

ENTROPY: FROM THERMODYNAMICS TO SIGNAL PROCESSING

JORDÃO NATAL DE OLIVEIRA JÚNIOR* TALYSSON MANOEL DE OLIVEIRA SANTOS* RAFAEL RODRIGUES MENDES RIBEIRO* IVONETE ÁVILA† CARLOS DIAS MACIEL*

* *University of São Paulo, São Carlos, Brazil*

† *Unesp, Guaratinguetá, Brazil*

Email: jordao.oliveira@usp.br talyssonsantos@usp.br rafael.mendes.ribeiro@usp.br
iavila@feg.unesp.br maciel@sc.usp.br

Abstract— Entropy is a concept that remote to the 19th century and it was associated with the work realized by a thermal machine in the context of the Industrial Revolution. The 20th century saw an unprecedented scientific revolution and one of the most important innovations from this time was Information Theory, which also has a concept of Entropy. It can be argued that this is one of the most misused scientific terms and researchers of different areas have been using it wrong. In this paper, a historical background for the evolution of the concept of “entropy” is presented, as well as mathematical proofs and logical arguments for the interconnection of the concept in different areas of science, and how it is related with complexity.

Keywords— Entropy, Information Theory, Thermodynamics, Kolmogorov Complexity.

1 Introduction

Entropy is a measure largely used in science and engineering (Cover and Thomas, 2012). Having been first introduced in thermodynamics by Clausius (Greven et al., 2014) and improved by Boltzmann and Gibbs still in nineteenth century (Wehrl, 1978). The concept was generalized by Shannon in the twentieth century (Shannon, 1948). Today, its applications can be found in biology (Brooks et al., 1988; De Martino and De Martino, 2018; Caro et al., 2018), cosmology, in which it is the center of one of the biggest open problems of science (Maldacena, 2018; Xiao et al., 2018; Bousso, 2018), economics (Bossomaier et al., 2018; Gu et al., 2015), engineering (Zeeshan et al., 2017; Rostaghi and Azami, 2016; He et al., 2016) and even linguistics (Degaetano-Ortlieb and Teich, 2017; Reynar and Ratnaparkhi, 1997; Campbell, 1982).

With such a wide range of applications, it raises the natural question: what is the difference, if any, among the “entropies” used in each field? It can be observed in several papers a misunderstanding of the meaning of entropy when applied to different areas other than physics and information theory (Tame, 2019; Adami, 2016; Kovalev, 2016; Hayflick, 2007; Morowitz, 1986). However, sometimes even in these areas the concept is misused (Martyushev, 2013; Henderson, 2003) and university students have many misconceptions about the theme (Sozibilir, 2003).

There have been attempts to conciliate the entropy of thermodynamics with that of information theory. The most common approach is defining entropy as “disorder” (Wright, 1970; Schrodinger, 1968), something that is introduced as soon as the high school for thermodynamics. However, if it has, at first, a didactic appeal, it is not a good analogy since “order” is a subjective human con-

cept and “disorder” is not the measurement that can be obtained always with entropy (Soubane et al., 2018).

A most sophisticated way of relating these two concepts can be done using quantum states representations and associating the Shannon entropy with von Neumann entropy (Weilenmann et al., 2016). This approach, however, demands prior knowledge of theoretical quantum mechanics hard to be found in biology or economics researchers.

Another confusion is made by relating complexity to entropy. Computer science and statistics have boarded the problem of complexity by the means of Kolmogorov Complexity (Kolmogorov, 1963). Again, the abstraction level of the concept can make researchers from other areas, such as biologists and chemists (in which the study of complexity plays an important role) misunderstand concepts (Adami, 2002).

In this paper, a historical evolution of the concepts of “entropy”, “information” and “complexity” is presented, as well as some clarifications for the concepts, showing how entropies in different areas are fundamentally the same, and how entropy and complexity differ from each other. Also, it is present a computational framework in python to go through these different concepts in the form of Maxwell’s Demon.

