Mutual relationship between Entropies

Angel Garrido

Facultad de Ciencias de la UNED

Abstract

Our paper analyzes some new lines to advance on an evolving concept, the so-called Entropy. We need to model this measure by adequate conditions, departing from vague pieces of information. For this, it will be very necessary to analyze the relationship between some such measures, which may be of different types, with their very interesting applications, as Graph Entropy, Metric Entropy, Algorithmic Entropy, Quantum Entropy, and Topological Entropy.

Keywords: Measure Theory, Fuzzy Measures, Symmetry, Entropy.

Mathematics Subject Classification: 68R10, 68R05, 05C78, 78M35.

1. Introduction

The study of different concepts of Entropy will be very interesting now, and not only on Physics, but also on Information Theory [27] and other Mathematical Sciences, considered in its more general vision [6, 44]. Also may be a very useful tool on Biocomputing, for instance, or in many others, as studying Environmental Sciences. Because, among other interpretations, with important practical consequences, the law of Entropy means that energy cannot be fully recycled.

Many quotations are made until now referring to the content and significance of this fuzzy measure. Between them, we pick up

"Gain in Entropy always means loss of Information, and nothing more" (G. N. Lewis).

"Information is just known Entropy. Entropy is just unknown Information" (M. P. Frank, *Physical Limits of Computing*).

Mutual Information and Relative Entropy , also called Kullback-Leibler divergence, among other related concepts [6, 11-14, 17], have been very useful in Learning Systems, both on supervised and on unsupervised cases.

Our paper attempt to analyze the mutual relationship between the distinct types of entropies, as

- the *Quantum Entropy,* also called *Von Neumann Entropy;*

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

- the *KS-Entropy* (for Kolmogorov and Sinai), which is also called *Metric Entropy;*

- the *Topological Entropy*, or

- the *Graph Entropy,* among others.

2. Quantum Entropy

This entropy was first defined by the Hungarian mathematician *Janos Neumann* (the same people as John von Neumann) in 1927, with the purpose to shown the irreversible behavior of quantum measurement processes [22].

In fact, the *Quantum Entropy* (from now, denoted as *QE*) is an extension of the precedent Gibbs Entropy to the quantum realm [18, 23, 29]. It will be interpreted as the average information the experimenter obtains, when he make many copies of a series of observations, on an identically prepared mixed state.

It plays a very important role for studying correlated systems, and also for defining entanglement measures. Recall that *"Entanglement"* is one of the properties of Quantum Mechanics that caused Einstein to dislike this theory. But from then, Quantum Mechanics has reached high success predicting experimental results, and also the correlation predicted by the theory of such entanglement have been proved.

We can aply the notion of QE to Networks [28]. As QE is defined for quantum states, we need a method to map graphs into states. Such states for a quantum mechanical system are described by a *density matrix.* Usually, it is denoted as ρ. In fact, it is a positive semidefinite matrix with unitary trace

$$tr\left(\rho\right) = 1$$

But there exist many different ways to associate graphs to density matrices. Until now, we dispose of certain interesting results [24-27, 30], but still remains many open problems.

Between the known results, we can see that

the entropy for a d-regular graph tends, in the limit, when $n \to \infty$*,*
to the entropy of K_n*, the complete n-graph*

Another result may be that

the entropy of graphs increase as a function of the cardinality of their edges

Between the open problems, we can list some of them, as the relative to a very related matrix, called the *Normalized Laplacian.* It is defined by

$$\pounds\ (G) = \Delta^{-1/2}\ L\left(G\right)\ \Delta^{-1/2}$$

The *Combinatorial Laplacian Matrix of G* (abridgedly *Laplacian of G*) is given as

$$L(G) = \Delta(G) - A(G)$$

being computable by the difference between the matrix degree, $\Delta(G)$, and the adjacency matrix, $A(G)$.

