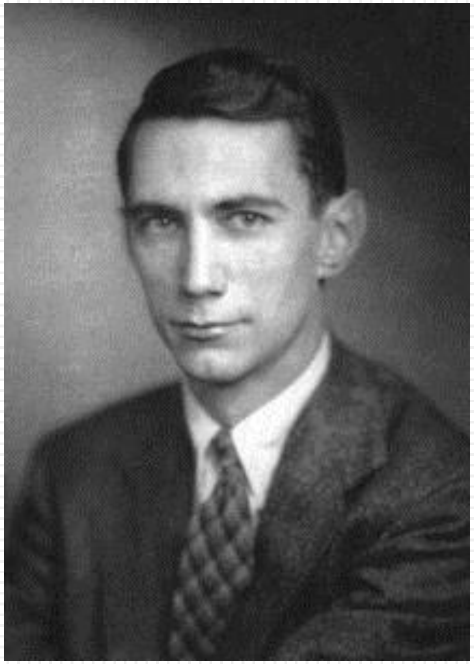
UR

Informationstheorie

Informationswissenschaft
Universität Regensburg
Jürgen Reischer

Einführung

- Die moderne Informationstechnologie basiert letztlich auf Arbeiten eines der wichtigsten Wissenschaftler des 20. Jahrhunderts:
  - *Claude Elwood Shannon* und sein Werk mit dem Titel "*A Mathematical Theory of Communication*" (veröffentlicht 1948/1949, geschrieben während des 2. Weltkriegs);



http://en.wikipedia.org/wiki/File:Claude_Elwood_Shannon_(1916-2001).jpg (23.8.2011)

Einführung

- Shannons Arbeit enthält zwei grundlegende Konzepte, die auch heute noch in der Informatik/ Informationswissenschaft und angrenzenden Wissenschaften relevant sind:
  - eine mathematisch fundierte *Informationstheorie*;
  - die binäre *Einheit 'bit'*, die unmittelbar aus der Theorie folgt.
- In der Kommunikations- bzw. Informationstheorie wird das technische Problem der Übertragung von Daten (Informationen) mittels Signalen (Zeichen und Zeichenketten) thematisiert und gelöst.

Informationsbegriff

- Im Vordergrund steht der *quantitative* Aspekt von Information:
  - Wie kann Information möglichst *effizient und effektiv* von einer *Quelle zu einem Ziel über einen Kanal* übertragen werden?
    - *effizient*: zeit- und energiesparend (Nutzung der maximalen Übertragungskapazität eines Kanals);
    - *effektiv*: störungssicher kodiert bezüglich der zu übermittelnden Botschaft/Nachricht (Vermeidung von Rauschen und Verfälschungen der Nachricht auf dem Kanal).

Informationsbegriff

- Wie viel und welche Information enthält ein Zeichen (Signal) bzw. eine Kette/Sequenz von Zeichen?
  - nummerisch erfasste Länge einer Botschaft: Anzahl/Menge übermittelter Informationseinheiten;
  - nummerisch erfasster Gehalt an Information einer Botschaft: Höhe des Informationsgehalts bzw. der Informativität (Aussagekraft) von Informationseinheiten.
- Welche syntaktischen Aspekte besitzt Information?
  - Frage nach der Kodierung einzelner Signale/Zeichen;
  - Frage nach der Struktur(ierung)/Anordnung einzelner Signale zu komplexeren Nachrichten (Informationen).

Informationsbegriff

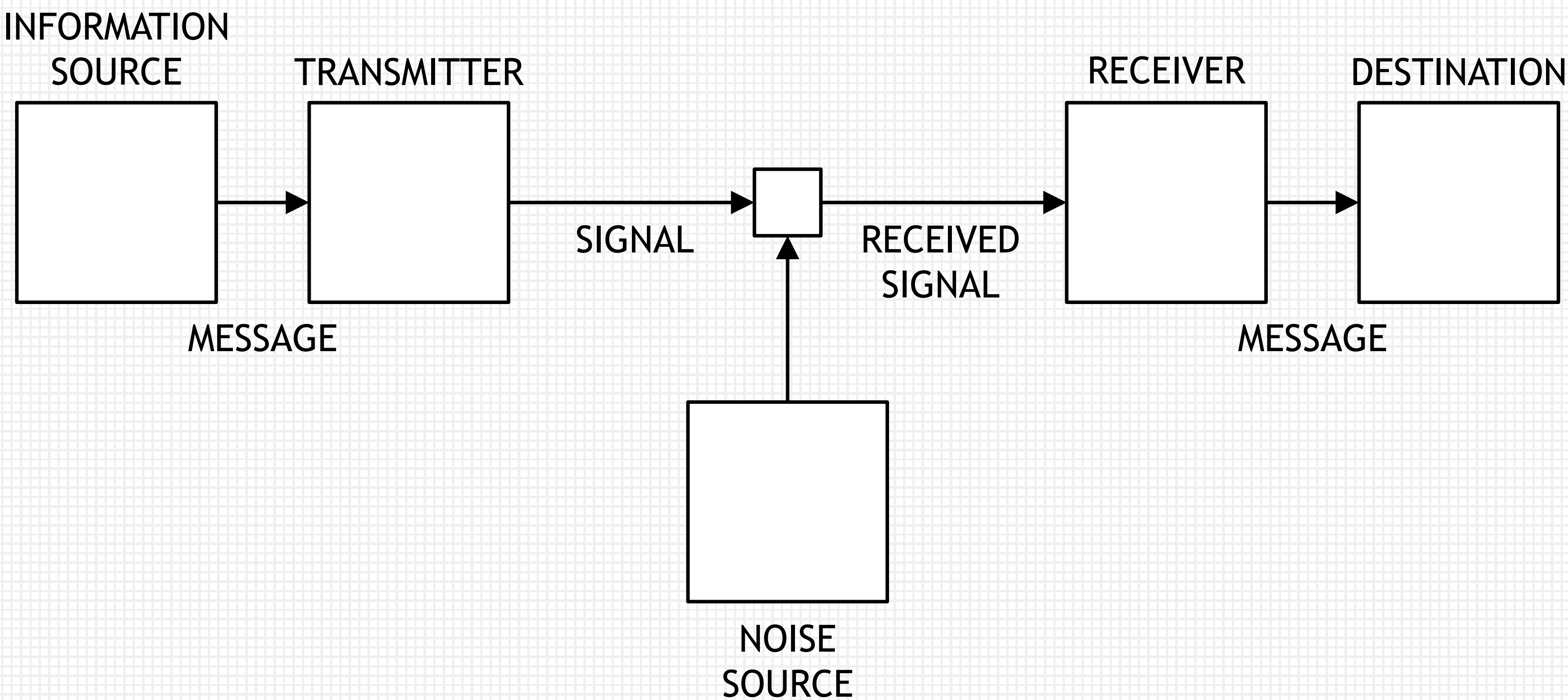
- Der *qualitative* Aspekt der Information tritt in den Hintergrund:
  - Was ist Information (seinem inneren Wesen nach)?
  - Welche semantischen Aspekte besitzt Information (im Sinne einer Nachricht):
    - Bezug eines Zeichens bzw. einer Nachricht zu Objekten der Welt (Abbildungswert/Repräsentationsaspekt und Sinngehalt/Bedeutung eines Zeichens)?
    - Bezug einer Nachricht zu den abgebildeten Sachverhalten der Welt (Wahrheitstreue/Faktizität einer Nachricht)?

Informationsbegriff

- Welche pragmatischen Aspekte besitzt Information?
  - Bezug eines Zeichens bzw. einer Botschaft zum Zeichennutzer (Mensch und/oder Maschine):
    - Zweck einer Information auf Seiten des Zeichenproduzenten;
    - Wirkung einer Information auf Seiten des Zeichenrezipienten.
  - Bezug eines Zeichens zum Kontext des Zeichennutzers:
    - Einbettung in die physikalische, physiologische, psychologische und soziale Situation des Zeichennutzers (konkrete Umstände der Zeichennutzung);
    - Einbettung in den sprachlich-kommunikativen, gesellschaftlichen, politischen, wirtschaftlichen usw. Kontext der Zeichennutzer (diskursiver und kultureller Hintergrund).

Kommunikationsmodell

- Im 'Kommunikationsmodell' nach Shannon sind die Komponenten einer Nachrichtenübertragung vereint:



Kommunikationsmodell

Erläuterungen:

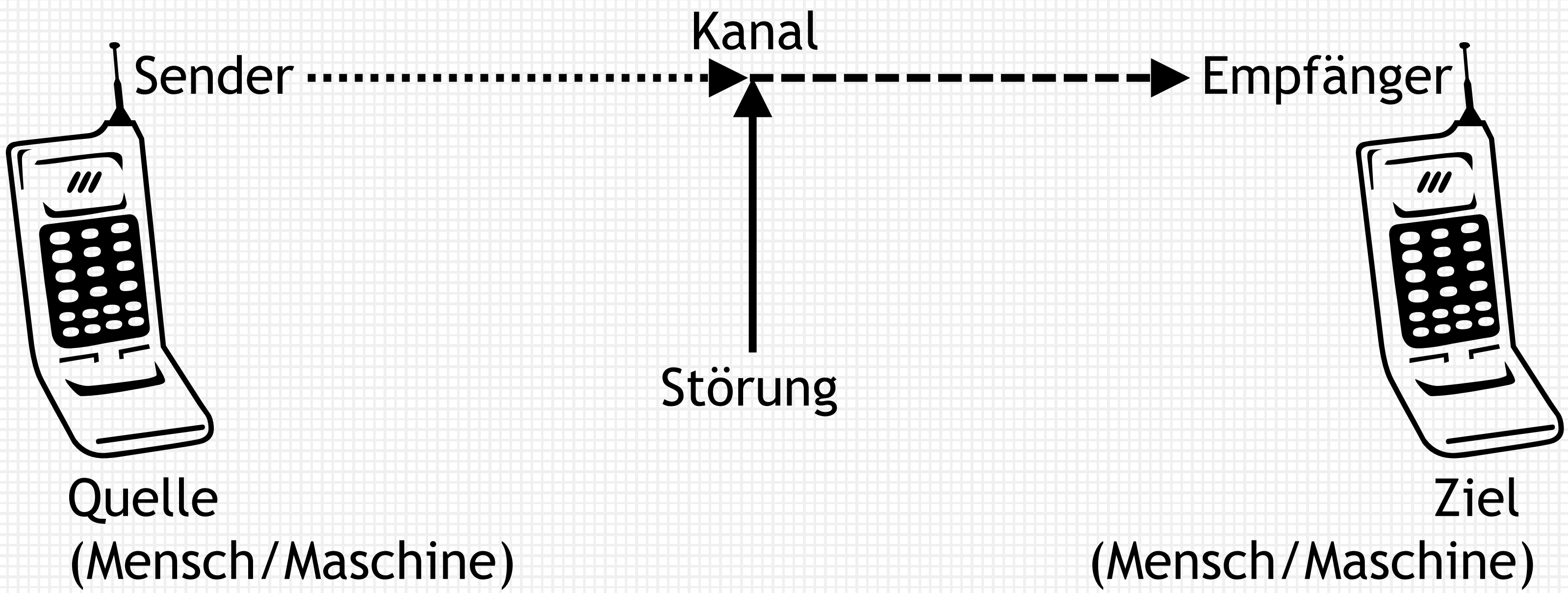
- Quelle ('Source') und Ziel ('Destination'): Mensch oder Maschine als gezielter oder zufälliger Informationsproduzent bzw. –rezipient:
  - Enkodierung/Synthese einer Botschaft auf Seiten der Nachrichtenquelle: Auswahl von Zeichen aus einer Zeichenmenge (Alphabet, Lexikon) und Kombination der Zeichen zu Nachrichten (Grammatik);
  - Dekodierung/Analyse der Botschaft auf Seiten des Nachrichtenziels: Klassifikation der Zeichen und Interpretation der ankommenden Nachrichten.

