Entropy demystified

RSE24

Prof. Dr. Haye Hinrichsen Theoretical Physics III University of Würzburg

Entropy demystified...

- 1) Historical outline
- 2) Entropy as a measure of information
- 3) Entropy and probability
- 4) Entropy in Physics
- 5) What exactly is temperature?
- 6) Sampling
- 7) A funny story

1) Historical outline

Rudolf Clausius (1822-1888)

- 1855 ETH Zürich
- 1867 University of Würzburg
- 1869 University of Bonn

... was one of the first Theoretical Physicists



Wikimedia

ÜBER DEN ZWEITEN HAUPTSATZ DER **MECHANISCHEN WÄRMETHEORIE** Ein Vortrag, gehalten in einer allgemeinen Sitzung der 41. Versammlung deutscher Naturforscher und Aerzte zu Frankfurt a. M. am 23. September 1867 von Enemanuel redalph Julius Enemanuel Professor der

Entropy = Internal Transformability

- en = internal
- tropé = transformation

S





Clausius and the <u>Second Law of Thermodynamics</u>

• The change of the entropy S is



• In isolated systems the entropy can only increase:

 $\Delta S > 0$

The Second Law of Thermodynamics



Original paper by Clausius

44 Abhandlung IX.

man daneben zugleich den anderen, seiner Bedeutung nach einfacheren Begriff der *Encryie* anwendet, man die den beiden Hauptsätzen der mechanischen Wärmetheorie entsprechenden Grundgesetze des Weltalls in folgender einfacher Form aussprechen kann:

- 1) Die Energie der Welt ist constant.
- 2) Die Entropie der Welt strebt einem Maximum zu.

The Energy of the world is constant.
 The Entropy of the world tends to a maximum.

The Second Law comes as a surprise, because...

• all mechanical systems are invariant under **time reversal**, meaning that they can also run backward in time.

$$F = ma$$

• but Entropy increases only with increasing time.

This defines the "arrow of time".

Ludwig Boltzmann (1840-1906)

- 1863 Vienna
- 1869 Graz
- 1873 Vienna (Mathematics)
- 1876 Graz (Experimental Physics)
- 1890 Munich

Co-founder of "atomism"



Wikimedia

Boltzmann's View

"The question of whether matter is **composed of atoms or continuous** is therefore reduced to whether those properties most accurately represent the observed properties of the matter when the number of particles is assumed to be extraordinarily large."

Boltzmann's definition of entropy

The entropy of an *isolated* system is



What are "complexions"?



Today we would probably call it "configurations".

Boltzmann's definition of entropy

$S = k_B \log \Omega_{\bullet}$

Number of possible configurations the gas molecules can be in.

What are "complexions"?



This raises the question of how to "count" the number of configurations

Claude Shannon (1916-2001)

- Study of mathematics/electrical engineering
- 1941 AT&T Bell Labs
- 1958 MIT

... Father of the information age



Wikimedia

Shannon's most important work

What is the maximum amount of information that can sent through a cable with a given bandwidth?

Shannon's most important work



ERUSAEMNOITAMROFNI

Channel (data stream)

E	12.7%	
т	9.1%	
Α	8.2%	
0	7.5%	
I	7.0%	
Ν	6.7%	
S	6.3%	
н	6.1%	
R	6.0%	
D	4.3%	

Shannon's most important work



Shannon's most important result



Shannon's key result - characteristics

(1) The average information stream is in fact an average:

$$H = -\sum_{i} p_i \, \log p_i = -\log p$$

 $-\log p_i$ is the information carried by the letter *i*

Shannon's key result - characteristics

(2) If all characters were **equally likely**, that is, if $p_i = \frac{1}{N}$ we would get:

$$H = -\sum_{i=1}^{N} \frac{1}{N} \log \frac{1}{N} = \log N$$

Boltzmann-Entropy

Shannon Information



$S = k_B \log \Omega$

Number of "complexions"

 $H = \log N$

Number of characters

Claude Shannon asks John von Neumann: How should I name this quantity?





Call it entropy!

No one knows what entropy really is, so in a debate you will always have an advantage.

Wikimedia



Heat transfer

Second Law

Number of configurations

 ΔS \overline{T}

 $\Delta S \ge 0$

$$S = k_B \log \Omega$$

Shannon-Information $H = -\sum_{i} p_{i} \log p_{i}$ \blacksquare $H = \log N$





Entropy = Information

2) Entropy = Information

What exactly is information?

Information

Article	Talk		Read
			_

From Wikipedia, the free encyclopedia

For other uses, see Information (disambiguation).