2 Theory

2.1 Thermodynamics approach

Clausius coined the word “entropy” (from the Greek word for “change”) in the 1850s to associate the inevitable generation of heat when work is done in a system, changing its temperature (Clausius, 1960). The relationship derived by him is showed in Eq. 1

$$dS = \frac{\delta Q}{T} + \delta S_{ger} \quad (1)$$

with unit J/K. It is a convention to use δQ to indicate an inexact differential, in which the integration depends not only on the starting and ending states but also on the process path in between. On the other hand, entropy is a thermodynamic property and, therefore, dS is an exact differential and the integration not depends on the process path between the starting and ending states.

The entropy generation amount, δS_{gen} , is due to irreversible phenomena occurring inside the system and this value is never negative. However, the entropy change of a system, dS , could be either positive or negative, depending on the direction of heat transfer (to or from the system). For reversible processes $\delta S_{gen} = 0$. It is noted that for an adiabatic process, $\delta Q = 0$, the increase in entropy is always associated with the irreversible paths.

Contrary to energy, the total entropy increases and is then not conserved. The reversible process is an ideal process and it never really occurs. Therefore, an amount of the irreversibility is always there in the system, i.e., the entropy of the isolated system still goes on increasing; it never reduces. This concept is referred to as the principle of the increase of entropy: the entropy of a closed system (a) never decreases and (b) tends to grow up by the system irreversibilities.

According to (Nag, 2013), an irreversible process always tends to take the isolated system to a state of greater disorder, and this system always tends to a state of greater entropy. Thus, there is a relation between entropy and chaos.

In the late 1800s, Boltzmann derived an statistical mechanical interpretation for entropy: the entropy S of a ideal gas is a state function of the number of microstates W possible for the molecules given the macrostate (defined by temperature, volume and pressure). Thus, his entropy is defined as

$$S = k \log W \quad (2)$$

In which k is the Boltzmann's constant. Gibbs (Jaynes, 1965) extended the concept of Boltzmann entropy to the cases in which the microstates are not equally likely:

$$S = -k \sum_i p_i \log p_i \quad (3)$$

where p_i is the probability of the i -nth microstate (if the W microstates are equally likely, then $p_{i=(1,2,3,\dots,n)} = 1/W$ and Eq. 19 is the same as the Boltzmann entropy).

2.2 Information theory approach

Shannon (Shannon, 1948) developed in 1948 the Information Theory. One of the main concepts of this theory is the entropy of a discrete probability distribution. Let X be a discrete random variable which can assume n different values (states). Then, the entropy H of X is

$$H(X) = - \sum_{i=1}^n p_i(x) \log p_i(x) \quad (4)$$

Generally, the base of the logarithm is 2, and the entropy is measured in bits (however, this is just a way to informing in which base it is, since entropy is dimensionless).

Shannon obtained the name “entropy” from von Neumann itself, as he related (Tribus and McIrvine, 1971):

“My greatest concern was what to call it. I thought of calling it ‘information’, but the word was overly used, so I decided to call it ‘uncertainty’. When I discussed it with John von Neumann, he had a better idea. Von Neumann told me, ‘You should call it entropy, for two reasons. In the first place your uncertainty function has been used in statistical mechanics under that name, so it already has a name. In the second place, and more important, no one really knows what entropy really is, so in a debate you will always have the advantage.’ ”

The motivation of Shannon was associated with the generation and transmission of symbols, like letters (encoded in some type of electrical signal). As letters do not appear with the same probability in any language (in English, the most common letter is “e” and the most common word is “the” (Newman, 2005)), it is possible to compress messages. For example (Cover and Thomas, 2012), suppose we have to send 8 letters with frequencies 1/2, 1/4, 1/8, 1/16, 1/64, 1/64, 1/64 and 1/64. Using a binary coding, initially, one could assume it is needed 3 bits (000, 001, 010, 011, 100, 101, 110, 111). However, since their frequencies differ, it is possible to encode as 0, 10, 110, 1110, 11101, 111101, 111110 and 111111. The average number of bits would be 2 bits. Therefore, one of interpretations of entropy in information theory is the measurement of the ultimate data compression.

This result is exactly what is obtained using Eq. 4, which give the following interpretation for entropy (Shannon): in any alphabet of n symbols with frequencies distributed as X , a message with m symbols encodes n^m states. Then, let $-\log p_i(x)$ be the information content of the event i . Then, the entropy is the average information for the frequencies:

$$H(X) = E[-\log p(X)] \quad (5)$$

The entropy of a system (or a message, or an event) in bits can be thought of as the number of yes/no questions needed to fully describe it. This fact was also known in chemistry in 1930 (Ben-Naim, 2008). The Gibbs entropy then measures the number of these questions needed to fully specify the microstate, given the macrostate.