Kommunikationsmodell

- Sender ('Transmitter') und Empfänger ('Receiver'):
  - Verschicken einer von der Informationsquelle konstruierten Nachricht in Form von Impulsen über einen Kanal zum Informationsempfänger (z. B. via Sendeantenne);
  - Abnehmen der vom Informationssender auf den Kanal gegebenen Impulse (z. B. via Empfangsantenne) und Rekonstruktion der Nachricht durch das Informationsziel.
- Rausch-/Störquelle ('Noise Source'):
  - Störung des Signals durch eine externe Rauschquelle (bei prinzipieller Rekonstruierbarkeit der Originalsignale);
  - Verfälschung des Signals (ohne Rekonstruierbarkeit).

Kommunikationsmodell

- Beispiel für alltägliche technische Kommunikation:



Das Shannonsche Modell ist als Modell der *technischen* Nachrichtenübermittlung gedacht; es eignet sich nicht zur Darstellung *menschlicher* Kommunikationsprozesse.

Kommunikationsmodell

- Bei der Betrachtung menschlicher Kommunikation wären nicht der *technische* Sender und Empfänger von Interesse, sondern Sprecher und Hörer (Personen) als reale Informationsquellen/-ziele:
  - *Semantik*:
    - Über welche *Inhalte* zu welchem *Thema* tauschen sich Sprecher und Hörer aus?
    - Welche Zeichen mit welcher *Bedeutung* werden hierzu benutzt?
  - *Pragmatik*:
    - Mit welchen *Absichten* und intendierten *Wirkungen* reden Sprecher und Hörer miteinander?
    - In welchem aktuellen Kontext tun sie dies?

Kommunikationsmodell

- Voraussetzung für eine sinnvolle Kommunikation ist, dass Informationsquelle und -ziel (bzw. Sprecher und Hörer) über einen gemeinsamen Zeichenvorrat verfügen und die Nachricht dem Ziel nicht schon vorab bekannt ist (vgl. [Rechenberg & Pomberger 2006: 214]):

Nachricht	redundant	nicht-redundant
relevant	keine Information, da Nachricht bekannt oder vorhersagbar ist (z. B. weiß das Ziel, dass auf 'q' ein 'u' folgt)	Information im eigentlichen Sinne, da Nachricht weder bekannt noch vorhersagbar ist
irrelevant	Quelle und Ziel verfügen nicht über einen gemeinsamen Zeichenvorrat (d. h. verstehen sich nicht)	

Zeichenmengen

- Die zu übertragende Nachricht wird aus einer endlichen Zeichenmenge ('Lexikon') konstruiert:
  - Eine ungeordnete Zeichenmenge wird als *Zeichenvorrat* bezeichnet.
  - Eine geordnete oder ordenbare Zeichenmenge heißt *Alphabet*:
    - jede Menge von Zeichen, die nach einer bestimmten Logik (an)geordnet oder sortiert werden kann;
    - hierzu ist eine Ordnungsrelation nötig, die festlegt, wie die Zeichen logisch anzuordnen sind (z. B. in üblicher alphabetischer Reihenfolge: 'a' < 'b' < 'c' usw.).

Zeichenmengen

- Beispiel für Zeichenvorräte (ungeordnet):
  - *Zeichen des Internationalen Phonetischen Alphabets IPA* (das eigentlich kein *Alphabet* ist!);
  - *Handzeichen*: ✌, 👌, 👍, 👎, ☜, ☞, ☝, ☟, ✋;
  - *Währungszeichen*: €, £, $, ¥, ₣, ₧;
  - *Genetischer Kode*: Adenin (A), Cytosin (C), Guanin (G), Thymin (T);
  - *Geschlechterzeichen*: ♀, ♂;
  - *Einzelzeichen (Sonderfall)*: ¤.

Zeichenmengen

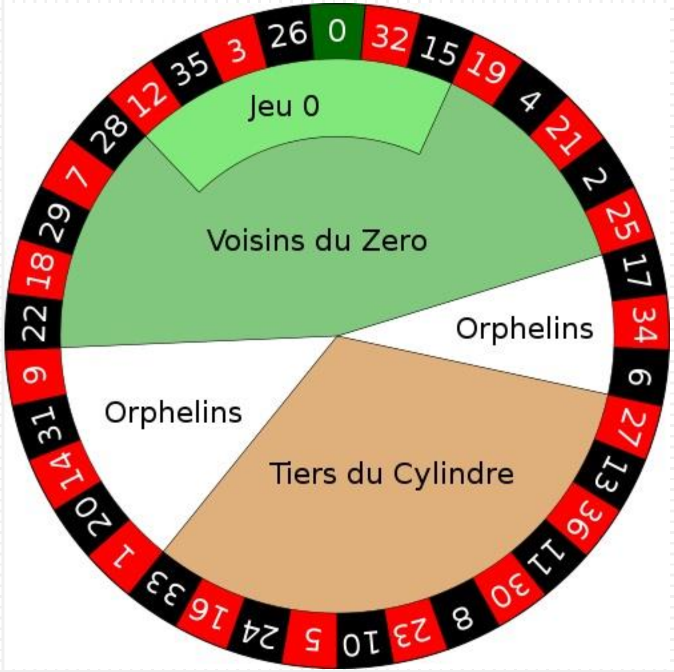
- Beispiele für Alphabete (geordnet):
  - *Ziffern*: 0, 1, 2, 3, 4, 5, 6, 7, 8, 9;
  - *Buchstaben*: A–Z, a–z (bzw. griechische, kyrillische Buchstaben usw.);
  - *Kartenwerte*: (Ass,) 2–10, Bube, Dame, König, Ass;
  - *Kartenfarben (beim Skat)*: ♣, ♠, ♥, ♦ (Kreuz, Pik, Herz, Karo in dieser Abfolge);
  - *Tierkreiszeichen*: ♈, ♉, ♊, ♋, ♌, ♍, ♎, ♏, ♐, ♑, ♒, ♓;
  - *Augenzahlen*: 1, 2, 3, 4, 5, 6 (Würfel).

Zeichenmengen

- Jedem Zeichen aus der Zeichenmenge wird dabei eine bestimmte Wahrscheinlichkeit p zugeordnet, von der Informationsquelle ausgewählt zu werden (Zeichenauswahl):
  - Die Zeichen einer Zeichenmenge können
    - gleich- oder ungleichwahrscheinlich verteilt sein,
    - zufällig oder gezielt ausgewählt werden.
  - Das Auftreten oder Erscheinen eines Zeichens stellt ein *Zeichenereignis* oder eine einfache Nachricht dar (komplexe Nachricht bei mehreren Zeichen).

Zeichenmengen

- Beispiel für Zeichen aus einer Zeichenmenge mit ihren Wahrscheinlichkeiten p:
  - Gleiche Wahrscheinlichkeiten (meist berechenbar):
    - Münzseite:
      p = 1/2 = 0.5 = 50%;
    - Würfelauge:
      p = 1/6 = 0.1666... ≈ 16.67%;
    - Roulettewert:
      p = 1/37 ≈ 0.027027... ≈ 2.7%.



http://upload.wikimedia.org/wikipedia/commons/thumb/7/7d/European_roulette_wheel.svg/500px-European_roulette_wheel.svg.png (7.10.2011)

Zeichenmengen

- Ungleiche Wahrscheinlichkeiten (oftmals nur beobachtbar, aber nicht vorausberechenbar):
  - Relative Auftretenshäufigkeiten von Buchstaben im Englischen in Prozent (vgl. Tabelle rechts);
  - gezinkte Münze, gezinkter Würfel, gezinktes Roulette usw.

🗎	p_k	🗎	p_k	🗎	p_k
e	12.702	d	4.253	p	1.929
t	9.056	l	4.025	b	1.492
a	8.167	c	2.782	v	0.978
o	7.507	u	2.758	k	0.772
i	6.966	m	2.406	j	0.153
n	6.749	w	2.360	x	0.150
s	6.327	f	2.228	q	0.095
h	6.094	g	2.015	z	0.074
r	5.987	y	1.974		

Informationsgehalt

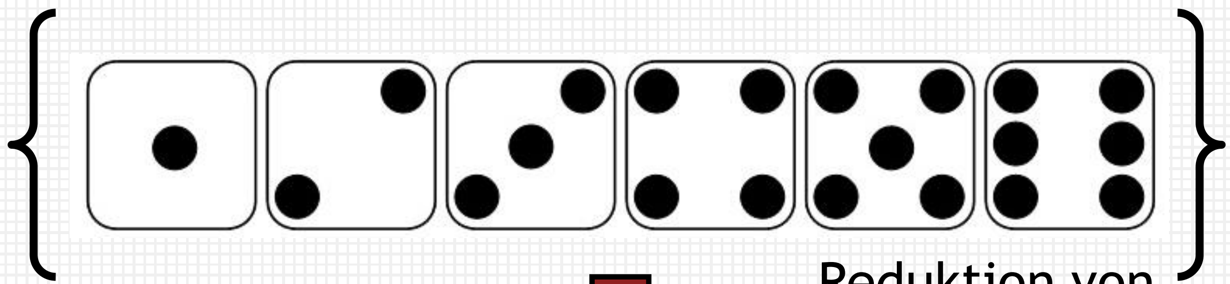
- Der Gehalt an Information (Informationsgehalt), der in einem Zeichenereignis steckt, wird mit der Wahrscheinlichkeit der Auswahl (im Sinne des Auftretens) des Zeichens korreliert:
  - *Vor* dem Zeichenereignis herrscht zunächst Unwissen (Ungewissheit/Unsicherheit) über den Ausgang des Ereignisses (vgl. Münz-/Würfelwurf, Buchstabe).
  - *Nach* dem Zeichenereignis wird das Unwissen um einen bestimmten Faktor reduziert, da der Ausgang des (Zufalls-)Ereignisses nun bekannt ist: *Hinterher hat man mehr Wissen = Information als vorher*.

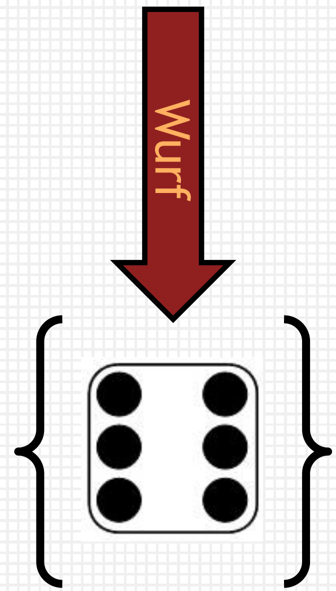
Informationsgehalt

- Information im Sinne der Informationstheorie kann als *beseitigtes oder reduziertes Nicht-/Unwissen* bzw. *Ungewissheit/Unsicherheit* im Informationsziel verstanden werden:
  - Information im statischen Sinne:
    Zustand geringeren/r Nichtwissens/Ungewissheit = mehr Wissen/Gewissheit (im Informationsziel);
  - Information im dynamischen Sinne:
    Prozess der Informierung bzw. Informationshinzu-gewinnung (im Informationsziel).

