustande.
- Um nun die Übertragungszeiten gering zu halten, beschließen sie, die beliebtesten Treffpunkte im Fanø-Code zu kodieren.





Die Gaststätten sollen in "billig" und "teuer" eingeteilt werden. Wie groß ist der Informationsgehalt der Aussage "billig", wenn "billig" Preise bis max. € 3,00 beinhaltet?

Gaststätte	Bevorzugt von Anzahl an Studenten	Kosten pro Getränk / €
Brasil	1	4,00
Dom	2	4,00
Krokodil	1	3,50
Miro	2	3,00
Titanic	6	2,50
Ubu	6	3,00
Wien	5	3,50
Zero	3	3,50



Die Gaststätten sollen in "billig" und "teuer" eingeteilt werden. Wie groß ist der Informationsgehalt der Aussage "billig", wenn "billig" Preise bis max. € 3,00 beinhaltet?

$$p = \frac{\sum \text{ billige Kneipen}}{\sum \text{ aller Kneipen}} = \frac{3}{8} = 0,375$$

$$H_e = ld (1/p) = ld (1/0,375) = 1,42 bit$$

Shannon-Fanø-Code



- Konstruktionsvorschrift:
 - Zeichenvorrat nach aufsteigender Wahrscheinlichkeit linear sortieren → lineare Ordnung
 - Zwei Teilmengen so konstruieren, dass die Summenwahrscheinlichkeiten der beiden Teilmengen möglichst gleich groß sind, d.h. die addierten Auftrittswahrscheinlichkeiten der in beiden Teilmengen enthaltenen Zeichen ergeben möglichst die gleiche Summe
 - Die erste Teilmenge erhält das Codierungszeichen "0" und die zweite Teilmenge eine "1"
 - Mit den jeweils resultierenden Teilmengen rekursiv in gleicher Weise fortfahren
 - → bis nur noch 1 Zeichen in resultierenden Teilmengen enthalten ist

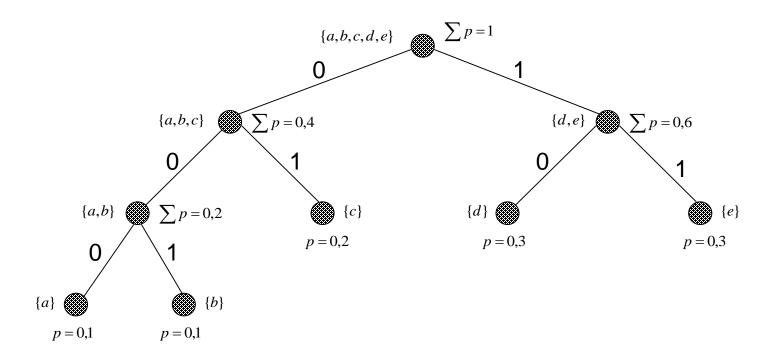


Shannon-Fanø-Code



Beispiel:

Zeichen	а	b	С	d	е
Wahrscheinlichkeit	0,1	0,1	0,2	0,3	0,3



Codiorung	Zeichen	а	b	С	d	е
Codierung:	Kodierung	000	001	01	10	11



Berechnen Sie die Wahrscheinlichkeit für den Besuch jeder einzelnen Gaststätte. Gehen Sie davon aus, dass die Wahrscheinlichkeit von der Beliebtheit abhängig ist.

Gaststätte	Bevorzugt von	Kosten pro Getränk / €	Auftrittswahrscheinlichkeit: p = Einzel El. / Gesamt
Brasil	1	4,00	$p(Brasil) = 1 / 26 \approx 0.0385$
Dom	2	4,00	$p(Dom) = 2 / 26 \approx 0.077$
Krokodil	1	3,50	$p(Krokodil) = 1 / 26 \approx 0.0385$
Miro	2	3,00	$p(Miro) = 2 / 26 \approx 0.077$
Titanic	6	2,50	p(Titanic) = $6 / 26 \approx 0.231$
Ubu	6	3,00	$p(Ubu) = 6 / 26 \approx 0.231$
Wien	5	3,50	p(Wien) = $5 / 26 \approx 0.192$
Zero	3	3,50	$p(Zero) = 3 / 26 \approx 0.115$
	Σ = 26	ı	$\Sigma = 1$



