

The background features a blurred, high-angle view of a modern chair with a complex, geometric frame. Overlaid on this is a semi-transparent layer of binary code (0s and 1s) that appears to be moving or vibrating, creating a sense of digital activity and data flow.

Grundlagen der Informationstheorie

Hanna Rademaker und Fynn Feldpausch



1. Thema

Informationstheorie

- geht zurück auf Claude Shannon
 - „*The Mathematical Theory of Communication*“
- beschäftigt sich mit
 - Information
 - Informationsübertragung
 - Datenkompression
 - Kodierung

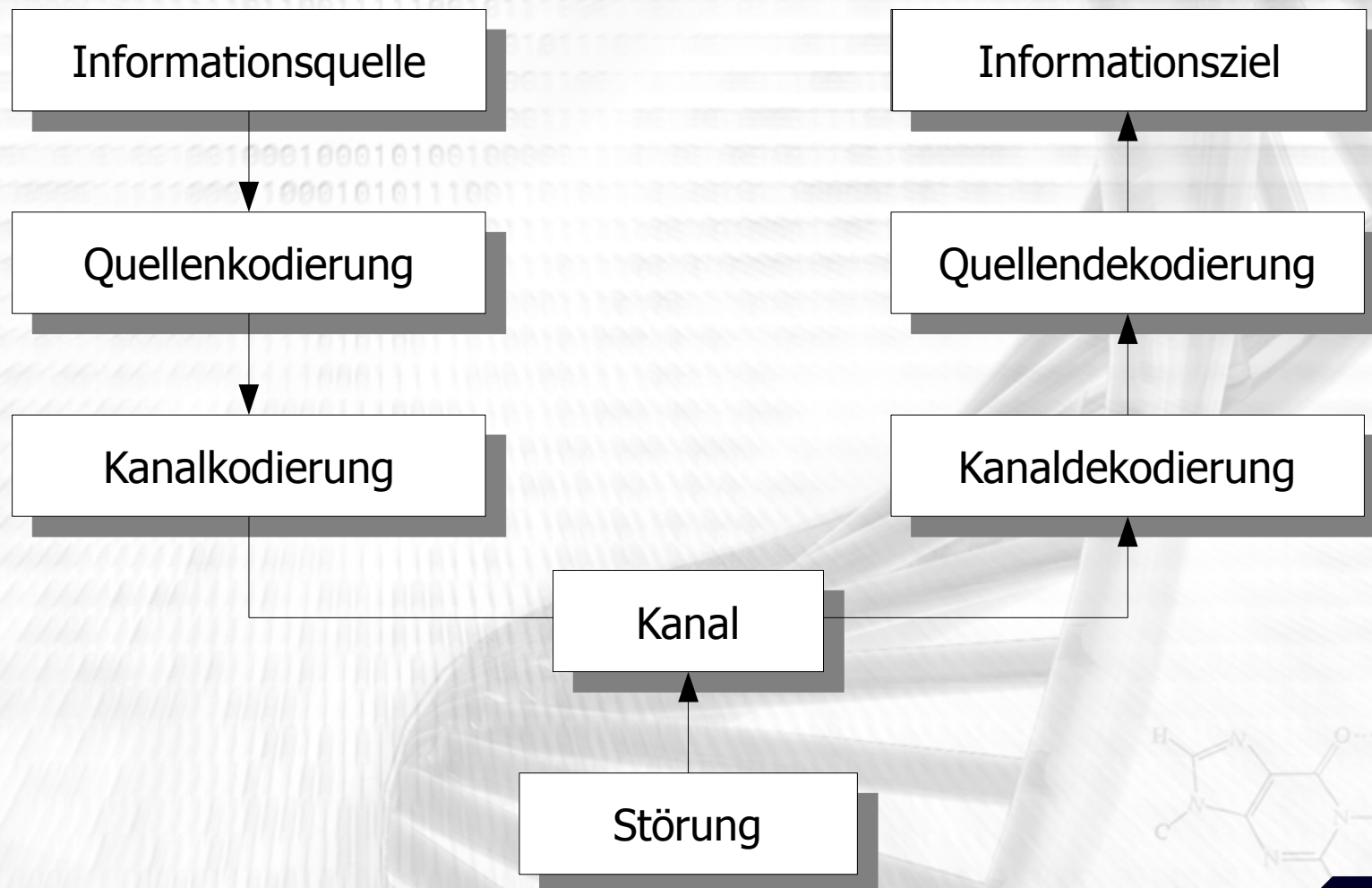


Claude Shannon
(1916 – 2001)

Quelle: <http://www.cl.cam.ac.uk/>

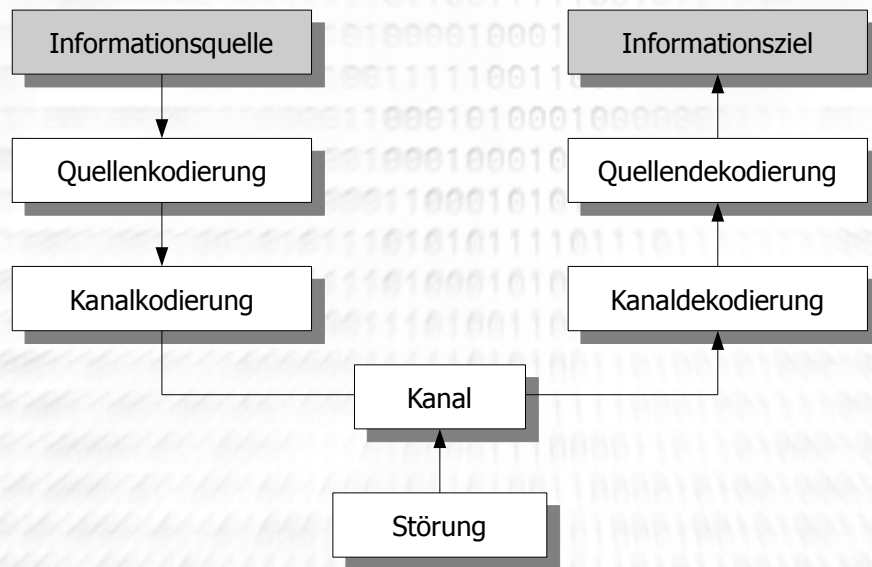
2. Überblick

Informationsübertragung



2. Überblick

Wichtige Fragestellungen

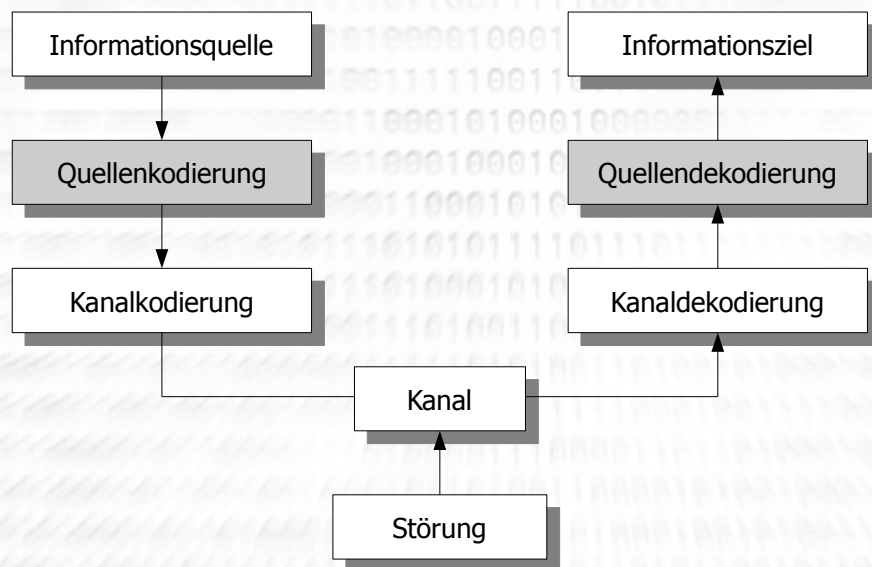


- Was ist Information?
- Wie kann man Information messen?



2. Überblick

Wichtige Fragestellungen

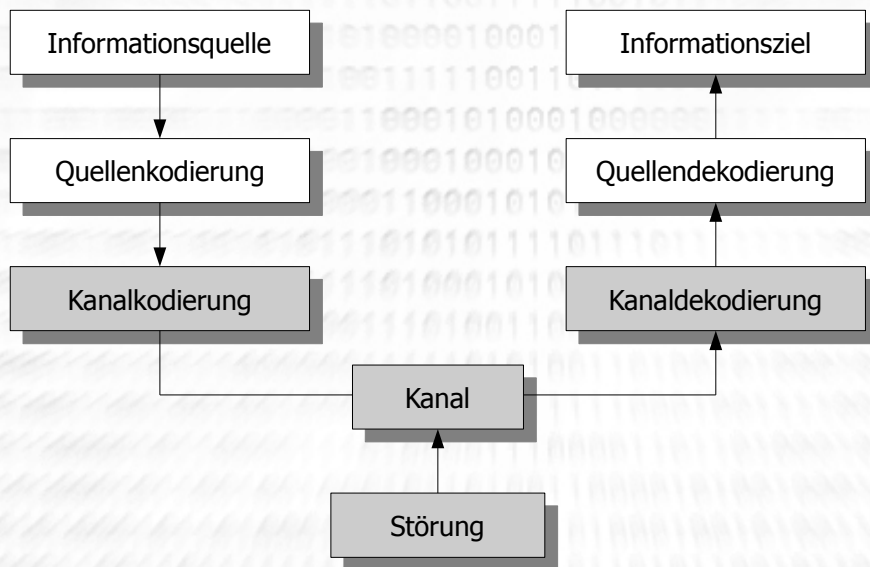


- Wie können Daten komprimiert werden?
- Wie viel Speicher wird mindestens benötigt?
- Ist es möglich Daten verlustfrei zu komprimieren?

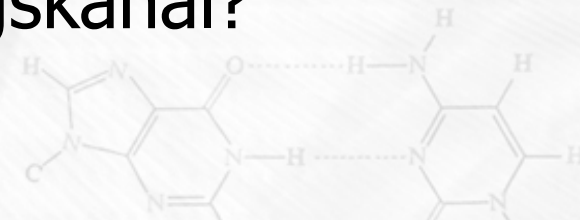


2. Überblick

Wichtige Fragestellungen

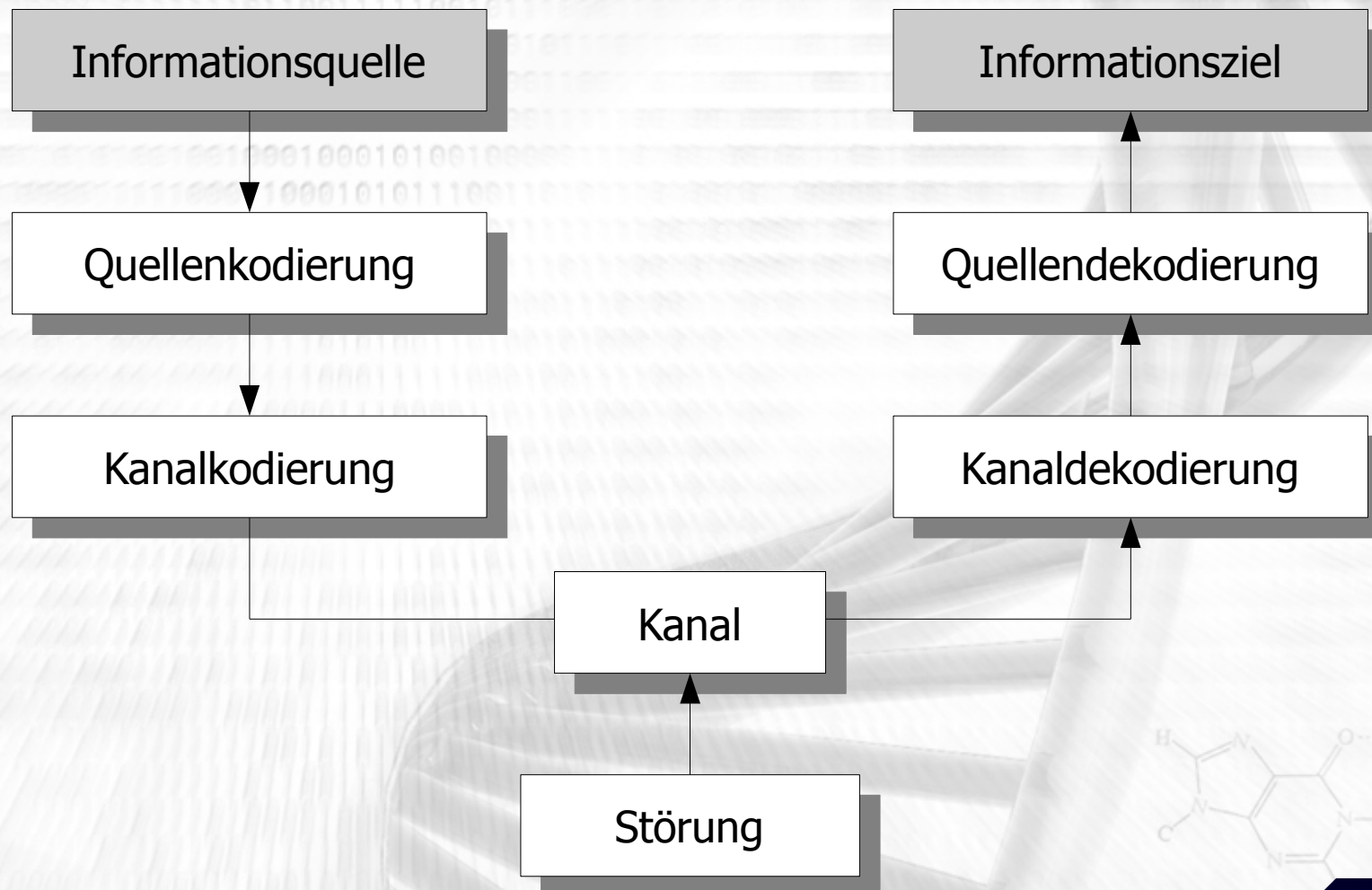


- Wie können Übertragungsfehler entdeckt werden?
- Ist es möglich sie zu korrigieren?
- Wie viel Daten passen durch einen Übertragungskanal?