Information is an abstract concept that refers to that which has the power to inform. At the most fundamental level, information pertains to the interpretation (perhaps formally) of that which may be sensed, or their abstractions. Any natural process that is not completely random and any observable pattern in any medium can be said to convey some amount of information. Whereas digital signals and other data use discrete signs to convey information, other phenomena and artifacts such as analogue signals, poems, pictures, music or other sounds, and currents convey information in a more continuous form.^[1] Information is not knowledge itself, but the meaning that may be derived from a representation through interpretation.^[2]

The concept of *information* is relevant or connected to various concepts,^[3] including constraint, communication, control, data, form, education, knowledge, meaning, understanding, mental stimuli, pattern, perception, proposition, representation, and entropy.

Information is often processed iteratively: Data available at one step are processed into information to be interpreted and processed at the next step. For example, in written text each symbol or letter conveys information relevant to the word it is part of, each word conveys information relevant to the phrase it is part of, each phrase conveys information relevant to the sentence it is part of, and so on until at the final step information is interpreted and becomes knowledge in a given domain. In a digital signal, bits may be interpreted into the symbols, letters, numbers, or structures that convey the information available at the next level up. The key characteristic of information is that it is subject to interpretation and processing.

By conveying information, ignorance about an object or issue is reduced.

Flow of information



- Physical system
- Fact
- Data set

- Human
- Measurement device
- Interacting system

Verbal definition of information / entropy

The information / entropy of an object is the minimal size of a data set that has to be transmitted to fully describe the object in a given context.

Verbal definition of information



Verbal definition of information

Information is the length of an optimally compressed file that is required to fully describe an object or a fact.

Unit of information: bit

Byte	В	1 Byte = 8 bit
Kilobyte	kВ	1 kB = 1024 Byte
Megabyte	MB	1 MB = 1024 kB
Gigabyte	GB	1 GB = 1024 MB
Terabyte	ТВ	1 TB = 1024 GB

Contextuality of entropy / information

The information of an object depends on the chosen **context**.



Wikimedia

• Context: switch position on/off Information: 1 bit

• Context: position of all molecules Information: huge

Configuration space

• By Ω we denote the set of all possible configurations in a chosen context, called **configuration space**.



• Let $|\Omega|$ be the number of all possible configurations

$$H = \log_2 |\Omega|$$
 ${}_{\leftarrow}$ measured in *bit*

Example: Entropy of a die

Context: number of pips

$$\Omega = \{1, 2, 3, 4, 5, 6\}$$

Cardinality: $|\Omega| = 6$

Wikimedia

$$\Rightarrow H = \log_2 6 = \frac{\ln 6}{\ln 2} \approx 2.585 \text{ bit}$$
Non-integer number of bits ???
Non-integer number of bits?

Number of dice	Number of configurations	Required number of bits	Bits per die
1	6	3	3
2	36	6	3
3	216	8	2,666
4	1296	11	2,75
5	7776	13	2,6
∞	~	~	~ 2,585



Definition of entropy in various disciplines



How many bits of entropy does a balloon contain?

$S \approx 126 \,\mathrm{J/K}$

chemistrytable.webs.com/enthalpyentropyandgibbs.htm



How many bits of entropy does a balloon contain?

$S \approx 126 \,\mathrm{J/K}$

chemistrytable.webs.com/enthalpyentropyandgibbs.htm

$$H = \frac{1}{k_B \ln 2} S \approx 1.3 \times 10^{25} \,\text{bit}$$



How many bits of entropy does a balloon contain?

$S \approx 126 \,\mathrm{J/K}$

chemistrytable.webs.com/enthalpyentropyandgibbs.htm

$$H = \frac{1}{k_B \ln 2} S \approx 1.3 \times 10^{25} \,\text{bit}$$

$$\implies \frac{H}{N_A} \approx 22 \, \text{bit}$$



3) Entropy and probability

What exactly is probability?

The train comes with a probability of about 50%



Probabilities express **partial knowledge**.

Configuration space Ω equipped with probabilities

Context: Arrival of the train

A: in time $\Omega = \{A, B, C\}$ C: cancelled

$$p_A = \frac{1}{2}, \ p_B = \frac{1}{3}, \ p_C = \frac{1}{6}$$

$$\Omega = \{A, B, C\}$$

$$p_A = \frac{1}{2}, \ p_B = \frac{1}{3}, \ p_C = \frac{1}{6}$$

Trick: Fictitious configuration space in which all elements are equally likely.