Gaststätte	Bevorzugt von	Kosten pro Getränk / €	р	_
Brasil	1	4,00	0.0385	
Dom	2	4,00	0.077	
Krokodil	1	3,50	0.0385	
Miro	2	3,00	0.077	sortieren nach p!
Titanic	6	2,50	0.231	cornoron naon p.
Ubu	6	3,00	0.231	
Wien	5	3,50	0.192	
Zero	3	4,50	0.115	

Gast- stätte	Titanic	Ubu	Wien	Zero	Dom	Miro	Brasil	Krokodil
р	0,231	0,231	0,192	0,115	0,077	0,077	0,0385	0,0395

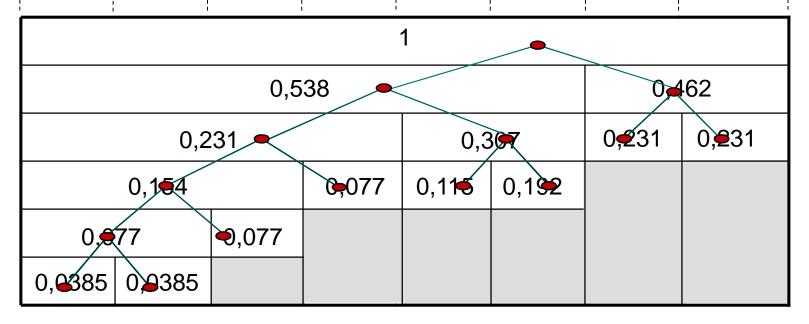


Gast- stätte	Krokodil	Brasil	Miro	Dom	Zero	Wien	Ubu	Titanic
р	0,0385	0,0385	0,077	0,077	0,115	0,192	0,231	0,231
	1	_	-				- - - -	

	1								
			0,462						
	0,2	231		0,3	307	0,231	0,231		
	0,154		0,077	0,115	0,192				
0,077 0,0		0,077							
0,0385	0,0385								



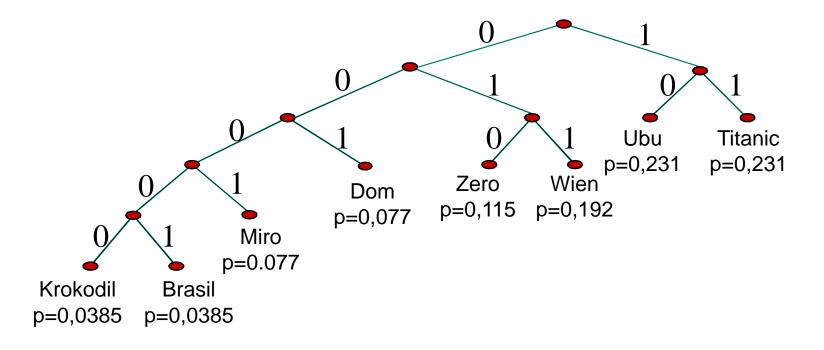
Gast- stätte	Krokodil	Brasil	Miro	Dom	Zero	Wien	Ubu	Titanic
р	0,0385	0,0385	0,077	0,077	0,115	0,192	0,231	0,231
	!	:	:				 	







Gast- stätte	Krokodil	Brasil	Miro	Dom	Zero	Wien	Ubu	Titanic
Codierung	00000	00001	0001	001	010	011	10	11





Wie hoch ist die mittlere Codewortlänge und welchen Wert weist das theoretische Minimum auf?

