Physik - Information - Informationssysteme

G. Vojta, W.Eisenberg, U.Renner
University of Leipzig -
Leipzig
Germany

1 *Entropie und Information*

Die erste Frage einer Informationswissenschaft ist die nach dem Informationsbegriff. Der akzeptierte Grundbegriff der Informationstheorie wurde von Shannon 1948 formuliert und multidisziplinär angewendet: in der Informationstheorie, der Physik, der Biologie, der Stochastik, der Linguistik, der Musik, der technischen Wissenschaften usw.

Mit dem Informationsproblem ist die Physik z. B. im Zusammenhang mit dem Maxwellschen Dämon (1871) konfrontiert worden. Dieser Dämon wirkt entgegengesetzt zum natürlichen Prozessablauf (2. Hauptsatz der Thermodynamik): er sortiert die schnellen und langsamen Moleküle in die Räume der Doppelkammer (verbunden mit einem Molekül durchgang - siehe Abbildung 1). Er schafft eine Differenzierung im Ortsraum hinsichtlich der Geschwindigkeiten, nachdem er sich über die Geschwindigkeiten der Moleküle virtuell informiert und die Unkenntnis hinsichtlich der Geschwindigkeiten jener Moleküle beseitigt hat. Der Dämon erzeugt dabei einen Nichtgleichgewichtszustand, entgegengesetzt zur prozessualen Tendenz zum Gleichgewicht (2. Hauptsatz), die mit einem Teilchenensemble mit einer mittleren Geschwindigkeit im thermodynamischen Gleichgewicht (oder dem Schwankungsgleichgewicht) verknüpf ist.

Abb. 1: Maxwellscher Dämon (siehe Artikelende)

Vertieft wurde diese Dämonologie von Szilard (1929). Einen formalen Zusammenhang zwischen dem Informationsmaß "Unkenntnis" und der Entropie stellte R. Becker her:
Entropie = k ln "Unkenntnis". Damit waren die Fragen verknüpft:

1. Ist die Information identisch mit der physikalischen Entropie ?
2. Oder ist sie identisch mit der Neg-Entropie ?

2 Informationstheorie, Thermodynamik und Statistische Physik

Das Informationsproblem soll nun am Beispiel eines diskreten Wahrscheinlichkeitsfeldes konkretisiert werden: Den n Elementarereignissen $w_1, w_2, ..., w_n$ (z.B. ein Würfel: $n = 6$) werden die Elementarwahrscheinlichkeiten $p_1, p_2, ... p_n$ (z. B. idealer Würfel: $Pi = 1/6$) zugeordnet. Diese nichtnegativen Wahrscheinlichkeiten $p_i \geq 0$ seien normiert:

$$\sum_{i=1}^{n} p_i = 1$$ (Sicherheit = Wahrscheinlichkeit 1).

Durch einen zufälligen Versuch (hier Wurf) erhält man *ein* zufälliges Elementarereignis w_i)
(z. B. die Augenzahl, die Würfelfarbe) mit der bekannten Wahrscheinlichkeit p_i.
Das Informationsproblem wird im Versuch durch die Reduktion der Unbestimmtheit und den Informationsgewinn gelöst. Die zentralen Fragen sind also:

1.Wie groß ist die Unbestimmtheit des Ereignisses (vor dem Versuch)?
2.Welchen Informationsgewinn erhält man durch das eingetretene Ereignis (nach dem Versuch)?

1. Schritt: Elementare Überlegungen zur Unbestimmtheit $U(w_n)$ des Ereignisses legen das Maß $1/\ p_n$ nahe. Denn als Folgerungen ergibt sich daraus, dass die Unbestimmtheit des Ereignisses klein ist, wenn dessen Wahrscheinlichkeit groß ist. Außerdem ist der Informationsgewinn $I(w_n)$ des eingetretenen Ereignisses in diesem Fall klein!

