Entropy in Thermodynamics and Information Systems

Introduction

In 1865 Clausius found dS=dq(rev)T for measuring entropy change in thermodynamic processes quantitatively. He proposed what we call the second law of thermodynamics. With these thermodynamic laws we are able to characterize thermodynamic systems for their temperature, energy, and entropy.

For more than a century eminent physicists, mathmematicians, and chemists who were capable of explaining entropy failed to do so. Following comments from them are more revealing about their inability to help their readers:

"…the notion of entropy, …may repel beginners as obscure and difficult of comprehension." (Gibbs, 1873).

"…entropy…it is to be feared that we shall have to be taught thermodynamics for several generations before we can expect beginners to receive as axiomatic the theory of entropy." (Maxwell, 1878).

"You should call it entropy, because nobody knows what entropy really is …John von Neumann to Claude Shannon." (ca. 1948) (Tribus and McIrvine, 1971).

In this paper we will present explanations of Ludwig Boltzmann's Entropy and Claude E. Shannon's Entropy. Boltzmann's Entropy will be giving us the role of entropy in thermodynamics systems while Shannon's Entropy will be explaining limits of compression of any communication. We will be seeing how these entropies are similar even though at first glance they seem unrelated.

Brief History of Entropy

In 1783 Lazarne Carnot published *Essai sur les machines en général.* He proposed for the first time in history the idea that in a mechanical apparatus shocks and frictions dissipated a sizable fraction of usable work. He left us a famous theorem (Carnot's theorem) and gave his nephew Sadi a first glimpse of entropy.

Sadi Carnot wrote in 1824 a booklet entitled *Réflexions sur la puissance motrice du feu et sur les machines propres à développer cette puissance*. In this book, the Carnot Cycle, the Carnot Heat Engine, Carnot's theorem and thermodynamic efficiency are all explained in detail for the first time.

In 1850 Rudolf Clausius, a German professor, re-re-discovered Sadi Carnot's work reading a paper written by the French engineer Clapeyron. Clausius stated in his paper called *Über die bewegende Kraft der Wärme* that there was a contradiction between the Carnot's theorem and the conservation of energy. Before anything, he classified heat in three categories:

1. Heat used to rise the temperature of a body
2. Heat used to produce "exterior" work (typically exerted by a piston)
3. Heat used to increase the "interior" energy of a body

After stating the first law of thermodynamics in the modern fashion, Clausius went further proving that in a transformation the passage of a quantity of heat *Q* from a temperature T1 to a temperature T2 the "equivalence-value" (now called entropy) has a value Q(1/T2 – 1/T1). According to Carnot's theorem, in a cycle this value must be equal or less than zero. The equivalence is obtained if the cycle is made of reversible transformations. In formulae:

$$\int \frac{\delta Q}{T} \leq 0$$

Later on, he called this quantity *entropy* (Greek: “inner transformation”) in analogy to the word *energy* (Greek: “inner work”).

In 1859,James Clark Maxwell, a young and brilliant Scottish physicist, after reading Clausis’ papers put forward a formula for the evaluation of molecular velocities, which gives the proportion of molecules having a certain velocity in a specific range. This is still considered one of the most important formulae ever conceived. In fact, not only Maxwell provided a way to relate the number of molecules, the volume and the pressure of a gas, but he even introduced *statistics* into classical mechanics. In fact, given the ridiculous large number of particles, it is too hard to follow the evolution of the system completely.

Boltzmann's Entropy

In 1871, the Austrian physicist Ludwig Boltzmann generalized Maxwell’s formula into the now-called Maxwell-Boltzmann distribution.

Probably, one of the most important contribution of Boltzmann is in fact his explanation of entropy. At the turn of the century, he proposed a statistical interpretation of this quantity. Boltzmann’s idea was to use statistical mechanics to describe the average properties of a system, namely a gas or a solid. When observed, a system is found in a certain configuration, called *state*. Now, many slightly different states might share on average the same values of pressure and temperature. It was Max Planck who based on Boltzmann results formulated of what was later called Boltzmann expression: $S = \mathrm{k} \ln W$

Here is S the entropy, k is Boltzmann constant, ln is the natural logarithm and W is the amount of realization possibilities the system has. The value of W is basically a measure of how likely a system can exist given certain characteristics. The constant of proportionality k is now called Boltzmann constant and in the kinetic theory of gases it relates the temperature to the average velocity of the molecules: $\frac{1}{2}m < v^2 >= \frac{3}{2}k_B T$

Imagine you have a deck of cards with 4 identical cards. The deck as a total can be described with parameters such as the number of cards, thickness of the deck, weight and so on. With four cards we have 4x3x2x1 = 24 possible configurations that all lead to the same (in terms of the parameters above) deck of cards. Therefore in this case $W = 24$. The Boltzmann constant, k, equals to 1.4 10^{-23} J/K and the entropy S is then kln24 = 4.4 10^{-23} J/K. The more possibilities a given system has to establish itself (and with the many atoms we have in one gram of material there are many possibilities!) the more likely it will be that we will indeed observe that system and the higher the entropy will be.