Gast- stätte	Titanic	Ubu	Wien	Zero	Dom	Miro	Brasil	Krokodil
р	0,231	0,231	0,192	0,115	0,077	0,077	0,0385	0,0385
m	2	2	3	3	3	4	5	5

Mittlere Codewortlänge:

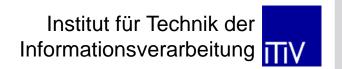
$$\overline{m} = \sum_{i=1}^{n} p(x_i) \cdot m(x_i) = 0.231 \cdot 2 + ... + 0.0385 \cdot 5 \approx 2,77 \text{ bit}$$

Das theoretische Minimum (durchschnittlicher Informationsgehalt):

$$H = \sum_{i=1}^{n} p(x_i) \cdot ld \frac{1}{p(x_i)} = 0.231*ld \frac{1}{0.231} + ... + 0.0385*ld \frac{1}{0.0385} \approx 2,72 \text{ bit}$$







Felix Pistorius

pistorius@kit.edu

Karlsruher Institut für Technologie (KIT)





 Ergänzen Sie folgende Tabelle durch Konvertierung der angegebenen Zahlen in die jeweiligen Zahlensysteme.

dual	oktal	dezimal	hexadezimal
1011 1001			
1101 0110			
	363		
	1021		
		317	
		1150	
			ED3
			C8E



- Beispiel:
 - Binär → dezimal (11010110)_R

$$= (1*2^{7} + 1*2^{6} + 0*2^{5} + 1*2^{4} + 0*2^{3} + 1*2^{2} + 1*2^{1} + 0*2^{0})_{D}$$

$$= (128 + 64 + 0 + 16 + 0 + 4 + 2 + 0)_{D}$$

$$= 214_{D}$$



Dez	Bin	Hex
0	0000	0
1	0001	1
2	0010	2
3	0011	3
4	0100	4
5	0101	5
6	0110	6
7	0111	7

Dez	Bin	Hex
8	1000	8
9	1001	9
10	1010	А
11	1011	В
12	1100	С
13	1101	D
14	1110	E
15	1111	F





- Beispiel:
 - Binär \rightarrow oktal $(11010110)_{\rm R}$

$$= (11\ 010\ 110)_{B}$$

$$= (3\ 2\ 6)_{O}$$

■ Binär → hexadezimal (11010110)_R

$$=(1101\ 0110)_{\rm B}$$

$$= (D 6)_{H}$$



dual	<u>oktal</u>	dezimal	<u>hexadezimal</u>
1011 1001			
1101 0110	326	214	D6
	363		
	1021		
		317	
		1150	
			ED3
			C8E