2. Schritt: Fortgesetzt werden diese Überlegungen mit der Durchführung eines Doppelversuches, z. B. durch das zweimalige Werfen des Würfels. Zu den Ereignissen w_1 und w_2 (z. B. Augenzahl 3, dann 5) gehört die Gesamtwahrscheinlichkeit $P(w_1, w_2 = p_1 p_2)$. Aber der dabei erzielte Informationsgewinn ist offensichtlich additiv: $I(w_1, w_2)$. Also lautet der konkretisierte Ansatz für *U* und *J* : $U(w_n) = I(w_n) = \log 1/\ p_n = -\log\ p_n > 0$.

3. Schritt: Für die weiteren Betrachtungen ist die Frage nach den Mittelwerten weiterführend:
1. Wie groß ist die mittlere Unbestimmtheit eines Versuchsergebnisses, d.h. Ereignisses?
2. Welchen mittleren Informationsgewinn erhält man nach dem Versuch?
Allgemein wird der Mittelwert einer Größe *f* (Pi) als gewichteter arithmetischer Mittelwert eingeführt $< f >= \sum_i p_i\, f(p_i)$ Hier ist $f(p_i) = \log 1/\ p_i = -\log\ p_i$ und man erhält:

$$< I >= H(p_1, p_2, ..., p_n) = \sum_i p_i I(w_i) = -\sum_i p_i \log\ p_i \geq 0$$

Die Information oder Entropie (Shannon-Entropie) oder Unsicherheitsfunktion (Greiner) wird wie folgt definiert:

$$H = -\sum_i p_i \log\ p_i$$

In der Physik gilt speziell log = $\log_e$ = ln (logarithmus naturalis), in der Informationstheorie hingegen log = $\log_2$ = ld (logarithmus dualis). In der Informationstheorie ist die Größe $\log_2 2$ = ld2 = 1 bit (binary digit) ein übliche Einheit. Die Alternative mit den beiden Wahrscheinlichkeiten p_1 = p_2 = 1/2 besitzt die Entropie

H = -½ log ½ - ½ log ½ = -log ½ = log 2 .

Im Extremfall des Eintreffens des Ereignisses *k* mit der Wahrscheinlichkeit p_k = 1 und des der anderen Fälle mit $p_{i \neq k} = 0$ erhält man die minimale Entropie *H* = -log l = 0. Im Fall der Gleichbesetzung von z. B. *n* Zuständen mit p*i* = 1/ *n* ist die Unkenntnis maximal und die Entropie hierfür beträgt: $H = -\sum_{i=1}^{n} \frac{1}{n} \log \frac{1}{n} = -n\left(\frac{1}{n} \log \frac{1}{n}\right) = \log n$ Diese Aussage ist die Basis des Jaynes-Prinzips.

Eine Verallgemeinerung des Informationsbegriffs ist die Renyi-Information mit der Kastenlänge l:

$$I_q(l) = -\frac{1}{q-1} \ln \sum_i (p_i)^q$$

Für *q* = I erhält man den Speziallfall der Shannon-lnformation. Die Renyi-Information bildet die Grundlage für die Multifraktaltheorie und findet Anwendung in der Chaostheorie oder der Mustererkennung.

Einen anderen Zugang zum Verständnis der Analogie von Information und Entropie bietet die Thermodynamik speziell mit der Formulierung des 2. *Hauptsatzes*.
Möchte man z. B. unterschiedliche Substanzen mit den Molzahlen n_i, die sich ursprünglich in den i. A. verschiedenen Volumina V_i im thermodynamischen Gleichgewicht befanden, miteinander vermischen, so wird das Ergebnis dieses Ausgleichsvorgangs durch die Mischungsentropie beschrieben. Die Entropie erhöht sich um:

$$\Delta S = R\sum_i n_i \ln \frac{V}{V_i}$$

R: allgemeine Gaskonstante. Setzt man das Volumen- bzw. Teilchenzahlverhältnis (Clausius,1864)

$$\frac{1}{V} = \frac{N_i}{N} = p_i$$

und verwendet wegen $N_A n_i = N_i$ anstelle der n_i die Teilchenzahl N_i (N_A, Avogadro-Konstante) so erhält man unter Verwendung von

$$k_B = k = \frac{R}{N_A}$$

(k_B, Boltzmann-Konstante) schließlich.