Note. Let $G = (V, E)$ be an *UG (undirected graph)*, with set of nodes

$$V(G) = \{1,\ 2, ...,\ n\}$$

and set of edges

$$E(G) \subseteq [V(G) \times V(G)] - \{(v, v)\}_{v \in V(G)}$$

Then, the *Adjacency Matrix of G,* denoted by $A(G)$, will be defined by

$$[A(G)]_{u,v} = \begin{cases} 1,\ if\ \ \{u, v\} \in E(G) \\ \quad\quad 0,\ otherwise, \end{cases}$$

The *degree of a node,* v, is the number of edges adjacent to v.

Usually, it is denoted by $d(v)$.

The degree sum of the graph G is d_G, and it will be given by

$$d_G = \sum d(v)$$

The *average degree of G* is expressed as

$$d_G^* = m\ \sum d(v)$$

where m is the number of non-isolated nodes.

A graph, G, is *d-regular,* if $d(v) = d$, for all $v \in V(G)$.

The *degree matrix of G* is a $(n \times n) - matrix$ with entries given as

$$[\Delta(G)]_{u,v} = \begin{cases} d(v),\ if\ u = v \\ \quad 0,\ otherwise \end{cases}$$

So, the *Laplacian of a graph, G,* scaled by its degree-sum is a density matrix,

$$\rho_G = \frac{L(G)}{d_G} = \frac{L(G)}{tr(\Delta(G))} = \frac{L(G)}{m\ d_G^*}$$

With the well-known expression for the *entropy of a density matrix,* ρ,

$$S(\rho) = -tr\ \ (\rho \log_2 \rho)$$

Hence, departing from the concept of Laplacian of a Graph, we can say that $S(\rho_G)$ is the QE of G.

If we suppose two decreasing sequences of eigenvalues of $L(G)$ and ρ_G, respectively given by

$$\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_n = 0$$
$$and$$
$$\mu_1 \geq \mu_2 \geq ... \geq \mu_n = 0$$

mutually related by a scaling factor, i.e.

$$\mu_i = \frac{\lambda_i}{d_G} = \frac{\lambda_i}{m\ d_G^*}$$

Therefore, the *Entropy of a density matrix* ρ_G can also be written as

$$S(G) = -\sum \mu_i\ \log_2 \mu_i$$

with the usual convention

$$0 \log 0 = 0$$

Since its rows sum up to 0, we can conclude that the smallest eigenvalue of the density matrix must also be equal to zero. And the number of connected components of the graph is given by the multiplicity of 0 as an eigenvalue.

The QE is a very useful tool for problems [17, 18, 22, 24-27, 29] such as when it is applied to the *Enumeration of Spanning Trees.*

3. Algorithmic Entropy

Algorithmic Entropy is the size of the smallest program that generates a string.

It is denoted by $K(x)$, *or* AE.

It receive many different names [4, 5, 36-38], as may be, for instance, *Kolmogorov-Chaitin Complexity,* or only *Kolmogorov Complexity.* But also *Stochastic Complexity,* or *Program-size complexity.*

AE is a measure of the amount of information in an object, x. Therefore, it also measures its randomness degree [4, 5, 39].

The AE of an object is a measure of the computational resources needed to specify such object. I. e. the AE of a string is the length of the shortest program that can produce this string as its output.

So, the *Quantum Algorithmic Entropy* (QAE), also called *Quantum Kolmogorov Complexity* (QKC) is the length of the shortest quantum input to a Universal Quantum Turing Machine ($UQTM$) that produces the initial qubit string with high fidelity.

Hence, the concept is very different of the Shannon Entropy, because whereas this will be based on probability distributions, the AE is based on the size of programs.

All strings used may be elements of $\Sigma^* = \{0,1\}^*$, being ordered lexicographically.

The length of a string x is denoted by $|x|$.