Informationsgehalt

- Beispiel:
  - Ungewissheit (Unsicherheit, Unwissen) über den Ausgang eines Würfelwurfs (6 Möglichkeiten):





Reduktion von Möglichkeiten = Reduktion von Ungewissheit bzw. Unwissen = Wissenszuwachs

  - Gewissheit (Sicherheit, Wissen) über den Ausgang des Wurfs (1 Möglichkeit):

Informationsgehalt

- Der Umfang des Gehalts an Information eines Zeichenereignisses ist im Informationsziel umso größer, je mehr Unwissen über den Ausgang dort reduziert werden kann:
  - *Der Informationsgehalt ist 0*, wenn das Ergebnis bereits vorab bekannt ist, da kein neues Wissen erlangt wird (z. B. ein Würfel, der *stets* auf 6 fällt oder ausschließlich 6en trägt, d. h. $p_{1-5} = 0$, $p_6 = 1$).
  - *Der Informationsgehalt ist größer als 0*, wenn das Ergebnis vorab nicht bekannt ist (z. B. $p_{1-6} = 1/6$).

Informationsgehalt

- Die Größe der übermittelten Information lässt sich auch am so genannten *Überraschungswert* (*Neuigkeitswert*) oder *Prognosewert* eines Zeichenereignisses im Empfänger verdeutlichen:
  - Je geringer/höher die Wahrscheinlichkeit des Eintretens eines Ereignisses ist, desto mehr/weniger *überrascht* es uns bzw. desto schlechter/besser kann es *prognostiziert* werden:
    - schlechte Prognose: Lotto-Sechser (ca. 1:14000000);
    - gute Prognose: keine Eins beim Würfeln (5:6 = 0.83).

Informationsgehalt

- Bei Wahrscheinlichkeit 1 bzw. 0 für ein Ereignis ist die Überraschung über das Eintreten minimal bzw. maximal und die Prognose best- bzw. schlechtest-möglich:
  - *Wahrscheinlichkeit p = 1*:
    - Zeichenereignis tritt immer ein, ein anderes Zeichenereignis tritt hingegen nie ein;
    - der Überraschungsfaktor des Eintretens ist daher Null.
  - *Wahrscheinlichkeit p = 0*:
    - Zeichenereignis tritt nie ein, irgendein anderes Zeichenereignis tritt hingegen immer ein;
    - der Überraschungsfaktor des Eintretens wäre hier Unendlich.

Informationsgehalt

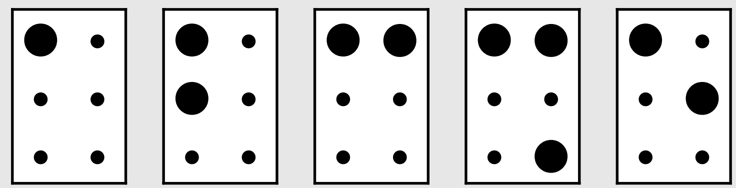
- Zwei Parameter haben Einfluss auf den Überraschungswert bzw. Informationsgehalt eines Zeichenereignisses:
  - *Anzahl Auswahlmöglichkeiten in der Zeichenmenge*: Je mehr Zeichen in der Menge, desto größer ist der Überraschungswert/Informationsgehalt:
    - Die Prognoserate sinkt, da grundsätzlich mehr Möglichkeiten (zum Raten) zur Auswahl stehen (unabhängig von der tatsächlichen Wahrscheinlichkeit einzelner Zeichen);
    - Beispiel: Das Geschlecht einer Person kann mit höherer Wahrscheinlichkeit erraten werden als ihre Haarfarbe.

Informationsgehalt

- *Verteilung der Wahrscheinlichkeiten der einzelnen Zeichen in der Zeichenmenge* (bei gleichbleibender Anzahl Möglichkeiten):
  - *Nicht-gleichwahrscheinliche Ereignisse*:
    - Da einige Zeichen häufiger bzw. seltener sind als andere, lassen sich diese besser bzw. schlechter prognostizieren, d. h. der Überraschungswert und Informationsgehalt sinkt bzw. steigt.
    - Beispiel: Würde die Zero beim Roulette viel häufiger erscheinen als alle anderen Zahlen, würde jeder auf die Zero setzen.
  - *Gleichwahrscheinliche Ereignisse*:
    - Da keines der Zeichen bei der Auswahl heraussticht, sind alle gleichermaßen gut/schlecht prognostizierbar;
    - daher 'worst case' für jede Prognose (z. B. faires Roulette-Spiel).

Übungsaufgaben

- Wie viele mögliche Zustände oder verschiedene Zeichen (Zeichenmenge) können folgende Objekte theoretisch darstellen?

<table>
<tr><th>Objekt</th><th></th><th>Zustände</th></tr>
<tr><td>Blindenschrift-Matrix von 2×3 Elementen (tastbare Erhebung vs. Nicht-Erhebung).
Beispiel-Zeichen rechts:</td><td>
die ersten Buchstaben A bis E</td><td></td></tr>
<tr><td>7-Segment-Anzeige für Ziffern auf Taschenrechnern oder Digitaluhren (s. rechts).</td><td>
jedes Segment kann einzeln an oder aus sein</td><td></td></tr>
<tr><td>Tic-Tac-Toe-Spiel mit zwei Arten von Spielsteinen.
Beispiel-Spielfeld rechts:</td><td><table><tr><td>o</td><td></td><td>o</td></tr><tr><td>o</td><td>x</td><td>x</td></tr><tr><td>x</td><td></td><td></td></tr></table>
jedes Feld kann unabhängig besetzt oder leer sein</td><td></td></tr>
</table>

Übungsaufgaben

Objekt	Zustände
zwei Hände mit zehn Fingern (jeder kann einzeln gehoben oder gesenkt sein)	
4×4-Raster von Schwarz-Weiß- oder Hell-Dunkel-Pixeln	

- Bei welchen Zeichenmengen handelt es sich um ein Alphabet, bei welchen um einen Zeichenvorrat?

Zeichenmenge	Alphabet oder Zeichenvorrat	
Blindenschrift-Matrix	☐ Alphabet	☐ Zeichenvorrat
7-Segment-Anzeige	☐ Alphabet	☐ Zeichenvorrat
zwei Hände mit zehn Fingern	☐ Alphabet	☐ Zeichenvorrat
Tic-Tac-Toe-Spiel	☐ Alphabet	☐ Zeichenvorrat
4×4-Raster von Pixeln	☐ Alphabet	☐ Zeichenvorrat

Übungsaufgaben

- Berechnen Sie die Wahrscheinlichkeiten folgender sechs Zeichenereignisse und ordnen Sie diese anschließend nach ihrem Überraschungs- bzw. Prognosewert an:
  - Auftreten einer Seite eines Polyeders:
    - Tetraeder (Vierflächler);
    - Oktaeder (Achtflächler);
    - Dodekaeder (Zwölfflächler);
    - Ikosaeder (Zwanzigflächler).
  - Erraten einer Karte aus einem Kartenspiel mit 52 Karten (4 Farben, 13 Werte);
  - Ermittlung der Position einer Schachfigur auf einem Schachbrett mit 64 Feldern.

Rang	Ereig-nis E	p_E
1.		
2.		
3.		
4.		
5.		
6.		

Informationsgehalt

- Zur Ermittlung des Ausgangs eines (Zeichen-)Ereignisses mit *gleichwahrscheinlichen* Möglichkeiten kann eine *optimale* Fragestrategie angewendet werden:
  1. Der Möglichkeitsraum (Auswahlmenge) M wird in zwei möglichst gleich große Mengen M_1 und M_2 unterteilt (idealerweise zwei Hälften gleicher Größe);
  2. durch eine Ja-Nein-Frage (Entscheidungsfrage) wird ermittelt, in welcher der beiden Mengen M_1 bzw. M_2 das Ereignis liegt: M wird entsprechend M_1 oder M_2 zugeordnet;
  3. Schritt 1. wird mit dem neuen M wiederholt, sofern noch kein eindeutiges Zeichenereignis identifiziert wurde.

Informationsgehalt

- Beispiele:
  - Münze (2 Möglichkeiten):
    - Frage: War es Kopf?
    - Antwort: Ja → Kopf, Nein → Zahl.
  - Genetischer Kode (4 Möglichkeiten):

<table>
<tr><th>Frage/Menge</th><th colspan="4">Aktuelle Auswahlmengen</th></tr>
<tr><td></td><td colspan="4">Ausgangsmenge:
{Adenin, Cytosin, Guanin, Thymin}</td></tr>
<tr><td>Adenin oder Cytosin?</td><td colspan="2">Ja:
{Adenin, Cytosin}</td><td colspan="2">Nein:
{Guanin, Thymin}</td></tr>
<tr><td>Endet Name auf "-nin"?</td><td>Ja:
{Adenin}</td><td>Nein:
{Cytosin}</td><td>Ja:
{Guanin}</td><td>Nein:
{Thymin}</td></tr>
</table>

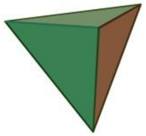
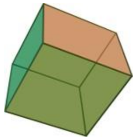
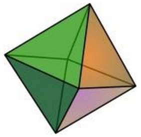
Informationsgehalt

- Planet des Sonnensystems:

<table>
<tr><th>Fragen</th><th colspan="8">Aktuelle Auswahlmengen</th></tr>
<tr><td></td><td colspan="8">{Merkur, Venus, Erde, Mars,
Jupiter, Saturn, Uranus, Neptun}</td></tr>
<tr><td>Gesteins-planet (oh-ne Ringe)?</td><td colspan="4">Ja:
{Merkur, Venus,
Erde, Mars}</td><td colspan="4">Nein:
{Jupiter, Saturn,
Uranus, Neptun}</td></tr>
<tr><td>Mondlos? / Über 60 Monde?</td><td colspan="2">Ja:
{Merkur,
Venus}</td><td colspan="2">Nein:
{Erde,
Mars}</td><td colspan="2">Ja:
{Jupiter,
Saturn}</td><td colspan="2">Nein:
{Uranus,
Neptun}</td></tr>
<tr><td>Innerer der beiden?</td><td>Ja:
{Me.}</td><td>Nein:
{Ve.}</td><td>Ja:
{Er.}</td><td>Nein:
{Ma.}</td><td>Ja:
{Ju.}</td><td>Nein:
{Sa.}</td><td>Ja:
{Ur.}</td><td>Nein:
{Ne.}</td></tr>
</table>

Informationsgehalt

- Weitere Beispiele:

Ereignis / Auswahl	Auswahlmöglichkeiten	Entscheidungsfragen
Geschlecht-Auswahl	2	1
Fußball-Ergebnis	3	1–2
Tetraeder-Wurf*	4 	2
Hexaeder-Wurf*	6 	2–3
Oktaeder-Wurf*	8 	3
Dodekaeder-Wurf*	12 	3–4

Informationsgehalt

Ereignis / Auswahl	Möglichkeiten	Entscheidungsfragen
Bundesland-Auswahl	16	4
Ikosaeder-Wurf*	20 	4–5
Zahn-Auswahl	32	5
Spielkarten-Auswahl	52	5–6
Schachbrett-Position	64	6
2-stellige Ziffernfolge	100	6–7
ASCII-Zeichen-Auswahl	128	7
Auswahl aus Gros	144	7–8
Farben-Auswahl GIF	256	8

* Polyeder-Grafiken via http://de.wikipedia.org/wiki/Spielwürfel (28.3.2012)

Informationsgehalt

- Offenbar besteht ein Zusammenhang zwischen der Anzahl gleichwahrscheinlicher Auswahlmöglichkeiten M und der (mittleren) Anzahl Entscheidungsfragen N, die man idealerweise stellen muss, um den Ausgang eines Ereignisses zu bestimmen:
  - Bei 2, 4, 8, 16, 32, 64, 128 und 256 Möglichkeiten beträgt die Anzahl der Fragen jeweils die ganzzahligen Werte 1, 2, 3, 4, 5, 6, 7 oder 8.
  - Bei 3, 6, 12, 20, 52, 100 und 144 ergeben sich nicht-ganzzahlige Werte von 1–2, 2–3, 3–4, 4–5, 5–6, 6–7 oder 7–8 Fragen (je nach Fragestrategie bzw. eingetretenem Ereignis).