A source of confusion between thermodynamical entropy and information entropy is the difference with the unit: while the latter is dimensionless, the former has units of J/K. Actually, the Boltzmann (Gibbs) entropy is the same as Shannon, except for the Boltzmann constant, which has units of J/K, giving the entropy unit J/K. However, it has been shown that this unit is historically connected with the definition of the Kelvin temperature system: the Lagrangian of temperature in statistical mechanics has units of energy (Leff, 1999), and is common in plasma physics to express temperature in eV (Bernard et al., 2018; Bagryansky et al., 2015). In this more generic approach, thermodynamical entropy is dimensionless, and the difference between the Shannon and the Gibbs entropies is merely the Boltzmann’s constant.

2.3 Maxwell’s Demon and the Landauer’s Principle

In 1867, Maxwell developed a mental experiment that supposedly violates the second law of thermodynamics. In his own words (Bennett, 1987):

“... if we conceive of a being whose faculties are so sharpened that he can follow every molecule in its course, such a being, whose attributes are as essentially finite as our own, would be able to do what is impossible to us. For we have seen that molecules in a vessel full of air at uniform temperature are moving with velocities by no means uniform, though the mean velocity of any great number of them, arbitrarily selected, is almost exactly uniform. Now let us suppose that such a vessel is divided into two portions, A and B, by a division in which there is a small hole, and that a being, who can see the individual molecules, opens and closes this hole, so as to allow only the swifter molecules to pass from A to B, and only the slower molecules to pass from B to A. He will thus, without expenditure of work, raise the temperature of B and lower that of A, in contradiction to the second law of thermodynamics.”

Several approaches to solve this paradox have been proposed. The most successful one is the Landauer’s principle (Landauer, 1961). It states

that every irreversible process (defining a logically irreversible device when its output does not uniquely define the inputs), such as the erasing of a logical bit, is accompanied by an increase of at least $kT \log 2$ J of heat. Therefore, by Eq. 1, there is an increase of entropy in the system, imposing fundamental limits to computation. Moreover, it can be used to solve the Maxwell’s Demon paradox: the demon has to, at some point, “forget” the state of the particles, to update the information. This process of forgetting generates heat and entropy, since it is an irreversible phenomena.

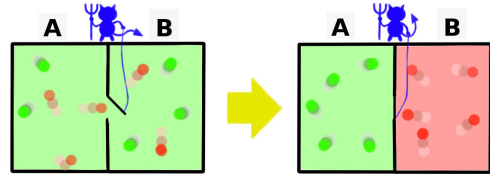


Figure 1: The Maxwell’s Demon: a being who knows the velocity of every particle in the box, and can select their passages using a opening in the wall that divides it could separate those with high energy from those with low energy without realizing work, violating the second law of thermodynamics. Actually, the demon has to forget past states of the system, and by the Landauer’s Principle, this process generates heat (at least $kT \log 2$ J per bit erased) and entropy.

2.4 Kolmogorov Complexity (KC)

Complexity is not the same as entropy (Carroll, 2017). Entropy measures the number of different combinations of microstates needed to specify the macrostate; complexity is the measure of the description of an object. Mathematically, this measure is realized by means of the Kolmogorov Complexity (KC) in a universal Turing Machine (a finite state machine that has an input of symbols of a finite alphabet and a process then, returning a new set of symbols). The KC $K(s)$ of the string s is the number of units of information (bits, for example) of the smallest algorithm in a language that can reproduce the object. This measure of complexity has in its core an interrogation about randomness. If a string is deterministic, then its KC is low, since the code that generate it is simple. For example, the string “001001001001001” and the string “011001101111011” both have 15 bits, but the first one can be coded as “repeat (001) 5 times”, and the second one seems to be random, and the code to generate it will have to contain the entire string. One could use the string to represent physical objects, and the KC would be a measure of complexity of a physical system.