Let U be a fixed prefix-free *Universal Turing Machine.*

For any string x of $\Sigma^* = \{0,1\}^*$, the *Algorithmic Entropy of* x will be defined by

$$K(x) = \min_p\{|p| :\ U(p) = x\}$$

From these concept, we can introduce the *t-time- Kolmogorov Complexity,* or *t-time-bounded algorithmic entropy.*

For any time constructible *t,* we introduce a refinement by

$$K^t(x) = \min_p\{|p| :\ U(p) = x,\ in\ at\ most\ t(|x|)\ \ steps\}$$

From these, we may obtain that for all x and y,

$$i)\ K(x) \leq K^t(x) \leq |x| + O(1)$$
$$and\ also$$
$$ii)\ K^t(x/y) \leq K^t(x) + O(1)$$

The AE or KC as a new tool have many applications, in fields as diverse as may be Combinatorics, Graph Theory, Analysis of Algorithms, or Learning Theory, among others.

4. Metric Entropy

We may consider the *Metric Entropy,* also called *Kolmogorov Entropy,* or *Kolmogorov-Sinai Entropy,* in acronym *K-S Entropy* [2, 3, 32].

Its name is given in hommage to Andrei N. Kolmogorov, and its disciple, Yakov Sinai [36-38].

Let $(X, \Omega,\ \mu)$ be a probability space, or in a more general way, a fuzzy measurable space.

Recall that a *measurable partition* of X is such that each of their elements is a measurable set; therefore, an element of the fuzzy $\sigma-$algebra, Ω.

And let I^X be the set of mappings from X to the closed unit interval, $I = [0, 1]$.

A fuzzy σ-algebra, Σ, on a nonempty set, X, is a subfamily of I^X satisfying that

(1) $1 \in \Sigma$.

(2) If $\alpha \in \Sigma,$ then $1 - \alpha \in \Sigma$.

(3) If $\{\alpha_i\}$ is a sequence in Σ, then

$$\vee_{i=1}^{\infty} \alpha_i = \sup \alpha_i \in \Sigma$$

A *fuzzy probability measure,* on a fuzzy $\sigma-$algebra, Σ, is a function

$$m : \Sigma \to [0, 1]$$

which holds

[1] $m(1) = 1$

[2] for all $\alpha \in \Sigma,\ m\left(1 - \alpha\right) = 1 - m(\alpha)$

[3] for all $\alpha,\ \beta \in \Sigma,\ m\left(\alpha \vee \beta\right) + m\left(\alpha \wedge \beta\right) = m\left(\alpha\right) + m\left(\beta\right)$

[4] If $\{\alpha_i\}$ is a sequence in Σ, such that $\alpha_i \uparrow \alpha$, with $\alpha \in \Sigma$, then

$$m\left(\alpha\right) = \sup\ m\left(\alpha_i\right)$$

We call $(X, \Omega,\ \mu)$ a *fuzzy-probability measure space,* and the elements of Ω are called *measurable fuzzy sets.*

The notion of *"fuzzy partition"* was introduced by E. Ruspini.

Given a finite measurable partition, $\wp$, we can define its *Entropy* by

$$H_\mu\left(\wp\right) = \sum\nolimits_{p \in \wp} -\mu\left(p\right) \log \mu\left(p\right)$$

As usually in these cases, we take as convention that 0 log 0 = 0.

Let $T : X \to X$ be a measure-preserving transformation. Then, the *Entropy of* T w.r.t. a finite measurable partition, $\wp$, is expressed as

$$h_\mu\left(T,\ \wp\right) = \lim\nolimits_{n\to\infty} H_\mu\left(\vee_{k=0}^{n-1} T^{-k} \wp\right)$$

with H_μ the entropy of a partition, and where $\vee$ denotes the join of partitions.

Such limit always exists.

Therefore, we may define the *Entropy of T* by

$$h_\mu (T) = \sup_\wp h_\mu (T,\ \wp)$$

taking the supremum over all finite measurable partitions.

Many times $h_\mu (T)$ is named the *Metric Entropy of T.* So, we may to differentiate this mathematical object from the well-known as Topological Entropy [2, 43, 45].

We may to investigate the mutual relationship between the Metric Entropy and the Covering Numbers.

Let $(X,\ d)$ a metric space, and let $Y \subseteq X$ a subset of X.