3. Information

Informationsübertragung



3. Information

„Das Wort Information wird in dieser Theorie in einem besonderen Sinn verwendet, der nicht mit dem gewöhnlichen Gebrauch verwechselt werden darf. Insbesondere darf Information nicht der Bedeutung gleichgesetzt werden. [...] Tatsächlich können zwei Nachrichten, von denen die eine von besonderer Bedeutung ist, während die andere bloßen Unsinn darstellt, in dem von uns gebrauchten Sinn genau die gleiche Information enthalten.“

Shannon & Weaver (1949)

3. Information

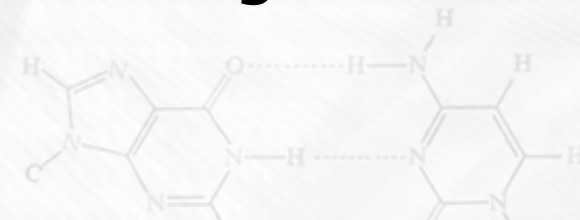
PRÜFUNG BESTANDEN TNGDR PNASUNEBFÜE



3. Information

PRÜFUNG BESTANDEN TNGDR PNASUNEBFÜE

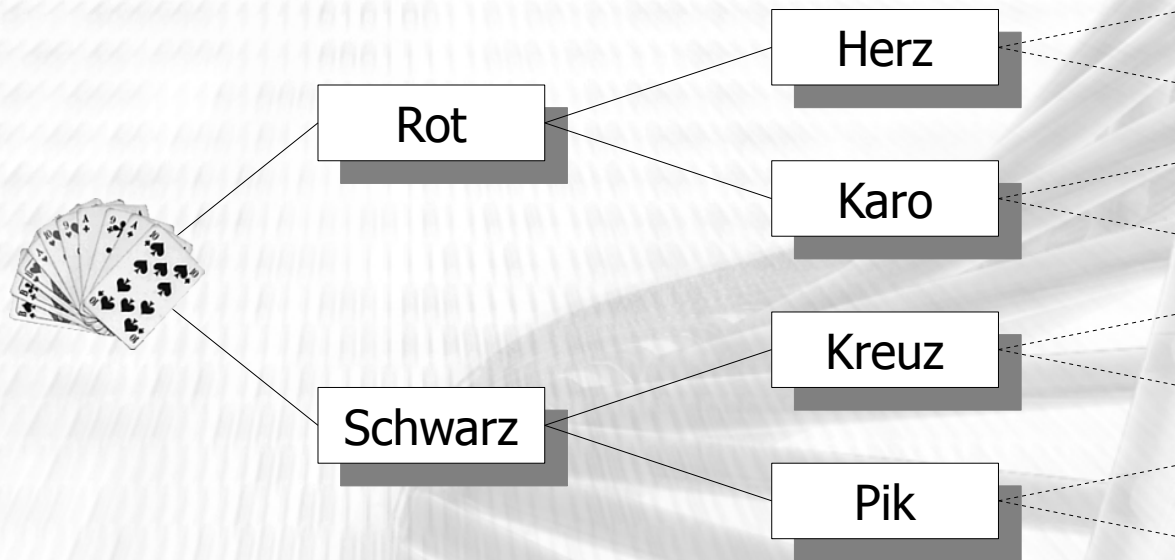
Beide Nachrichten haben denselben Informationsgehalt!



3. Information

Der Shannon-Informationsgehalt

- Spieler A wählt eine Spielkarte
- Spieler B soll diese durch „Ja-Nein“-Fragen erraten
- Optimale Strategie:



3. Information

Der Shannon-Informationsgehalt

- Wahrscheinlichkeit „ja“ zu erhalten: $p(j) = \frac{1}{2}$
- Def. Informationsgehalt [bit]:

$$h(x) = \log_2 \frac{1}{p(x)}$$



3. Information

Der Shannon-Informationsgehalt

- Wahrscheinlichkeit „ja“ zu erhalten: $p(j) = \frac{1}{2}$
- Def. Informationsgehalt [bit]:

$$h(x) = \log_2 \frac{1}{p(x)}$$

$$h(j) = \log_2 \frac{1}{p(j)} = \log_2 2 = 1$$

- 1 Antwort = 1 bit \rightarrow 5 Antworten = 5 bit



3. Information

Der Shannon-Informationsgehalt

- Informationsgehalt einer naiven Strategie:

- $h_1(n) = \log_2 \frac{32}{31} = 0,0458 \text{ bit}$

- $h_2(n) = \log_2 \frac{31}{30} = 0,0473 \text{ bit}$

- ...



3. Information

Der Shannon-Informationsgehalt

- Informationsgehalt einer naiven Strategie:

- Kein Treffer nach 16 Fragen: $\sum_{i=1}^{16} h_i(n) = 1 \text{ bit}$

- Treffer bei 17. Frage: $h_{17}(j) = \log_2 \frac{16}{1} = 4 \text{ bit}$



3. Information

Der Shannon-Informationsgehalt

- Informationsgehalt einer naiven Strategie:

- Kein Treffer nach 16 Fragen: $\sum_{i=1}^{16} h_i(n) = 1 \text{ bit}$

- Treffer bei 17. Frage: $h_{17}(j) = ld \frac{16}{1} = 4 \text{ bit}$

- Treffer bei 1. Frage: $h_1(j) = ld \frac{32}{1} = 5 \text{ bit}$



3. Information

Die Entropie

- entspricht dem durchschnittlichen Shannon-
Informationsgehalt:

$$\begin{aligned} H(X) &= \sum_x p(x) h(x) \\ &= \sum_x p(x) \log_2 \frac{1}{p(x)} \\ &= - \sum_x p(x) \log_2 p(x) \end{aligned}$$

- Einheit: bit



3. Information

Die Entropie

- $X = \{a, b, c, d\}$

x	a	b	c	d
p(x)	0,5	0,25	0,125	0,125



3. Information

Die Entropie

- $X = \{a, b, c, d\}$

x	a	b	c	d
p(x)	0,5	0,25	0,125	0,125

$$h(x) = \lg \frac{1}{p(x)}$$

$$h(a) = \lg 2 = 1$$

$$h(b) = \lg 4 = 2$$

$$h(c) = h(d) = \lg 8 = 3$$



3. Information

Die Entropie

- $X = \{a, b, c, d\}$

x	a	b	c	d
p(x)	0,5	0,25	0,125	0,125

$$h(x) = \lg \frac{1}{p(x)}$$

$$h(a) = \lg 2 = 1$$

$$h(b) = \lg 4 = 2$$

$$h(c) = h(d) = \lg 8 = 3$$

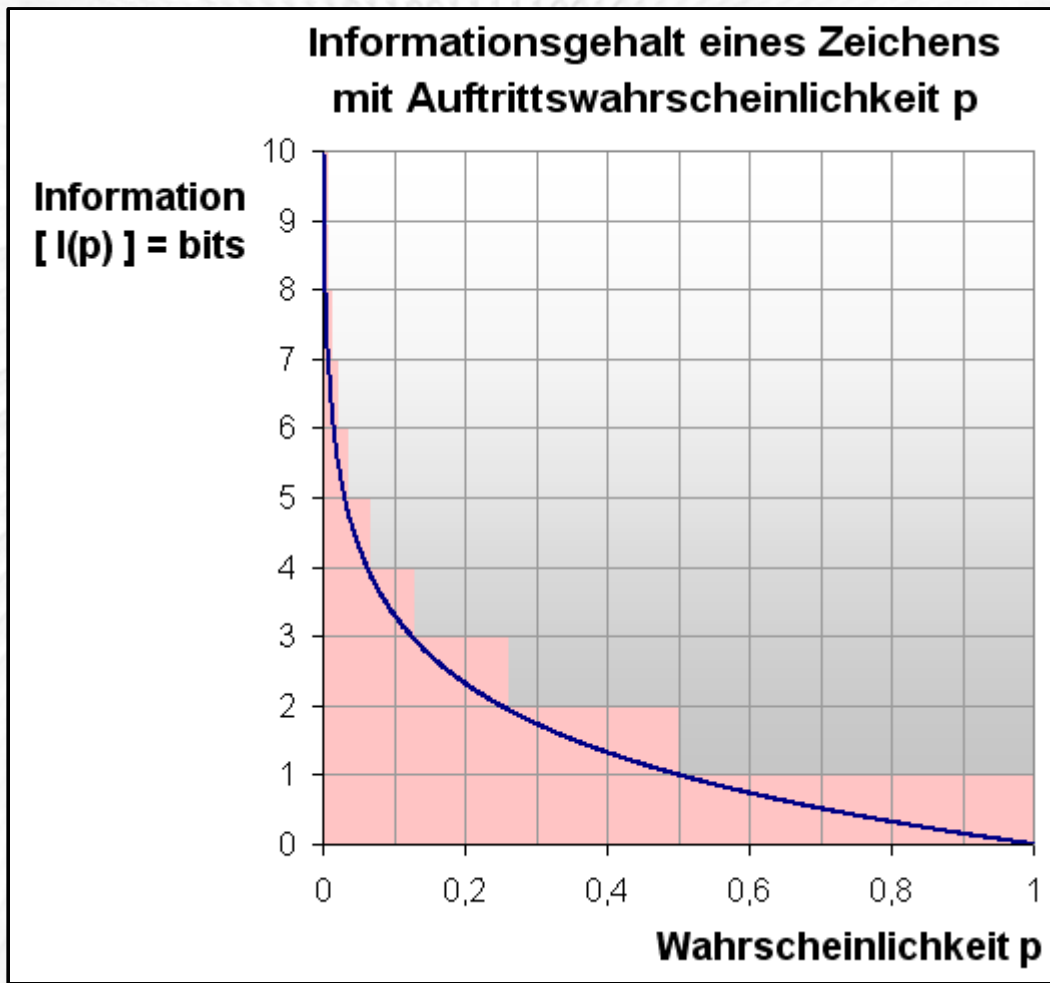
$$H(X) = \sum_x p(x) \lg \frac{1}{p(x)}$$

$$= \frac{1}{2} \cdot 1 + \frac{1}{4} \cdot 2 + \frac{1}{8} \cdot 3 + \frac{1}{8} \cdot 3$$