One classical example used in KC is the Mandelbrot’s set. Fig. 2 shows the graph for $f(z) =$

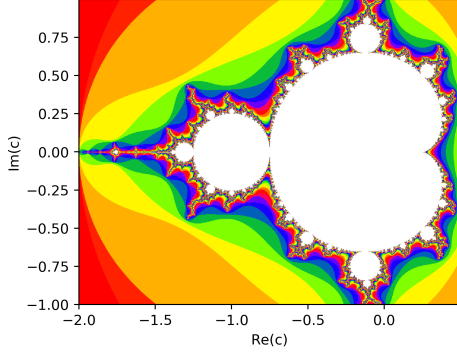


Figure 2: A Mandelbrot's set is a recursive complex-domain sequence, that can be zoomed infinitely and still show the same patterns. However, it is not complex in the sense of Kolmogorov, since it can be generated by recursion. Therefore its KC is very low. The python code that generates this figure has only 507 bytes.

$z_{n+1} = z_n^2 + c$. The figure seems to be very complex and of hard description, with several colors and a intricate pattern for the perimeter. However, since it can be generated through iterations in the complex plane, its KC turns out to be small and the figure itself is very simple.

Shannon's entropy and KC holds a remarkable relationship. Using the Kraft inequality, it can be show that (Cover and Thomas, 2012):

$$E \left[\frac{1}{n} K(X^n|n) \right] \rightarrow H(X) \quad (6)$$

and therefore, the compressibility of KC in the universal computer goes to the entropy limit.

However, in a physical sense, it is important to notice that, even though the entropy is always increasing, complexity follows another pattern (Fig. 3): it was low in the early ages of the universe, and will be low again in the far future, when star production ends (roughly 100 trillion years from now) and the cosmos goes to heat death (Frautschi, 1982). It is now, in this intermediary epoch, in which complexity exists, with structures like planets, galaxies and living beings. Entropy, however, has always been increasing, since the Big Bang (Fig. 3).

These complex structures, that uses energy to decrease entropy locally, can only exist in a time when entropy has not reached its maximum (heat death), and therefore, it is possible to use the gradient of energy (consequence of the fact that the universe is not in equilibrium) to generate spontaneously complex systems.

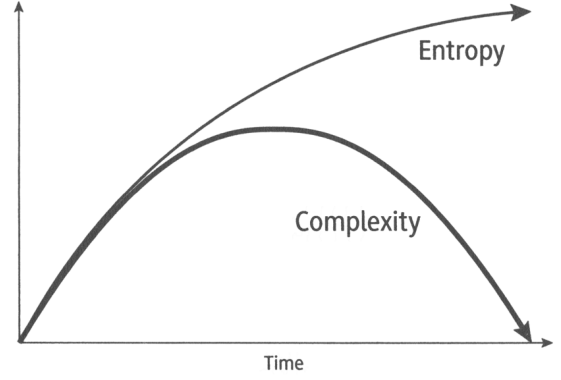


Figure 3: Complex structures need to exchange energy with the environment to reduce their entropy and increase the entropy of the environment. Therefore, they can only form in times in which entropy can increase. Entropy, however, is always increasing by the second law of thermodynamics. The early universe and the far future universe are both simple, uniform hot and dense state in the beginning and empty space in the end, but the entropy was low in the past and will be at its maximum at the end by heat death. Source of figure: (Carroll, 2017).

3 Results

3.1 Deriving Boltzmann's and Gibbs entropy

Lets take the system for the ideal gas and divide it in two parts. It is know that $S = S_1 + S_2$ and $W = W_1 W_2$. Therefore, the deduction is as follows:

$$S(W_1) + S(W_2) = S(W_1 W_2) \quad (7)$$

Deriving both sides with respect to W_1 and keeping W_2 constant results in

$$S'(W_1) = W_2 S'(W_1 W_2) \quad (8)$$

Deriving now in W_2 keeping W_1 constant, applying the chain rule:

$$0 = S'(W_1 W_2) + W_1 W_2 S''(W_1 W_2) \quad (9)$$

$$0 = S'(W) + W S''(W) \quad (10)$$

Replacing $S'(W) = f(W)$:

$$f(W) + W \frac{df(W)}{dW} = 0 \quad (11)$$

$$f(W)dW + Wdf(W) = 0 \quad (12)$$

$$(fW)' = 0 \quad (13)$$

Integrating both sides, returns

$$fW = k \quad (14)$$

which is the same as

$$W \frac{dS}{dW} = k \quad (15)$$

$$\int dS = k \int \frac{dW}{W} \quad (16)$$

$$S = k \log W + c \quad (17)$$

But it is known that a crystal at 0 K has null entropy (Nernst, 1907), and only one microstate. Replacing this fact in the equation 17:

$$0 = k \log 1 + c \quad (18)$$

from which we conclude that $c = 0$ and $S = k \log W$ is the entropy of an ideal gas, where k is the Boltzmann constant. Gibbs (Jaynes, 1965) extended the concept of Boltzmann entropy to the cases in which the microstates are not equally likely:

$$S = -k \sum_i p_i \log p_i \quad (19)$$

where p_i is the probability of the i -nth microstate (if the W microstates are equally likely, then $p_{i=(1,2,3,\dots,n)} = 1/W$ and Eq. 19 is the same as the Boltzmann entropy).

3.2 Information theoretic proof that Gibbs entropy is the same as Clausius

With the development of information theory in the twentieth century and the concept of maximum entropy for statistical mechanics (Jaynes, 1957), which states by the second law of thermodynamics that a system in thermodynamic equilibrium has reached its maximum entropy (and therefore, it is in the macrostate that has the most microstates, corresponding to gas velocities), it is possible to show that Shannon entropy is the same as Clausius entropy as well.

Using Eq. 19, and the unitarity principle, $\sum_i p_i = 1$, in which i is the i -nth state, we can write the average energy of a system is

$$\langle E \rangle = \sum_i p_i E_i = U \quad (20)$$

Applying Lagrange multipliers, we have

$$\begin{aligned} \mathcal{L} = & -k \sum_i p_i \log p_i + \lambda_1 \left(\sum_i p_i - 1 \right) \\ & + \lambda_2 \left(\sum_i p_i E_i - U \right) \end{aligned} \quad (21)$$

Differentiating and equating zero

$$-k \log p_i - k + \lambda_1 + \lambda_2 E_i = 0 \quad (22)$$

Isolating p_i

$$p_i = \exp \left(\frac{-k + \lambda_1 + \lambda_2 E_i}{k} \right) \quad (23)$$

Using the canonical partition function Z (Baxter, 2016), defined as

$$Z = \sum_i \exp \left(\frac{\lambda_2}{k} E_i \right) \quad (24)$$

The partition function combines state functions, such as temperature and energy for the microstates, and has a central role in statistical mechanics (Bo-sture et al., 1985). Differentiating $\log Z$ with respect to λ_2 returns

$$\frac{\partial \log Z}{\partial \lambda_2} = \frac{\langle E \rangle}{k} \quad (25)$$

Using unitarity again, Eq.23 can be written as

$$\exp \left(\frac{-k + \lambda_1}{k} \right) Z = 1 \quad (26)$$

Therefore,

$$\log \left(\frac{1}{Z} \right) + 1 = \frac{\lambda_1}{k} \quad (27)$$

Rewriting Eq. 19 in terms of Z , results in

$$S = -k \sum_i p_i \left(\frac{\lambda_2}{k} E_i - \log Z \right) \quad (28)$$

Using 25 in 28, give us

$$\begin{aligned} S = & -\lambda_2 \sum_i p_i E_i + k \log Z \sum_i p_i \\ = & -\lambda_2 U + k \log Z \end{aligned} \quad (29)$$

Using the definition of thermodynamics temperature (Callen, 1998)

$$\frac{1}{T} = \frac{\partial S}{\partial U} \quad (30)$$

Since $\frac{\partial S}{\partial U} = -\lambda_2$, Eq. 19 can be written as

$$S = \frac{U}{T} + k \log Z \quad (31)$$

Now, lets change the energy of the system by δQ . Every microstate will increase its energy by q_i . Calculating the change in the entropy results in

$$dS = \frac{\delta U}{T} + k \delta \log Z \quad (32)$$

Calculating the second term

$$\delta \log Z = \frac{d \log Z}{dZ} \delta Z = \frac{\delta Z}{Z} \quad (33)$$

Noticing that $Z = \sum_i \exp(-E_i/kT)$, the new partition function can be written as

$$Z = \sum_i \exp\left(-\frac{E_i + q_i}{kT}\right) \quad (34)$$

Applying Taylor expansion to $e^{-q_i/kt}$, since q_i is infinitesimal, the partition function is

$$Z = \sum_i \exp\left(\frac{E_i}{kT}\right) \left(1 - \frac{q_i}{kT}\right) = Z_0 + \delta Z \quad (35)$$