We says that $Y^* \subseteq X$ is an $\epsilon - cover\ of\ Y$, if for each $y \in Y$, there exists a $y^* \in Y^*$ such that

$$d\ \ (y,\ y^*) \leq \varepsilon$$

It is clear that there are many different covers of Y. But we are specially interested here in one which contains the lesser number of elements. We call [2] the cardinal, or size, of such a cover its *Covering Number.*

Mathematically expressed, the $\epsilon - covering\ number\ of\ Y$ is

$$N (\varepsilon,\ Y,\ d) = \min\{card\ (Y^*) :\ Y^*\ is\ an\ \epsilon - cover\}$$

A *proper cover* is one where $Y^* \subseteq Y$.

And a *proper covering number* is defined in terms of the cardinality of the minimum proper cover.

Both, covering numbers and proper covering numbers are related by

$$N (\epsilon,\ Y) \leq N_{proper} (\epsilon,\ Y) \leq N \left(\tfrac{\varepsilon}{2},\ Y\right)$$

Furthermore, we recall that the Metric Entropy, $H (\epsilon, Y)$, is a natural representation of the cardinal of the set of bits needed to send, in order to identify an element of the set up to precision ε.

It will be expressed by

$$H (\epsilon,\ Y) = \log N (\epsilon,\ Y)$$

In a dynamical system, the metric entropy is equal to zero for nonchaotic motion. And it is strictly greater than zero for chaotic motion. So, it will be interpreted as a simple indicator of the complexity of a dynamical system.

5. Topological Entropy

Let $(X,\ d)$ be a compact metric space, and let $f : X \to\ X$ be a continuous map [1, 39, 43, 44, 45].

For each $n > 0$, we define a new metric, d_n, by

$$d_n(x,\ y) = max\{d(f^i(x),\ f^i(y)) : 0 \leq i < n\}$$

Two points, x and y, are *close* with respect to (w. r. t.) this metric, if their first n iterates (given by f^i, i=*1,2,...*) are close.

For $\epsilon > 0$, and $n \in N^*$, we say that $S \subset X$ is an $(n,\epsilon) -$ *separated set*, if for each pair, $x,\ y$, of points of S, we have

$$d_n(x, y) > \epsilon$$

Denote by $N(n, \epsilon)$ the *maximum cardinality of a (n, ϵ)-separated set.*

It must be finite, because X is compact. In general, this limit may exists, but it could be infinite.

A possible interpretation of this number [3] is as a measure the average exponential growth of the number of distinguishable orbit segments. So, we could say that

the higher the topological entropy is,

the more essentially different orbits we have.

From an analytical viewpoint, the topological entropy is a continuous, and monotonically increasing function.

This concept was introduced, in 1965, by Adler, Konheim and McAndrew [1].

$N(n, \epsilon)$ shows the number of “distinguishable” orbit segments of length n, assuming we cannot distinguish points that are less than ϵ apart.

The *topological entropy of* f is then defined by

$$H_{top} = \lim_{\epsilon \to 0} \lim\ \sup_{n \to \infty} \left[\frac{1}{n} \log N\left(n, \epsilon\right)\right]$$

Therefore, TE is a non-negative number measuring the complexity degree of the system. So, it gives the exponential growth of the cardinality for the set of distinguished orbits, according the time advances.

6. Graph Entropy and Chromatic Entropy

A *system* [7-9] can be defined as a set of components functioning togheter as a whole. A systemic point of view allows us to isolate a part of the world, and so, we can focus on those aspect that interact more closely than others.

The entropy of a system represent [4, 35, 43] the amount of uncertainty one observer has about the state of the system. The simplest example of a system will be a random variable, which can be shown by a node into the graph, being their edges the representation of the mutual relationship between them. Information measures the amount of correlation between two systems, and it reduces to a mere difference between entropies.

So, the Entropy of a Graph (from now, denoted by *GE*) is a measure of graph structure, or lack of it. Therefore, it may be interpreted as the amount of Information, or the degree of "surprise", communicated by a message.

And as the basic unit of Information is the bit, Entropy also may be viewed as the number of bits of "randomness" in the graph, verifying that

the higher the entropy,
the more random is the graph

We consider a functional on a graph, $G = (V, E)$, with P a probability distribution on its node (or vertex) set, V. These mathematical construct will be denoted by *GE*. It will be defined as

$$H(G, P) = min \sum p_i \ \log \ p_i$$

Observe that such H is a convex function.