$$\Delta S = -kN\sum_i p_i \ln p_i$$

Der Entropiebegriff in der *Statistischen Thermodynamik* wurde von Ludwig Boltzmann im Rahmen der kinetischen Gastheorie (1877) eingeführt:

$$S = .kN\int\int f \ln f d^3r d^3v .$$

Die Wahrscheinlichkeitsdichtefunktion $f = f(\vec{r}, \vec{v}, t)$ ist eine Funktion der Ortes $\vec{r}$, der Geschwindigkeit $\vec{v}$ und der Zeit *t*. Nach dem Boltzmann-Prinzip (Boltzmann-Planck-Prinzip, Max Planck um 1900) ist die Entropie gegeben durch:

$$S = k \ln W.$$

Verwendet man für die Wahrscheinlichkeit $W = \frac{N!}{\prod N_i!}$ die Näherung von Stirling:

$N! = N^N e^{-N}$, so ist $\qquad S = -kN\sum_i p_i \ln p_i$.

Die Bedeutung der Shannon-Entropie $H = -\sum_i p_i \log p_i$ lässt sich zusammenfassend wie folgt hervorheben:

1. In der Statischen Physik ist sie ein Maß für die Unordnung eines Systems im Gleichgewicht.
2. Bei *gegebenen* Wahrscheinlichkeiten p_j ist sie ein Maß für die Information (syntaktische Information, keine Semantik).

3. Bei *unbekannten* Wahrscheinlichkeiten p_j ist sie Ausgangspunkt für deren Berechnung.
   Hierzu wird das Prinzip der maximalen Entropie, das Jaynes-Prinzip,
   angewendet: $H + \sum l_i$.(Nebenbedingungen)$_i$ = Max. Die Ableitungen nach den Wahrscheinlichkeiten sind 0. Es ist auch für Nichtgleichgewichte und in der mathematischen Statistik anwendbar.
4. *Nachrichtentheorie:* Produziert der Sender *A* die Nachricht *a* mit der Wahrscheinlichkeit

$p(a)\left(\sum_a p(a) = 1\right)$, dann ist die Entropie $H(A) = -\sum_a p(a) \log_2 p(a)$,

nach dem rauschfreien Codierungstheorem von Shannon die untere Schranke für die benötigte Länge *L* einer Bit-Folge zur Darstellung der Nachricht *a:* $L \geq H(A)$.

5. Unter der Annahme eines festen "Kostenbetrages" für jedes Bit (Energie, Raum, Zeit, Geld,..) ist *H(A)* ein Maß für die "Kosten" einer Darstellung der von einem Sender *A* erzeugten Information.

Man beachte, dass bisher die Entropie als "statische" Größe betrachtet wurde. Eine mögliche Verallgemeinerung wäre die Formulierung einer dynamischen Entropie. Auch wurden nur klassische Systeme behandelt. Eine weitere Verallgemeinerung könnte in Hinblick auf Quantensysteme erfolgen.

Das Neg-Entropie-Prinzip (verallgemeinerter 2. Hauptsatz) wurde 1953 von L. Brillouin zur Diskussion des Messproblems eingesetzt. Das System ändert sich durch die Messung (vor der Messung: Entropie S_0; Information I_0 =0; nach der Messung: Entropie S_1; Information I_1 >0).
Es gilt allgemein: $S_1 \geq S_0 + \Delta S = S_0 - I_1$, dabei ist $\Delta S = S_1 - S_0 < 0$,jedoch der
Informationsgewinn positiv: $\Delta I = I_1 - I_0 > 0$.