Therefore, the variation of the partition function is given by

$$\delta Z = -\frac{1}{kT} \sum_i q_i \exp\left(-\frac{E_i}{kT}\right) \quad (36)$$

Using the first law of thermodynamics, the change in U can be expressed as

$$\delta U = \sum_i \delta E_i p_i + \sum_i q_i p_i = \delta Q + \delta W \quad (37)$$

Calculating $\delta \log Z$, replacing 36 in 33:

$$\delta \log Z = -\frac{1}{kT} \sum_i p_i q_i \quad (38)$$

This value is exactly $\delta W/kT$. Replacing this relation in Eq. 32, we get

$$dS = \frac{\delta Q}{T} \quad (39)$$

which is the Clausius first definition of entropy.

3.3 Simulation of the Maxwell's Demon

The algorithm that generates the simulation is available in (David, 2017). It was developed using *Tkinter* python library, and simulates the idea proposed by Maxwell. There is a wall controlled by a demon (in this case, the software), with the power to, analysing the velocities and trajectories of every particle, it can separate the hot ones (the reds, with high kinetic energy) from the cold ones (the blues, with low kinetic energy).

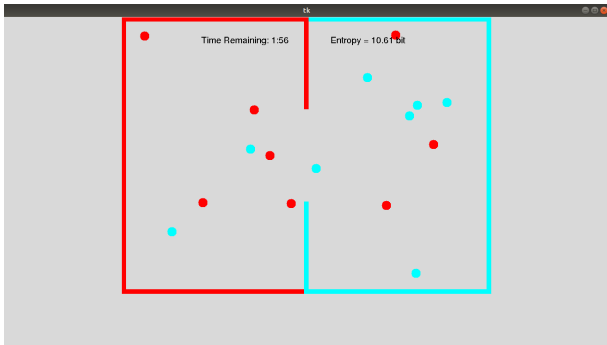


Figure 4: Initial state of the simulation, with the hot and cold molecules evenly distributed (high entropy state).

The demon must track all velocities and positions for every particle in the system, in order to change their positions in the boxes, and in every instant, he must have in his memory this information. Let's represent the information as in a computer, using bits. So, in the first instant, the demon has part of his memory used, and part erased, say "00000...10110". For every iteration of the particles, more memory needs to be allocated, until the state "11111...11111". In this moment, to still having tracking abilities, the demon has to erase his memory, exchanging heat with the world (at least $kT \log 2$ J/bit, but the current technology generates millions times this value) and increasing his entropy, solving the "paradox" created by Maxwell.

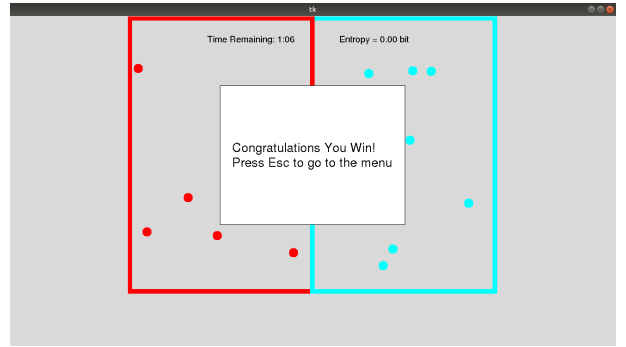


Figure 5: Final state of the simulation, with the hot and cold molecules separated by the wall controlled by the Maxwell's Demon (low entropy state).

4 Discussion and conclusion

Entropy, as a concept, is misused in great part due to misunderstanding its mathematical or physical meaning. It is expected that, with this paper, the reader can clearly understand what the "different" types of entropy are, how they are deeply connected and its relationship with complexity.

Readers working with complex systems can use the two terms, entropy and complexity, to relate the objects being studied with the physical notions here deducted, and therefore merge different areas of science, such as biology, economics, linguistics and computer science. Being a concept so fundamental to science, due to its own nature, emerging from counting problems, it is not strange that so many fields of knowledge are using it.

