Informationstheorie

- Wozu das Ganze?
- Information
- Nachrichtenübertragungssysteme
- Codierung
- Kanalkapazität
- .... Signalerkennung

Wozu das Ganze?

- Begründet:
  „A Mathematical Theory of Communication“
  by C. E. Shannon (1948)
- Wofür:
  Bessere Analyse von
  Telekommunikationssystemen

Was ist Information?

- Der Begriff ≠ Alltagsbedeutung
- Eine Einheit
    - Münzwurf
    - das Bit
    - Würfeln
    - $\log_2(n)$
- Information und Entropie

Exakte Berechnung

- Eigenschaften

  $f \quad n \quad \dagger f \quad n \; {}_1$

  $n \quad \dagger n \; {}_1 n \; {}_2$

  stetig

  monoton wachsend

  $f \quad n \quad \dagger c \; \mathrm{log}$

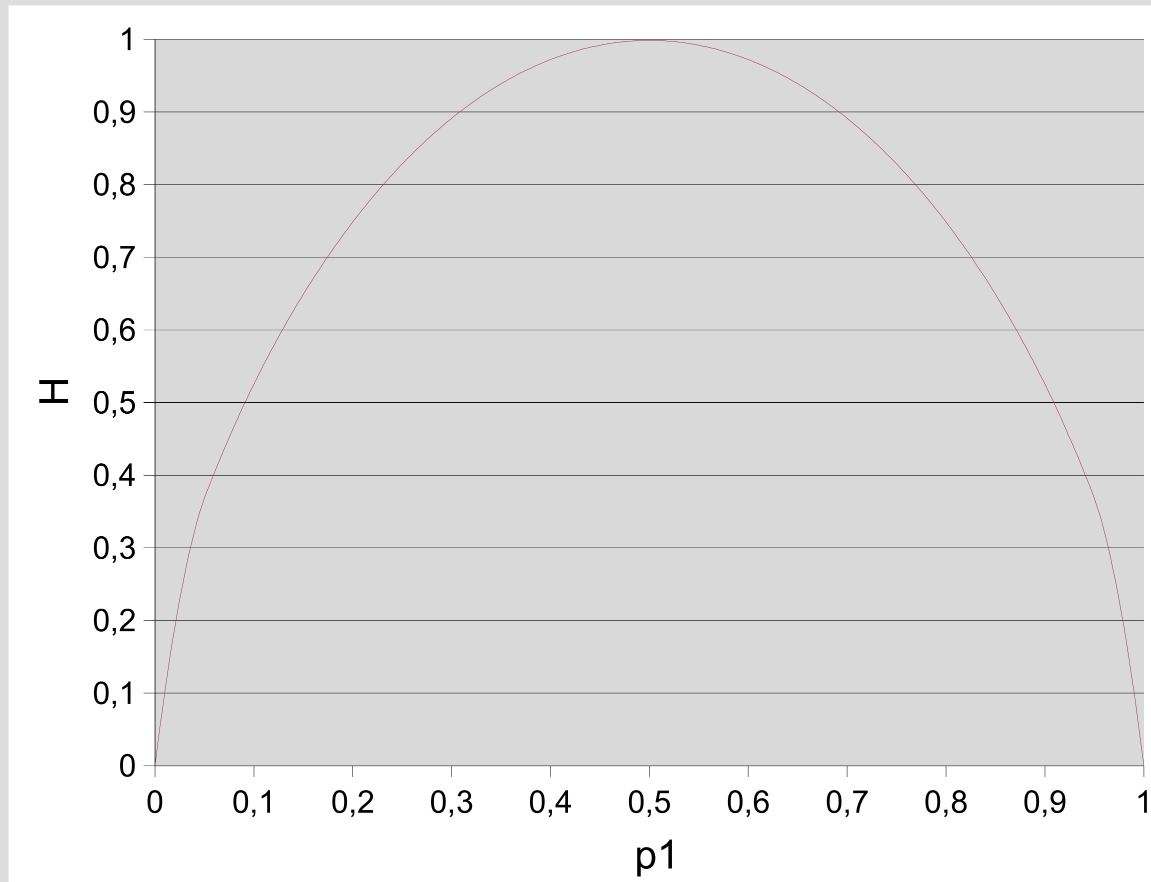
- Experiment mit n Ausgängen, in n_1 und n_2 unterteilt $\quad H \quad \dagger \log_2 \quad n \quad -$
- Erwartungswert an Information

Maximum an Information

- Maximum für

$k \; : \; p_k = \frac{1}{n} \quad H = \log_2$

- Bsp. mit 2 Ausgängen



$H = - \qquad p$

Anwendungen

- Wie gewinnt man das „20-Fragen“-Spiel?

