Was ist Information?

https://iuk.one/1010-1005

Clemens H. **Cap**
ORCID: 0000-0003-3958-6136

Department of Computer Science
University of **Rostock**
Rostock,Germany
clemens.cap@uni-rostock.de

Version 5

Ziele

Informatik gilt auch als **Wissenschaft von der Information.**

Was ist Information **genau**?

Inhaltsübersicht

Information

Information ist ein Konzept zur Auflösung von Unsicherheit

Information repräsentiert die **Auswahl des Konkreten aus dem Möglichen.**

Information ist ein **Mechanismus, um Dinge zu konstruieren.**

Weitere mögliche Sichtweisen:

1. Ein Maß für
2. Ein Maß für empirische Häufigkeit.
3. Ein Maß für die Bereitschaft, eine Wette in bestimmter Höhe abzuschließen.

1. Wahrscheinlichkeit

Ziele: Wahrscheinlichkeit ist die Grundlage des Informationsbegriffs nach Shannon. Wir wiederholen die Grundprinzipien der Wahrscheinlichkeit und wenden sie auf einfache Beispiele an.

Wahrscheinlichkeit

Ein **Experiment** hat einen präzise beschriebenen Ablauf.

Ein **Elementarereignis** ist das kleinste, eigenständig mögliche Ergebnis eines Experiments.

Ein **Ereignis** ist eine Menge von Elementarereignissen.

Beispiele für Wahrscheinlichkeits-Experimente

1

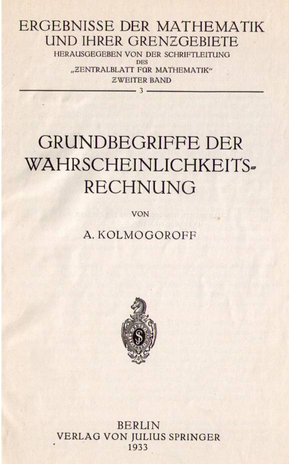
Beispiel: Münzwurf

Experiment: *Einmal* mit *einer* Münze werfen und das Ergebnis notieren.
Menge der Elementarereignisse: $M = \{\texttt{Kopf}, \texttt{Zahl}\}$ (insgesamt 2)
Bsp. für ein Elementarereignis: $\texttt{Kopf}$
Menge aller Ereignisse: $2^M = \{\ \{\texttt{Kopf}, \texttt{Zahl}\}\ ,\ \{\texttt{Kopf}\}\ ,\ \{\texttt{Zahl}\}\ ,\ \emptyset\ \}$ (insgesamt 4)

Beispiel: Würfel

Experiment: *Einmal* mit *einem* Würfel würfeln und das Ergebnis notieren.
Menge der Elementarereignisse: $M = \{$⚀, ⚁, ⚂, ⚃, ⚄, ⚅$\}$
Bsp. für ein Elementarereignis: ⚂
Bsp. für ein Ereignis: "Ungerade Augenzahl", als Menge $U = \{$⚀, ⚂, ⚄$\}$.

Kolmogoroff



ERGEBNISSE DER MATHEMATIK
UND IHRER GRENZGEBIETE
HERAUSGEGEBEN VON DER SCHRIFTLEITUNG
DES
„ZENTRALBLATT FÜR MATHEMATIK“
ZWEITER BAND
3

GRUNDBEGRIFFE DER
WAHRSCHEINLICHKEITS-
RECHNUNG

VON

A. KOLMOGOROFF

BERLIN
VERLAG VON JULIUS SPRINGER
1933

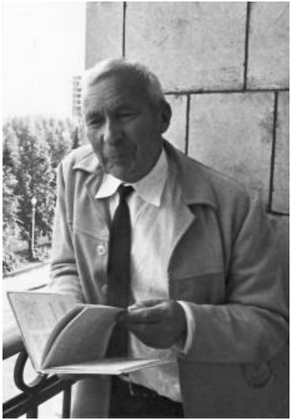


Abb. 1: Andrej **Kolmogoroff** (1942–1987), russischer Mathematiker, Erfinder der axiomatischen Wahrscheinlichkeitsrechnung, neben seinem Werk. Rechte s. Anhang.

4 Axiome von Kolmogoroff

1

Sei M eine Menge von Elementarereignissen und bezeichne 2^M die Potenzmenge von M.

Eine **Wahrscheinlichkeit** ist eine Funktion $P\colon 2^M \to [0,1]$ für die gilt:

1. Wahrscheinlichkeit des **sicheren Ereignisses** ist 1: $P(M) = 1$
2. Wahrscheinlichkeit des **unmöglichen Ereignisses** ist 0: $P(\emptyset) = 0$
3. Wahrscheinlichkeit des **Nichteintretens** ist 1 – Wahrscheinlichkeit des Eintretens. $P(\complement A) = 1 - P(A)$
4. Wird ein Ereignis aus **Teilereignissen** so zusammengesetzt, daß je zwei den leeren Durchschnitt haben, so ergibt sich die Wahrscheinlichkeit des Ereignisses als Summe der Wahrscheinlichkeiten der Teilereignisse.

$$\text{Wenn} \quad A_1, A_2, \ldots, A_n \subseteq M \text{ mit } \forall i,j : i \neq j \Rightarrow A_i \cap A_j = \emptyset \quad \text{dann}$$

$$P(A_1 \cup A_2 \cup \ldots \cup A_n) = P(A_1) + P(A_2) + \ldots + P(A_n)$$

Beispiele: 4. Axiom von Kolmogoroff (1)

Frage: Was besagt das 4. Axiom am Würfelexperiment mit den Elementarereignissen $M = \{\epsdice{1}, \epsdice{2}, \epsdice{3}, \epsdice{4}, \epsdice{5}, \epsdice{6}\}$.