$$= \frac{14}{8} = 1,75$$

3. Information

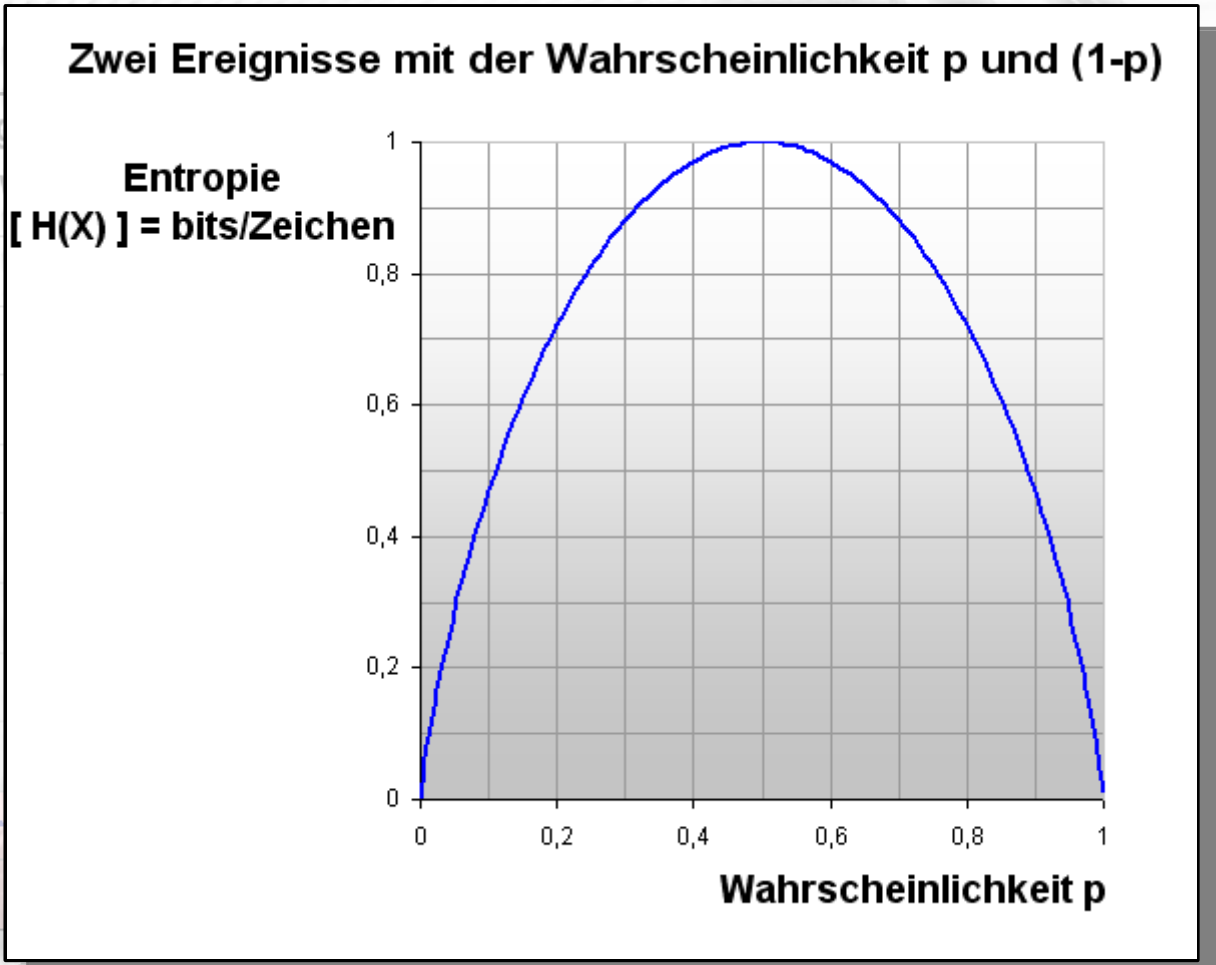
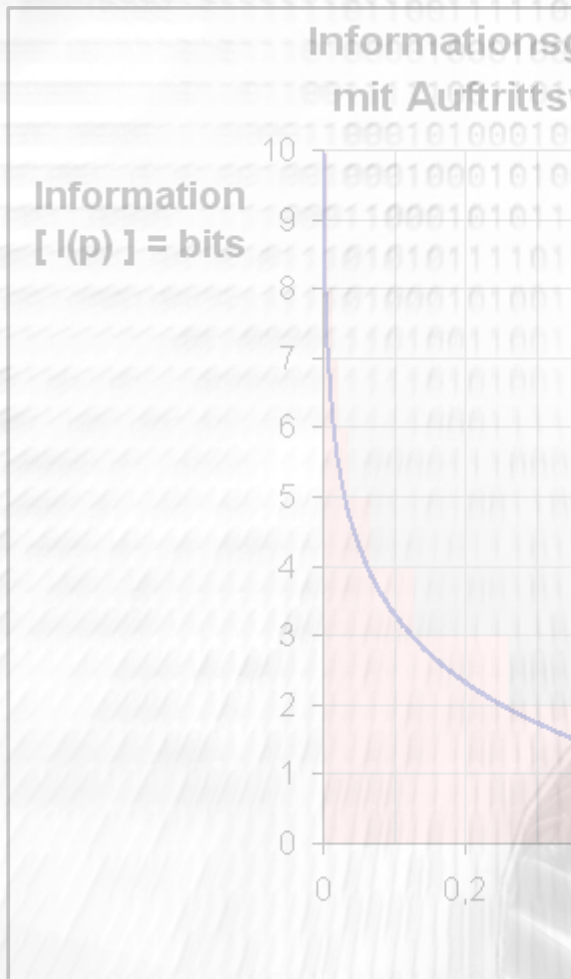
Die Entropie



Quelle: [http://de.wikipedia.org/wiki/Entropie_\(Informationstheorie\)](http://de.wikipedia.org/wiki/Entropie_(Informationstheorie))

3. Information

Die Entropie

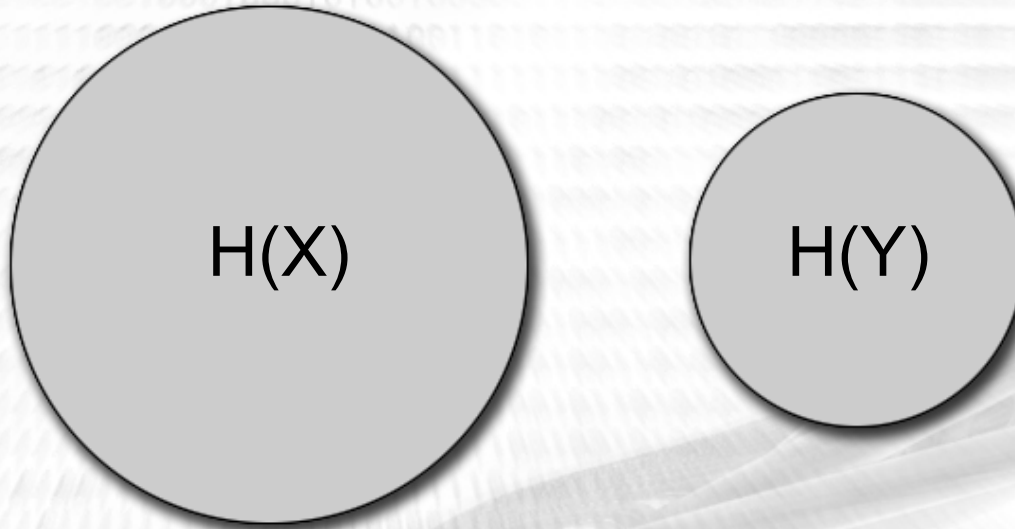


Quelle: [http://de.wikipedia.org/wiki/Entropie_\(Informationstheorie\)](http://de.wikipedia.org/wiki/Entropie_(Informationstheorie))