- Beispiel:
 - oktal → binär → hexadezimal → dezimal
 1021₀

=
$$(001\ 000\ 010\ 001)_{B}$$

= $(0010\ 0001\ 0001)_{B}$
= $(2\ 1\ 1)_{H}$
= $(2*16^{2}+1*16^{1}+1*16^{0})_{D}$
= $(512+16+1)_{D}$
= 529_{D}

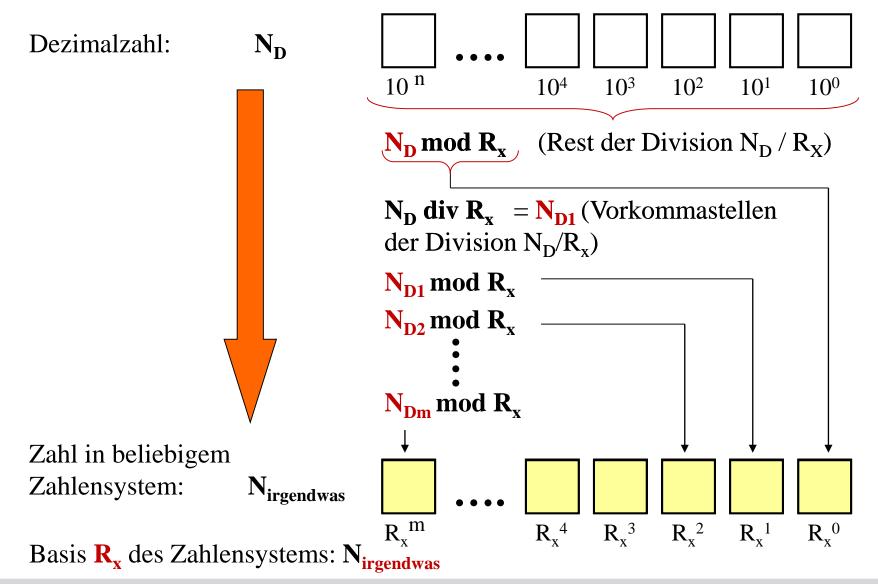


dual	<u>oktal</u>	<u>dezimal</u>	<u>hexadezimal</u>
1011 1001			
1101 0110	326	214	D6
	363		
10 0001 0001	1021	529	211
		317	
		1150	
			ED3
			C8E



Wandlung der Dezimalzahl in andere Systemen



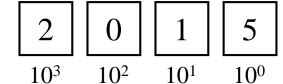


Wandlung der Dezimalzahl in andere Systemen



Dezimalzahl:

$$N_{\rm D} = 2015$$

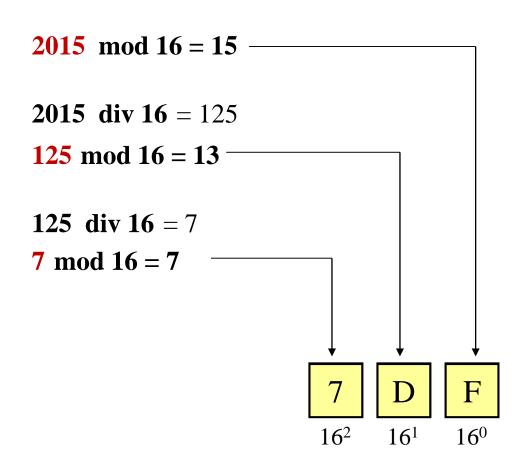


Beispiel:

Wandlung der Dezimalzahl 2004 in eine Hexadezimalzahl.



Hexadezimalzahl: N_H





Beispiel: dezimal → oktal 1150_D

$$1150_{D}$$

= 2176_{O}
= $010\ 001\ 111\ 110_{B}$
= $47E_{H}$





<u>dual</u>	<u>oktal</u>	dezimal	<u>hexadezimal</u>
1011 1001	271	185	B9
1101 0110	326	214	D6
1111 0011	363	243	F3
10 0001 0001	1021	529	211
0001 0011 1101	475	317	13D
0100 0111 1110	2176	1150	47E
1110 1101 0011	7323	3795	ED3
1100 1000 1110	6216	3214	C8E







Felix Pistorius

pistorius@kit.edu

Karlsruher Institut für Technologie (KIT)





- Bei einer Alarmanlage sollen alle Einbrüche und Einbruchsversuche in einem nichtflüchtigen Speicher protokolliert werden. Da bei diesem Baustein jede Speicherzelle nur 1x geschrieben werden kann, soll dieser möglichst gut ausgenutzt werden.
- Aus den Statistiken der Polizei ist bekannt, dass Einbrüche durch eine Tür 4x häufiger sind als durch ein Bad- oder Küchenfenster. Weiterhin ist bekannt, dass Einbrüche durch eine Tür 8x häufiger sind als durch ein sonstiges Fenster im Erdgeschoss. Schließlich zeigt die Statistik, dass ein Einbruch durch ein Fenster im 1. Stockwerk 16x seltener auftritt, als ein Einbruch durch eine Tür.



Berechnen Sie zuerst die Auftrittswahrscheinlichkeiten für einen Einbruchsort in einer Wohnung mit 1 Wohnungstür, 1 Balkontür, 1 Badfenster, 1 Küchenfenster, 1 Schlafzimmerfenster und 2 Wohnzimmerfenstern im Erdgeschoss sowie 2 Kinderzimmerfenster und 2 Atelierfenster im 1. Stock.