Information needed to specify one of the 6 elements: $H = \log_2 6$ Information needed to specify A: $H_A = \log_2 6 - \log_2 3$

Individual Entropy / Information of the event A:



$$H_A = \log_2 6 - \log_2 3 = \log_2 \frac{6}{3} = \log_2 \frac{1}{p_A} = -\log_2 p_A$$

Every event $\omega \in \Omega$ with probability $p_{\omega} \in (0, 1]$ carries an **individual entropy** *I* **information**

$$H_{\omega} = -\log_2 p_{\omega}$$

Every event $\omega \in \Omega$ with probability $p_{\omega} \in (0, 1]$ carries an **individual entropy** *I* **information**

$$H_{\omega} = -\log_2 p_{\omega}$$

Average:

$$H = \overline{H_{\omega}} = \sum_{\omega \in \Omega} p_{\omega} H_{\omega} = -\sum_{\omega \in \Omega} p_{\omega} \log_2 p_{\omega}$$





Entropy = Information

4) Entropy in Physics

Cartoon of a physical system



Center of Nanoscience, München



Configuration Space

Spontaneous jumps in configuration space



Probabilities per unit time for spontaneous jumps in Ω :

 $w_{i \to j} \ge 0$

Isolated physical systems: Time reversal symmetry



Perfect isolation

In **isolated** systems the jump probabilities are symmetric:

$$w_{i \to j} = w_{j \to i}$$

Gibbs' postulate



In **isolated** physical systems, after long equilibratrion time, each configuration is equally likely!

$$\lim_{t \to \infty} p_{\omega} = \frac{1}{|\Omega|}$$

Second Law of Thermodynamics



The system performs an unbiased random walk in its own configuration space.

$$\lim_{t \to \infty} p_{\omega} = \frac{1}{|\Omega|}$$

$$H_{\rm max} = \log_2 |\Omega|$$

 $\Delta H \ge 0$





Entropy = Information

5) What exactly is temperature?

Perfectly isolated physical systems...



...are unlikely in Nature



Heat bath = Energy reservoir



Entropy of a heat bath





Entropy of a heat bath



Entropy of a heat bath



Energy is traded for Entropy

Price =
$$\frac{\text{Entropy}}{\text{Energy}} = \frac{1}{T}$$







Trade of Energy for Entropy



The "correct" unit of temperature...

$$\beta = \frac{\Delta S}{\Delta E} = \frac{1}{T}$$
$$[\beta] = \frac{\text{bit}}{\text{Joule}}$$
$$[T] = \frac{\text{Joule}}{\text{bit}}$$

The "correct" unit of temperature

$$\beta = \frac{\Delta S}{\Delta E} = \frac{1}{T}$$
$$[\beta] = \frac{\text{bit}}{\text{Joule}}$$
$$[T] = \frac{\text{Joule}}{\text{bit}}$$

Computer Science: $H = \log_2 N$ $H = \ln |\Omega|$ **Mathematics:** $S = k_B \ln \Omega$ Physics / Chemistry: $k_B = 1.38 \times 10^{-23} \frac{\text{Joule}}{K \text{ bit}}$ $\frac{\partial E}{\partial H} = k_B 293 K \approx 4 \times 10^{-21} \frac{\text{Joule}}{\text{hit}}$ Room Temperature





Entropy = Information

6) Sampling

Random data source

Random bit generator

0110100001010111101010111010011101011101

Random data source

Random bit generator

We would like to estimate its entropy.

0110100001010111101010111010011101011101

Random data source

Random bit generator

Let *p* be the probability of a 1 and 1-*p* the probability of a 0

0110100001010111101010111010011101011101

$$H = -p \log_2(p) - (1-p) \log_2(1-p)$$


0110100001010111101010111010011101011101

$$H = -p \log_2(p) - (1-p) \log_2(1-p)$$





Random bit generator

0110100001010111101010111010011101011101

Sample of size *N* with *k* times '1'



"Naive" estimation of *H*:

• Take a sample of size N

• Estimate
$$p \approx \frac{k}{N}$$

• Plug this estimate into the entropy formula

$$H = -p \log_2(p) - (1-p) \log_2(1-p)$$

• Evaluate to get an estimate of *H*



• Take a sample of size N

• Estimate
$$p \approx \frac{k}{N}$$

• Plug this estimate into the entropy formula

$$H = -p \log_2(p) - (1-p) \log_2(1-p)$$

• Evaluate to get an estimate of H

Non-linear estimators lead to systematic errors!

Statistical error



Statistical error



Sampling error distribution



Binomial distribution:

$$P_{k} = \binom{N}{k} p^{k} (1-p)^{N-k}$$

Sampling error distribution



Binomial distribution:

$$P_{k} = \binom{N}{k} p^{k} (1-p)^{N-k}$$

$$\langle P \rangle = \sum_{k=1}^{N} P_k \frac{k}{N} = p = 0.6$$

Angled brackets = Average over many samples

Sampling error distribution



Estimation of *p*: No systematic error





Nonlinear estimators like the entropy lead to <u>systematic</u> errors.