3. Information

Die gemeinsame Entropie $H(X,Y)$

- Für **unabhängige** Variablen X, Y



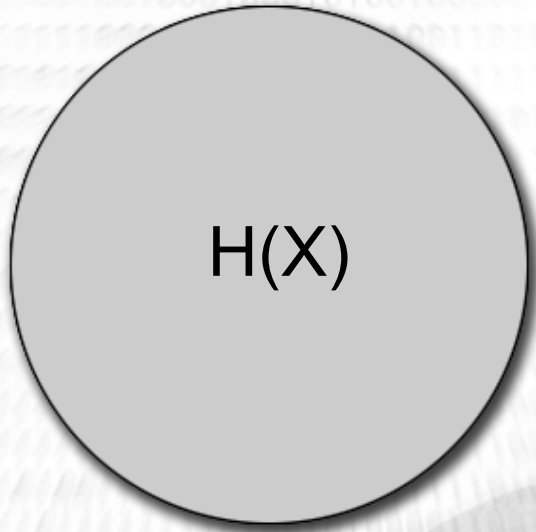
- $H(X,Y) = H(X) + H(Y)$



3. Information

Die gemeinsame Entropie $H(X,Y)$

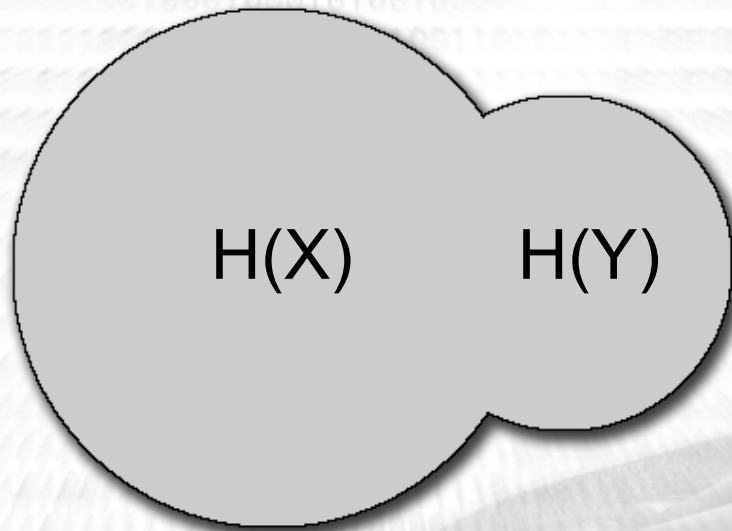
- Für **abhängige** Variablen X, Y



3. Information

Die gemeinsame Entropie $H(X,Y)$

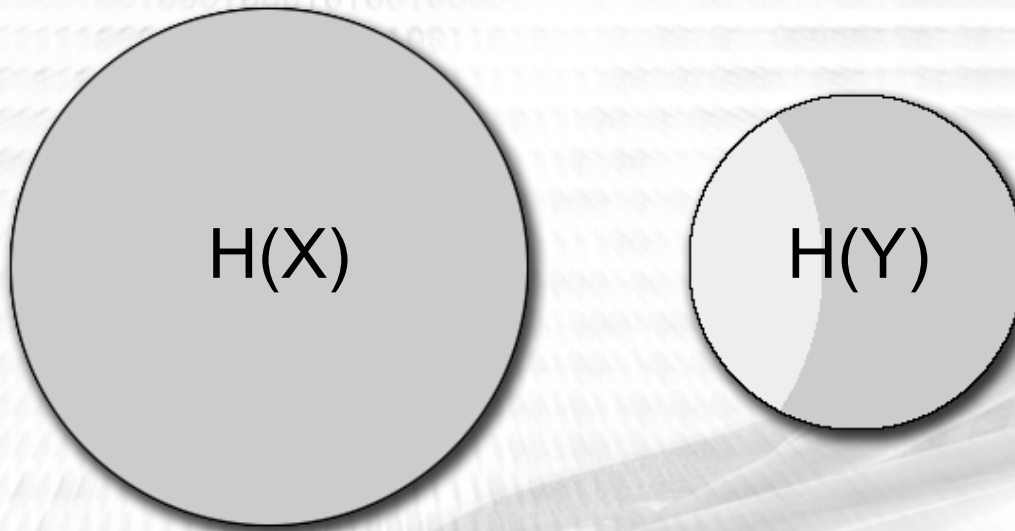
- Für **abhängige** Variablen X, Y



3. Information

Die gemeinsame Entropie $H(X,Y)$

- Für **abhängige** Variablen X, Y



- $H(X,Y) < H(X) + H(Y)$



3. Information

Die bedingte Entropie $H(Y|X)$

- $H(Y|X)$ bezeichnet die Entropie von Y unter Kenntnis von X



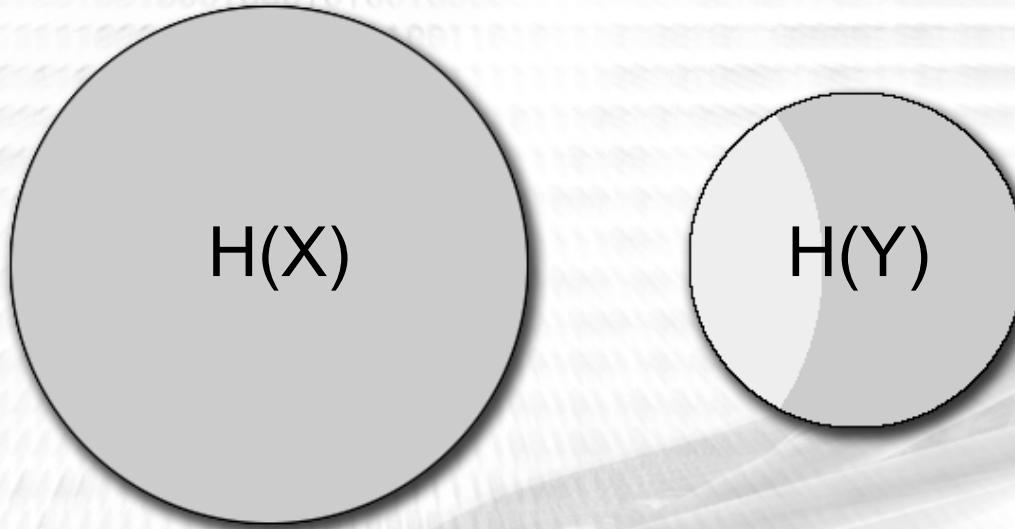
$H(Y|X)$



3. Information

Die gemeinsame Entropie $H(X,Y)$

- Für **abhängige** Variablen X, Y



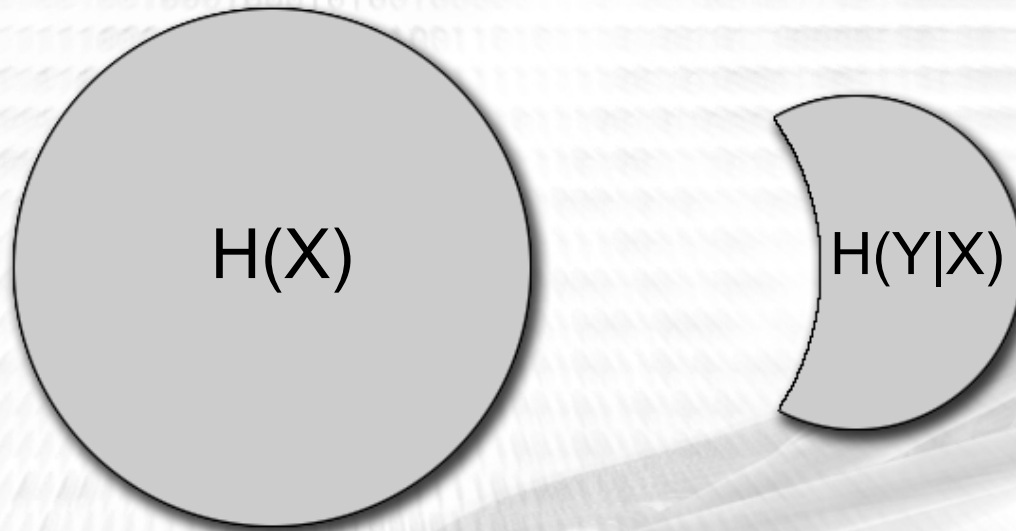
- $H(X,Y) < H(X) + H(Y)$



3. Information

Die gemeinsame Entropie $H(X,Y)$

- Für **abhängige** Variablen X, Y

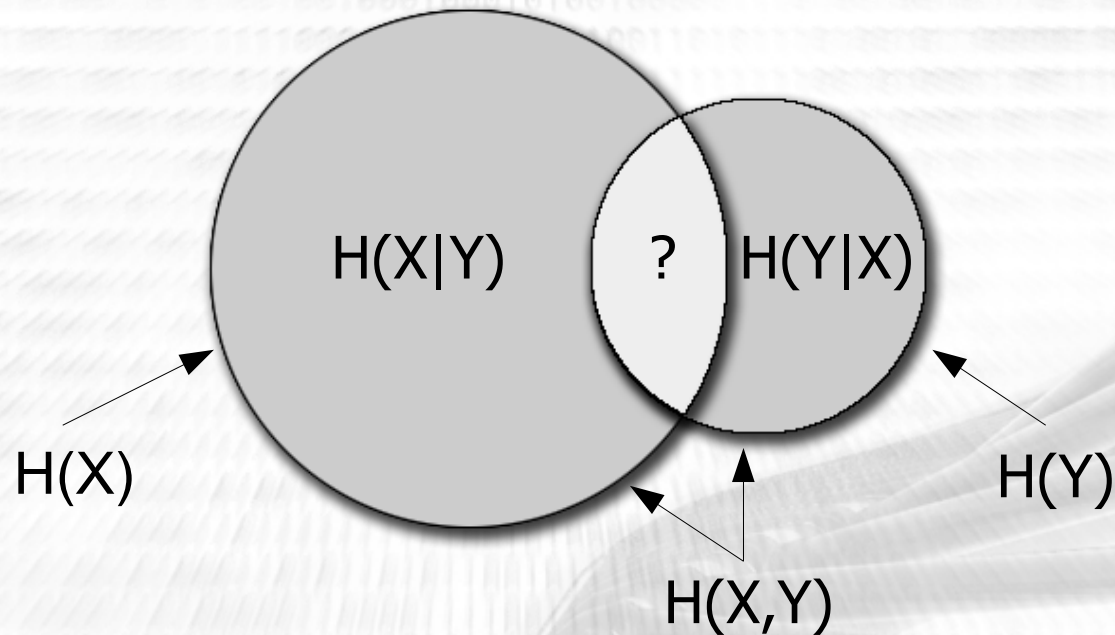


- $H(X,Y) < H(X) + H(Y)$
- $H(X,Y) = H(X) + H(Y|X) = H(Y) + H(X|Y)$

3. Information

Die gemeinsame Entropie $H(X,Y)$

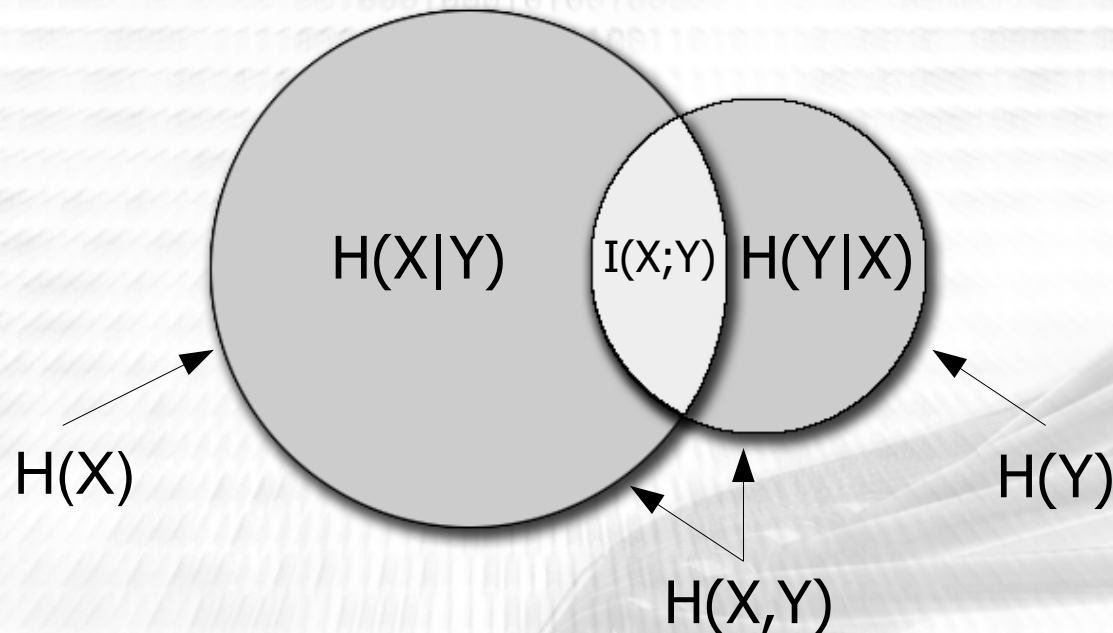
- Für **abhängige** Variablen X, Y



3. Information

Die wechselseitige Information $I(X;Y)$

- $I(X;Y)$ bezeichnet die Information, die von X und Y geteilt wird



- Maß für die Abhängigkeit der Variablen



3. Information

Zusammenfassung

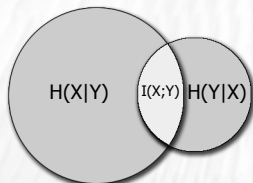
- **Shannon-Informationsgehalt** eines Ereignisses

$$h(x) = \log_2 \frac{1}{p(x)}$$

- **Entropie:** „Selbstinformation“ einer Variablen X

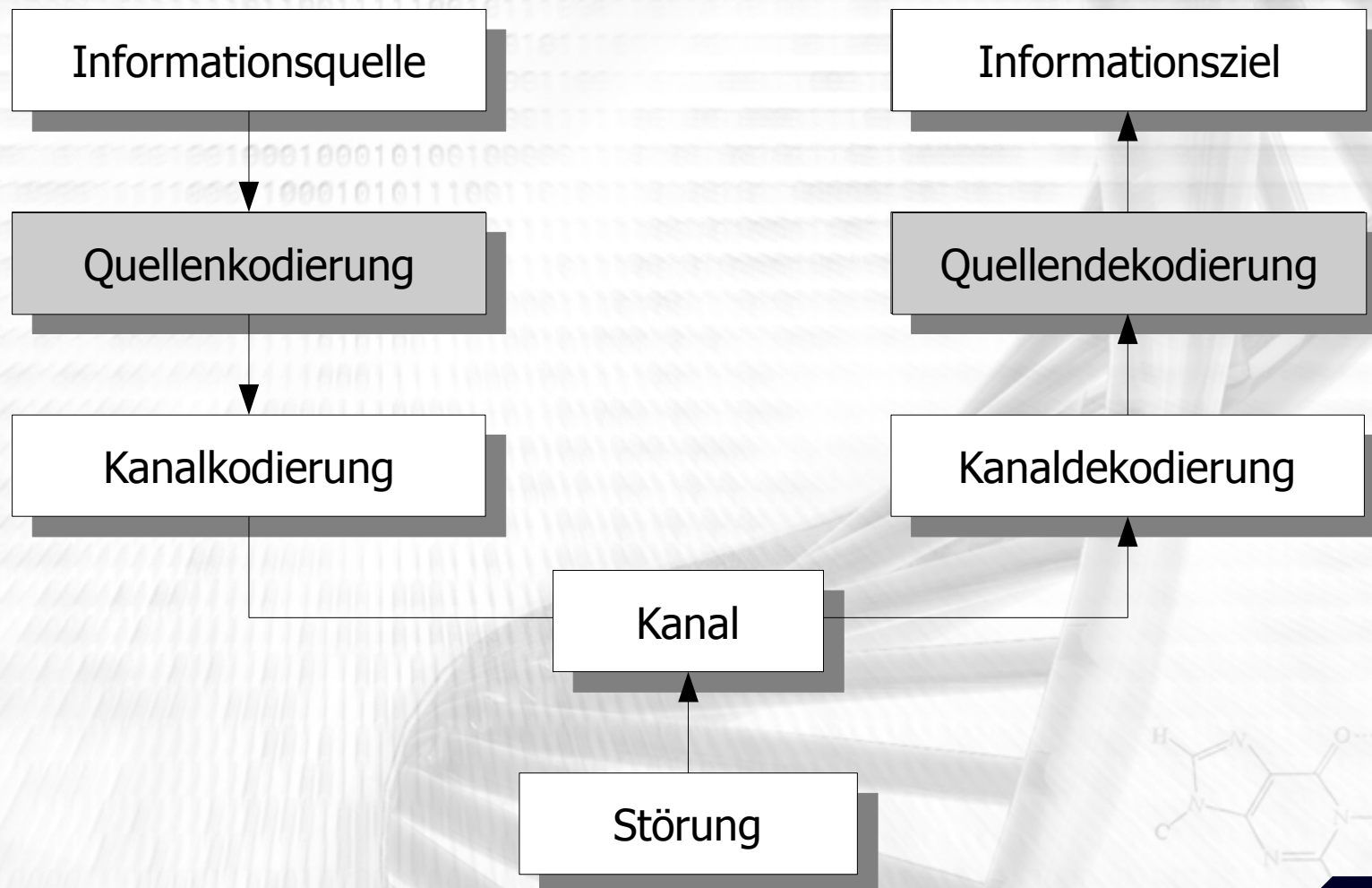
$$H(X) = \sum_x p(x) \log_2 \frac{1}{p(x)}$$

- **Bedingte Entropie, Wechselseitige Information**



4. Quellenkodierung

Informationsübertragung



4. Quellenkodierung

Quellenkodierung

- **Ziel:** Reduktion des Speicherbedarfs einer gegebenen Datenmenge
 - Datenkompression
- **Vorgehen:** Suchen einer kompakteren Repräsentation der Daten
 - Kodierung



4. Quellenkodierung

Grundbegriffe

- Alphabet **A**
- Wörter **A***
- Codewort **c(aba) = 01101**
- durchschnittliche Wortlänge **L**
- Redundanz
 - Information mehrfach enthalten
 - Weglassen ohne Informationsverlust



4. Quellenkodierung

Verlustbehaftete Kompression

- Informationen weglassen, die...
 - ... selten auftreten
 - ... nur wenig wahrgenommen werden
 - ... **irrelevant** sind
- Daten nicht (originalgetreu) wiederherstellbar



4. Quellenkodierung

Verlustbehaftete Kompression

- $X = \{a, b, c, d\}$

x	a	b	c	d
$p(x)$	0,5	0,25	0,125	0,125

- mögliche Darstellung:
 - $c(a) = 00$
 - $c(b) = 01$
 - $c(c) = 10$
 - $c(d) = 11$



4. Quellenkodierung

Verlustbehaftete Kompression

- $X = \{a, b, c, d\}$

x	a	b	c	d
p(x)	0,5	0,25	0,125	0,125

- mögliche Darstellung:
 - $c(a) = 00$
 - $c(b) = 01$
 - $c(c) = 10$
 - $c(d) = 11$
- komprimierte Darstellung:
 - $c(a) = 0$
 - $c(b) = 1$
 - $c(c) = -$
 - $c(d) = -$



4. Quellenkodierung

Verlustbehaftete Kompression

- $X = \{a, b, c, d\}$

x	a	b	c	d
$p(x)$	0,5	0,25	0,125	0,125

- mögliche Darstellung:

- $c(a) = 00$

- $c(b) = 01$

- $c(c) = 10$

- $c(d) = 11$

$$L = 2, \delta = 0$$

- komprimierte Darstellung:

- $c(a) = 0$

- $c(b) = 1$

- $c(c) = -$

- $c(d) = -$

$$L = 1, \delta = 0,25$$

4. Quellenkodierung

Verlustbehaftete Kompression



Quelle: www.oekostation.de/gfx/tipps/Coccinella.jpg

- Übertragung des Durchschnitts
- Bildpunkte:
 - 100%



4. Quellenkodierung

Verlustbehaftete Kompression



- Übertragung des Durchschnitts
- Bildpunkte:
 - 100%
 - 1%



4. Quellenkodierung

Verlustbehaftete Kompression



- Übertragung des Durchschnitts
- Bildpunkte:
 - 100%
 - 1%
 - 0,25%



4. Quellenkodierung

Verlustbehaftete Kompression



- Übertragung des Durchschnitts
- Bildpunkte:
 - 100%
 - 1%
 - 0,25%
 - 0,11%



4. Quellenkodierung

Verlustbehaftete Kompression



- Übertragung des Durchschnitts
- Bildpunkte:
 - 100%
 - 1%
 - 0,25%
 - 0,11%
 - 0,06%



4. Quellenkodierung

Verlustbehaftete Kompression



- Übertragung des Durchschnitts
- Bildpunkte:
 - 100%
 - 1%
 - 0,25%
 - 0,11%
 - 0,06%
 - 0,04%

4. Quellenkodierung

Verlustfreie Kompression

- Daten weglassen, die...
 - ... keine weiteren Informationen enthalten
 - ... **redundant** sind
- Daten originalgetreu wiederherstellbar