Bsp 1: $A_1 = \{\epsdice{1}, \epsdice{2}, \epsdice{3}\}$ und $A_2 = \{\epsdice{4}, \epsdice{5}, \epsdice{6}\}$. Voraussetzungen erfüllt? Ja!
Also: $P(\{\epsdice{1}, \epsdice{2}, \epsdice{3}, \epsdice{4}, \epsdice{5}, \epsdice{6}\}) = P(\{\epsdice{1}, \epsdice{2}, \epsdice{3}\}) + P(\{\epsdice{4}, \epsdice{5}, \epsdice{6}\})$.

Bsp 2: $A_1 = \{\epsdice{1}, \epsdice{2}\}$ und $A_2 = \{\epsdice{3}\}$. Voraussetzungen erfüllt? Ja!
Also: $P(\{\epsdice{1}, \epsdice{2}, \epsdice{3}\}) = P(\{\epsdice{1}, \epsdice{2}\}) + P(\{\epsdice{3}\})$.

Bsp 3: $A_1 = \{\}$ und $A_2 = \{\epsdice{3}\}$. Voraussetzungen erfüllt? Ja!
Also: $P(\{\epsdice{3}\}) = P(\{\}) + P(\{\epsdice{3}\})$.
Dann muß auch $P(\{\}) = 0$ sein?!
Das ist nach dem 2. Axiom der Fall!

Beispiele: 4. Axiom von Kolmogoroff (2)

Bsp. 4: $A_1 = \{⚀\}$ und $A_2 = \{⚀, ⚁\}$. Voraussetzungen nicht erfüllt, da $A_1 \cap A_2 = \{⚀\}$. Also ist auch nichts weiter zu tun.

Bsp. 5: $A_1 = \{⚀, ⚁\}$ $A_2 = \{⚂, ⚃\}$ $A_3 = \{⚄, ⚅\}$. Voraussetzungen erfüllt?
Hier ist mehr zu prüfen: $A_1 \cap A2 = \{\}$, $A_2 \cap A_3 = \{\}$ und $A_1 \cap A_3 = \{\}$. Ja!
Also: $P(\{⚀, ⚁, ⚂, ⚃, ⚄, ⚅\}) = P(\{⚀, ⚁\}) + P(\{⚂, ⚃\}) + P(\{⚄, ⚅\})$

Bsp. 6: $A_1 = \{⚀\}$ und $A_2 = \{⚁\}$. Voraussetzungen erfüllt? Ja.
Also: $P(\{⚀, ⚁\}) = P(\{⚀\}) + P(\{⚁\})$

Bsp. 7: $A_1 = \{⚀\}$ und $A_2 = \{⚁\}$ und $A_3 = \{⚅\}$. Voraussetzungen erfüllt? Ja.
Also: $P(\{⚀, ⚁, ⚅\}) = P(\{⚀\}) + P(\{⚁\}) + P(\{⚅\})$

Bedeutung des 4. Axioms von Kolmogoroff

2

Satz über endliche Wahrscheinlichkeiten

Die Wahrscheinlichkeit eines Ereignisses $E = \{e_1, e_2, \ldots e_k\}$, das die *endlich vielen* **Elementar**ereignisse $e_1, e_2, \ldots, e_k$ umfaßt, ist die **Summe** der Wahrscheinlichkeiten aller dieser **Elementar**ereignisse.

$$P(\{e_1, e_2, \ldots, e_k\}) = P(\{e_1\}) + P(\{e_2\}) + \ldots + P(\{e_k\}) = \sum_{j=1}^{k} P(\{e_j\})$$

Folge: Bei Experimenten mit endlich vielen möglichen Ergebnissen ist die Wahrscheinlichkeit aller Ereignisse bereits durch die Wahrscheinlichkeiten der **Elementarereignisse eindeutig** festgelegt.

Beschreibung eines Experiments

Gegeben: Ein Würfel mit Aufschrift: ⚀ ⚁ ⚂ ⚃ ⚄ ⚅.

Gesucht: Wie groß ist die Wahrscheinlichkeit, eine ⚀ zu werfen?

Formal: Menge der Elementarereignisse ist $M = \{$⚀, ⚁, ⚂, ⚃, ⚄, ⚅$\}$
Gesuchtes Ereignis ist $E = \{$⚀$\}$.
Berechne $P(E)$.

Vorgehen: Finde eine Funktion $P : P(M) \to [0, 1]$,
welche die Axiome von Kolmogoroff erfüllt.
Bilde dann $P(E)$.

Wissen: Es genügt, die Elementarereignisse zu betrachten !

1. Lösung

$$P(\{\text{⚀}\}) = 1$$

$$P(\{\text{⚁}\}) = 0$$

$$P(\{\text{⚂}\}) = 0$$

$$P(\{\text{⚃}\}) = 0$$

$$P(\{\text{⚄}\}) = 0$$

$$P(\{\text{⚅}\}) = 0$$

Interpretation:	Würfel, der immer auf der ⚀ landet.
Antwort:	$P(\{\text{⚀}\}) = 1$
Erklärung:	Gegenseite mit Blei beschwert.