Eigentlich ist das Messproblem aber mit der Quanteninformatik zu beschreiben. Dabei stellen sich aber die Fragen nach der Nützlichkeit des informationstheoretischen Zugangs (was zu bejahen ist) und der Wesensgleichheit der thermodynamischen Entropie und der Information (offenes Problem). Letzteres soll an einem Beispiel des konstanten Stromflusses durch einen Leiter erläutert werden. Im stationären Zustand ist die Stromstärke und die Entropie konstant, weil die Entropieproduktion mit einer Entropieabfuhr verbunden ist.

Das Jaynes-Prinzip liefert mit einem Variationskalkül (Optimierungstheorie) die unbekannte Wahrscheinlichkeit, und zwar maximal vorurteilsfrei.
Im nun diskutiertem Anwendungsbeispiel seien die Wahrscheinlichkeiten *Pi* für die Zustände *i* eines lokalisierten Systems gesucht, d. h. es ist das Maximum von $S = -\sum_i p_i \ln p_i$
zu bestimmen.

Neben der Normierung der Wahrscheinlichkeiten durch $\sum_i p_i = 1$ sind der Erwartungswert der Energie $< E >= \sum_i p_i E_i >$ und der weiteren Größen G_a , z.B. der Stromdichte, durch $< G^a >= \sum_i p_i G_i^a$

vorgegeben, *a=1,2,...r*

Daraus ergibt sich die Wahrscheinlichkeit $p_i = \exp\left(-\Omega - bE - \sum_a l_a G_i^a\right)$, wobei $\exp\Omega$ die Zustandssumme und $l_a = \frac{1}{kT_a}$ ist.

Die T_a bezeichnen verallgemeinerte Temperaturen (Ingarden).

Sind die Größen G_a Nichtgleichgewichtsgrößen, dann sind die p_i durch Dichteoperatoren *p* zu ersetzen ($l_a[t]$; Zubarev – Formalismus, 1965). Im Robertson – Formalismus werden die Zeitabhängigen Projektionsoperatoren p(t) verwendet (relevanter Anteil: P8t) *p*.

3. Quanteninformatik

Als Vorläufer der Quanteninformatik gelten Hellström u. a. (etwa ab 1970), die die quantenmechanische Systemtheorie, Signaltheorie und Kommunikationstheorie entwickelten.

Diese Theorienentwicklung umfasste zwei Konzeptionen:
1. Einführung quantenmechanischer Wahrscheinlichkeiten in die klassische Shannonsche Informationstheorie (präzisierte Theorie; keine neue Theorie);
2. Entwicklung der Quanteninformatik durch volle Ausnutzung der quantenmechanischen Gesetze (Vorläufer: u. A. Armin Uhlmann).

Die letztere Konzeption der Quanteninformatik setzt ausreichende Kenntnisse über die Quantenmechanik voraus (Wellenfunktion als Wahrscheinlichkeitsamplitude Ψ ; Wahrscheinlichkeitsaussagen: Wahrscheinlichkeit $=\Psi^*\Psi$; Zustandsfunktionen Ψ_i im Hilbertraum; Hamilton - Operator **H** usw.). Die Grundgleichung ist die Schrödinger - Gleichung:

$$ih\frac{dy}{dt} = Hy$$

Neben den Zustandsfunktionen y_i ist auch die Uberlagerung dieser Zustände $\sum_i c_i y_i$ ein zulässiger Zustand (verschränkt, rein). Der gemischte Zustand hingegen wird durch den Dichteoperator $p(x,x`) = \sum_i c_i y_i y_i^*$ (Diagonalelemente:Wahrscheinlichkeitsdichte) definiert.