  Max. 2^{20} = 1 048 576 Möglichkeiten
- Finden einer falschen Münze unter 9 mit dreimaligen Wiegen

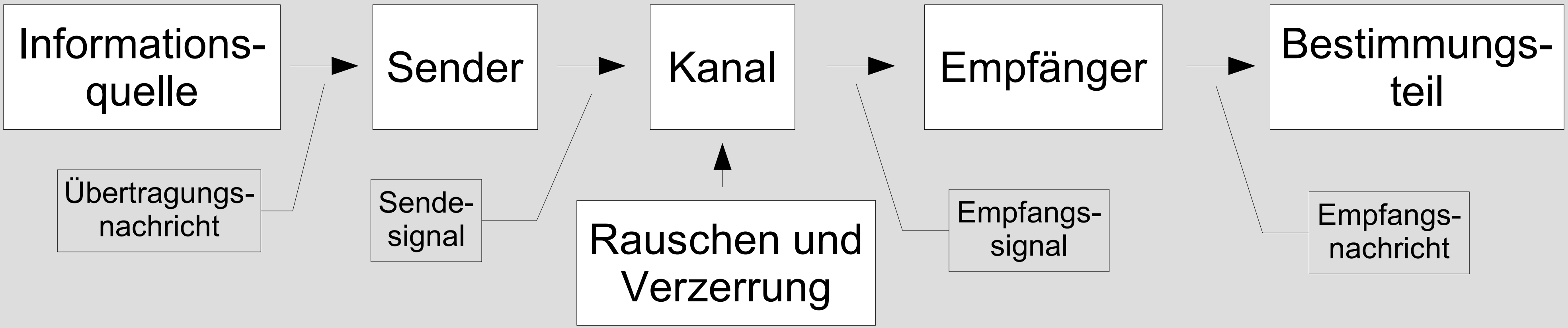
  Problem besitzt: $\log_2 \; 2 \;\; 9 \quad †4{,}16\,l$

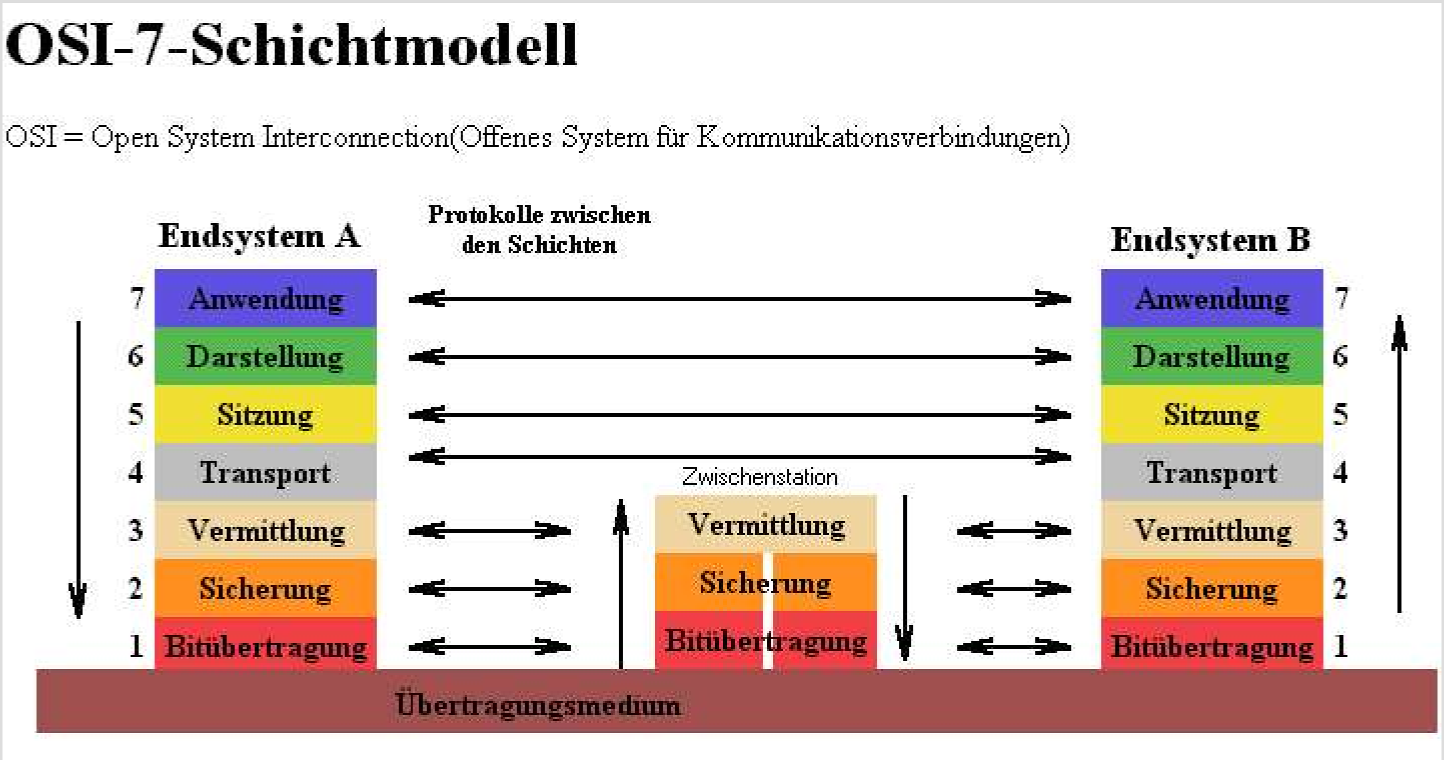
  Max. durch Wiegen zu gewinnen: $3 \; \log_2 \;\; 3 \quad †4{,}74\,l$