2. Lösung

$$P(\{\epsdice{1}\}) = \frac{1}{2}$$

$$P(\{\epsdice{2}\}) = 0$$

$$P(\{\epsdice{3}\}) = 0$$

$$P(\{\epsdice{4}\}) = 0$$

$$P(\{\epsdice{5}\}) = 0$$

$$P(\{\epsdice{6}\}) = \frac{1}{2}$$

Interpretation: Würfel, der immer auf der $\epsdice{1}$ oder $\epsdice{6}$ landet.

Antwort: $P(\{\epsdice{1}\}) = \frac{1}{2}$

Erklärung: Achse $\epsdice{1}$—$\epsdice{6}$ mit Blei beschwert.

3. Lösung

$$P(\{\epsdice{1}\}) = \frac{1}{6}$$

$$P(\{\epsdice{2}\}) = \frac{1}{6}$$

$$P(\{\epsdice{3}\}) = \frac{1}{6}$$

$$P(\{\epsdice{4}\}) = \frac{1}{6}$$

$$P(\{\epsdice{5}\}) = \frac{1}{6}$$

$$P(\{\epsdice{6}\}) = \frac{1}{6}$$

Interpretation: Fairer Würfel.

Antwort: $P(\{\epsdice{1}\}) = \frac{1}{6}$

Erklärung: Keine besondere.

Analyse und Experiment

Alle diese Lösungen (und weitere) sind korrekt und in der Würfelpraxis möglich!

Wenn wir nur **ein** Ergebnis haben wollen, so ist eine zusätzliche Annahme erforderlich!

Unser **Bauchgefühl** sagt uns: $P(\{⚀\}) = \frac{1}{6} = 0,1666\ldots = 16,666\ldots\%$

Wir machen **Experimente:**

1. Experiment mit 10 Würfen ergibt: 20% aller Würfe ergeben ⚀.
2. Experiment mit 10 Würfen ergibt: 0% aller Würfe ergeben⚀.
3. Experiment mit 100 Würfen ergibt: 15% aller Würfe ergeben ⚀.
4. …

Experimente bedürfen der Interpretation und können Schwankungen aufweisen.

Zusätzliche Annahme

Wenn wir **annehmen**, daß der Würfel exakt symmetrisch gebaut ist,
dann ist die **Schlußfolgerung** sinnvoll, daß alle Seiten gleich wahrscheinlich sind.

$$P(\{\epsdice{1}\}) = P(\{\epsdice{2}\}) = P(\{\epsdice{3}\}) = P(\{\epsdice{4}\}) = P(\{\epsdice{5}\}) = P(\{\epsdice{6}\}) := \lambda$$

Wir wissen:

$$P(\{\epsdice{1}, \epsdice{2}, \epsdice{3}, \epsdice{4}, \epsdice{5}, \epsdice{6}\}) = 1$$

Wir wissen ferner:

$$\begin{aligned} P(\{\epsdice{1}, \epsdice{2}, \epsdice{3}, \epsdice{4}, \epsdice{5}, \epsdice{6}\}) &= \text{ 4. Axiom von Kolmogoroff} \\ P(\{\epsdice{1}\}) + P(\{\epsdice{2}\}) + P(\{\epsdice{3}\}) + P(\{\epsdice{4}\}) + P(\{\epsdice{5}\}) + P(\{\epsdice{6}\}) &= 6 \cdot \lambda \end{aligned}$$

Also $6 \cdot \lambda = 1$ und somit $\lambda = \frac{1}{6}$ und daher **Lösung:** $P(\{\epsdice{1}) = \frac{1}{6}$.

Beobachtung: Erst die Zusatzannahme ergab eine eindeutige Lösung.

Laplacesches Indifferenzprinzip

3

Prinzip: Wenn wir eine Situation mit genau n verschiedenen Elementarereignissen betrachten und keinen Grund haben, anzunehmen, daß ein bestimmtes davon wahrscheinlicher ist als ein anderes, dann *nehmen wir an*, daß sie *alle gleichwahrscheinlich* sind.

Beachte: Das ist eine sinnvolle Annahme – aber *keine* Aussage, die aus Axiomen folgt.

Anwendung: Sei n die Anzahl aller möglichen Ereignisse. Dann ist die Wahrscheinlichkeit $P(E)$ eines **Elementarereignisses** E

$$P(E) = \frac{1}{n}$$

Anwendung: Sei nun E ein Ereignis **bestehend aus** k **Elementarereignissen**. Dann ist seine Wahrscheinlichkeit:

$$P(E) = \frac{k}{n}$$ sogenannte **Laplacesche Formel**

Abb. 2: Pierre-Simon **Laplace** (1749-1827), französischer Mathematiker, Physiker und Astronom, der auch etliche Beiträger zur Wahrscheinlichkeitstheorie geliefert hat. Rechte s. Anhang.

Wissenschaftliche Methodik (1)

Wissenschaft arbeitet mit einer Kombination von

1. **Intuition und Bauchgefühl** als Ideenlieferant.
   Achtung: Das kann helfen oder in furchtbare Sackgassen führen.
   Und: Viele "seriöse" Wissenschaftler werden das nur "ungern" zugeben.
2. **Kritischen rationalen Überlegungen**, in der Form von Theorien und Modellen.
3. **Praktischen Beobachtungen**, in der Form von Experimenten.

"Halbgare" Vorstellung:

1. Verändere die Theorie, bis ihre Prognosen "halbwegs" zum Experiment "passen".
2. Behaupte, "die richtige" Theorie gefunden zu haben und kassiere Applaus.