Die Entropie der Quantenmechanik ist nach John von Neumann definiert durch:

$$S = -k\ \mathbf{Sp}(pln\ p)\ .$$

Weiterhin ist zu beachten, dass durch Beobachtung des Systems, d.h. durch Messung des Ortes, das System gestört wird und somit sich der Systemzustand ändert. Nicht vertauschbare Größen sind nicht gleichzeitig messbar.

Problematisiert wurde der Begriff der Quanteninformation durch das Einstein-PodolskyRosen-Paradoxon (EPRP), das in der Spinsprache ($\uparrow\downarrow$) klar formulierbar ist. Nach der Trennung der Spins wird durch eine Messung der Spinzustand $\uparrow$ gefunden und z. B. 10 km davon entfernt ist der Zustand des getrennten Spins ohne Messung fixiert $\downarrow$. Nach einer weiteren Messung (Änderung der Magnetfeldrichtung) wird ein veränderter Spinzustand angetroffen $\rightarrow$ und 10 km davon entfernt ist paradoxerweise der zweite Spinzustand wieder ohne Messung fixiert $\leftarrow$. Das EPRP impliziert natürlich bestimmte Fragen, z. B. diese:
Ist die Signalgeschwindigkeit größer als die Lichtgeschwindigkeit c?

Es ist auch die Basis möglicher Sicherheitstechnologien für den Informationstransport, wie eine Analyse des Quantenkanals ergibt:

Sender A →→→→→→→→→→→→→→→→→→ Empfänger B
ALICE ⇑ BOB

Abhören, Stören durch EVA wird entdeckt !

Dieses analysierte Modellsystem ist ein quantenmechanisches Zweizustandssystem und wird durch den Begriff Quantenbit (q-bit, qubit, Schumacher, 1995) beschrieben (weiteres siehe Vortrag von Prof. K. Kreher).

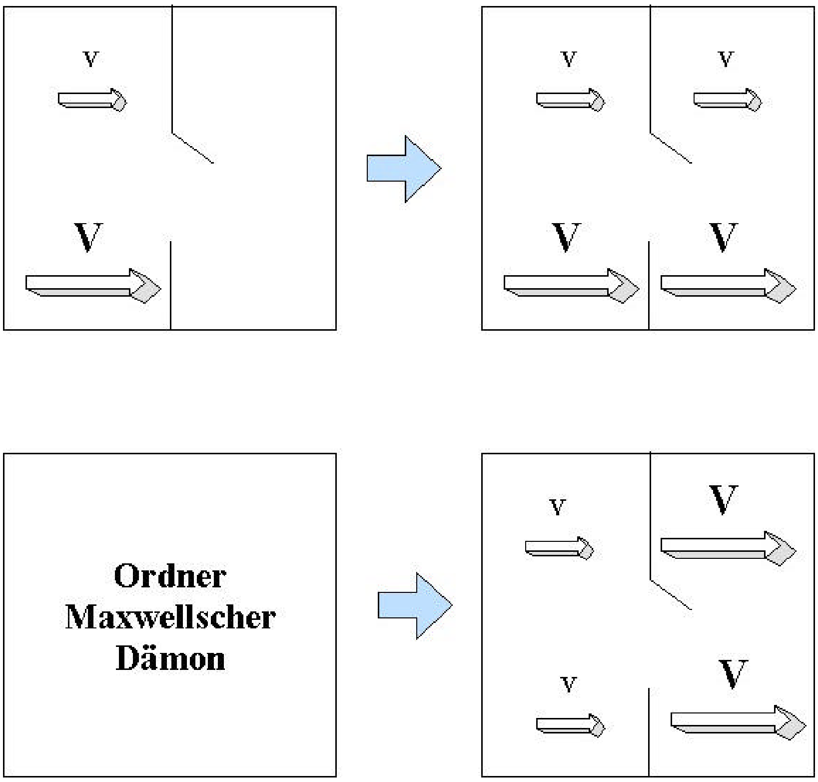

Abb.1 Maxwellscher Dämon