Informationsgehalt

- Mathematisch kann der Zusammenhang zwischen gleichwahrscheinlichen Möglichkeiten M und optimalen Entscheidungsfragen N als einfache Funktion beschrieben werden:
  - *Schritt 1*: Hat man N Entscheidungsfragen mit jeweils genau 2 Antworten Ja und Nein zur Verfügung, so lassen sich damit 2^N Möglichkeiten erfragen: $2^N = M$.
  - *Schritt 2*: Will man die Anzahl notwendiger Fragen N für eine bestimmte Auswahlmenge an Möglichkeiten M ermitteln, so ist die Gleichung per Logarithmisierung nach N aufzulösen: $2^N = M \Leftrightarrow \mathrm{ld}\ 2^N = \mathrm{ld}\ M \Leftrightarrow N \cdot \mathrm{ld}\ 2 = \mathrm{ld}\ M \Leftrightarrow N \cdot 1 = \mathrm{ld}\ M \Leftrightarrow$ $N = \mathrm{ld}\ M$ (ld = 'logarithmus dualis' mit Basis 2).

Informationsgehalt

- Der Logarithmus (log) ist allgemein die Umkehrfunktion zur Exponentierung ($\_^{exp}$):
  - $b^x = a \Leftrightarrow x = \log_b a$:
    - $b^x = a$: x ist der Exponent, mit der die Basis b potenziert werden muss, um a zu erhalten;
    - $x = \log_b a$: a ist das Argument, von dem der Logarithmus zur Basis b genommen werden muss, um x zu erhalten.
  - Für unseren Fall ist die Basis b = 2:
    - $2^x = a \Leftrightarrow x = \log_2 a = \mathrm{ld}\ a = \mathrm{lb}\ a$ ('logarithmus binaris');
    - $2^N = M \Leftrightarrow N = \log_2 M = \mathrm{ld}\ M = \mathrm{lb}\ M$ (mit x = N, a = M).

Informationsgehalt

- Der *Informations-* oder *Entscheidungsgehalt* H eines Zeichens, das den Ausgang eines Zeichenereignisses bei M *gleichwahrscheinlichen* Auswahlmöglichkeiten p = 1/M (⇨ M = 1/p, p·M = 1) darstellt, wird definiert als:

$$H = \mathrm{ld}\ M = \mathrm{ld}\ 1/p \text{ bit} \qquad (\text{mit } M = 1/p \text{ wie oben})$$

  - 'bit' ('*bi*nary un*it*' oder '*b*asic *i*nformation un*it*') ist die *mathematische Einheit* zum Zahlenwert von H, die den Informationsgehalt charakterisiert (ähnlich wie die physikalischen Einheiten 'kg' und 'cm' die Zahlenwerte für Gewicht oder Größe charakterisieren);
  - statt 'bit' wird auch die Einheit 'Shannon' ('Sh') verwendet.

Informationsgehalt

- Anstelle von N wird hier die Bezeichnung H verwendet, wobei dies nur unterschiedlichen Interpretationen des Wertes geschuldet ist:
  - N verweist auf die (durchschnittliche) Anzahl Ja-Nein-Fragen, die zur Ermittlung eines bestimmten Zeichens/ Zeichenereignisses der Auswahlmenge notwendig ist;
  - H verweist auf die Größe des Entscheidungs- bzw. Informationsgehalts, den die Informationsquelle als Ganzes bzw. ein einzelnes Zeichenereignis besitzt.
  - ➡ *Der Informationsgehalt wird identifiziert mit der (durchschnittlichen) Anzahl Ja-Nein-Fragen!*

Informationsgehalt

- Alltagssprachlich formuliert heißt dies:
  - Information ist letztlich definiert als *Antwort auf eine (Entscheidungs-)Frage* (bzw. mehrere Antworten auf mehrere Fragen, wenn eine allein hierzu nicht ausreicht).
  - Information steht in Zusammenhang mit Wissen bzw. Unwissen/Unsicherheit:
    - Solange man Fragen stellen muss, ist man sich einer Sache nicht sicher und weiß zu wenig;
    - konnten alle Fragen beantwortet werden, ist man sich seiner Sache sicher und weiß genug.

Informationsgehalt

- Mit jeder Antwort auf eine Entscheidungsfrage wird der Möglichkeitsraum idealerweise halbiert oder weiter eingeschränkt:
  - Der Interpret der Antwort (das Informationsziel) besitzt weniger Unkenntnis/Ungewissheit über den Ausgang eines Ereignisses (den Zustand der Welt, das Ergebnis eines Zufallsexperiments usw.).
  - Die Antwort "Ja" oder "Nein" abstrahiert dabei vom jeweiligen Inhalt der Frage, d. h. die Antwort "Ja" oder "Nein" kann auf *jede* Entscheidungsfrage unabhängig von deren tatsächlichem Inhalt (Bedeutung, Sinn) gegeben werden.

Informationsgehalt

- Der Informationsgehalt kann nun auch für solche Ereignisse berechnet werden, wo die Auswahlmenge keine 2er-Potenz 2^M ist:

Ereignis / Berechnung	H = ld M	H = ld 1/p (mit p = 1/M)	Ergebniswert
Fußballergebnis	H = ld 3	ld 1/(1/3)	1.585 bit
Würfel-Wurf	H = ld 6	ld 1/(1/6)	2.585 bit
Dodekaeder-Wurf	H = ld 12	ld 1/(1/12)	3.585 bit
Ikosaeder-Wurf	H = ld 20	ld 1/(1/20)	4.322 bit
Spielkarten-Zug	H = ld 52	ld 1/(1/52)	5.700 bit
Doppelziffer-Auswahl	H = ld 100	ld 1/(1/100)	6.644 bit
Gros-Auswahl	H = ld 144	ld 1/(1/144)	7.170 bit

Die Werte bemessen dabei die *durchschnittliche* (*mittlere*) Anzahl nötiger Entscheidungsfragen bei optimaler Fragestrategie.

Übungsaufgaben

- Machen Sie sich Gedanken über folgende Sonderfälle von Auswahlmengen:

Fall	Frage	Antwort
1	Wie viele Fragen benötigt man, um eine Auswahl aus einer Menge mit genau *einem* Element zu treffen?	
	Was heißt dies für den Informationsgehalt des Ereignisses?	
0	Wie viele Fragen benötigt man, um eine Auswahl aus einer leeren Menge zu treffen (0 Alternativen)?	
	Wie würde sich entsprechend der Informationsgehalt berechnen?	

Übungsaufgaben

- Wie hoch ist der Informationsgehalt bei folgenden Ereignissen, d. h. wie viele Entscheidungsfragen benötigt man im Mittel zur Identifizierung eines (Zeichen-)Ereignisses (vgl. Anhang für Logarithmus):

Ereignis	Informationsgehalt
Auswahl eines Tages der Woche	
Auswahl eines Feldes beim Tic-Tac-Toe (3×3)	
Auswahl einer Stunde des Tages	
Auswahl einer Lottozahl aus 49 Kugeln	
Auswahl einer Sekunde oder Minute	
Auswahl einer Person aus 1048576	

Informationsgehalt

- Für M *nicht-gleichwahrscheinliche* Ereignisse kann H nicht einfach aus H = ld M (bit) und p nicht einfach aus p = 1/M ermittelt werden.
- Stattdessen müssen die verschiedenen Wahrscheinlichkeiten der einzelnen Zeichenereignisse der Auswahlmenge *empirisch* (per Beobachtung) ermittelt werden:
  - Zunächst sind die absoluten *Häufigkeiten* der Ereignisse zu ermitteln (Auftretensfrequenzen);
  - daraus lassen sich die relativen Häufigkeiten bezogen auf die betrachtete Gesamtmenge und relativ zu den anderen möglichen Ereignissen ermitteln (z. B. Frequenz pro 1000).

Informationsgehalt

- Beispiele hierfür:
  - Wahrscheinlichkeiten von Buchstaben und Wörtern, wie sie in den Texten einer Schriftsprache auftreten;
  - Wahrscheinlichkeit der Blutgruppenzugehörigkeit zu einer der Blutgruppen A, B, AB, 0 mit jeweils Rhesusfaktor positiv oder negativ;
  - Wahrscheinlichkeit einer Mädchen- vs. Jungengeburt bzw. eines Todesfalles Mann vs. Frau;
  - Wahrscheinlichkeit der Fähigkeit zum Zungenrollen oder Ohrenwackeln.

Informationsgehalt

- Entsprechend muss die Berechnung des Entscheidungsgehalts H einer Zeichenquelle für M verschiedene Wahrscheinlichkeiten p_1 bis p_M von Zeichenereignissen der Auswahlmenge verallgemeinert werden:

$$H = p_1 \cdot \mathrm{ld}\ 1/p_1 + p_2 \cdot \mathrm{ld}\ 1/p_2 + \ldots + p_M \cdot \mathrm{ld}\ 1/p_M \text{ bit} = \Sigma_{K=1..M}\ p_K \cdot \mathrm{ld}\ 1/p_K \text{ bit}$$

- Für gleichwahrscheinliche Ereignisse ergibt sich wieder die erste Formel von oben (mit $p_1 = p_2 = \ldots = p_M = p$):
  $H = p \cdot \mathrm{ld}\ 1/p + p \cdot \mathrm{ld}\ 1/p + \ldots + p \cdot \mathrm{ld}\ 1/p$ (M Mal) $=$
  $M \cdot p \cdot \mathrm{ld}\ 1/p = 1 \cdot \mathrm{ld}\ 1/p$ $(M \cdot p = 1!)$ $= \mathrm{ld}\ 1/p = \mathrm{ld}\ M.$

Informationsgehalt

- H bezeichnet hierbei den *mittleren Informationsgehalt* einer Zeichenquelle, d. h. wie viele Ja-Nein-Entscheidungsfragen sind *im Mittel* zur Identifizierung eines bestimmten Zeichen(ereignisse)s notwendig:
  - Jedes Zeichen Z_K besitzt nun *seinen eigenen* Informationsgehalt $I(Z_K) = \mathrm{ld}\ 1/p_K$ bit (wogegen bei gleichwahrscheinlichen Ereignissen $I(Z) = \mathrm{ld}\ 1/p = \mathrm{ld}\ M$ bit war, da alle p_K für alle Z_K identisch waren);
  - aufgrund der unterschiedlichen Wahrscheinlichkeiten der Zeichenereignisse muss H daher als gewichtetes Mittel berechnet werden: $H = p_1 \cdot I(Z_1) + p_2 \cdot I(Z_2) + \ldots + p_M \cdot I(Z_M)$ bit (mit $I(Z_K) = \mathrm{ld}\ 1/p_K$ von oben).

Informationsgehalt

- Beispiele:
  - Überraschungswert einer 'krummen' Münze:
    - $p_{Kopf} = p_1 = 0.51$, $p_{Zahl} = p_2 = 0.49$;
    - $H = 0.51 \cdot ld\ 1/0.51 + 0.49 \cdot ld\ 1/0.49 \approx 0.9997$.
  - Informationsgehalt eines gezinkten Würfels:
    - $p_1 = p_2 = p_3 = p_4 = p_5 = 0.15$, $p_6 = 0.25$ (in 25% statt 16.67% aller Fälle erscheint eine Sechs);
    - $H = (0.15 \cdot ld\ 1/0.15) + (0.15 \cdot ld\ 1/0.15) + (0.15 \cdot ld\ 1/0.15) + (0.15 \cdot ld\ 1/0.15) + (0.15 \cdot ld\ 1/0.15) + (0.25 \cdot ld\ 1/0.25) = 5 \cdot (0.15 \cdot ld\ 1/0.15) + (0.25 \cdot ld\ 1/0.25) \approx 2.5527$.