Wissenschaftliche Methodik (2)

4
5

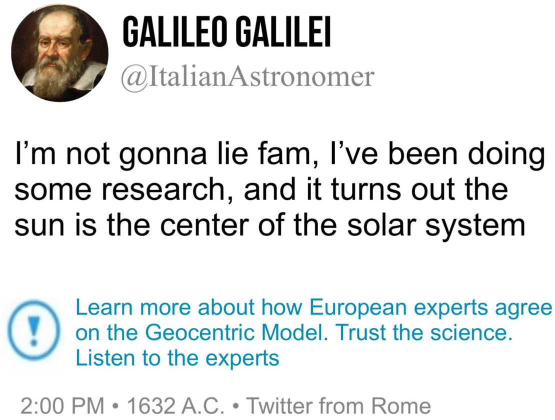
Problem:

1. Es kann mehrere, unterschiedliche Theorien geben, die ~~stimmen~~"halbwegs passen".
2. Diese können sogar formal zueinander im Widerspruch stehen!

Daher bessere Vorstellung:
Wissenschaft sucht prognostisch gut geeignete Theorien
ohne die Korrektheit einer bestimmten Theorie zu behaupten
und untersucht unvoreingenommen alle Gründe,
warum, wo und wie welche Theorien an ihre Grenzen stoßen.

Ingenieurwissenschaften wenden diese Erkenntnisse auf praktische Probleme an.

Abb. 3: Eines der Grundprobleme wissenschaftlicher Methodik illustriert an einem historischen Beispiel, humorvoll eingekleidet und bezogen auf heutige Formen sogenannten "Fact Checking" und sogenannter "Richtigstellungen" auf "sozialen" Medien. Rechte s. Anhang.



2. Information (nach Shannon)

Ziel: Wir lernen den Informationsbegriff von Shannon kennen, der eine zentrale Rolle in der Informatik spielt.

Motivation

2

Idee: Je unwahrscheinlicher ein Ereignis, desto höher sein Informationsgehalt.
Information ist ein **Maß für Überraschung**.

Beispiel: `"Eine ⚅ geworfen"` hat mehr Information als `"gerade Augenzahl"`

Beispiel: Der kurze Zufallstext `"Babalilu"` hat weniger Info als Goethes Faust.

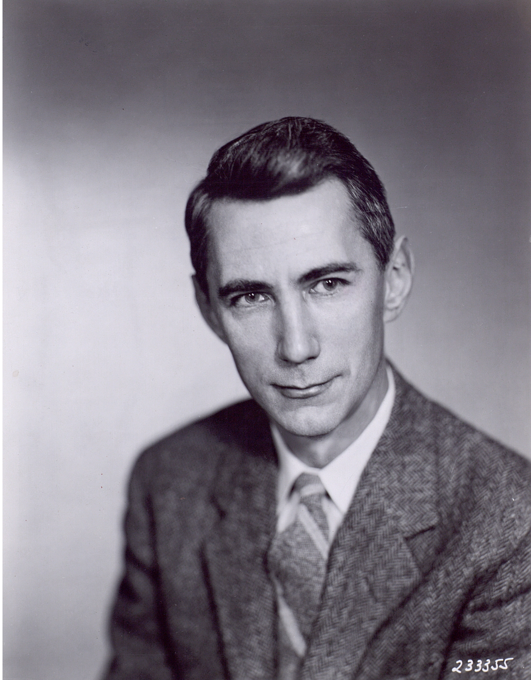
Problem 1: **Kenntnis** der Wahrscheinlichkeiten.
Wahrscheinlichkeiten unbekannt, Annahmen nötig.

Problem 2: **Subjektivität** der Wahrscheinlichkeiten.
Persönliche Aspekte beeinflussen die Annahmen.

Problem 3: Formale Schwierigkeiten bei unendlichen vielen Möglichkeiten.

Beispiel: "Nennen Sie eine zufällige reelle Zahl!"
Wie groß ist die Wahrscheinlichkeit, daß Sie die Zahl π nennen?

Abb. 4: Claude **Shannon** (1916-2001), US-amerikanischer Mathematiker, der als Erfinder der statistischen Informationstheorie gilt. Rechte s. Anhang.



1. Axiom der Informationstheorie

6

Axiom 1: Das **Informationsmaß** I_P eines Ereignisses E ist eine **stetige** und **injektive Funktion** seiner Wahrscheinlichkeit P.

$$f : [0, 1] \to \mathbb{R}_0^+ \cup \{+\infty\} = [0, +\infty] \qquad I_P(E) = f(\ P(E)\ )$$

Das bedeutet: Zu jedem Wahrscheinlichkeitsmaß P gehört ein Informationsmaß I_P.

Motivation des Axioms:

1. **Funktion:** Gleiche Wahrscheinlichkeit führt auf gleichen Informationsgehalt.
2. **Injektiv:** Gleicher Informationsgehalt erfordert gleiche Wahrscheinlichkeit.
3. **Stetig:** Kleine Änderung in Wahrscheinlichkeit führt zu kleiner Änderung im Informationsgehalt.

Unabhängigkeit: Motivation (1)

Experiment: Mit einer Euro- und einer Dollarmünze werfen und das Ergebnis notieren.
Genauer: Das Euro-Ergebnis zuerst, danach das Dollar-Ergebnis.

Menge der **Elementarereignisse:** $M = \{(K,K),\ (K,Z),\ (Z,K),\ (Z,Z)\}$.