The interdisciplinarity of the topic has his challenges, due to the different mathematical level of each field, but with the theoretical exposition, together with the mathematical proofs, we expect to reach different target audience with the same discussion about the meaning of entropy, and how to use it in their areas.

5 Acknowledgement

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001.

References

- Adami, C. (2002). What is complexity?, *BioEssays* **24**(12): 1085–1094.
- Adami, C. (2016). What is information?, *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* **374**(2063): 20150230.
- Bagryansky, P., Shalashov, A., Gospodchikov, E., Lizunov, A., Maximov, V., Prikhodko, V., Soldatkina, E., Solomakhin, A. and Yakovlev, D. (2015). Threefold increase of the bulk electron temperature of plasma discharges in a magnetic mirror device, *Physical review letters* **114**(20): 205001.
- Baxter, R. J. (2016). *Exactly solved models in statistical mechanics*, Elsevier.
- Ben-Naim, A. (2008). *A Farewell to Entropy: Statistical Thermodynamics Based on Information: S*, World Scientific.
- Bennett, C. H. (1987). Demons, engines and the second law, *Scientific American* **257**(5): 108–117.
- Bernard, T. N., Shi, E. L., Gentle, K., Hakim, A., Hammett, G. W., Stoltzfus-Dueck, T. and Taylor, E. I. (2018). Gyrokinetic continuum simulations of plasma turbulence in the texas helimak, *arXiv preprint arXiv:1812.05703*.
- Bo-sture, S. et al. (1985). *Coherent states: applications in physics and mathematical physics*, World scientific.
- Bossomaier, T., Barnett, L., Steen, A., Harré, M., d’Alessandro, S. and Duncan, R. (2018). Information flow around stock market collapse, *Accounting & Finance* **58**: 45–58.
- Bousso, R. (2018). Black hole entropy and the bekenstein bound, *arXiv preprint arXiv:1810.01880*.
- Brooks, D. R., Wiley, E. O. and Brooks, D. (1988). *Evolution as entropy*, University of Chicago Press Chicago.
- Callen, H. B. (1998). *Thermodynamics and an introduction to thermostatistics*.
- Campbell, J. (1982). *Grammatical man: Information, entropy, language, and life*, Simon and Schuster New York.
- Caro, J. A., Valentine, K. G. and Wand, A. J. (2018). Role of conformational entropy in extremely high affinity protein interactions, *Biophysical Journal* **114**(3): 67a.
- Carroll, S. (2017). *The big picture: on the origins of life, meaning, and the universe itself*, Penguin.
- Clausius, R. (1960). *On the Motive Power of Heat, and on the Laws which Can be Deduced from it for the Theory of Heat*, Annalen der Physik, Dover.
- URL:** 1
- Cover, T. M. and Thomas, J. A. (2012). *Elements of information theory*, John Wiley & Sons.
- David, J. (2017). Entropy game: Maxwell’s demon game in python, <https://github.com/davidj11/entropy-game>.
- De Martino, A. and De Martino, D. (2018). An introduction to the maximum entropy approach and its application to inference problems in biology, *Heliyon* **4**(4): e00596.
- Degaetano-Ortlieb, S. and Teich, E. (2017). Modeling intra-textual variation with entropy and surprisal: topical vs. stylistic patterns, *Proceedings of the Joint SIGHUM Workshop on Computational Linguistics for Cultural Heritage, Social Sciences, Humanities and Literature*, pp. 68–77.
- Frautschi, S. (1982). Entropy in an expanding universe, *Science* **217**(4560): 593–599.
- Greven, A., Keller, G. and Warnecke, G. (2014). *Entropy*, Vol. 47, Princeton University Press.
- Gu, R., Xiong, W. and Li, X. (2015). Does the singular value decomposition entropy have predictive power for stock market?—evidence from the shenzhen stock market, *Physica A: Statistical Mechanics and its Applications* **439**: 103–113.
- Hayflick, L. (2007). Entropy explains aging, genetic determinism explains longevity, and undefined terminology explains misunderstanding both.
- He, D., Wang, X., Li, S., Lin, J. and Zhao, M. (2016). Identification of multiple faults in rotating machinery based on minimum entropy deconvolution combined with spectral kurtosis, *Mechanical Systems and Signal Processing* **81**: 235–249.
- Henderson, L. (2003). The von neumann entropy: A reply to shenker, *The British journal for the philosophy of science* **54**(2): 291–296.