Example of a simple non-linear estimator: $F(p)=p^2$



Example of a simple non-linear estimator: $F(x)=x^2$



Example of a simple non-linear estimator: $F(x)=x^2$



Mathematical origin of this discrepancy:

- Averaging of a random variable X is a linear operation $\langle X \rangle = \sum_{i} p_{i} X_{i}$
- Linear functions L(x) commute with averaging $\langle L(X) \rangle = L(\langle X \rangle)$
- Nonlinear functions F(x) do <u>not</u> commute with averaging $\langle F(X) \rangle \neq F(\langle X \rangle)$

Systematic errors

How can we detect systematic errors?

Take the smallest sample size!

To see systematic errors, go to small sample sizes

• Take a sample of size N=1

• Estimate
$$p \approx \frac{k}{N} = 0,1$$

• Plug the estimate into the entropy formula

$$H = -p \log_2(p) - (1-p) \log_2(1-p) = 0$$

• Evaluate to get an estimate of *H*

The entropy is systematically under-estimated.

Systematic error underestimating the entropy



Systematic error underestimating the entropy





• Systematic errors are to some extent predictable.

• You can find compensation formulas in the literature.

• The more systematic errors are compensated for, the more this comes at the expense of statistical error.

Compensating systematic errors of entropy estimates

• Miller (so so simple, no excuse for not using this one)

G. Miller (1955), Information theory in Psychology II-B ed. H. Quastler, Glencoe, Illinois: Free Press 95

Grassberger

Finite sample corrections to entropy and dimension estimates, Phys. Lett. A 128 (1988) 369

• Bonachela / H.H. / Munoz

Entropy estimates of small data sets, J. Phys. A 41 (2008) 202001

Compensating systematic errors of entropy estimates

• Miller (so so simple, no excuse for not using this one)

G. Miller (1955), Information theory in Psychology II-B ed. H. Quastler, Glencoe, Illinois: Free Press 95

• Grassberger

Finite sample corrections to entropy and dimension estimates, Phys. Lett. A 128 (1988) 369

• Bonachela / H.H. / Munoz

Entropy estimates of small data sets, J. Phys. A 41 (2008) 202001

The Miller correction

• Naive entropy estimate

$$H_{naive} = -\sum_{j} \frac{k_{j}}{N} \ln \frac{k_{j}}{N}$$
 Simply replace p_{j} by $\frac{k_{j}}{N}$

The Miller correction

• Naive entropy estimate

$$H_{naive} = -\sum_{j} \frac{k_j}{N} \ln \frac{k_j}{N}$$
 Simply replace p_j by $\frac{k_j}{N}$

• Miller-correted entropy estimate

$$H_{Miller} = -\sum_{j} \frac{k_{j}}{N} \ln \frac{k_{j}}{N} + \underbrace{\frac{1}{2N}}_{\text{Miller correction}}$$

Miller-corrected entropy estimation



7) Last but not least: A funny story



The frequencies of an aurally tuned piano deviate significantly from the mathematical formula.

Well-tuned pianos are always "out of tune"



Inharmonicity of a realisitic piano string

Ideal string

Ideal rod

 $f_n = n \cdot f_1$

 $f_n = n^2 \cdot f_1$

Realistic string

 $f_n = n \cdot f_1 \cdot \sqrt{1 + B n^2}$ Inharmonicity coefficient







(a fifth higher)



Both together (well tuned)



Both together (out of tune)


Well tuned = Low entropy of the spectrum





Entropy-based piano tuning



UNIVE WÜRZ	IS-Maximilians- RSITÄT BURG				Logir	n Kontakt Impre
	Universität	Fakultäten	Forschung	Lehre	Einrichtungen	Zielgruppen
Archiv		Wenn P	hysiker ein K	lavier stin	nmen	

Ein Würzburger Physikprofessor veröffentlicht einen Aufsatz. Englischsprachige Medien greifen das Thema mit schmissigen Überschriften auf - und sorgen damit für Unruhe in Musikerkreisen. Die Geschichte eines Missverständnisses.