Wir haben 4 Elementarereignisse. Wir wenden die Laplace-Regel an:

$$P(\{(K,K)\}) = \frac{1}{4} \qquad P(\{(K,Z)\}) = \frac{1}{4} \qquad P(\{(Z,K)\}) = \frac{1}{4} \qquad P(\{(Z,Z)\}) = \frac{1}{4}$$

Wir berechnen:

$$P_{\text{Euro-Kopf } \textbf{und} \text{ Dollar-Zahl}} = P(\{(K,Z)\}) = \frac{1}{4}$$

$$P_{\text{Euro-Kopf}} = P(\{(K,K),\ (K,Z)\}) = P(\{(K,K)\}) + P(\{(K,Z)\}) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$$

$$P_{\text{Dollar-Zahl}} = P(\{(K,Z),\ (Z,Z)\}) = P(\{(K,Z)\}) + P(\{(Z,Z)\}) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$$

Unabhängigkeit: Motivation (2)

Wir haben also *zwei identische Ergebnisse:*

$$\underbrace{P_{\text{Euro-Kopf }\textbf{und}\text{ Dollar-Zahl}}}_{\frac{1}{4}} = \frac{1}{4} = \underbrace{P_{\text{Euro-Kopf}}}_{\frac{1}{2}} \cdot \underbrace{P_{\text{Dollar-Zahl}}}_{\frac{1}{2}}$$

Es kann aber in einem **anderen Fall** auch so sein:
Wenn der Euro Kopf zeigt, **dann** zeigt der Dollar Zahl (magnetische Koppelung).
Euro-Kopf und Dollar-Kopf treten niemals gemeinsam aus.
Es gibt eine Art wechselseitiger Abhängigkeit der Ergebnisse.
Wir finden eine andere Funktion, zur Unterscheidung nennen wir sie Q.
Nach der Laplace-Regel erhalten wir:

$$Q(\{(K,K)\}) = 0 \qquad Q(\{(K,Z)\}) = \frac{1}{3} \qquad Q(\{(Z,K)\}) = \frac{1}{3} \qquad Q(\{(Z,Z)\}) = \frac{1}{3}$$

Unabhängigkeit: Motivation (3)

Wir berechnen:

$$Q_{\text{Euro-Kopf }\textbf{und}\text{ Dollar-Zahl}} = Q(\{(K,Z)\}) = \frac{1}{3}$$

$$Q_{\text{Euro-Kopf}} = Q(\{(K,K),\ (K,Z)\}) = Q(\{(K,K)\}) + Q(\{(K,Z)\}) = 0 + \frac{1}{3} = \frac{1}{3}$$

$$Q_{\text{Dollar-Zahl}} = Q(\{(K,Z),\ (Z,Z)\}) = Q(\{(K,Z)\}) + Q(\{(K,Z)\}) = \frac{1}{3} + \frac{1}{3} = \frac{2}{3}$$

Wir haben also *zwei verschiedene Ergebnisse:*

$$Q_{\text{Euro-Kopf }\textbf{und}\text{ Dollar-Zahl}} = \frac{1}{3} \qquad \text{und} \qquad \underbrace{Q_{\text{Euro-Kopf}}}_{\frac{1}{3}} \cdot \underbrace{Q_{\text{Dollar-Zahl}}}_{\frac{2}{3}} = \frac{2}{9}$$

Unabhängigkeit: Definition

7

Zwei Ereignisse A und B heißen **unabhängig**, wenn gilt:

$$P(A \textbf{ und } B) = P(A) \cdot P(B)$$

Mathematisch **präziser**, wenn

$$P(A \cap B) = P(A) \cdot P(B)$$

Anschaulich bedeutet Unabhängigkeit: Ob das **eine** Ereignis eingetreten ist, hat keinen Einfluß auf die Wahrscheinlichkeit des **anderen** Ereignisses.

Hinweis: Man kann das noch viel tiefer begründen.

2. Axiom der Informationstheorie

Axiom 2: Sind A und B unabhängige Ereignisse, dann ist der Informationsgehalt des gemeinsamen Ereignisses $A \cap B$ gleich der Summe der Informationen der Einzelereignisse.

Anschauliche Motivation des Axioms:

1. Wenn ich zwei von einander *unabhängige* Zeitungen lese, etwa die New York Times vom 1. 1. 2012 und die FAZ vom 1. 1. 2019, dann ist die erhaltene Information die Summe der Informationen beider Zeitungen.
2. Wenn die Zeitungen aber nicht unabhängig sind, etwa weil sie voneinander abschreiben oder über denselben Tag berichten, dann ist die aus beiden Zeitungen erhaltene Information weniger als die Summe der Informationen beider Zeitungen.

3. Axiom der Informationstheorie

Das **3. Axiom** normiert zwei extreme Situation:

Axiom 3(a): Das **sichere Ereignis** enthält überhaupt keine Information:

$$I(M) = 0$$

Axiom 3(b): Das **unmögliche Ereignis** enthält unendlich viel Information.

$$I(\emptyset) = +\infty$$

Formel für die Information

8
3
4
5
9

Satz vom Informationsgehalt

Der **Informationsgehalt** ist ein **negativer Logarithmus der Wahrscheinlichkeit**.