Informationsgehalt

- Es ist zu beobachten, dass obige Werte beide *unterhalb* der Werte für entsprechend gleich-wahrscheinliche Ereignisse liegen:
  - 2.585 (fairer Würfel) > 2.5527 (gezinkter unfairer Würfel);
  - 1.0 (faire Münze) > 0.9997 (krumme unfaire Münze).
- *Allgemein gilt, dass M gleichwahrscheinliche Ereignisse immer informationshaltiger sind als M ungleichwahrscheinliche.*

Informationsgehalt

- Grund hierfür:
  - Sobald auch nur einer der Wahrscheinlichkeitswerte größer oder kleiner ist als alle anderen,
    - kann bezüglich dieses 'Ausreißers' eine bessere Prognose über das (Nicht-)Eintreten dieses Ereignisses abgegeben werden als über die anderen,
    - da das Vorabwissen über den Ausgang des Ereignisses (etwas) größer ist als bei gleichwahrscheinlichen Ereignissen.
  - Nur bei gleichwahrscheinlichen Ereignissen ist die Ungewissheit über den Ausgang eines 'Experiments' (z. B. Münz-/Würfelwerfen usw.) am größten, d. h. keines der Ereignisse 'sticht hervor'.

Übungsaufgaben

- Berechnen Sie für folgende Auswahl-Ereignisse den mittleren Informationsgehalt/Überraschungswert:

Ereignis	Informationsgehalt/Überraschungswert
Farbenblindheit (Rot-Grün-Verwechsler) mit einem Anteil von etwa 10% an der Bevölkerung	
Geburt an einem Arbeits- vs. Wochenendtag (gerechnet auf eine einzelne Woche ohne Feiertage)	

Übungsaufgaben

Aufschlag eines Meteoriten im Wasser bzw. auf Land bei einem Wasseranteil von 70.8% an der Erdoberfläche	
Eintreten eines Vorfalls A, B oder C, wenn B doppelt so wahrscheinlich ist wie A und C doppelt so wahrscheinlich wie B	
Geburt in einem der vier Quartale eines normalen Jahres mit 365 Tagen	

Übungsaufgaben

Weitere Aufgaben zur Vertiefung:

- Verlieren eines Teils einer bestimmen Sorte eines zerlegten 3×3×3-Zauberwürfels mit folgenden Sorten von Teilen:

Sorte des Teils	Eckstück	Kantenstück	Mittelstück	Innenstück
Anzahl/Chance				

- Wurf mit unfairem Würfel:

Auge	1	2	3	4	5	6
Chance	20%	5%	12.5%	12.5%	15%	35%

- Blutgruppenzugehörigkeit in Deutschland:

Blutgruppe	0	A	B	AB
Rhesus +	35%	37%	9%	4%
Rhesus –	6%	6%	2%	1%

Informationskodierung

- Anwendungen der Informationstheorie finden sich bei der Kodierung von Werten oder Inhalten in Computern:
  - Die Informationsmenge in *bit* als *mathematische* Einheit bildet die Grundlage der *informatischen* Einheit *Bit*.
  - Ein Bit kodiert eine Ja-Nein-Antwort auf eine (beliebige) Entscheidungsfrage in einem binären (zweistelligen) Ziffernsystem: 'Bit' = '*Bi*nary Dig*it*' (vs. 'bit' = '*bi*nary un*it*'!).

Informationskodierung

- Die Kodierung eines Bits erfolgt technisch durch quantitativ unterschiedliche Ausprägungen einer Größe, die genau *zwei* Werte/Zustände annehmen kann:
  - Strom an vs. Strom aus (z. B. in einer Leiterbahn);
  - Widerstand hoch vs. Widerstand niedrig (z. B. in einem Transistor als Schaltelement);
  - Ladung vorhanden vs. Ladung nicht-vorhanden (z. B. auf einem Kondensator);
  - Magnetisierung plus vs. Magnetisierung minus (z. B. Magnetpartikel-Ausrichtung auf einer Festplatten-Scheibe);
  - Licht an/hell vs. Licht aus/dunkel (z. B. auf einer DVD-Oberfläche oder auf Papier [Strichkode]);
  - Loch/Erhebung vorhanden vs. Loch/Erhebung nicht-vorhanden (z. B. auf einer Lochkarte).

Informationskodierung

- Die qualitative Interpretation der quantitativen Größen erfolgt durch den Menschen oder die Maschine je nach Verwendungszweck:

Interpretation eines Bits als	Ausprägungen/Werte
Zahlenwert	0 oder 1
Wahrheitswert	wahr oder falsch (true/false)
Pixel	schwarz/dunkel oder weiß/hell, nicht-gesetzt oder gesetzt
Ton	aus oder an, hoch oder niedrig
Tastendruck/Touch-Input	vorhanden oder nicht-vorhanden

Informationskodierung

- Mehrere Bits können zu größeren Einheiten zusammengeschlossen werden, um mehr als zwei verschiedene Zustände kodieren zu können:
  - 2 Bits ($2 \cdot 2 = 2^2$): Darstellung der gleichen oder verschiedenen Geschlechter zweieiiger Zwillinge:
    - Fall 1: <männlich, männlich>;
    - Fall 2: <männlich, weiblich>;
    - Fall 3: <weiblich, männlich>;
    - Fall 4: <weiblich, weiblich>.
  - 4 Bits ($2 \cdot 2 \cdot 2 \cdot 2 = 2^4$): Darstellung eines der 16 Bundesländer;
  - 8 Bits ($2 \cdot 2 \cdot 2 \cdot 2 \cdot 2 \cdot 2 \cdot 2 \cdot 2 = 2^8$): Darstellung von Ganzzahlen im Bereich von 0 bis 255 oder −128 bis +127.

Informationskodierung

* Übersicht zur Unterscheidung bit vs. Bit:

Einheit	bit	Bit
Art der Einheit	mathematische Größeneinheit	informatische Kodierungseinheit
Zahlenwerte	ganzzahlig oder gebrochenzahlig	nur ganzzahlig (Aufrundung von X.Y bit auf X+1 Bits, wenn Y ≠ 0)
Realitätsstatus	abstrakt (nicht physikalisch realisierbar/realisiert)	konkret physikalisch darstellbare und verarbeitbare Einheit
Pluralisierbarkeit	nicht pluralisierbar: "~~bits~~"	pluralisierbar: "Bits"
Gruppierbarkeit	8 bit ≠ 1 Byte	8 Bits = 1 Byte (Daten)

Informationskodierung

- Die möglichst effiziente Kodierung der Möglich-keiten einer Auswahlmenge als Kette von Bits basiert auf folgenden Überlegungen:
  - Der mittlere Informationsgehalt eines 'Experiments' in bit sagt etwas darüber aus, wie viele Bits man zur Kodierung des Ergebnisses *theoretisch* benötigt.
  - Aufgrund der Tatsache, dass technisch keine gebrochenen (halben, viertelten usw.) Bits dargestellt werden können, kann eine *optimale* Kodierung nur angenähert werden.

Informationskodierung

- Ein einfaches Verfahren zur effizienten Kodierung liegt mit dem Huffman-Kode vor:
  - Jedem möglichen Zeichenereignis Z_k einer Auswahlmenge M wird ein Kodewort gemäß der Wahrscheinlichkeit des Eintretens von Z_k zugeordnet.
  - Die Grundidee ist, dass häufiger auftretende Ereignisse eine kürzere Kodierung erhalten als seltener auftretende.
  - Dies rechnet sich dann, wenn mehrere Ereignisse kodiert werden müssen (Daten-Komprimierung!).

Informationskodierung

- Ausgangspunkt zur Erzeugung eines Huffman-Kodes:
  - Die M verschiedenen Ereignisse mit ihren jeweiligen Wahrscheinlichkeiten p_1 bis p_M werden gemäß diesen Werten in aufsteigender Sortierung nebeneinander aufgeschrieben;
  - im Anschluss daran wird nach einem gegebenen Verfahren ein so genannter 'Kodebaum' generiert, an dem sich abschließend die Kodewörter für die einzelnen Zeichen ablesen lassen.

Informationskodierung

* Verfahrensvorschrift zur Huffman-Kodierung:

1) Wiederhole die folgenden Schritte solange, wie noch Ereignisse unkodiert sind:
   a) Wähle die zwei Zeichenereignisse X und Y mit den aktuell *niedrigsten* Wahrscheinlichkeiten (bei mehreren gleichen Möglichkeiten kann einer davon beliebig gewählt werden);
   b) addiere die Wahrscheinlichkeiten von X und Y zu Z, schreibe Z über die Mitte zwischen X und Y und verbinde die Werte von X und Y mit dem Wert Z;
   c) streiche X und Y aus der Liste der verwendbaren Werte und verwende dafür jetzt Z.
2) Gehe zu 1).

Informationskodierung

- Beispiel:
  - Konstruktion des Kodebaums für einen gezinkten Würfel mit folgenden Wahrscheinlichkeiten für die einzelnen Augenzahlen (Zeichenereignisse):

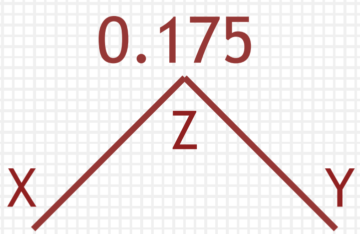
Auge	1	2	3	4	5	6
$p_{1/2/3/4/5/6}$	0.200	0.050	0.125	0.125	0.150	0.350

  - Nach Wahrscheinlichkeiten p_k aufsteigend sortierte Augenzahlen (Zeichenereignisse) Z_k:

p_k:	0.050	0.125	0.125	0.150	0.200	0.350
Z_k:	2	3	4	5	1	6

Informationskodierung

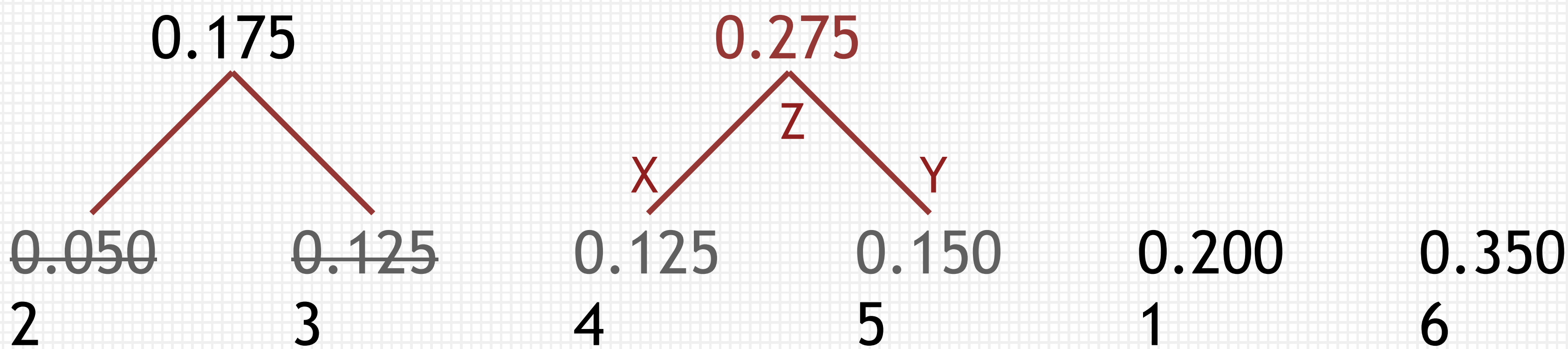
- Konstruktion Kodebaum Durchgang 1:



0.050	0.125	0.125	0.150	0.200	0.350
2	3	4	5	1	6

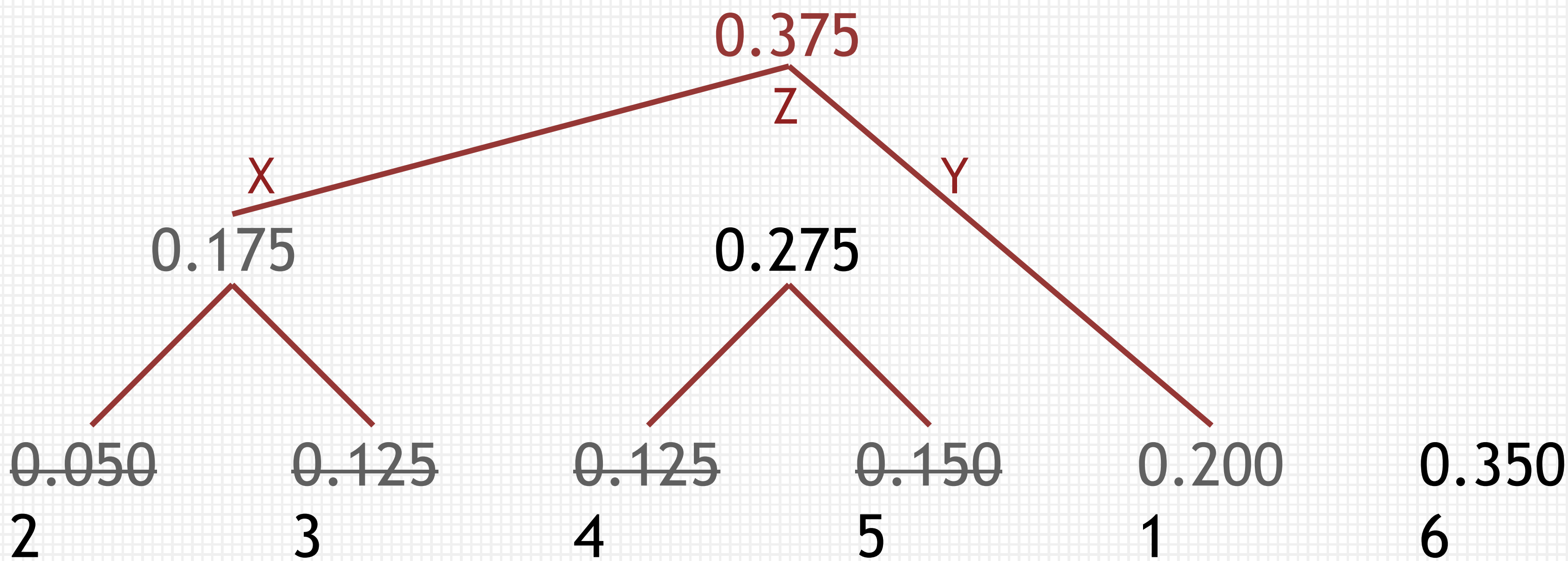
Informationskodierung

- Konstruktion Kodebaum Durchgang 2:



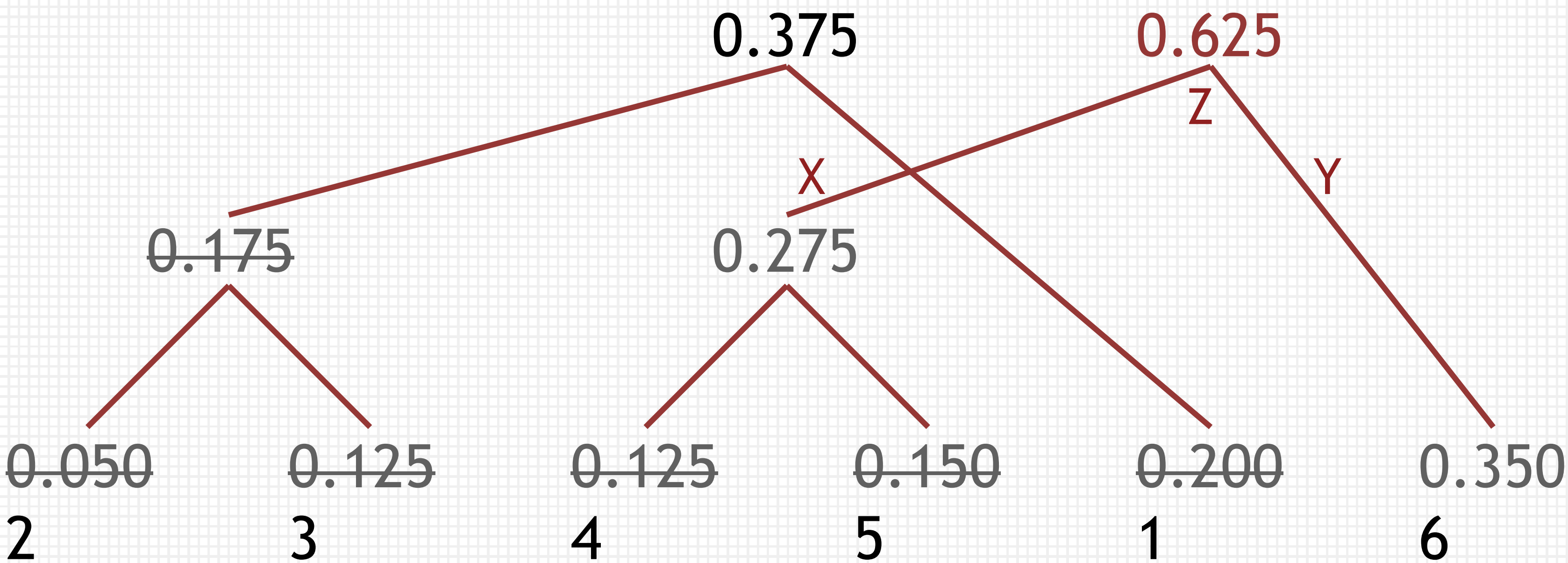
Informationskodierung

- Konstruktion Kodebaum Durchgang 3:



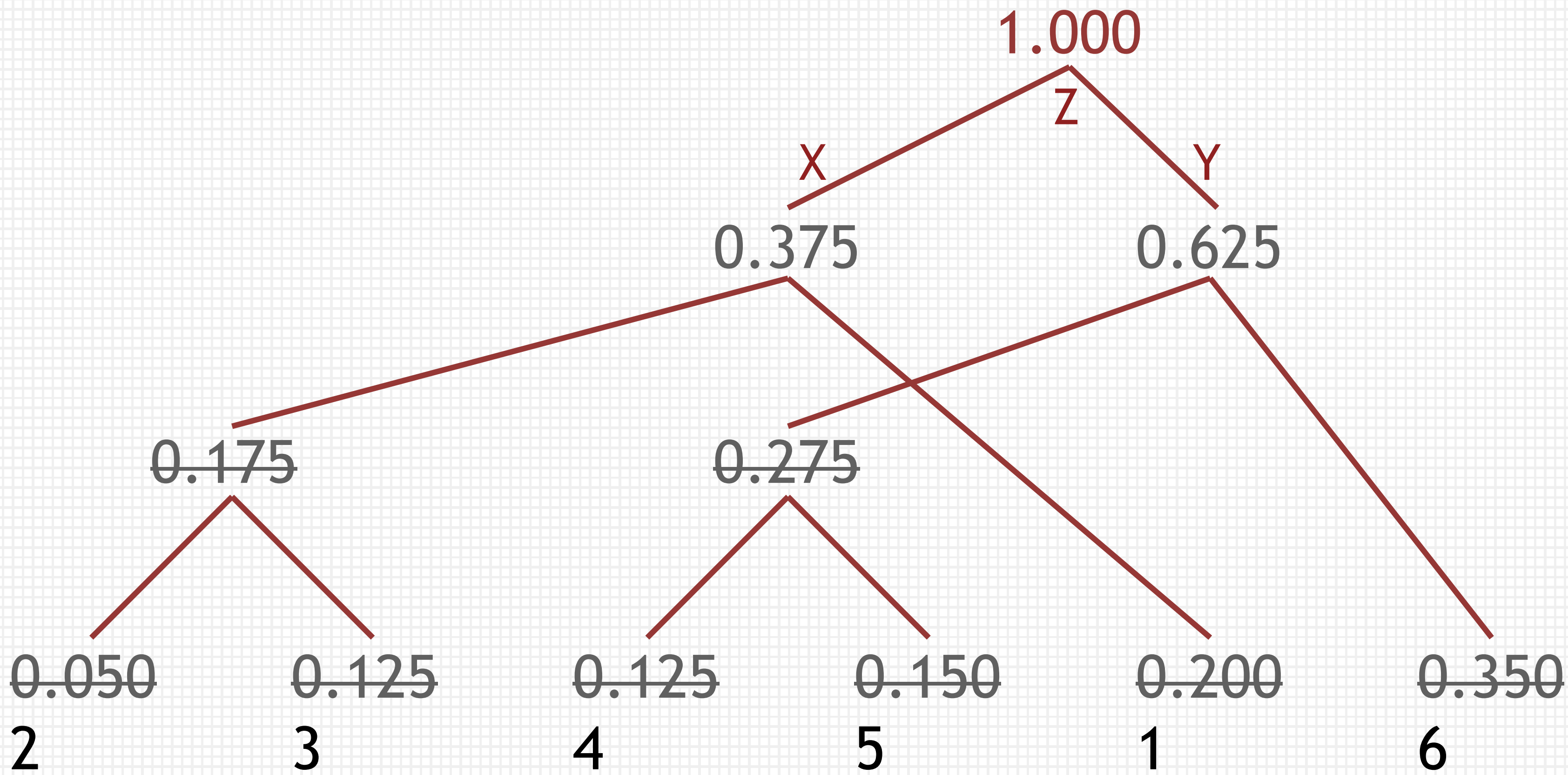
Informationskodierung

- Konstruktion Kodebaum Durchgang 4:



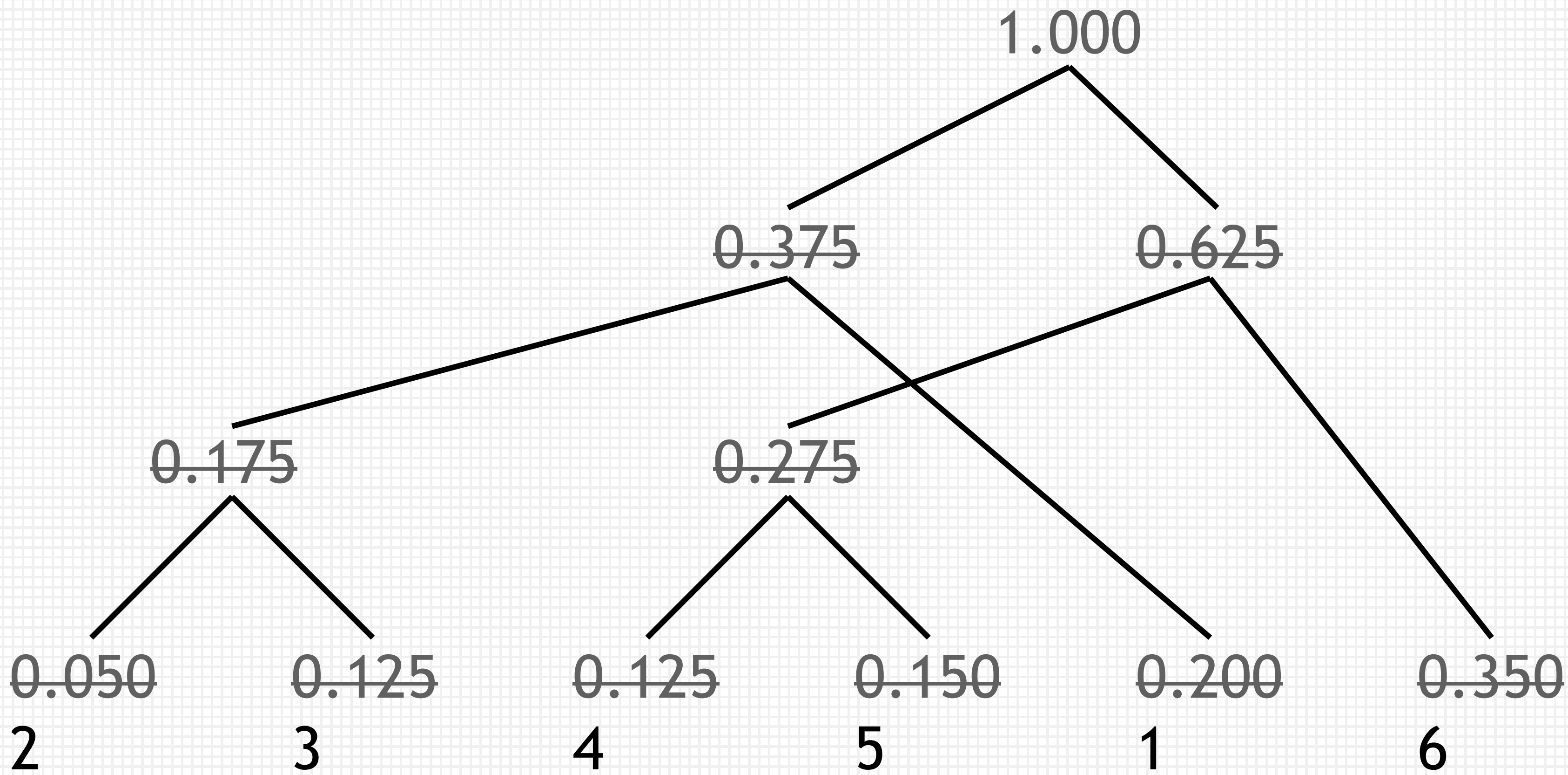
Informationskodierung

- Konstruktion Kodebaum Durchgang 5:



Informationskodierung

- Konstruktion Kodebaum Ergebnis:



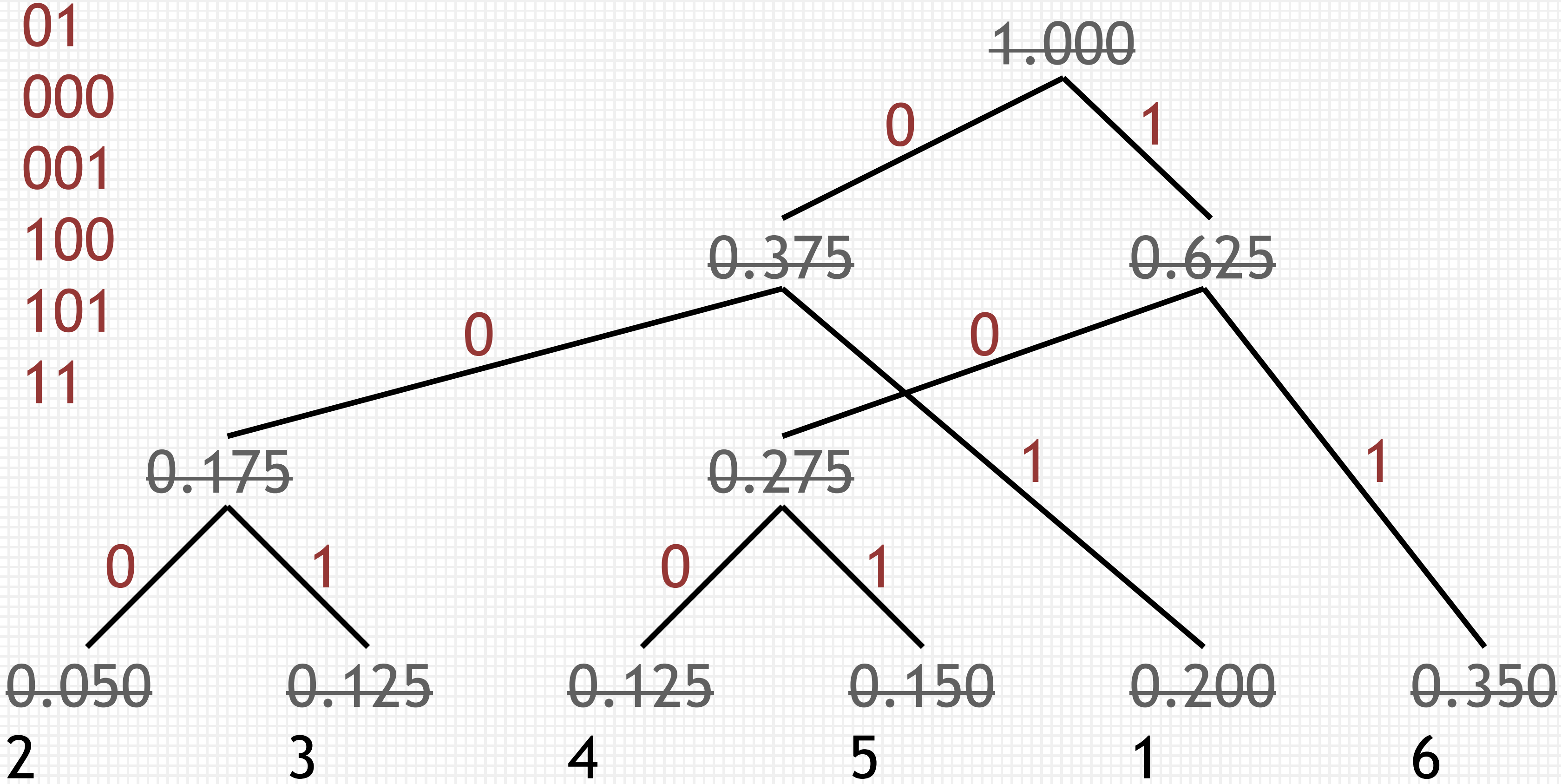
Informationskodierung

- Nach der vollständigen Konstruktion des Kodebaums muss noch der Binärkode für jedes Zeichenereignis ermittelt werden:
  - Von der Wurzel oben ausgehend wird jeweils allen linken Verzweigungen eine 0, allen rechten eine 1 (oder umgekehrt) solange zugeordnet, bis man jeweils auf der untersten Ebene der Zeichen angelangt ist.
  - Der Kode für jedes Zeichenereignis ergibt sich, wenn man von der Wurzel ausgehend alle Nullen und Einsen nacheinander aufschreibt, bis man beim entsprechenden Ereignis Z_K angelangt ist.

Informationskodierung

- Kodierung:

1: 01
2: 000
3: 001
4: 100
5: 101
6: 11



Informationskodierung

- Beobachtung:
  - Zeichen mit höherer Auftretenswahrscheinlichkeit erhalten kürzere Kodes (hier 1 und 6).
  - Kein kürzeres Kodewort ist der Anfang eines längeren Kodeworts, d. h. alle Kodes sind von Beginn an eindeutig (analog Telefonnummern):
    - 0100000110010111 kodiert eindeutig für 123456;
    - 1110110000100001 kodiert eindeutig für 654321.
  - Die Wahl von '0' für links und '1' für rechts ist willkür-lich und kann prinzipiell auch umgekehrt werden.

Übungsaufgaben

- Konstruieren Sie jeweils einen Kodebaum für einen fairen Tetraeder, Hexaeder, Oktaeder und Dodekaeder.

  Tetraeder: Hexaeder:

Übungsaufgaben

Oktaeder:

Übungsaufgaben

Dodekaeder:

Übungsaufgaben

- Frage und Antwort:

Frage	Antwort
Wie lange ist der Kode eines Zeichens Z aus einer Auswahlmenge von $M \geq 2$ Zeichen, wenn Z eine Auftretenswahrscheinlichkeit von $p_1 > 0.5$ besitzt (d. h. alle anderen Zeichen zusammen besitzen eine Wahrscheinlichkeit von $p_{2-M} < 0.5$)?	
Warum beginnt man bei der Konstruktion eines Huffman-Kodebaums bei den Zeichen mit den geringsten Wahrscheinlichkeiten? Was würde passieren, wenn man mit den wahrscheinlichsten Zeichen beginnen würde?	

Übungsaufgaben

- Folgender Huffman-Kode ist gegeben:

Zeichen	A	B	D	E	K	M	O	U
Kode	00	01	100	101	1100	1101	1110	1111

  - Dekodieren Sie folgende Bitkette in Klartext:

Kode	1 1 0 0 1 1 1 0 1 0 0 1 0 1 0 1 0 0 1 1 1 1 1 1 0 1
Klartext	

  - Enkodieren Sie folgendes Wort als Bitkette:

Kode								
Klartext	B	A	D	E	M	O	D	E

Übungsaufgaben

- Konstruieren Sie einen entsprechenden Kodebaum für den oben gegebenen Huffman-Kode:

Übungsaufgaben

- Erzeugen Sie jeweils einen Huffman-Kode für folgende Zeichenketten:
  - "MISSISSIPPI":

Übungsaufgaben

- "ERSCHEINUNGSFORM" und "MEINUNGSFORSCHER":

Übungsaufgaben

- Ein Farbbild aus acht Farben weist die rechts stehende Verteilung an Farben/Farbpixeln auf:
  - Erzeugen Sie einen Huffman-Kode für die acht Bildfarben, d. h. jeder Bildfarbe wird letztlich eine Bitkette (via Kode-baum) zugeordnet, die ein Pixel der entsprechenden Farbe kodiert.

Farbe	Anteil	Kodebits
Schwarz	30%	
Weiß	25%	
Rot	10%	
Grün	10%	
Blau	10%	
Gelb	5%	
Cyan	5%	
Magenta	5%	

Übungsaufgaben

'Nebenrechnung': Huffman-Kodebaum

Übungsaufgaben

- Berechnen Sie die Gesamtkodelänge in Bits für ein Bild mit 1024×768 Pixeln unter zwei Bedingungen:

Szenario 1	Szenario 2
wenn für jede Pixelfarbe eine Gleichverteilung unterstellt würde (12.5% für jede Farbe)	wenn jede Pixelfarbe wie in der oben stehenden Tabelle verteilt ist (Ungleichverteilung)

Übungsaufgaben

- Berechnen Sie jeweils den Informationsgehalt in bit für die beiden obigen Szenarien (jeweils acht Wahrscheinlichkeiten) und multiplizieren Sie jeweils mit 1024•768:

Szenario 1	Szenario 2

Informationsmengen

- Die Anzahl Bits, die ein einzelnes (Zeichen-) Ereignis zu seiner Kodierung benötigt, stellt eine bestimmte *Menge an Information* dar:
  - Da meist mehrere Zeichen/Ereignisse zu kodieren sind (z. B. 1000 Würfelwürfe, 1024×768 Pixel mit 8 Farben), ergeben sich meist längere Ketten von Bits.
  - Die Menge an Information, die durch die Bitketten jeweils dargestellt wird, wird zu bestimmten Informationsmengen mit entsprechenden Bezeichnungen zusammengefasst.