Haye Hinrichsen ist Professor am Physikalischen Institut der Universität Würzburg. Und er ist Familienvater. Seine Kinder spielen gerne Klavier, und darum schafft die Familie ein solches Instrument an. Ein Klavierstimmer bringt es zum vollendeten Wohlklang, "Wie Physiker eben sind: Bei dieser Aktion kam ich auf die Idee, einfach



Triday Apr 13 2

Find a Job | MLS Wee | Our Papers | Feedback

Cohendata

Online

Home III U.S. | Sport | TVAShoebiz | Femail | Health | Science | Money | Right

lew tridings from a German scientist may be music to a musician's ears - but it pould put those bith priced human tubers out of

Arbeitslosigkeit für Klavierstimmer prognostiziert diese Überschrift der Daily Mail. Schuld daran soll der Würzburger Physikprofessor Haye Hinrichsen sein. Screenshot: Robert Emmerich



Christoph Wick

3.1. Average density matrix

Let us now compute the average density matrix

$$\langle \rho(t) \rangle_{\text{GUE}} = \sum_{i,j=1}^{4} \left\langle e^{-i(E_i - E_j)t} \right\rangle_E \underbrace{\left\langle |\phi_i\rangle\langle\phi_i|\rho(0)|\phi_j\rangle\langle\phi_j| \right\rangle_2}_{=T_{ij}}.$$
 (23)

Computing the average over the energies

$$\left\langle e^{-i(E_i - E_j)t} \right\rangle_E = \frac{1}{N} \int_{-\infty}^{+\infty} dE_1 \cdots dE_4 P_{\text{GUE}}(E_1, \dots, E_4) e^{-i(E_i - E_j)t},$$
 (24)

where $\mathcal{N} = \int_{-\infty}^{+\infty} dE_1 \cdots dE_4 P_{\text{GUE}}(E_1, \dots, E_4) = \frac{9\pi}{2A^8}$ is the normalization factor, we obtain the result

$$\left\langle e^{-i(E_i - E_j)t} \right\rangle_E = g(\tau) + (1 - g(\tau)) \,\delta_{ij} = \begin{cases} 1 & \text{for } i = j \\ g(\tau) & \text{for } i \neq j \end{cases}, \quad (25)$$

where we defined the scaled time $\tau:=t/\sqrt{A}$ and the function

$$g(\tau) = \frac{(1152 - 2304\tau^2 + 1104\tau^4 - 256\tau^6 + 25\tau^8 - \tau^{10})}{1152}e^{-\frac{\tau^2}{2}}.$$
 (26)

Thus Eq. (23) reduces to

$$\langle \rho(t) \rangle_{\text{GUE}} = g(\tau) \sum_{i,j=1}^{4} T_{ij} + (1 - g(\tau)) \sum_{i=1}^{4} T_{ii}$$
 (27)

What remains is to determine the operators

$$T_{ij} = \left\langle U_{\alpha}^{\dagger} | i \rangle \langle i | U_{\alpha} \rho(0) U_{\alpha}^{\dagger} | j \rangle \langle j | U_{\alpha} \right\rangle_{\alpha}.$$
(28)

Obviously, the first sum in Eq. (27) is given by

$$\sum_{i=1}^{4} T_{ii} = \left\langle U_{\alpha}^{\dagger} U_{\alpha} \rho(0) U_{\alpha}^{\dagger} U_{\alpha} \right\rangle = \left\langle \rho(0) \right\rangle = \rho(0). \tag{29}$$

≰ ⊙ <u>F</u>ile <u>T</u>ools <u>H</u>elp

🔒 🔏 % 🍂 🔀 🔳

9 🕒

lo 🛈 😥



Recording keystroke

Double-blind test at the University of Music, Würburg





er

TP3 @ University of Würzburg



2 2 2 2 × Volume control Signal analyzer 9 ູ 🔾 Idle Idle Quality A0 I Off On - Hz Record Record PI PI Calculate Calculate 1 Tune Tune

Idle <t

A REPORT OF A REPORT OF A DESCRIPTION OF A DESCRIPTIONO OF A DESCRIPTION O

Support für diese App

Mehr von TP3 @ University of \rightarrow Würzburg



Perzeptron TP3 @ University of Würzburg 4,6 *



Tonhöhenmonitor für Chöre TP3 @ University of Würzburg

ñ₩

Just Intonation TP3 @ University of Würzburg

Über diese App \rightarrow

Stimmen Sie Ihr Klavier oder Ihren Flügel mit Hilfe eines neuen Verfahrens, das auf Entropieminimierung beruht und die individuellen Eigenschaften Ihres Instruments berücksichtigt. > Trailer



- 1) Clausius Boltzmann Gibbs Shannon von Neumann
- 2) Entropy = Information
- 3) Partial knowledge = Probability
- 4) Random walk in configuration space → 2nd Law of Thermodynamics
- 5) Temperature is the price of Energy in terms of Entropy
- 6) Sampling and the correction of systematic errors
- 7) Entropy and Piano Tuning

Thank you !