Genauer: Ist I ein Maß für den Informationsgehalt und P ein dazugehöriges Maß für die Wahrscheinlichkeit, dann gibt es eine positive Zahl $b \in \mathbb{R}^+$ so, daß für alle Ereignisse E gilt:

$$I(E) = -\log_b(P(E)) \qquad I = -\log_b \circ P$$

Die Wahl der Basis b des Logarithmus legt die Einheit der Information fest:

Basis b	Logarithmus	Name der Einheit
2	binärer Logarithmus	[bit] (übliche Wahl der Informatik)
e=2.71828 ...	Natürlicher Logarithmus	[nat]
10	dekadischer Logarithmus	[Hartley]

Tab. 1: Einheiten für das Maß der Information

Genauere Analyse des 3. Axioms

Axiom 3(a) paßt zur Formel wegen:

$$I(M) = 0 \qquad \text{weil } P(M) = 1 \text{ und } \log(1) = 0$$

Axiom 3(b) paßt zur Formel wegen:

$$I(\emptyset) = +\infty \qquad \text{weil } P(\emptyset) = 0 \text{ und } \log(0) = -\infty$$

Streng genommen ist $\log(0)$ nicht definiert. Wir wissen aber: $\lim_{h \to +0} \log(h) = -\infty$.
Aufgrund der Stetigkeit (Axiom 1) ist das also sinnvoll.

Beispiel: Informationsgehalt gezogener Kugel

Situation: In einem Becher befinden sich eine rote, eine blaue und eine grüne Kugel.
Alice zieht eine Kugel.
Alice ruft aus *"Die Kugel ist rot"*.

Aufgabe: Bestimme den Informationsgehalt dieser Nachricht?

Lösung: Es gibt drei Elementarereignisse: rot, grün, blau.
Wir haben keine weiteren Hinweise.
Daher ist die Regel von Laplace anwendbar.

$$P(\{\text{rot}\}) = P(\{\text{grün}\}) = P(\{\text{blau}\}) = \frac{1}{3}$$

Ergebnis: Der Informationsgehalt beträgt:

$$I_P(\{\mathit{rot}\}) = -\log_2\left(\frac{1}{3}\right) = \log_2(3) = 1.584\ldots\ \text{[bit]}$$

Bits und Bit: Codierung und Informationsgehalt

11

Bits: Um aus n verschiedenen Optionen genau eine Option **eindeutig mit binären Symbolen zu kodieren** werden

$$\textbf{ceil}\left(\log_2(n)\right) = \lceil \log_2(n) \rceil$$

binäre Symbole ("Bits") benötigt.

bit: Der **Informationsgehalt einer Auswahlentscheidung** aus n verschiedenen Optionen ist bei Annahme von Gleichwahrscheinlichkeit aller Optionen für jede einzelne Option gegeben durch:

$$-\log_2\left(\frac{1}{n}\right) = \log_2(n) \ \ [\text{bit}]$$

Hinweis: Die Funktion $\textbf{ceil}(\cdot) = \lceil \cdot \rceil$ ist die Rundung nach oben.

3. Information (nach Chaitin)

Ziel: Neben dem Informationsbegriff von Shannon gibt es noch weitere Informationsbegriffe. Wir wollen hier den Informationsbegriff nach Chaitin kennenlernen.

Motivation

Idee: Je mehr ich "erzählen" muß, um ein Ereignis zu beschreiben, um so höher ist sein Informationsgehalt.

Definition: Der Informationsgehalt eines Wortes ist die Anzahl Buchstaben *eines* kürzesten Programms, das dieses Wort ohne jede Eingabe ausdruckt.

Beispiel 1: `"aidnezrtcvwgemqptkgudhrfbendsf"` hat mehr Information als
`"111111111111111111111111111111"`
Den zweiten Text kann ich knapper beschreiben als den ersten.

Beispiel 2: Der Informationsgehalt der Zahl $\pi = 3.1415926\ldots$ ist die Länge *eines* kürzesten Programms, das **alle** Dezimalstellen von π ausgibt.

Nutzanwendung

Beachte: Abhängig von der gewählten Beschreibungs- oder Programmiersprache wären Einheiten [C], [Pascal], [Javascript] denkbar.
Umrechnung ist möglich, aber nicht mit multiplikativen Faktoren.

Problem: Die "Länge des kürzesten Programms" ist nicht sinnvoll berechenbar.
Details dazu finden sich in einem Kurs zu Berechenbarkeit.

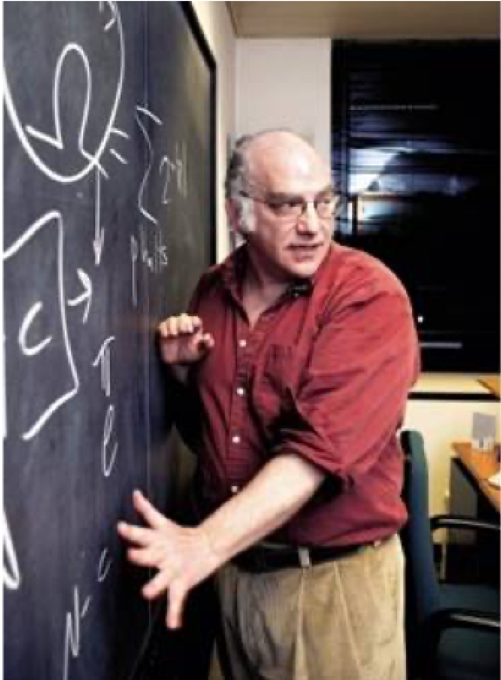


Abb. 5: Gregory **Chaitin** (geb. 1947), argentinischer Mathematiker, einer der Erfinder der algorithmischen Informationstheorie. Rechte s. Anhang.

Quizfrage 1: Gamblers Fallacy

12

Alice spielt gerne Roulette. Sie setzt nur auf die Farben rot und schwarz.

Bob meint: "Jetzt ist sieben Mal hintereinander rot gekommen. Du mußt jetzt auf schwarz setzen. Jetzt kommt sicher bald schwarz."

Frage 1: Hat **Bob** recht?