Informationsmengen

- Gängige Bezeichnungen für Informationsmengen:

Bezeichnung	Bits	Bemerkung
Bit ('Binary Digit')	1	"a bit" = "ein Bisschen" ('ein kleiner Bissen')
Dibit	2	Doppelbit
Tribit	3	Dreifachbit
Quadbit	4	Vierfachbit, Halbbyte
Byte	8	"a bite" = "ein Biss(en)"
Wort	16	Achtung: Je nach Rechnerarchitektur wird der Begriff 'Wort' auch für 32-bittige Einheiten verwendet; entsprechend ist ein Doppelwort dann 64 Bit, ein Quadwort 128 Bit usw.
Doppelwort	32	
Quadwort	64	
Oktawort	128	

Informationsmengen

- Anzahl jeweils darstellbarer verschiedener Zustände:

Bezeichnung	Bits	Anzahl Zustände
Bit	1	2
Dibit	2	4
Tribit	3	8
Quadbit	4	16
Byte	8	256
Wort	16	65 536
Doppelwort	32	4 294 967 296
Quadwort	64	18 446 744 073 709 551 616
Oktawort	128	$3.40282366920938463463374607431770 \cdot 10^{38}$

Informationsmengen

- Noch größere Gruppen von Bits und Bytes werden durch Vorsätze für Größenordnungen ausgedrückt:

Vorsatz	Größe	Beispiel
Kilo (K)	2^{10} ($\approx 10^{3}$)	1 KByte/KB = 1 024 Bytes = 8 192 Bits
Mega (M)	2^{20} ($\approx 10^{6}$)	1 MByte/MB = 1 048 576 Bytes = 8 388 608 Bits
Giga (G)	2^{30} ($\approx 10^{9}$)	1 GByte/GB = 1 073 741 824 Bytes
Tera (T)	2^{40} ($\approx 10^{12}$)	1 TByte/TB = 1 099 511 627 776 Bytes
Peta (P)	2^{50} ($\approx 10^{15}$)	1 PByte/PB = 1 125 899 906 842 624 Bytes
Exa (E)	2^{60} ($\approx 10^{18}$)	1 EByte/EB = 1 152 921 504 606 846 976 Bytes
Zetta (Z)	2^{70} ($\approx 10^{21}$)	1 ZByte/ZB = 1 180 591 620 717 411 303 424 B.
Yotta (Y)	2^{80} ($\approx 10^{24}$)	1 YByte/YB = 1 208 925 819 614 629 174 706 176

Informationsmengen

- Solche Maßzahlen werden bei der Bemessung von Speichergrößen verwendet, wobei in zwei unterschiedlichen Zahlensystemen gemessen wird:
  - *Basis 2*: Größenangaben in KB, MB, GB, TB usw. werden als Größenordnung $2^{10/20/30/40/\text{usw.}}$ verstanden (üblicher in der Informationstechnik);
  - *Basis 10*: Größenangaben in KB, MB, GB, TB usw. werden als Größenordnung $10^{3/6/9/12/\text{usw.}}$ verstanden (unüblicher in der Informationstechnik).

Informationsmengen

- Beispiele zur Bemessung von Informations-mengen für die Basen 2 und 10:
  - Bemessung zur Basis 2 (d. h. alle Größen berechnen sich zu 2^N):
    - Speicherkomponenten, bemessen in *Bytes*:
      - Flash-Speicher: 1 – 1024 GB (USB-Sticks, SSDs);
      - Hauptspeicher: 1 – 64 GB (DIMMs, SIMMs);
      - Prozessor-Cache: 1 – 16 MB (innerhalb des Prozessors).
    - Speicherbausteine, bemessen in *Bits*:
      - Flash-Speicherchip: 128 – 256 GBits (16 – 32 GBytes);
      - Hauptspeicher-Speicherchip (DRAM): 2 – 4 GBits (0.25 – 0.5 GBytes).

Informationsmengen

- Bemessung zur Basis 10 (Größenberechnung zu 10^N):
  - Scheibenspeicher in Bytes:
    - Diskette: 1.44 MB (ca. $1.44 \cdot 10^6$ = 1 440 000 Bytes);
    - CD ('Compact Disk'): 700 MB (ca. $700 \cdot 10^6$ = 700 000 000 Bytes);
    - DVD ('Digital Versatile Disk'): 4.7 – 17 GB (ca. $4.7 \cdot 10^9$ – $17 \cdot 10^9$ Bytes);
    - BD ('Blu-ray Disk'): 25 – 50 GB (ca. $25 \cdot 10^9$ – $50 \cdot 10^9$ Bytes);
    - HDD ('Hard Disk Drive' = Festplatte): 250 GB – 4 TB ($250 \cdot 10^9$ – $4 \cdot 10^{12}$).
  - Bandlaufwerke in Bytes:
    - Audio-Kassetten (Datasette): ca. 100 KB pro 30 Minuten-Seite;
    - Cartridge: ca. 500 GB.

Übungsaufgaben

- Berechnen Sie, wie viele Bytes die Informationsmengen 1 GB (1 GByte) bzw. 1 TB (1 TByte) aufweisen:
  - wenn die Informationsmenge zur Basis 2 ermittelt wird;
  - wenn die Informationsmenge zur Basis 10 ermittelt wird.

Informations-menge in Bytes	Informationsmenge in Bytes zur Basis 2	Informationsmenge in Bytes zur Basis 10
1 GB (Gigabyte)		
1 TB (Terabyte)		

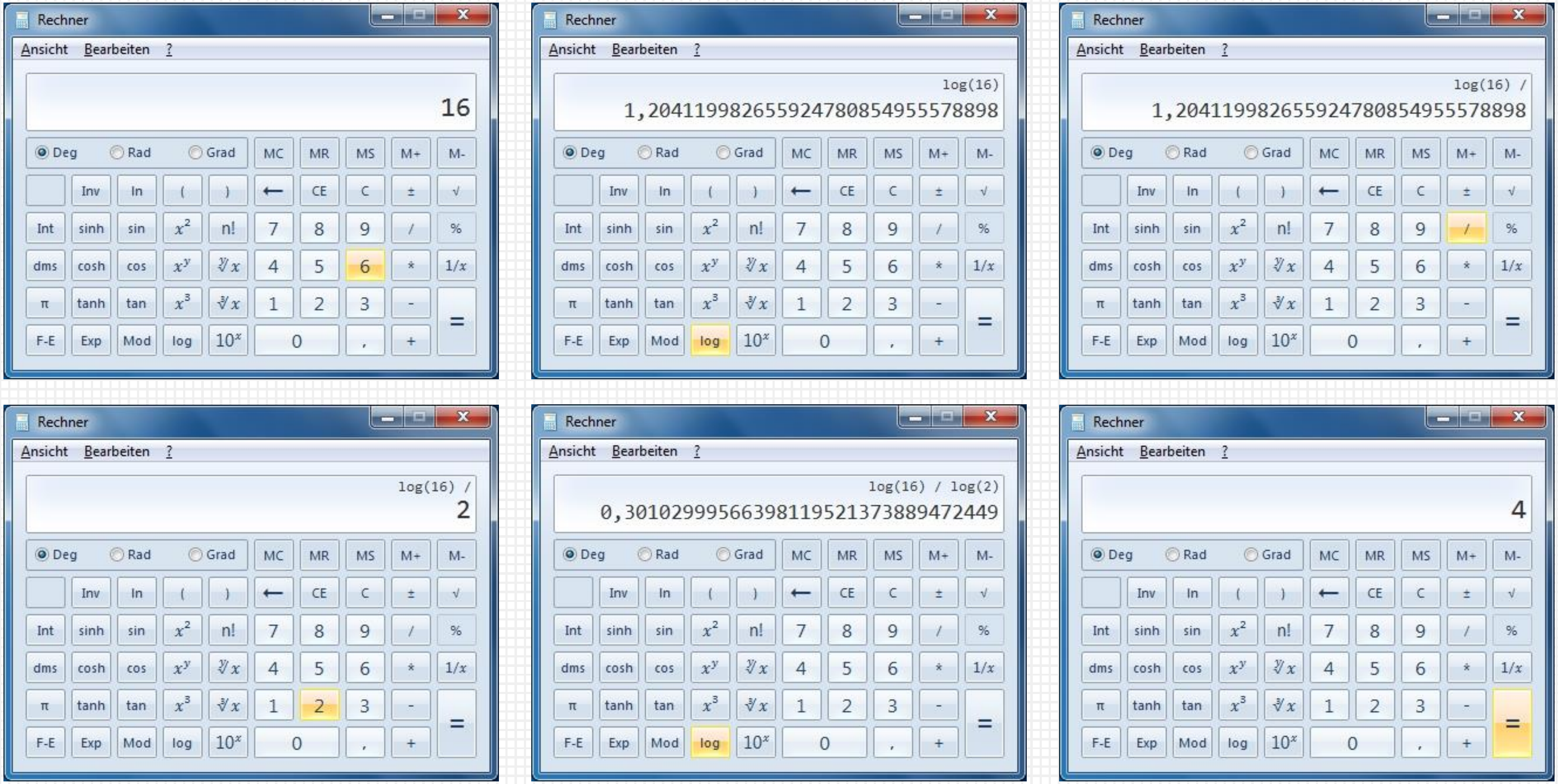
Warum werden Speichergrößen bei Festplatten und anderen Speichermedien gerne in GB/TB zur Basis 10 statt 2 angegeben?

Anhang

- Logarithmus:
  - *Beschränkungen*:
    - $x = \log_b a$ (Auflösung nach x für $b^x = a \Leftrightarrow b^{\log_b a} = a$);
    - $a \in \mathbb{R}^+$, $b \in \mathbb{R}^+\backslash\{1\}$ ($\mathbb{R}$ = Menge der reellen Zahlen)
      (d. h. a darf nicht 0 sein, b darf nicht 0 oder 1 sein).
  - *Basisumrechnung*:
    - Allgemein: $\log_b a = \log_d a / \log_d b$ $(b, d \in \mathbb{R}^+\backslash\{1\})$;
    - Speziell: $\mathrm{ld}\ a = \log_e a / \log_e b = \ln a / \ln b$ $(e \approx 2.7183)$;
      $\mathrm{ld}\ a = \log_{10} a / \log_{10} b = \lg a / \lg b$.
  - *Rechengesetze*:
    - $\log_b c/a = \log_b c - \log_b a$; $\log_b 1/a = -\log_b a$ (da $\log_b 1 = 0$);
    - $\log_b c \cdot a = \log_b c + \log_b a$; $\log_b a^c = c \cdot \log_b a$.

Anhang

- Beispiel: Berechnung von ld 16 = lg 16 / lg 2 = 4 mit Win-TR:



Literatur

- [Attneave 1974] Attneave, F. (³1974): *Informationstheorie in der Psychologie*. Bern u. a.: Hans Huber.
- [Herold & al. 2007] Herold, H. & Lurz, B. & Wohlrab, J. (2007): *Grundlagen der Informatik*. München u. a.: Pearson.
- [Rechenberg 2003] Rechenberg, P. (2003): Zum Informationsbegriff der Informationstheorie. *Informatik Spektrum*, 26(5), S. 317–326.
- [Rechenberg & Pomberger 2006] Rechenberg, P. & Pomberger, G. (⁴2006; Hrsg.): *Informatik Handbuch*. München & Wien: Hanser.
- [Reischer 2006] Reischer, J. (2006): *Zeichen, Information, Kommunikation. Analyse und Synthese des Zeichen- und Informationsbegriffs*. http://epub.uni-regensburg.de/10483/.
- [Shannon & Weaver 1949] Shannon, C. E. & Weaver, W. (1949): *The Mathematical Theory of Communication*. Urbana u. a.: The University of Illinois Press.