4. Quellenkodierung

Verlustfreie Kompression

- **Ansatz:** Kürzere Codewörter für häufige Zeichen



4. Quellenkodierung

Verlustfreie Kompression

- **Ansatz:** Kürzere Codewörter für häufige Zeichen

- $X = \{a, b, c, d\}$

x	a	b	c	d
$p(x)$	0,5	0,25	0,125	0,125

- mögliche Darstellung:

- $c(a) = 00$
- $c(b) = 01$
- $c(c) = 10$
- $c(d) = 11$



4. Quellenkodierung

Verlustfreie Kompression

- **Ansatz:** Kürzere Codewörter für häufige Zeichen

- $X = \{a, b, c, d\}$

x	a	b	c	d
$p(x)$	0,5	0,25	0,125	0,125

- mögliche Darstellung:

- $c(a) = 00$
- $c(b) = 01$
- $c(c) = 10$
- $c(d) = 11$

- optimierte Darstellung:

- $c(a) = 0$
- $c(b) = 10$
- $c(c) = 110$
- $c(d) = 111$



4. Quellenkodierung

Verlustfreie Kompression

- **Ansatz:** Kürzere Codewörter für häufige Zeichen

- $X = \{a, b, c, d\}$

x	a	b	c	d
$p(x)$	0,5	0,25	0,125	0,125

- mögliche Darstellung:

- $c(a) = 00$

- $c(b) = 01$

- $c(c) = 10$

- $c(d) = 11$

$$L = 2, \delta = 0$$

- optimierte Darstellung:

- $c(a) = 0$

- $c(b) = 10$

- $c(c) = 110$

- $c(d) = 111$

$$L = 1,75, \delta = 0$$

4. Quellenkodierung

Verlustfreie Kompression

- *Was ist die optimale Länge eines Codewortes?*
- *Wann ist ein Code optimal?*



4. Quellenkodierung

Verlustfreie Kompression

- *Was ist die optimale Länge eines Codewortes?*
- *Wann ist ein Code optimal?*
- Shannons **Source Coding Theorem:**
 - $H(X) \leq L$



4. Quellenkodierung

Verlustfreie Kompression

- *Was ist die optimale Länge eines Codewortes?*
- *Wann ist ein Code optimal?*
- Shannons **Source Coding Theorem:**
 - $H(X) \leq L < H(X) + 1$



4. Quellenkodierung

Huffman Code

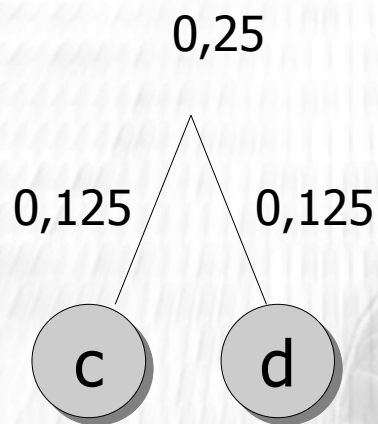
x	a	b	c	d
p(x)	0,5	0,25	0,125	0,125



4. Quellenkodierung

Huffman Code

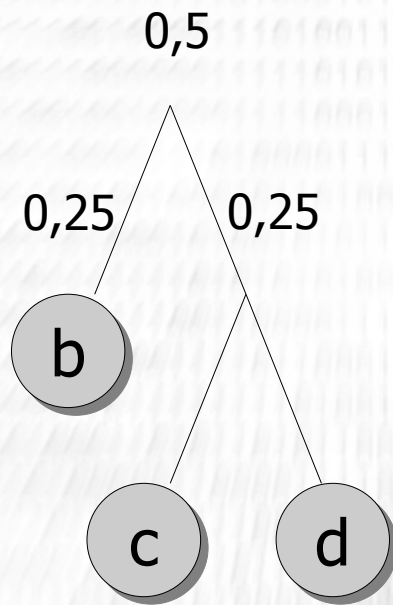
x	a	b	c	d
p(x)	0,5	0,25	0,125	0,125



4. Quellenkodierung

Huffman Code

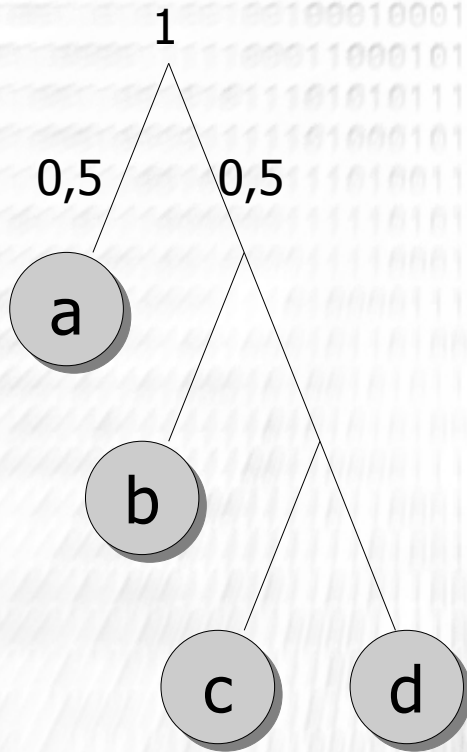
x	a	b	c	d
p(x)	0,5	0,25	0,125	0,125



4. Quellenkodierung

Huffman Code

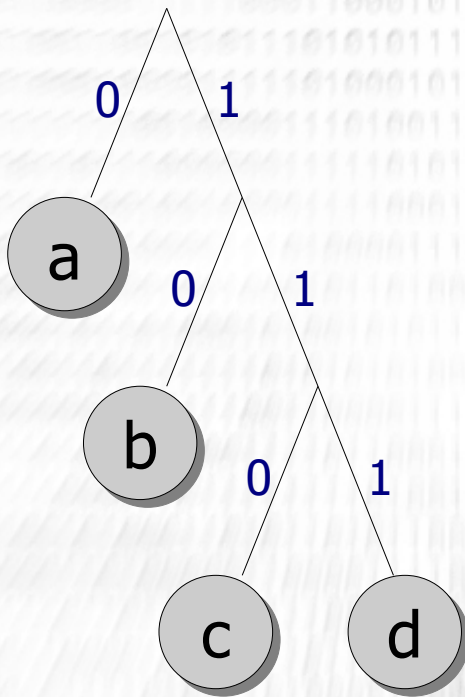
x	a	b	c	d
p(x)	0,5	0,25	0,125	0,125



4. Quellenkodierung

Huffman Code

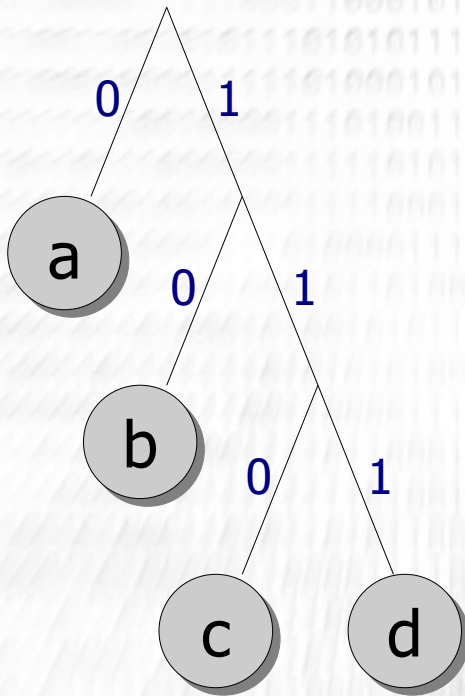
x	a	b	c	d
$p(x)$	0,5	0,25	0,125	0,125



4. Quellenkodierung

Huffman Code

x	a	b	c	d
p(x)	0,5	0,25	0,125	0,125



- Kodierung:
 - $c(a) = 0$
 - $c(b) = 10$
 - $c(c) = 110$
 - $c(d) = 111$

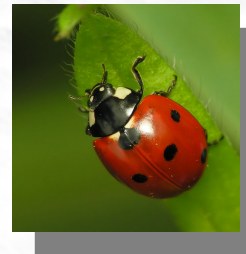


4. Quellenkodierung

Zusammenfassung

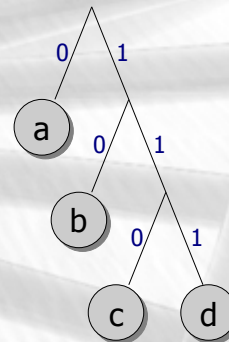
- **verlustbehaftete Kompression**

- Irrelevanzreduktion
- $\delta > 0$



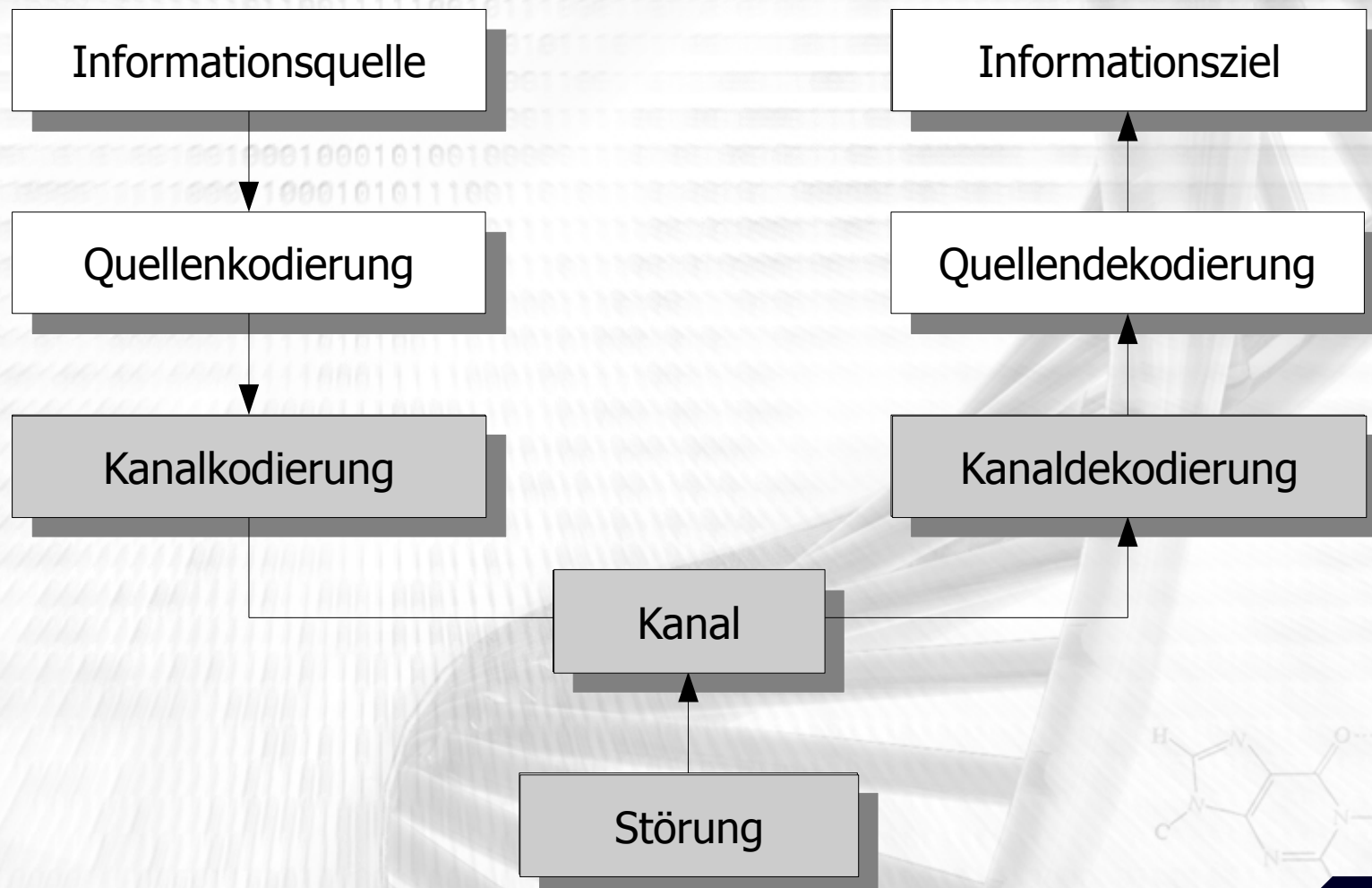
- **verlustfreie Kompression**

- Redundanzreduktion
- $H(X) \leq L < H(X) + 1$



5. Kanalkodierung

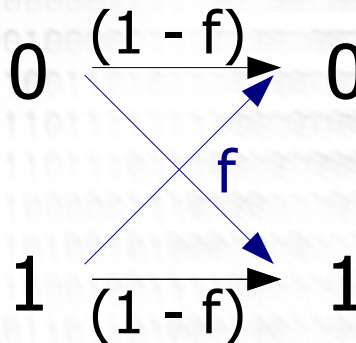
Informationsübertragung



5. Kanalkodierung

Fehlerhafte Übertragung

- Fehler entstehen durch das **Kippen** von Bits:



Quelle: MacKay 2003

- Beispiele:
 - verrauschte Radiosignale
 - Genmutationen der DNA
 - ungenaues Abtasten von CDs



5. Kanalkodierung

Lösungsansätze

- physikalische Verbesserungen
 - besseres Leitermaterial
 - höherer Energieaufwand
- fehlerkorrigierende Codes
 - Repetition Codes
 - Hamming Codes



5. Kanalkodierung

Repetition Codes

- Kodierung
 - Mehrfachübertragungen
 - a b c ---> 0 10 110 ---> 000 111 000 111 111 000