It tends to $+\infty$ on the boundary of the non-negative orthant of R^n , and monotonically to $-\infty$ along rays from the origin.

So, such minimum is always achieved and it will be finite.

Let G be now an arbitrary finite rooted *Directed Acyclic Graph* (*DAG,* in acronym).

For each node, v, we denote $i(v)$ the number of their edges that terminates at v.

Then, the *Entropy of the graph* is expresable as

$$H(G) = \sum_{\substack{v \in V \\ i(v) \geq 2}} [i(v) - 1] \ \log_2 \left(\frac{Card\ (E)\ -\ Card\ (V)\ +\ 1}{i(v) - 1} \right)$$

$H(X)$ may be interpreted in some different ways. For instance, given a random variable, X, it informs us about how random X is, how uncertainty we should about X, or how much variability X has.

In a variant of the "Graph Coloring Problem", we take as the objective function to minimize the Entropy of such coloring. So, it is called the *Minimum Entropy Coloring.*

As *Chromatic Entropy,* we understand the *minimum Entropy of a coloring.* Its role is essential in the problem of coding. If we consider this *problem* from a computational viewpoint, it is *NP-hard*; for instance, on Interval Graphs.

7. Mutual relationship between Entropies

In the mid 1950's, the Russian mathematician Kolmogorov imported Shannon's probabilistic notion of entropy into the theory of dynamical systems [36, 38], and showed how entropy can be used to tell whether two dynamical systems are non-conjugate, i.e., non-isomorphic.

His work inspired a whole new approach in which entropy appears as a numerical invariant of a class of dynamical systems. Because the Kolmogorov´s metric entropy is an invariant of measure theoretical dynamical systems, and thus, it is closely related to Shannon´s source entropy.

Ornstein [24-27] showed that metric entropy suffices to completely classify two-sided Bernoulli processes, a basic problem which for many decades appeared completely intractable. Recently, S. Tuncel [32, 34, 39, 40-42] has shown how to classify one-sided Bernoulli processes; this turns out to be quite a bit harder.

In 1961, Adler et al. [1] introduced the aforementioned *topological entropy,* which is the analogous invariant for topological dynamical systems.

There exist a simple relationship between these quantities, because *maximizing the metric* entropy, over a suitable class of measures defined on a dynamical system, *gives its topological entropy.*

The relationship between TE and the *Entropy in the sense of Measure Theory (K-S)* is given by the so-called *Variational Principle,* which established that

$$h(T) = \sup\{h_{\mu}(T)\}_{\mu \in P(X)}$$

This may be interpreted as that TE is equal to the supremum of Kolmogorov-Sinai (or *K-S) entropies,* $h_{\mu}(T)$, with μ belonging to the set of all T-invariant Borel probability measures on *X.*

The mutual relationship between Algorithmic Entropy and Shannon Entropy is that the expectation of the former gives us the latter, up to a constant depending on the distribution. Also we may expressed, departing of *P(x)* as a recursive probability distribution, that

$$0 \leq \sum P(x)\ K(x) - H(P) \leq K(P)$$

Finally, we recall that given a random variable, X, its *Shannon Entropy* is given by

$$H(X) = -\sum P(x) \ \log_2 P(x)$$

whereas the *Rényi Entropy of order* $\alpha \neq 1$ of such random variable is

$$H_\alpha(X) = \frac{1}{1-\alpha} \log_2 \left(\sum P(x)^\alpha\right)$$

The *Renyi Entropy of order* α *converges to the Shannon Entropy,* when α tends to one, i.e.

$$\lim_{\alpha \to 1} \left\{ \frac{1}{1-\alpha} \log_2 \left(\sum P(x)^\alpha\right) \right\} = -\sum P(x) \ \log_2 P(x)$$

So,

$$\lim_{\alpha \to 1} H_\alpha(X) = H(X)$$

Therefore, the Rényi Entropy may be considered as a generalization of the Shannon Entropy, or dually expressed, the Shannon Entropy will be a particular case of Rényi Entropy.