  1. Wiegen:
  - n Münzen auf linke / rechte Schale
  - $9 -$
  - Maximum bei n = 3

  ...

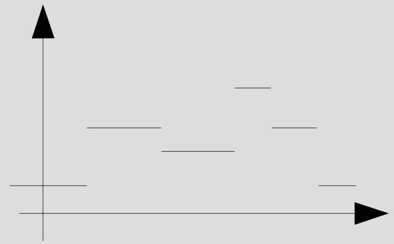
Nachrichtenübertragungs-systeme



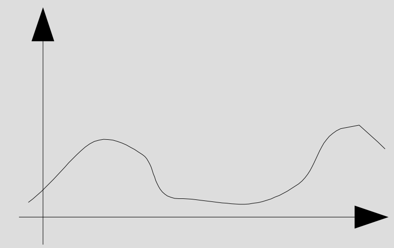


Der Kanal

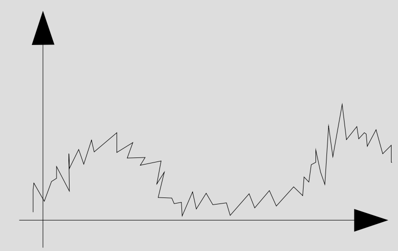
- Kanaltypen
  - Diskrete Kanäle
    - Nur endliche Zustandsmenge kann angenommen werden



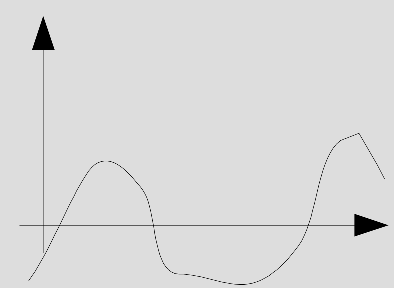
  - Kontinuierliche Kanäle
    - Jede reelle Zahl in einem Intervall ist möglich (unendlich viele Zustände)



- Störungen
  - Rauschen
    - Zufälliges Signal (weißes Rauschen)



  - Verzerrungen
    - Bijektive Abbildung
    - Umkehrung möglich



Codierung

- Was ist das?

  Umsetzung der Informationen auf anderes Alphabet

  - z.B. A -> 1, B -> 2, ...

  Präfixeigenschaft

  - z.B. A-> 0, B -> 10, C -> 110, ...
  - 01011001100 -> 0|10|110|0|110|0 -> ABCACA

- Optimaler Code?

  bestmögliche Ausnutzung durch Einbeziehung der Wahrscheinlichkeiten

Beispiel Optimales Codieren

- Quelle mit 80 Buchstaben pro Minute:
  A (80 %) -> 0
  B (20%) -> 1
  80 Ziffern / Minute

Buchstabe	W-keit	Ziffer	Gewichtete Ziffernanzahl
A	0,8	0	0,8
B	0,2	1	0,2
			1

- 1. Annäherung:
  62,4 Ziffern / Minute

Buchstabe	W-keit	Ziffer	Gewichtete Ziffernanzahl
AA	0,64	0	0,64
AB	0,16	10	0,32
BA	0,16	110	0,48
BB	0,04	111	0,12
			1,56