Alice entgegnet: "Nein, die Wahrscheinlichkeit von rot und schwarz ist in jeder Runde gleich. Ich wette 10 €, daß jetzt wieder rot kommt."

Bob antwortet: "Nein. Schwarz ist jetzt wahrscheinlicher. Ich wette 15 € auf schwarz."

Frage 2: Was wetten Sie?

Das Beispiel weist auch auf den Zusammenhang zwischen Wahrscheinlichkeiten und der Bereitschaft hin, Wetten zur selben oder zu einer anderen Höhe abzuschließen als der Wettpartner.

Quizfrage 2: Informationsgehalt

13

Betrachte die beiden unteren Symbolketten!

Welche hat nach **Chaitin** einen höheren Informationsgehalt?

Welche hat nach **Shannon** einen höheren Informationsgehalt (unter Annahme von Gleichverteilung auf $\{0, 1\}$)?

1. 101101000001101110100000010
2. 111111111111111111111111111

Anhang

Übersicht

Verzeichnis aller Abbildungen

Verzeichnis aller Tabellen

Rechtsnachweise

Abb. 1 Quelle: https://commons.wikimedia.org/wiki/File:Andrej_Nikolajewitsch_Kolmogorov.jpg, Konrad Jacobs, CC BY-SA 2.0 DE https://creativecommons.org/licenses/by-sa/2.0/de/deed.en

Abb. 2 Quelle: https://commons.wikimedia.org/wiki/File:Laplace,_Pierre-Simon,_marquis_de.jpg, James Posselwhite, Public domain

Abb. 3 Quelle: https://twitter.com/Chanand1erB0ng5/status/1327089499386638336/photo/1

Abb. 4 Quelle: https://www.flickr.com/photos/tekniskamuseet/6832884236, Nutzung nach https://creativecommons.org/licenses/by/2.0/

Abb. 5 Quelle: https://mindmatters.ai/2021/03/gregory-chaitins-almost-meeting-with-kurt-godel/

Rechtliche Hinweise (1)

Rechtliche Hinweise (2)

Disclaimer: Die sich immer wieder ändernde Rechtslage für digitale Urheberrechte erzeugt ein nicht unerhebliches Risiko bei der Einbindung von Materialien, deren Status nicht oder nur mit unverhältnismäßig hohem Aufwand abzuklären ist. Ebenso kann den Rechteinhabern nicht auf sinnvolle oder einfache Weise ein Honorar zukommen, obwohl deren Leistungen genutzt werden.

Daher binde ich gelegentlich Inhalte nur als Link und nicht durch Framing ein. Lt EuGH Urteil 13.02.2014, C-466/12 (Pressemitteilung, Blog-Beitrag, Urteilstext). ist das unbedenklich, da die benutzten Links ohne Umgehung technischer Sperren auf im Internet frei verfügbare Inhalte verweisen.

Wenn Sie diese Rechtslage stört, dann setzen Sie sich für eine Modernisierung des völlig veralteten Vergütungs- und Anreizsystems für urheberrechtliche Leistungen ein. Bis dahin klicken Sie bitte auf die angegebenen Links und denken Sie darüber nach, warum wir keine für das digitale Zeitalter sinnvoll angepaßte Vergütungs- und Anreizsysteme digital erbrachter Leistungen haben.

Zu Risiken und Nebenwirkungen fragen Sie Ihren Rechtsanwalt oder Gesetzgeber.

Weitere Hinweise finden Sie im Netz hier und hier oder hier.

Zitierweise dieses Dokuments

Wenn Sie Inhalte aus diesem Werk nutzen oder darauf verweisen wollen, zitieren Sie es bitte wie folgt:

Clemens H. Cap: Was ist Information?. Electronic document. https://iuk.one/1010-1005 11. 10. 2021.

Bibtex Information: https://iuk.one/1010-1005.bib

```
@misc{doc:1010-1005,
   author       = {Clemens H. Cap},
   title        = {Was ist Information?},
   year         = {2021},
   month        = {10},
   howpublished = {Electronic document},
   url          = {https://iuk.one/1010-1005}
}
```

Typographic Information:
Typeset on October 11, 2021
This is pdfTeX, Version 3.14159265-2.6-1.40.21 (TeX Live 2020) kpathsea version 6.3.2
This is pgf in version 3.1.5b
This is preamble-slides.tex myFormat©C.H.Cap

Index

Verzeichnis aller Folien

Legende:
- Fortsetzungsseite
- Seite ohne Überschrift
- Bildseite

Fragen und Kurzaufgaben (1/2)

1. Begriffsklärung: Was ist ein Axiom? Die Formalwissenschaften der Mathematik und Logik orientieren sich entlang der Strukturen von Axiomen, Definitionen, Sätzen, Theoremen, Beweisen, Lemmata und Korollaren. Was sind das für Dinger? Warum ist in diesen Wissenschaften das strikte Vorgehen in diesen Strukturen so wichtig? 9

2. Abgrenzung: Im “Satz über endliche Wahrscheinlichkeiten” ist der Wortteil “Elementar” zwei Mal fett gedruckt. Ist das denn wesentlich? Würde der Satz denn auch für Ereignisse gelten? Nein! Gib dazu ein Argument und ein Gegenbeispiel an. Überlege, ob das Argument logisch wirklich ausreicht (das hängt natürlich von der art des Arguments ab) oder ob ein Gegenbeispiel besser ist! 12

3. Abgrenzung: Bei der Binomialverteilung benutzen wir auch so etwas wie k Erfolge und das Formelzeichen k findet sich dort auch. Es bedeutet aber etwas anderes. Informieren Sie sich über die Binomialverteilung! 19