Ort	Häufigkeit	Wahrscheinlichkeit	Ermittelte Codierung	
Wohnungstür	16	16/ 50		WT:16/50
Balkontür	16	16/ 50		BT: 16/50
Badfenster	4	4/ 50		BF: 4/50
Küchenfenster	4	4/ 50		KF: 4/50
Schlazimmerfenster	2	2/ 50		SF: 2/50
Wohnzimmerfenster 1	2	2/ 50		WF1: 2/50
Wohnzimmerfenster 2	2	2/ 50		WF2: 2/50
Kinderzimmerfenster 1	1	1/50		KF1: 1/50
Kinderzimmerfenster 2	1	1/50		KF2: 1/50
Atelierfenster 1	1	1/50		AF1: 1/50
Atelierfenster 2	1	1/50		AF2: 1/50

$$\sum = 50$$





 Erstellen Sie eine optimale Codierung für die möglichen Orte und tragen Sie Ihre Ergebnisse in die oben vorgegebene Tabelle ein. Nutzen Sie hierzu das Huffmanverfahren.

Ort	Häufigkeit	Wahrscheinlichkeit	Ermittelte Codierung	
Wohnungstür	16	16/ 50		WT:16/50
Balkontür	16	16/ 50		BT: 16/50
Badfenster	4	4/ 50		BF: 4/50
Küchenfenster	4	4/ 50		KF: 4/50
Schlazimmerfenster	2	2/ 50		SF: 2/50
Wohnzimmerfenster 1	2	2/ 50		WF1: 2/50
Wohnzimmerfenster 2	2	2/ 50		WF2: 2/50
Kinderzimmerfenster 1	1	1/50		KF1: 1/50
Kinderzimmerfenster 2	1	1/50		KF2: 1/50
Atelierfenster 1	1	1/50		AF1: 1/50
Atelierfenster 2	1	1/50		AF2: 1/50

$$\sum = 50$$



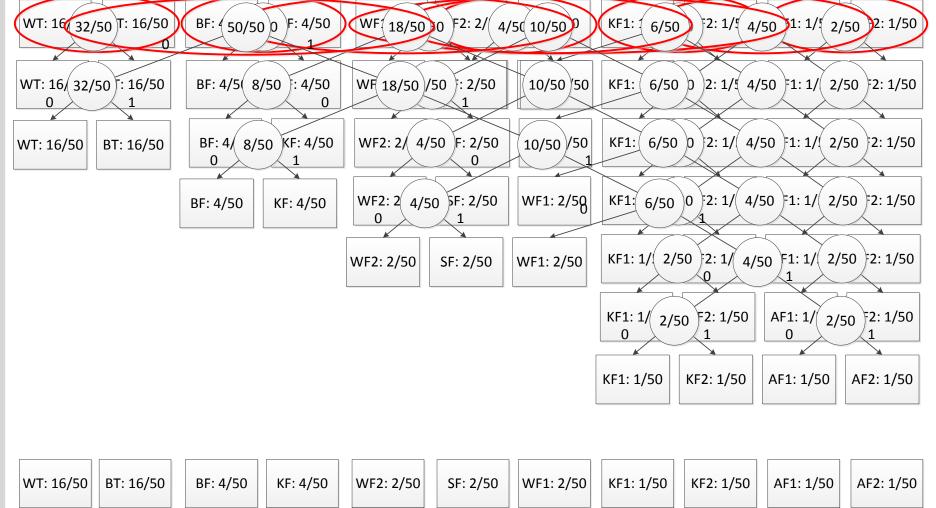
Huffman Kodierung



Huffman's Kodierungs-Algorithmus (HKA) (1)

- HKA konstruiert einen Baum T, der die Kodierung der einzelnen Zeichen repräsentiert
- Eine Queue Q wird dafür benutzt, um die zwei Zeichen mit der niedrigsten Auftrittswahrscheinlichkeit zu finden
- Diese zwei Zeichen werden zu einem Objekt verschmolzen, dessen Auftrittswahrscheinlichkeit die Summe der Einzelwahrscheinlichkeiten ist







 Erstellen Sie eine optimale Codierung für die möglichen Orte und tragen Sie Ihre Ergebnisse in die oben vorgegebene Tabelle ein. Nutzen Sie hierzu das Huffmanverfahren.