5. Kanalkodierung

Repetition Codes

- Kodierung
 - Mehrfachübertragungen
 - a b c ---> 0 10 110 ---> 000 111 000 111 111 000
- Rauschen
 - 000 111 000 111 111 000 ---> 00**1** 111 000 **0**11 111 0**1**0



5. Kanalkodierung

Repetition Codes

- Kodierung

- Mehrfachübertragungen

- a b c ---> 0 10 110 ---> 000 111 000 111 111 000

- Rauschen

- 000 111 000 111 111 000 ---> 00**1** 111 000 **0**11 111 0**1**0

- Dekodierung

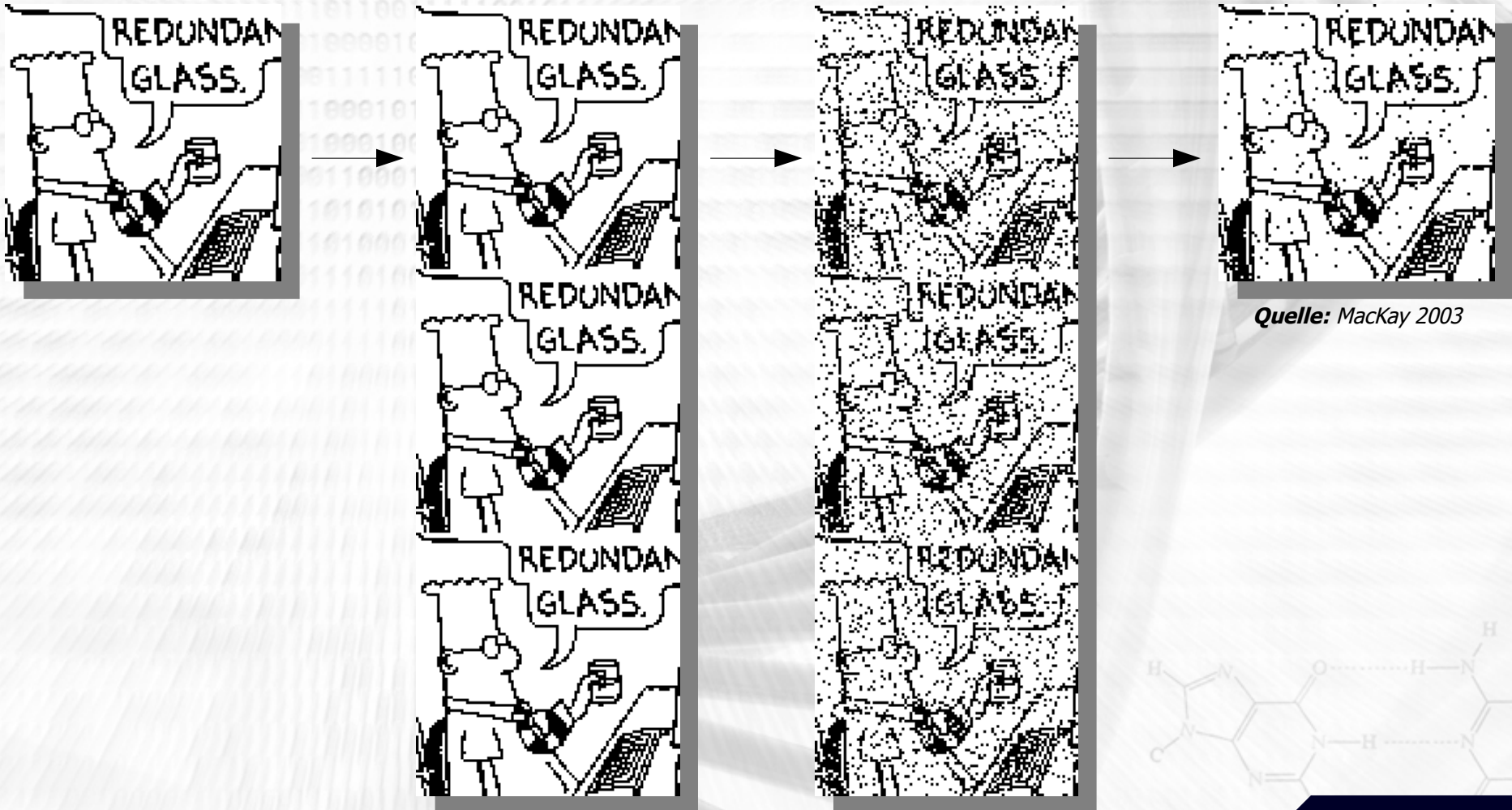
- Mehrheitsentscheidung

- 00**1** 111 000 **0**11 111 0**1**0 ---> 0 10 110 ---> a b c



5. Kanalkodierung

Repetition Codes

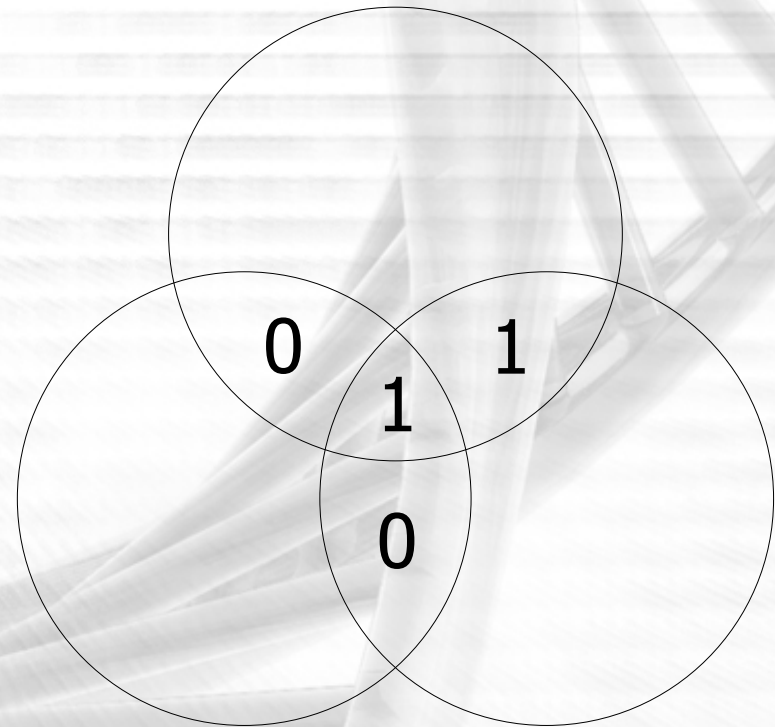


Quelle: MacKay 2003

5. Kanalkodierung

Hamming Codes

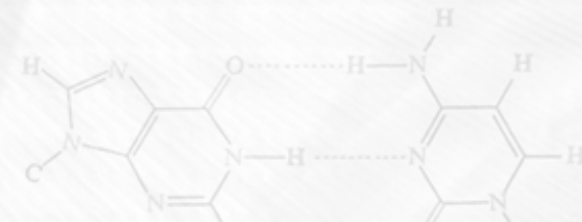
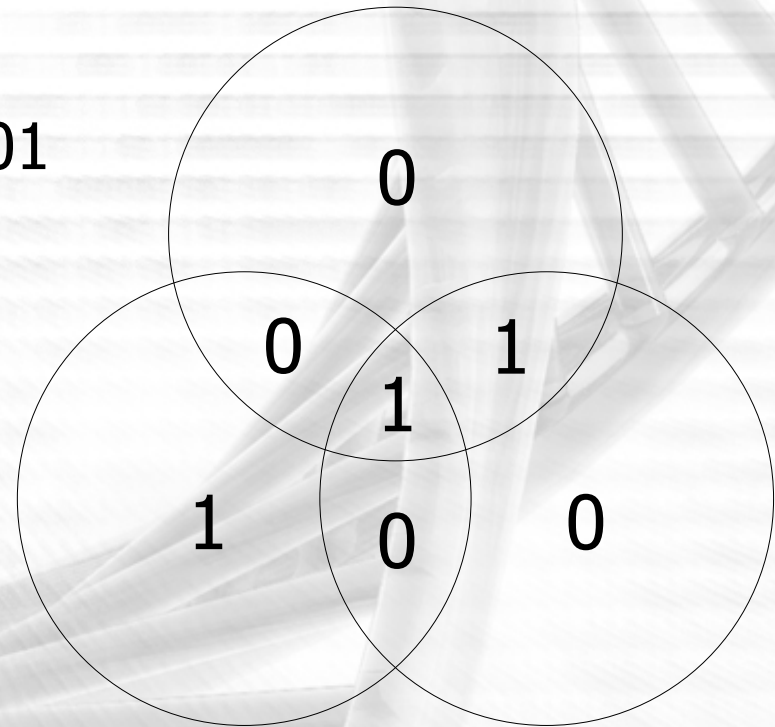
- Kodierung
 - a c ---> 0 1 1 0



5. Kanalkodierung

Hamming Codes

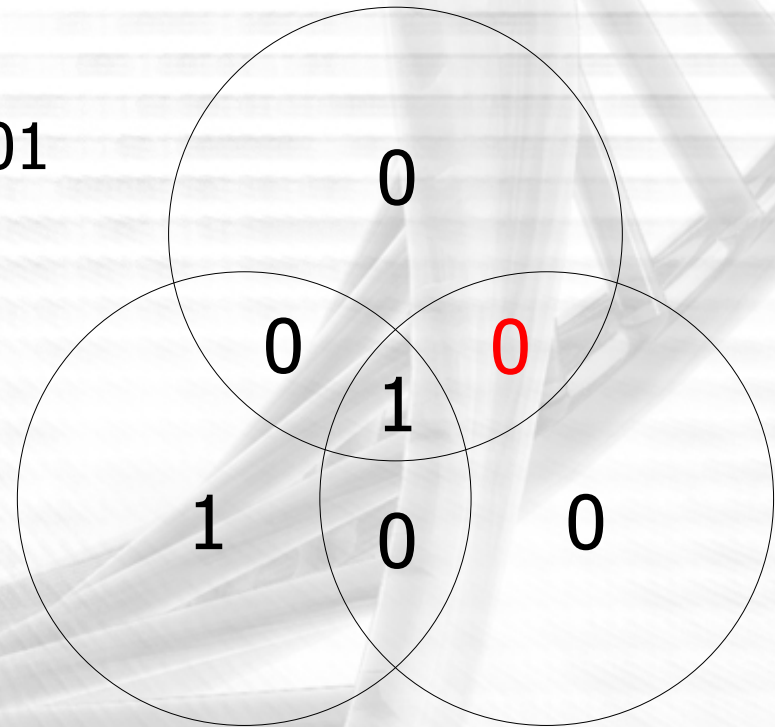
- Kodierung
 - a c ---> 0 110 ---> 0 110 001
 - Paritätsbits einfügen



5. Kanalkodierung

Hamming Codes

- Kodierung
 - a c ---> 0 110 ---> 0 110 001
 - Paritätsbits einfügen
- Rauschen
 - 0 110 001 ---> 0 **0**10 001



5. Kanalkodierung

Hamming Codes

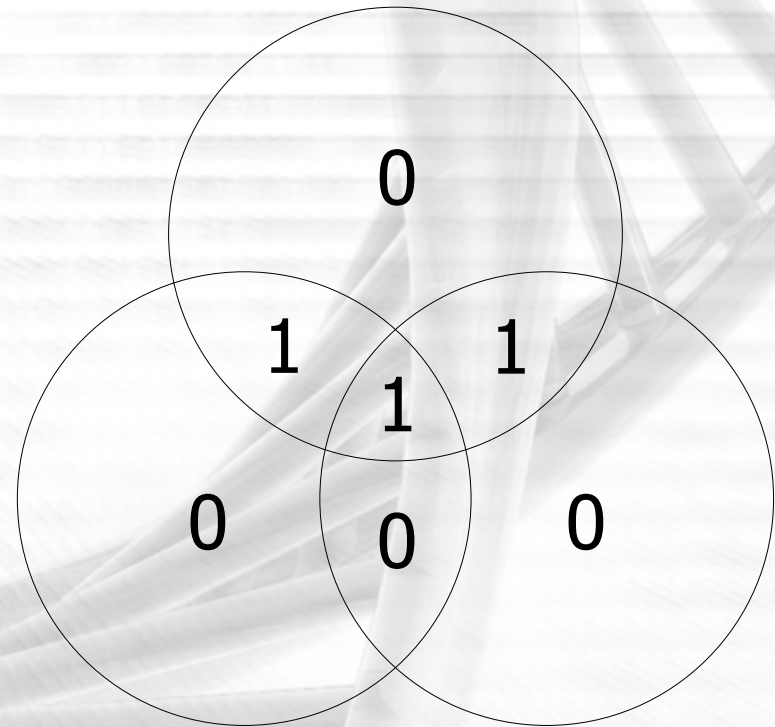
- Kodierung
 - a c ---> 0 110 ---> 0 110 001
 - Paritätsbits einfügen
- Rauschen
 - 0 110 001 ---> 0 **0**10 001
- Dekodierung
 - Paritätsbits auswerten
 - 0 **0**10 001 ---> 0 110 ---> a c



5. Kanalkodierung

Hamming Codes

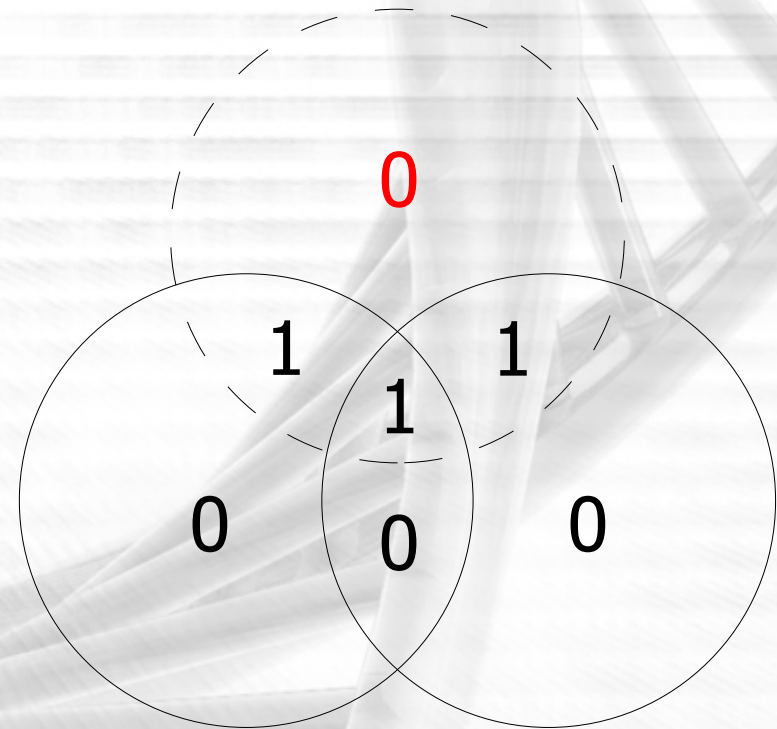
- empfangener Code:
1110000
- gesandte Nachricht: ?