4. Begriffsklärung: Warum ist hier das Wort “stimmen” durchgestrichen? Weshalb ist es unglücklich, und streng genommen unwissenschaftlich, eine Theorie als “richtig”, “stimmend” oder “wahr” zu bezeichnen? 22

5. Diskussion: Bob sagt: “Das ist doch Quatsch! Zwei gültige Theorien können niemals zueinander im Widerspruch stehen!”. Was halten Sie von diesem Argument? 22

6. Präzisiere: Was bedeutet injektiv denn genau? Und warum ergibt sich die hier angegebene Motivation? Leite das auf der mathematischen Definition der Injektivität ab! 27

7. Beispiele: Finde weitere Alltagsbeispiele von Ereignissen, die unabhängig bzw. nicht unabhängig sind. Ziel ist ein besseres intuitives Verständnis der Begrifflichkeit. 31

8. Wiederholung: Wiederholen Sie die Logarithmengesetze! 34

9. Dumm gefragt: Wenn wir hier schon Logarithmen mit beliebiger Basis betrachten: Was wäre denn der Logarithmus zur Basis 1? Und weil wir gerade dabei sind und neugierig geworden sind: Was wäre der Logarithmus zur Basis 1/2? Empfehlung: Rechnen Sie einige Werte über die Definition des Logarithmus aus und skizzieren Sie den Graphen. Sie können auch `https://www.mathway.com/` oder `https://www.wolframalpha.com/` fragen – aber Achtung: Es gibt Spezialfälle, in denen solche Systeme nicht immer eine sinnvolle Antwort geben. Ferner: Durch Rumtippen auf der Tastatur lernen Sie nichts! 34

10. Klarstellung: Wieso steht in der Formel $\lim_{h \to +0} \log(h)$ ein $+$? Ist das nicht das gleiche wie $\lim_{h \to 0} \log(h)$? Und warum nicht $\lim_{h \to -0} \log(h)$? 35

11. Recherche: Wir kennen jetzt die Bedeutung von `bit` und `Bit`. Was aber sind dann: `Byte`. `Kbit`. `Mbit`. `Gbit`. `Kibibit`. `Mebibit`. `Gibibit`. `Kibibyte`. `Mebibyte`. `Gigibyte`. Wozu hat man diese Einheiten eingeführt? Wo werden sie in der Praxis verwendet? 37

12. Analyse: Untersuchen Sie die Situation der Gamblers Fallacy. Welche Argumente finden Sie? Wodurch entsteht das psychologische Problem für uns? Wie können wir uns das erklären? 42

13. Lösen: Lösen Sie die Quizfrage 2! 43

Aufgabe 1: Zu Folie 7
Komplexere Experimente

Gegeben sei das Experiment “Dreimal hintereinander mit einer Münze werfen” und das Ergebnis in der Form `Wurf1`, `Wurf2`, Wurf3 aufschreiben.

Geben Sie die Menge aller Elementarereignisse an.

Geben Sie die Menge aller Ereignisse an.

Geben Sie ein Elementarereignis an.

Geben Sie ein Ereignis an.

Nun lösen Sie dieselbe Aufgabe für das Experiment “Einmal zugleich drei Münzen hochwerfen” und das Ergebnis in der Form “`Zahl der Köpfe`” und “`Zahl der Zahlen`” aufschreiben.

Was können wir zu den Wahrscheinlichkeiten dieser beiden Experimente sagen?

Aufgabe 2: Zu Folie 25

Bestimmen Sie den Informationsgehalt des Zufallstextes ‘‘`Babalilu`’’.

Bestimmen Sie den Informationsgehalt von Goethes Faust.

Diskutieren Sie an diesen Beispielen die Notwendigkeit von Zusatzannahmen.

Aufgabe 3: Zu Folie 34
Basiswechsel bei Logarithmen

Für die Umrechnung von Logarithmen gilt das Gesetz:

$$\log_a(x) = \frac{\log_b(x)}{\log_b(a)}$$

Eselsbrücke: Neuer Logarithmus ist bekannter Logarithmus dividiert durch Logarithmus[1] der Basis[2].

Aufgabe: Leiten Sie das Gesetz für den Basiswechsel aus den "übrigen" Logarithmengesetzen ab.

[1] und das ist dann natürlich der bekannte Logarithmus, denn den neuen können wir ja nicht berechnen

[2] und das ist dann natürlich die neue Basis, denn der bekannte Logarithmus von der bekannten Basis wäre 1.

Aufgabe 4: Zu Folie 34
Basiswechsel bei Exponentialfunktionen

Vermutung: Für Exponentialfunktionen gilbt es ebenso ein Gesetz, das die Exponentialfunktion einer Basis auf die Exponentialfunktion einer anderen Basis umwandelt.

Genauer: Wenn wir Exponentialfunktion und Logarithmus zur Basis b berechnen können, dann können wir daraus auch die Exponentialfunktion zu einer anderen Basis a berechnen.

Aufgabe: Finden Sie ein solches Gesetz.

Hinweis: Duckduckgo bedienen kann jeder. Die Herausforderung besteht darin, selber so ein Gesetz abzuleiten.

Hilfestellung: Beginnen Sie mal mit $b^x = a^y$. Nun können Sie auf beiden Seiten den Logarithmus $\log_a$ nehmen...

Umrechnung Informationsmaße

Füllen Sie die folgende Tabelle aus:

[bit]	[nat]	[Hartley]
1	____	____
____	1	____
____	____	1