8. Conclusions

Statistical entropy is a probabilistic measure of uncertainty [17, 31], or ignorance about data. Whereas *Information* is a measure of a reduction in that uncertainty [21, 33, 40-42]. And the Entropy of a probability distribution is just the expected value of the information of such distribution [26].

All these improved tools must permits to advance not only in essential fields as Optimization Theory, but also on many others, as Generalized Fuzzy Measures [6, 10-14]; Economics [15]; Machine Learning, or on A. I. in general [40-42, 43, 46]; constructing biological or ecological models; describing economical or psychological behavior, and so on.

With this paper, my clear purpose was attempt to reach an ever partial completion of a very long cycle on Analysis of Fuzzy Symmetry and Entropies.

References

[1] Adler, Konheim, and Mac Andrew: "Topological Entropy". *Transactions of the AMS.* Vol. 114, pp. 309-319, 1965.

[2] Bartlett: "Covering Numbers and Metric Entropy". *Statistical Learning Theory.* Lecture, Spring 2006.

[3] Burbea, and Rao: Entropy Differential Metric, Distance and Divergence Measures in Probabilistic Spaces: A Unified Approach. *J. Multi. Analysis,* Vol. 12, pp. 575-596, 1982.

[4] Chaitin: "Algorithmic Entropy of Sets". *Comp. and Maths with Appls.* Vol. 2, pp. 233-245, 1976.

[5] Devine: " The Insights of Algorithmic Entropy". *Entropy journal,* Vol.11, IssueNo. 1, pp. 85-110. MDPI Publishers, Basel, 2009.

[6] Dubois, and Prade: *Fundamentals of Fuzzy Sets.* Series: The Handbooks of Fuzzy Sets. Vol. 7. Springer-Kluwer, 2000.

[7] Dumitrescu, D.: Entropy of a fuzzy dynamical system, FSS, Vol. 70, pp. 45-57, 1995.

[8] Dumitrescu, D.: Entropy of a fuzzy process. FSS. Vol. 55, pp. 169-177, 1993.

[9] Dumitrescu, D.: Fuzzy measures and the entropy of fuzzy partitions. J. Math. Anal. Appl. Vol. 175, pp. 359-373, 1993.

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

[11] Garrido: "Additivity and Monotonicity in Fuzzy Measures". Plenary Talk in *ICMI45,* at Bacau University, then published at *SCSSM journal,* Vol. 16, pp. 445-458, 2006.

[12] Garrido: "Information and Entropy Measures". Plenary Talk in the *6th ICTAMI,* at Alba Iulia University, accepted; must to be published at *AUA journal,* Vol. 19 (special issue), pp. 18, 2009.

[13] Garrido: "Analyzing the concept of Entropy". *EIJ-AMO* journal, Vol. 11, Issue No. 4, pp. 525-530, 2009.

[14] Garrido: "Metric, Topological, and Graph Entropies". *EIJ-AMO* journal, Vol. 11, Issue No. 4, pp. 593-600, 2009.

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

[16] Hung and Yang: "Fuzzy entropy on Intuitionistic Fuzzy Sets". *Intern. J. of Intelligent Systems.* Vol. 21, pp. 443-451, 2006.

[17] Jaynes: "Information Theory and Statistical Mechanics". *Phys. Rev.* Vol. 108, pp. 171-190, 1956.

[18] Jozsa: "Quantum Information and Its Properties". Chapter in the book *Introduction to Quantum Computation and Information.* Eds. Lo, Popescu, and Spiller. World Scientific, 1998.

[19] Kammeyer, and Rudolph: "Restricted Orbit Equivalences of Discrete Amenable Groups". Cambridge University Press, 2002.

[20] Liu: "Entropy, distance and similarity measures of fuzzy sets and their relations. *FSS,* Vol. 52, pp. 305-318, 1992.

[21] Luca and Termini: "A definition of non-probabilistic entropy, in the setting of fuzzy theory. Information and Control. Vol. 20, pp. 301-318, 1972.

