Analizing the Entropy concept

Angel Garrido, Facultad de Ciencias de la UNED

Abstract

Our paper analyzes some aspects of the evolving concepts of a very important Uncertainty Measure, the so-called Entropy. We need to obtain new ways to model adequate conditions or restrictions, constructed from vague pieces of information. For this, it will be very necessary to classify the different types of measures; in particular, to consider certain fuzzy measures. And previously, fixing well the emplacement of such ideas. For such reason, we attempt to advance now some historical notes, perhaps interesting in the analysis of Entropy.

Keywords: Measure Theory, Fuzzy Measures, Entropy.

Mathematics Subject Classification: 68R10, 68R05, 05C78.

1. Introduction to Entropy

As you have surely heard, according to the story Tribus tells us [11], Shannon didn't know what to call his measure, so he asked Von Neumann,

"- 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, he had a better idea.

John 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, nobody knows what entropy really is, so in a debate you will always have the advantage ´".

2. Some historical quotations

"I propose to name the quantity S the entropy of the system after the Greek word *trope,* transformation. I have deliberately chosen the word entropy to be as similar as possible to the word energy: the two quantities to be named by these words are so closely related in physical significance that a certain similarity in their names appears to be appropriate". *Rudolph Clausius* (1865).

"The fundamental problem of communication is that of reproducing at one point, either exactly or approximately, a message selected at another point." *Claude E. Shannon* (1916-2001), "A Mathematical Theory of Communication", famous paper of 1948.

[1] AMO - Advanced Modeling and Optimization. ISSN: 1841-4311

Because the basic problem is to reconstruct, as closely as possible, the input signal after observing the received signal at the output.

The development of the idea of entropy of random variables and processes by Shannon, provided the foundations of *IT (Information Theory),* and also of *Ergodic Theory.* Recall that it is a branch of Mathematics that studies Dynamical Systems with an invariant measure.

3. Developing Entropy theories

Entropy and related information measures provide descriptions of the long term behavior of random processes, and that this behavior is a key factor in developing the

Coding Theorems of IT

The contributions of *Andrei N. Kolmogorov* (1903-1987) to the mathematical I T produces great advances to the Shannon formulations, proposing a new complexity theory, now translated to Computer Sciences. According such theory, the complexity of a message is given by the size of the program necessary to be possible the reception of such message.

From these ideas, Kolmogorov also analyzes the entropy of literary texts. More concretely, on Pushkin poetry. Such entropy appears as a function of the semantic capacity of the texts, depending of factors as their extension and also the flexibility of the corresponding language.

And we need mentionate to *Prof. Solomon Marcus,* tireless researcher in Computational Linguistics, and Information Theory. Because he has studied the Entropy in the poetical work of Mihai Eminescu, national Romanian poet.

Also may be mentioned *Norbert Wiener* (1894-1964), considered the founder of Cybernetics, who in 1948 also propose a similar vision of such problem.

But the approach used by Shannon differs from that of Wiener in the nature of the transmitted signal and in the type of decision made at the receiver.

In the Shannon model messages are first encoded and then transmitted, whereas in the Wiener model the signal is communicated directly through the channel without being encoded.

Another measure would be due to *R. A. Fischer* (1890-1962), the so called *Fisher Information (FI)*, which is another different concept, used applying statistics to estimation. It will be thought of as the amount of information that a message carries about an unobservable parameter.

The initial studies on IT were undertaken by *Harry Nyquist* (1889-1976) in 1924. And later by *Ralph Hartley* (1888-1970), who in 1928 recognized the logarithmic nature of the measure of information.

Later, it appears the key, with the essential *Shannon* and *Wiener* papers.

It is also very interesting the contribution of the Romanian mathematician and economist *Nicholas Georgescu-Roegen* (1906-1994), which studied in London with Karl Pearson.

Its great work was *The Entropy Law and the Economical Process* [4]. In such memorable book, he propose that the second law of thermodynamics governs economic processes. Such ideas permits the development of some new fields, as can be Bioeconomics, or Ecological Economics.

But also some others, as *Kerrige* (1961), studying a different kind of measure, the so called *inaccuracy measure,* involving two probability distributions.

Sibson (1969) studied another divergence measure, also involving two probability distributions, by using the concavity property of Shannon's entropy, calling *information radius.*

More later, *Burbea and Rao* (1982) studied extensively the information radius and its parametric generalization.

Yager (1979), and *Higashi and Klir* (1983), showed the entropy measure as the difference between two fuzzy sets. More concretely, it will be the difference between a fuzzy set and its complement, which is also a fuzzy set.

Pal and Pal (1989) introduce a definition for probabilistic entropy based on an exponential gain function, and used it as basis for defining this measure of fuzziness.

Bhandari and Pal (1993) use Renyi´s entropy to define a non-probabilistic entropy, or measure of fuzziness, for fuzzy sets.

Jumaire (1990) propose a definition, now for differentiable membership functions.

The *Shannon Entropy* is a measure of the average information content one is missing when one does not know the value of the random variable.

This concept proceeds from the Shannon´s 1948 aforementioned famous paper.

So, it represents an absolute limit on the best possible lossless compression of any communication, under constraints, treating messages to be encoded as a sequence of independent and identically distributed (i.i.d.) random variables.

The information that we receive from an observation is equal to the degree to which uncertainty is reduced.

So,

$$I = H\ (before) - H\ (after)$$

Among their main properties, we have

Continuity

The measure H should be continuous, in the sense that changing the values of the probabilities by a very small amount, should only change the H value by a small amount.

Maximality

The measure H will be maximal, if all the outcomes are equally likely. That is, the Uncertainty is highest when all possible events are equiprobable.