5. Kanalkodierung

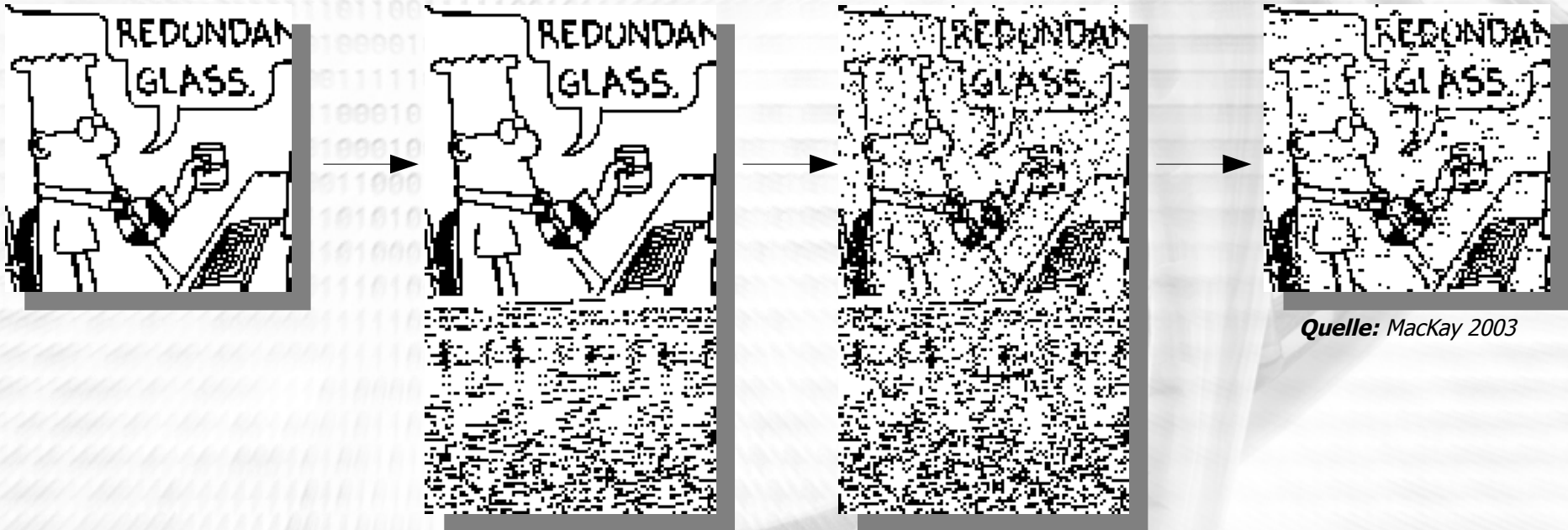
Hamming Codes

- empfangener Code:
1110**0**00
- gesandte Nachricht:
111 0 **0**00 ---> 111 0
---> d a



5. Kanalkodierung

Hamming Codes



Quelle: MacKay 2003



5. Kanalkodierung

Rate

- $Rate = \frac{\text{Informationsbits}}{\text{gesandte Bits}}$

- Repetition Code R3

- $Rate = \frac{1}{3} \approx 0,33$

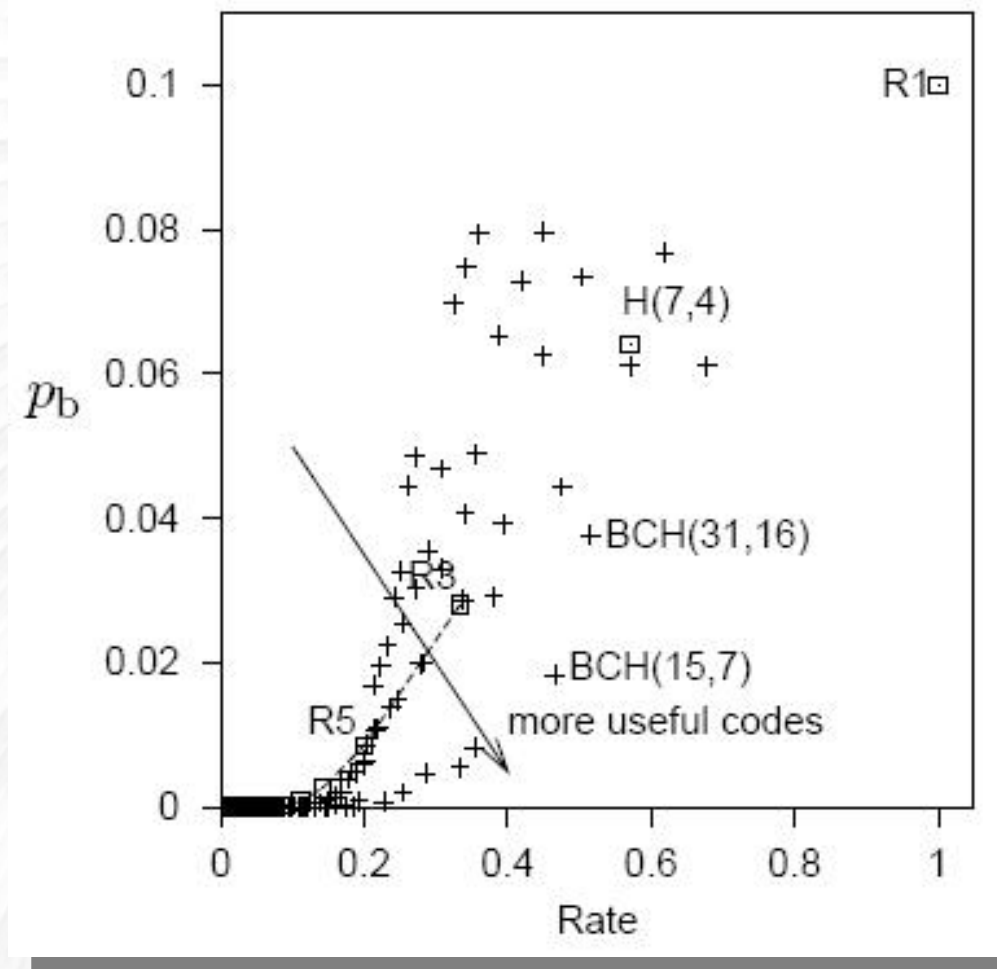
- Hamming Code (7,4)

- $Rate = \frac{4}{7} \approx 0,57$



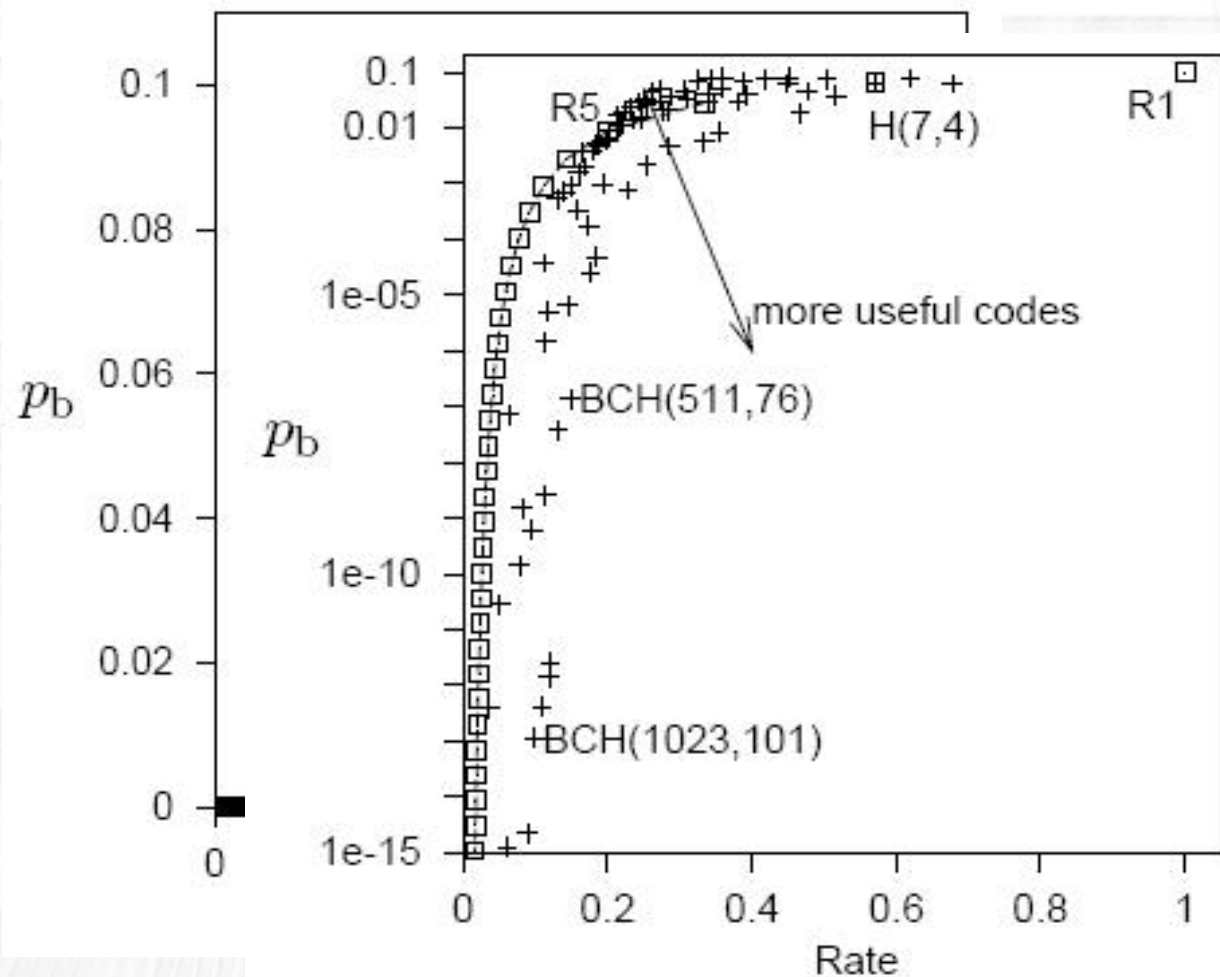
5. Kanalkodierung

Rate und Fehlerwahrscheinlichkeit



5. Kanalkodierung

Rate und Fehlerwahrscheinlichkeit



5. Kanalkodierung

Noisy-Channel Coding Theorem

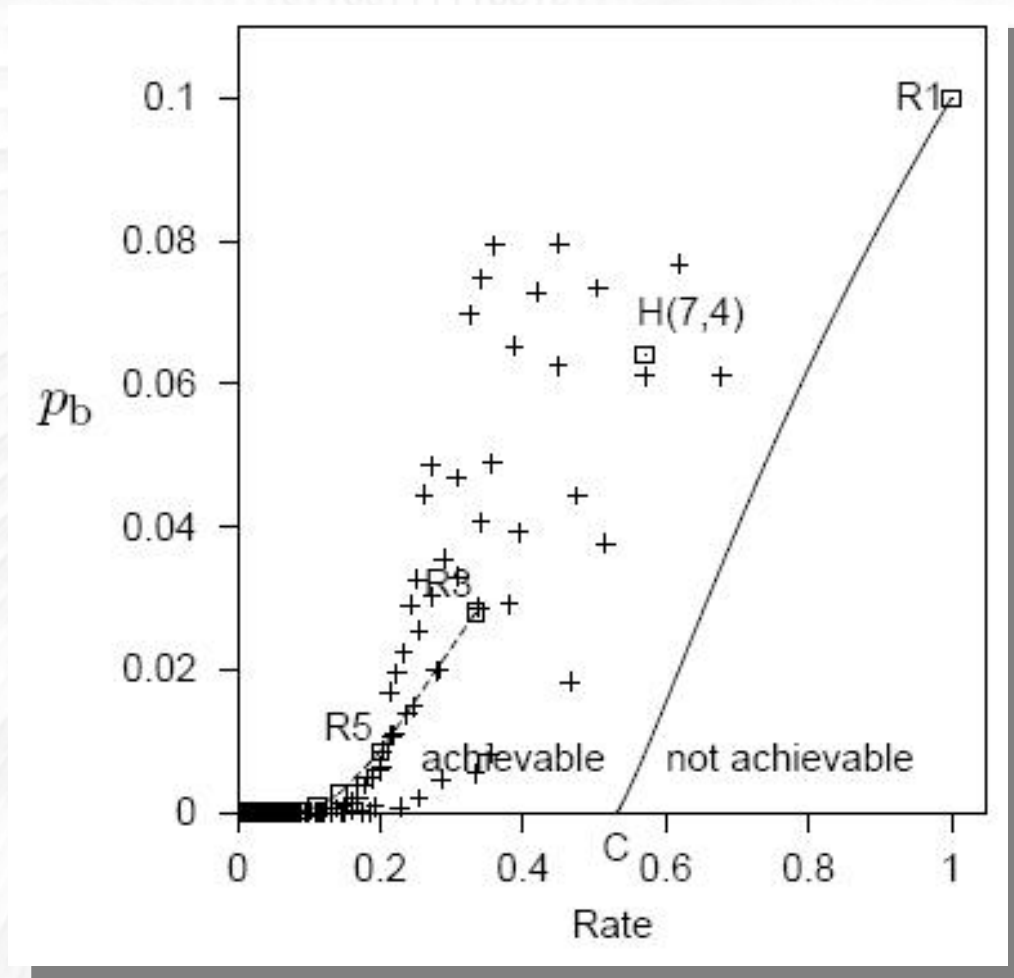
„Information can be communicated over a noisy channel at a non-zero rate with arbitrarily small error probability.“