Shannon's Entropy

In Shannon information theory, the entropy is a measure of the uncertainty over the true content of a message, but the task is complicated by the fact that successive bits in a string are not random, and therefore not mutually independent, in a real message. Also note that "information" is not a subjective quantity here, but rather an objective quantity, measured in bits.

For a random variable X with values in a finite set $\mathcal{X}$, *Shannon entropy* is defined as (Shannon 1948):

$$H(X) = -\sum_{x \in \mathcal{X}} p(x) \log_2 p(x) \geq 0$$

It quantifies the *unevenness* of the probability distribution *p*. In particular, the minimum $H(X)=0$ is reached for a constant random variable, a variable with a determined outcome, which reflects in a fully localized probability distribution $p(x_0) = 1$ and $p(x) = 0$ for $x \neq x_0$. At the opposite, $H(X)$ is maximal, equal to $\log_2(|X|)$, for a uniform distribution. $H(X)$ is also denoted:

$$S(p) = -\sum_{i=1}^{|\mathcal{X}|} p(x_i) \log_2 p(x_i)$$

which underlines the fact that entropy is a feature of the probability distribution p. Entropy does not depend on the graph $x \to p(x)$, it is not a feature of the random variable itself but only of the set of its probability values. This property reflects in a permutation invariance of $H(X)$: let the variable $\sigma.X$ obtained by a permutation of the states, namely $\text{Prob}(\sigma.X = x_{\sigma(i)}) = p(x_i)$, then $H(X) = H(\sigma.X)$.

A sequence of N coin tosses (of an assumed unbiased coin) has 2^N possible out- comes: there is an uncertainty in the outcome that we measure by the "entropy" which is the log of the number of possible equally likely outcomes:

$$S = \log 2^N = N \log 2.$$

We could use the sequences e.g. heads,heads,. . . tail to send a message, and we could send 2^N different messages. The "information capacity" of this scheme is again measured by the log of the number of possible messages

$$I = \log 2^N = N \log 2.$$

In this context we would often use base 2 for the log and say there are N bits of information. An alternative point of view is that the measurement of a particular result i.e. sequence of heads and tails has told us something about the system and we have learned N log 2 bits of information The ideas of uncertainty of outcome (entropy) and what has been learned from a measurement (information) are complementary.

Conclusion

The work done, primarily by Boltzmann & Gibbs, on the foundations of statistical mechanics, is of profound significance that can hardly be overestimated. In the hands of Clausius and his contemporarys, entropy was an important, but strictly thermodynamic property. Outside of physics, it simply had no meaning. But the mathematical foundations of statistical mechanics are applicable to any statistical system, regardless of its status as a thermodynamic system. So it is by the road of statistical mechanics, that we are able to talk about entropy in fields outside of thermodynamics.

In *A Mathematical Theory of Communication*, appendix 2, Shannon proves his Theorem 2, that this Boltzmann entropy is the only function which satisfy's the requirements for a function to measure the uncertainty in a message (where a "message" is a string of binary bits). In this case, the constant **k** is recognized as only setting the units; it is arbitrary, and can be set equal to exactly 1 without any loss of generality. In this case the probabilty $\mathbf{P_i}$ is the probability for the value of a given bit (usually a binary bit, but not necessarily).

Boltzmann constant in Boltzmann's entropy formula is different from p in Shannon's Entropy formula. The p_k's in Shannon Entropy formula may change because of transmission errors while Boltzmann's is constant.

References

- Lambert, Frank L. "The Conceptual Meaning of Thermodynamic Entropy in the 21st Century." *International Research Journal of Pure & Applied Chemistry* (n.d.): 1-2.
- "A Brief History of Entropy" Web log post. *Beyond the Standard Model Pub*. N.p., n.d. Web. <http://disipio.wordpress.com/tag/heat/>.

- Schmitz, John E.J. "What Is Entropy? Boltzmann’s Interpretation." Weblog post. *The Second Law of Life*. N.p., n.d. Web. <http://secondlawoflife.wordpress.com/2007/06/03/what-is-entropy-boltzmanns-interpretation/>.
- Thompson, Tim. "The Definitions of Entropy." *What Is Entropy?* N.p., n.d. Web.
- Shannon, Claude E. "A Mathematical Theory of Communication." (1948): n. pag. Web.

Tolga Gölbaşı 12400002006