Mathematically,

$$H_n\ (p_1,\ p_2,\ \ldots,\ p_n) \leq H_n\ (1/n,\ 1/n,\ \ldots,\ 1/n)$$

For equiprobable events, H will increase with the number of outcomes,

$$H_n\ (1/n,\ 1/n,\ \ldots,\ 1/n) < H_{n+1}\ (1/n{+}1,\ 1/n{+}1,\ \ldots,\ 1/n{+}1)$$

Additivity

The amount of entropy should be independent of how the process is considered, as being divided into parts.

Such functional relationship characterizes the entropy of a system respect to their sub-systems. It demands that the entropy of every system can be identified, and then, computed from the entropies of their sub-systems.

So, we can see now *some classical definitions of Entropy,* according to the different perspectives and applications:

1) Entropy is a measure of the probability of a particular event.

2) Entropy is a measure of the disorder of a system.

3) Entropy measures heat divided by absolute temperature
of a body, i.e.

$$S = Q\ /\ T$$

4) The known vision of Entropy as loss of information.

Entropy is related to the 2nd Law of Thermodynamics, which provides a definition of time's arrow. Such 2nd Law of Thermodynamics states that in a closed system, the entropy function (S) increases, and this increase occurs in time.

The physical notation of S for Entropy is a tribute to *Sadi Carnot* (1796-1832), which is very often described as "the father of Thermodynamics".

Whereas it is usually denoted as *H,* in Computation and Information Theory.

About some apparent "evidences" prescribing that the Shannon information measure is the only possible one, it must be clear that it will be only valid within the more restricted scope of coding problems which the own C. E. Shannon [10] had see in his time.

As pointed out by *Alfred Rényi (1961),* in his essential paper on generalized information measures [9], in other sort of problems other quantities may serve just as well, or even better, as measures of information.

This should be supported either by their operational significance or by a set of natural postulates characterizing them, or, preferably, by both.

Thus, the idea of generalized entropies arises in the scientific literature.

Also we may mencionate to the physicist *John Archibald Wheeler* (1911-2008), saying that:

"I think of my lifetime ... as divided into three periods. In the first, I was in the grip of the idea that

Everything is Particles

I call my second period

Everything is Fields

Now, I am in the grip of a new vision, that

Everything is Information".

4. Some previous ideas

R. Clausius (1865) gives the aforementioned well-known formula

$$S = Q \,/\, T$$

L. Boltzman (1862): *Kinetic Theory.* S as disorder in the energy space.

J. W. Gibbs (1880s): thermodinamical equilibrium corresponds to a maximum of entropy.

C. E. Shannon (1948): *Information Theory.* Entropy (denoted by H) measures the lack of information of a system [10].

Relationship with Thermodynamics

The name of Entropy proceeds indeed from the resemblance between Shannon´s formula and some similar formulae which are usual in Thermodynamics.
So, in Statistical Thermodynamics we will take the Gibbs Entropy.

The standard Boltzmann-Gibbs entropy may be generalized to the so-called *Tsallis Entropy* (1988).

It is also possible to translate the Gibss Entropy to Quantum Physics, giving us the

Von Neumann Entropy (1927)

5. Conclusions

Statistical entropy is a probabilistic measure of uncertainty, or ignorance about data.

Whereas *Information* is a measure of a reduction in that uncertainty.

But must be avoided, according to Frank Lambert (1918-), the interpretation of entropy as disorder. So, instead proposing [8] the notion as *"dispersal of energy".*

We need to develop a usable measure of the information we get from observing that occurs an event with probability p.

For this purpose, we must to ignore particular features of such event, only observing whether or not it happened. So, we can consider the event as the observance of a symbol whose probability of occurring is p.

Whereas the *Entropy of a probability distribution* is just the expected value of the information of such distribution.

All these ideas induces improved tools to advance in Optimization Theory, with tireless researchers as for instance, Neculai Andrei and its group, at CAMO.

References

[1] De Luca, and Termini: "A definition of non-probabilistic entropy in the setting of Fuzzy Set theory". *Inform. and Control*, Vol. 20, pp. 301-312, 1972.

[2] Fang and Tsao: "Entropy Optimization: Shannon Measure of Entropy and Its Properties". *Encyclopedia of Optimization.* (Floudas and Pardalos, eds.). 2nd edition. Springer, 2009.

[3] Garmendia: "The Evolution of the Concept of Fuzzy Measure". *Studies in Computational Intelligence,* Vol. 5, pp. 186-200, 2005.

[4] Georgescu-Roegen: *The Entropy Law and the Economic Process.* Harvard University Press, 1971.

[5] Jaynes: *Probability Theory: The Logic of Science.* Cambridge University Press, 2003.

[6] Jeffreys: *Theory of probability.* Oxford University Press, 1948.

[7] Kullback: *Information Theory and Statistics.* Wiley, 1959.

[8] Lambert: "Entropy is Simple, Qualitatively", *Journal of Chemical Ed.,* Vol. 79, pp. 1241-1246, 2002 [it is also disposable online].

[9] Rényi: "On measures of information and entropy". *Proc. of the 4th Berkeley Symposium on Mathematics, Statistics and Probability,* pp. 547-561, 1961.

[10] Shannon: "A Mathematical Theory of Communication", Bell System Technical Journal, pp. 379-423 and 623-656, 1948. Later, it appears as book, at Illinois Press, 1963. Also E-print.

[11] Tribus, and Irvine: "Energy and Information". *Scientific American* 225, No. 3: 179-188, 1971.

[12] Volkenstein: *Entropy and Information.* Series: Progress in Mathematical Physics, Vol. 57, Birkhäuser Verlag, 2009.