[22] Neumann: *Grundlagen der Quantenmechanik.* Springer Verlag, Berlin, 1955.

[23] Ohya, and Petz: *Quantum Entropy and Its Use.* Springer, 1993.

[24] Ornstein, and Weiss: "Entropy and Isomorphism Theorems for actions of amenable groups". *J. Analyse Math.* Vol. 48, pp. 1-141, 1987.

[25] Ornstein, Rudolph, and Weiss: "Equivalence of measure preserving transformations". *Mem. AMS,* Vol.262, 1984.

[26] Ornstein: *Ergodic Theory, Randomness, and Dynamical Systems.* Yale Math. Monographs. Yale University Press, New Haven, 1974.

[27] Ornstein: "Two Bernoulli shifts with infinite entropy are isomorphic". *Adv. in Maths.* Vol. 5, pp. 339-348, 1970.

[28] Passarini et al.: "The von Neumann entropy of networks". *MPRA,* Paper No. 12538. University of Munich, 2009.

[29] Peres: *Quantum Entropy: Concepts and Methods.* Kluwer, 1993.

[30] Petz: "Properties of Quantum Entropy". Chapter in the book *Quantum Probability and related topics VII,* pp. 275-297. World Scientific Publishers, 1992.

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

[32] Rudolph: *Fundamentals of Measurable Dynamics: Ergodic Theory of Lebesgue Spaces.* Clarendon/Oxford University Press, New York, 1990.

[33] 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 by E-print.

[34] Shields: *The Theory of Bernoulli Shifts.* University of Chicago Press, 1973.

[35] Simonyi: "Graph Entropy: A Survey". *DIMACS Series in Discrete Mathematics and Theoretical Computer Science.*

[36] Sinai: "On the concept of Entropy of a Dynamical System". *Doklady Akademii Nauk SSSR,* Vol. 124, pp. 768-771, 1959.

[37] Ibid.: "On the foundations of the Ergodic Hypothesis for a Dynamical System of Statistical Mechanics". *Doklady Akademii Nauk SSSR,* Vol. 153, pp. 1261-1264, 1963.

[38] Ibid.:"Gibbs measures in Ergodic Theory". *Russian Math. Surveys,* Vol. 27, pp. 21-69, 1972.

[39] Smorodinski: "Ergodic Theory, Entropy". *LNM (Lecture Notes in Mathematics).* Vol. 214, Springer, New York, 1971.

[40] Titchener, Nicolescu, Staiger, Gulliver, and Speidel: "Deterministic Complexity and Entropy". *Fundamenta Informaticae,* Vol. 64, Issue 1-4, Contagious Creativity-In Honor of the 80th Birthday of Prof. Solomon Marcus, 2004.

[41] Titchener: "A measure of information". *Proc. Data Compression Conference,* pp. 353-362, Snowbird, UT, 2000.

[42] Titchener: "A deterministic theory of Complexity, Information and Entropy". *IEEE Information Theory Workshop.* San Diego, CA, 1998.

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

[44] Wang, and Klir: *Generalized Measure Theory.* Springer Verlag, Berlin-New York, 2008.

[45] Wehrl: "General properties of entropy". *Rev. Mod. Phys.* Vol. 50, pp. 221-260, 1978.

[46] Yager: “On the measure of fuzziness and negation”. *Int. J. Gral. Systems,* Vol. 5, pp. 189-200, 1979.

[47] You, and Gao: “Maximum entropy membership functions for discrete fuzzy variables”. *Information Sciences,* 179, Issue No. 14, pp. 2353-2361, 2009.

[48] Yu, and Hu: “Entropies of fuzzy indiscernibility relation and its operations”. *Intern. J. of Uncertainty, Fuzziness and Knowledge-Based Systems.* Vol. 12, Issue No. 5, pp. 575-589, 2005.

[49] Zeng, and Li: "Relations between Similarity and Entropy of Interval Valued Fuzzy Sets". *FSS,* Vol. 157, pp. 1477-1484, 2006.