- 2. Annäherung
  58,24 Ziffern / Minute

Buchstabe	W-keit	Ziffer	Gewichtete Ziffernanzahl
AAA	0,512	0	0,512
AAB	0,128	100	0,384
ABA	0,128	101	0,384
ABB	0,128	110	0,384
BAA	0,032	11100	0,160
BAB	0,032	11101	0,160
BBA	0,032	11110	0,160
BBB	0,008	11111	0,040
			2,184

- theoretisches Optimum
  H = 0,72 Bit -> 57,6 Ziffern / Minute

Fehlererkennung und Korrektur

- Paritätsprüfung

  Blockweise Zusatzbit so, dass Anzahl Einsen gerade

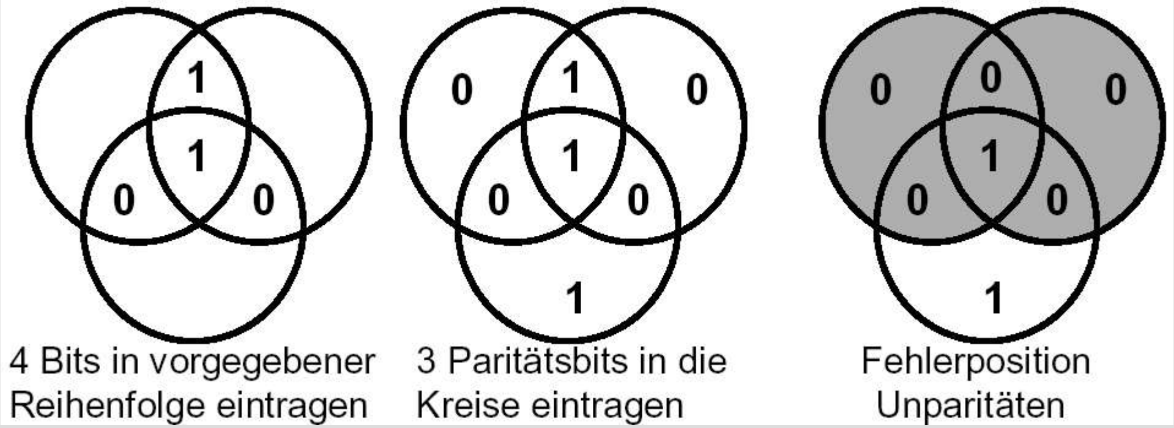
0	1	0	1	0

  Nur Erkennung von Fehlern
  - auch nur wenn deren Anzahl ungerade

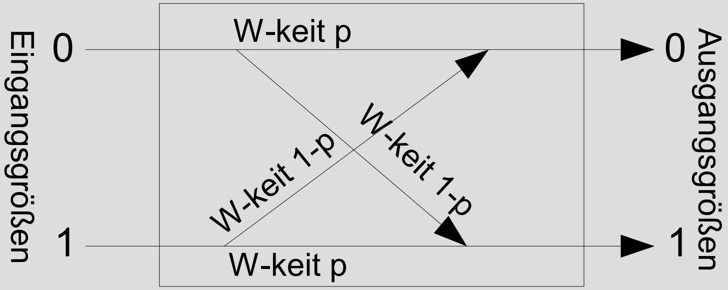
- Hamming Code

  Korrigiert einen Fehler

  Erkennt sogar noch 2



Kanalkapazität bei diskreten Kanälen



- Beispiel binärer Kanal

  Rauschen -> Bits kippen um

  z.B.

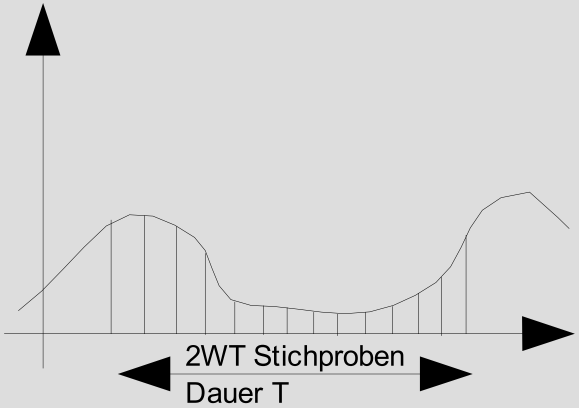
Eingang	0	1	0	1	0	1	0	1
Ausgang	0	1	1	0	0	0	0	1
	R	R	F	F	R	F	R	R