MackKay (2003)



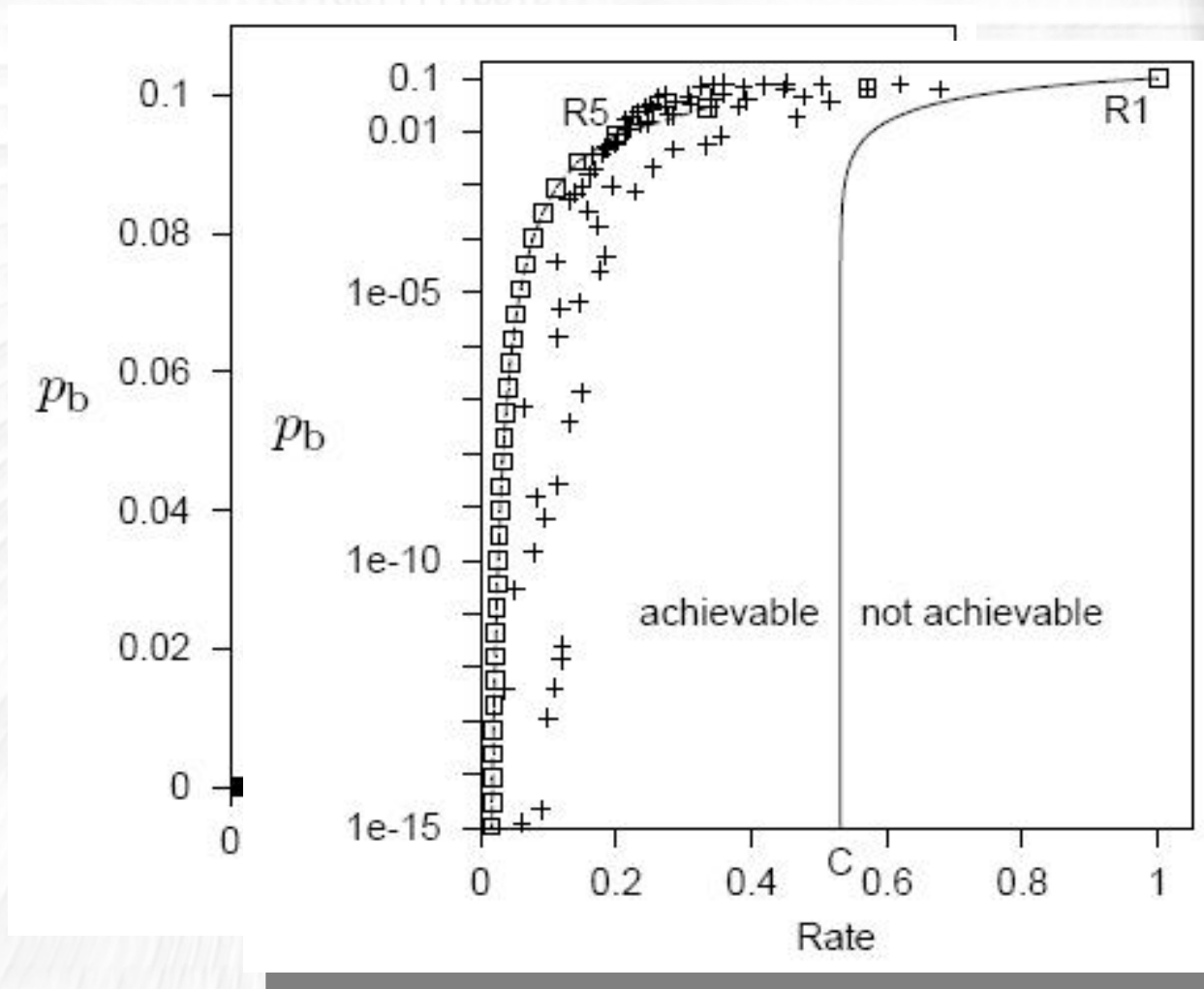
5. Kanalkodierung

Noisy-Channel Coding Theorem



5. Kanalkodierung

Noisy-Channel Coding Theorem



5. Kanalkodierung

Noisy-Channel Coding Theorem

- Kapazität C
 - maximale Übertragungsrate bei beliebig kleinem Fehler
 - $C > 0$
 - existiert für jeden Kanal



5. Kanalkodierung

Noisy-Channel Coding Theorem

- Beispiel:

Festplatte mit $f = 0,1$



gewünschte Fehlerwahrscheinlichkeit $p_b = 10^{-15}$



5. Kanalkodierung

Noisy-Channel Coding Theorem

- Beispiel:

Festplatte mit $f = 0,1$

gewünschte Fehlerwahrscheinlichkeit $p_b = 10^{-15}$



- mit Repetition Code:



5. Kanalkodierung

Noisy-Channel Coding Theorem

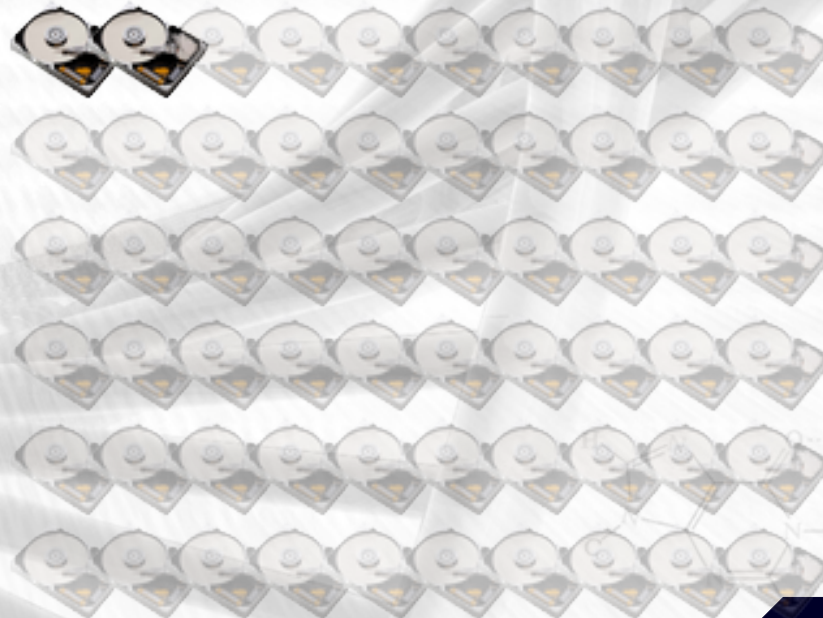
- Beispiel:

Festplatte mit $f = 0,1$



gewünschte Fehlerwahrscheinlichkeit $p_b = 10^{-15}$

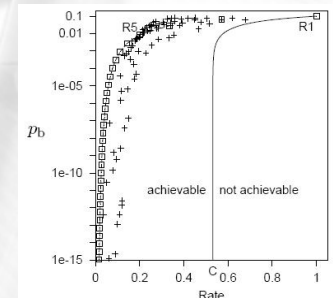
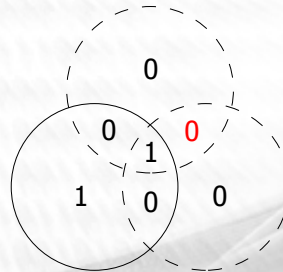
- mit Repetition Code:
- Kapazität $C \approx 0,53$
- theoretisch möglich:



5. Kanalkodierung

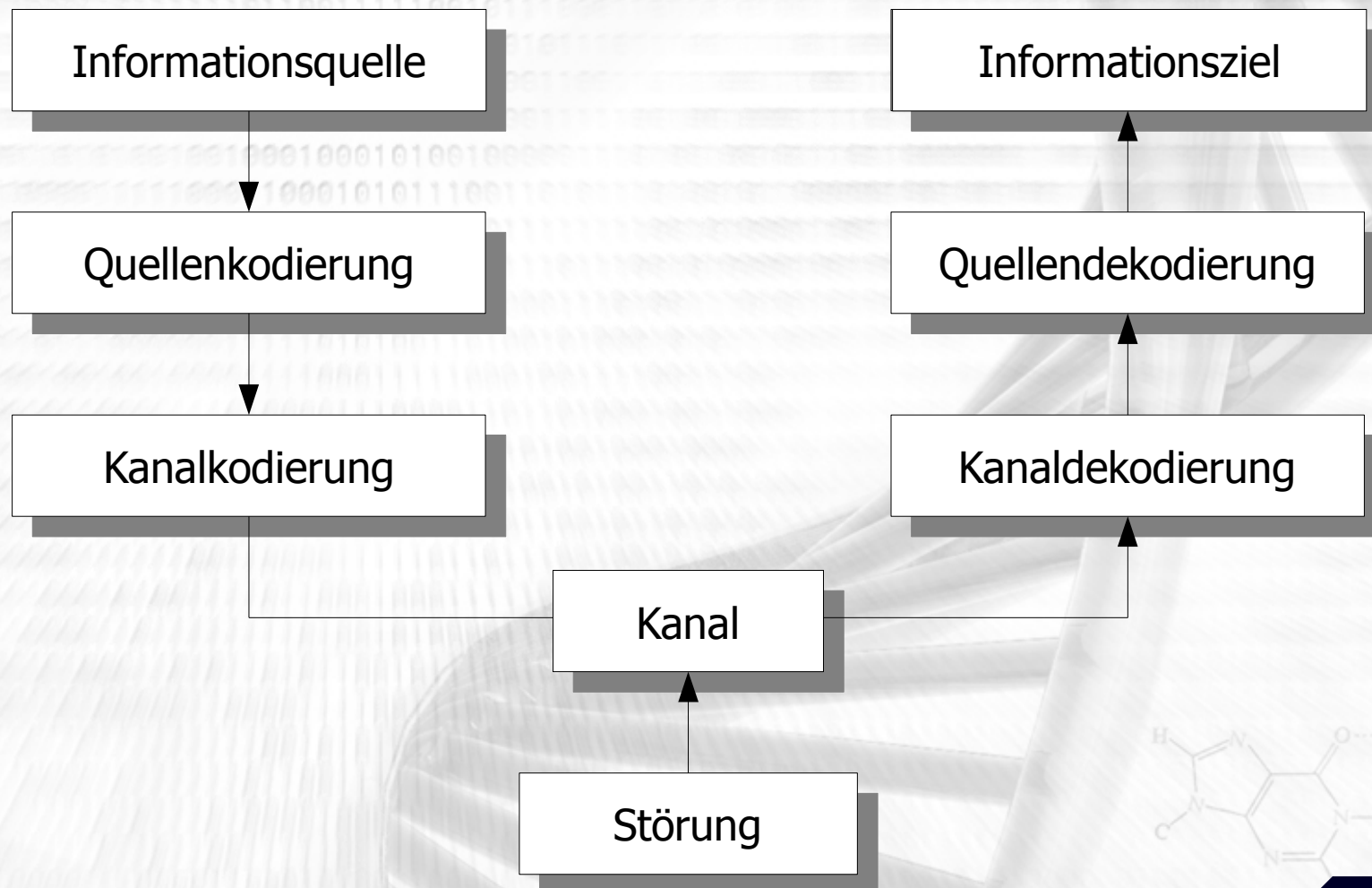
Zusammenfassung

- Kanäle sind verrauscht
- Fehlerkorrigierende Codes
 - Repetition Codes
 - Hamming Codes
- Noisy-Channel Coding Theorem
 - Kapazität C



6. Zusammenfassung

Informationsübertragung



The background features a blurred, grayscale image of a chair on the right side. Overlaid on this is a pattern of binary code (0s and 1s) that appears to be moving or vibrating, creating a sense of digital activity. The overall aesthetic is modern and technical.

Grundlagen der Informationstheorie

Hanna Rademaker und Fynn Feldpausch