Ort	Häufigkeit	Wahrscheinlichkeit	Ermittelte Codierung	
Wohnungstür	16	16/ 50	00	WT:16/50
Balkontür	16	16/ 50	01	BT: 16/50
Badfenster	4	4/ 50	100	BF: 4/50
Küchenfenster	4	4/ 50	101	KF: 4/50
Schlazimmerfenster	2	2/ 50	1101	SF: 2/50
Wohnzimmerfenster 1	2	2/ 50	1110	WF1: 2/50
Wohnzimmerfenster 2	2	2/ 50	1100	WF2: 2/50
Kinderzimmerfenster 1	1	1/50	111100	KF1: 1/50
Kinderzimmerfenster 2	1	1/50	111101	KF2: 1/50
Atelierfenster 1	1	1/50	111110	AF1: 1/50
Atelierfenster 2	1	1/50	111111	AF2: 1/50

$$\sum = 50$$





Wie viele Codeworte k\u00f6nnen Sie in einem Speicher mit einer Kapazit\u00e4t von 1024 Bit abspeichern? Berechnen Sie dazu zuerst die mittlere Codewortl\u00e4nge.

Ort	Häufigkeit	Wahrscheinlichkeit	Ermittelte Codierung
Wohnungstür	16	16/ 50	00
Balkontür	16	16/ ₅₀	01
Badfenster	4	4/ 50	100
Küchenfenster	4	4/50	101
Schlazimmerfenster	2	² / ₅₀	1101
Wohnzimmerfenster 1	2	² / ₅₀	1110
Wohnzimmerfenster 2	2	² / ₅₀	1100
Kinderzimmerfenster 1	1	1/50	111100
Kinderzimmerfenster 2	1	1/50	111101
Atelierfenster 1	1	1/50	111110
Atelierfenster 2	1	1/50	111111



Ort	Häufigkeit	Wahrscheinlichkeit	Ermittelte Codierung
Wohnungstür	16	16/ 50	00
Balkontür	16	16/ 50	01
Badfenster	4	4/ 50	100
Küchenfenster	4	4/ 50	101
Schlazimmerfenster	2	2/ 50	1101
Wohnzimmerfenster 1	2	² / ₅₀	1110
Wohnzimmerfenster 2	2	² / ₅₀	1100
Kinderzimmerfenster 1	1	1/50	111100
Kinderzimmerfenster 2	1	1/50	111101
Atelierfenster 1	1	1/50	111110
Atelierfenster 2	1	1/50	111111

Mittlere Codewortlänge: $\overline{m} = \sum_{i=1}^{n} m(x_i) \cdot p(x_i)$

$$\overline{m} = 2(2 \cdot 16 / 50) + 3(2 \cdot 4 / 50) + 4(3 \cdot 2 / 50) + 6(4 \cdot 1 / 50) = 136 / 50$$
 (= 2,72)





Ort	Häufigkeit	Wahrscheinlichkeit	Ermittelte Codierung
Wohnungstür	16	16/ 50	00
Balkontür	16	16/ ₅₀	01
Badfenster	4	4/ ₅₀	100
Küchenfenster	4	4/ ₅₀	101
Schlazimmerfenster	2	² / ₅₀	1101
Wohnzimmerfenster 1	2	2/ 50	1110
Wohnzimmerfenster 2	2	2/ 50	1100
Kinderzimmerfenster 1	1	1/50	111100
Kinderzimmerfenster 2	1	1/50	111101
Atelierfenster 1	1	1/50	111110
Atelierfenster 2	1	1/50	111111

Speicherkapazität:
$$n = \frac{Speicherkapazität}{\frac{m}{m}}$$

$$n = \frac{1024}{136/50} = \frac{51200}{136} (\approx 376,5) \text{ CWs}$$