  „Richtig- / Falsch-Information“ kann als Quelle mit dem Informationsgehalt $p \quad \log_2$ aufgefasst werden

  Kapazität: $H_c + 1 \quad p \quad \log_2 \quad p \quad 1 -$ bit / Symbol

Kanalkapazität bei kontinuierlichen Kanälen – 1

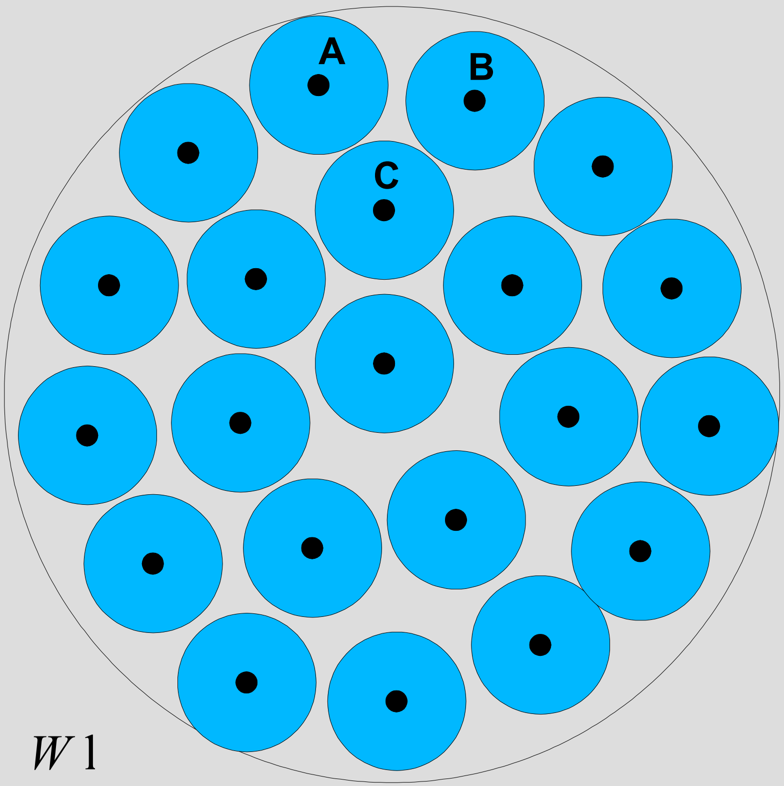
- Wie viele Signale sind unterscheidbar?
- Signal mit Maximalfrequenz W, weißes Rauschen, ohne Verzerrungen
- Sampling Theorem:
  - Abtastrate = 2W
  - Signal der Dauer T ist durch 2TW-dim. Vektor eindeutig bestimmt



- Signalenergie: $E = \frac{1}{2W} \sum_{n=1}^{2TW} x_n^2$
- Abstand zum Ursprung $d = \sqrt{\sum_{n=1}^{2TW} x_n^2} = \sqrt{2WE}$

Kanalkapazität bei kontinuierlichen Kanälen – 2

- Signalleistung: P – Rauschleistung: N
- Gesamtvolumen:
- Einzelvolumen:
- Anzahl unterscheidbarer Signale: $M \quad 1 \quad \frac{}{N}$
- Signal-Rausch-Verhältnis: P/N
- Informationsfluss:



$$W \log_2 \quad 1 \quad \frac{P}{N} \qquad \frac{1}{T} \log_2 E_{av} \qquad \frac{1}{T} \log_2 M \qquad W\,1$$

- mittlere Fehlerwahrscheinlichkeit: E_{av}

Der Mensch

- Lesen von 500 Worten / Minute
  - 1 Bit / Buchstabe
  - 5 Buchstaben / Wort
  - 42 bit / s
- Simultan-Blindschach
  - 40 Partien
  - 40 Züge / Spiel
  - 6 Stunden
  - Alle 14 s einen Zug auswendig lernen
  - 3 Bit / Zug
  - 0,2 bit / s aufnehmen und voll verwerten
- Abschätzen von Größen/Kurzzeitgedächtnis
  - 7 unterschiedliche Symbole -> 2,8 bit

Literatur

- F. Topsøe: Informationstheorie (B.G.Teubner, 1974)
- Gordon Raisbeck: Informationstheorie (Oldenbourg Verlag, 1970)