- Jaynes, E. T. (1957). Information theory and statistical mechanics, *Physical review* **106**(4): 620.
- Jaynes, E. T. (1965). Gibbs vs boltzmann entropies, *American Journal of Physics* **33**(5): 391–398.
- Kolmogorov, A. N. (1963). On tables of random numbers, *Sankhyā: The Indian Journal of Statistics, Series A* pp. 369–376.
- Kovalev, A. V. (2016). Misuse of thermodynamic entropy in economics, *Energy* **100**: 129–136.
- Landauer, R. (1961). Irreversibility and heat generation in the computing process, *IBM journal of research and development* **5**(3): 183–191.
- Leff, H. S. (1999). What if entropy were dimensionless?, *American Journal of Physics* **67**(12): 1114–1122.
- Maldacena, J. (2018). Black hole entropy and quantum mechanics, *arXiv preprint arXiv:1810.11492*.
- Martyushev, L. (2013). Entropy and entropy production: old misconceptions and new breakthroughs, *Entropy* **15**(4): 1152–1170.
- Morowitz, H. (1986). Entropy and nonsense, *Biology and Philosophy* **1**(4): 473–476.
- Nag, P. (2013). *Engineering thermodynamics*, Tata McGraw-Hill Education.
- Nernst, W. (1907). *Experimental and theoretical applications of thermodynamics to chemistry*, C. Scribner’s sons.
- Newman, M. E. (2005). Power laws, pareto distributions and zipf’s law, *Contemporary physics* **46**(5): 323–351.
- Reynar, J. C. and Ratnaparkhi, A. (1997). A maximum entropy approach to identifying sentence boundaries, *Proceedings of the fifth conference on Applied natural language processing*, Association for Computational Linguistics, pp. 16–19.
- Rostaghi, M. and Azami, H. (2016). Dispersion entropy: A measure for time-series analysis, *IEEE Signal Processing Letters* **23**(5): 610–614.
- Schrodinger, E. (1968). Order, disorder and entropy, *Modern Systems Research for the Behavioral Scientist. Chicago: Aldine* pp. 143–146.
- Shannon, C. E. (1948). A mathematical theory of communication, *Bell system technical journal* **27**(3): 379–423.
- Soubane, D., El Garah, M., Bouhassoune, M., Tirbiyine, A., Ramzi, A. and Laasri, S. (2018). Hidden information, energy dispersion and disorder: Does entropy really measure disorder?, *World* **8**: 197–202.
- Sozbilir, M. (2003). What studentsâ understand from entropy?: A review of selected literature, *Journal of Baltic Science Education* **2**(1): 21–27.
- Tame, J. R. (2019). On entropy as mixed-up-ness, *Approaches to Entropy*, Springer, pp. 153–170.
- Tribus, M. and McIrvine, E. C. (1971). Energy and information, *Scientific American* **225**(3): 179–188.
- Wehrl, A. (1978). General properties of entropy, *Reviews of Modern Physics* **50**(2): 221.
- Weilenmann, M., Kraemer, L., Faist, P. and Renner, R. (2016). Axiomatic relation between thermodynamic and information-theoretic entropies, *Physical review letters* **117**(26): 260601.
- Wright, P. (1970). Entropy and disorder, *Contemporary Physics* **11**(6): 581–588.
- Xiao, M., Du, P., Horne, K., Hu, C., Li, Y.-R., Huang, Y.-K., Lu, K.-X., Qiu, J., Wang, F., Bai, J.-M. et al. (2018). Supermassive black holes with high accretion rates in active galactic nuclei. vii. reconstruction of velocity-delay maps by the maximum entropy method, *The Astrophysical Journal* **864**(2): 109.
- Zeeshan, A., Hassan, M., Ellahi, R. and Nawaz, M. (2017). Shape effect of nanosize particles in unsteady mixed convection flow of nanofluid over disk with entropy generation, *Proceedings of the Institution of Mechanical Engineers, Part E: Journal of Process Mechanical Engineering* **231**(4): 871–879.