Review of: "Information Is Immanent Incongruence"

Karoly Heberger[1]

1 Research Centre for Natural Sciences

Potential competing interests: No potential competing interests to declare.

Preprint platform: Qeios

Title: Information Is Immanent Incongruence

Qeios ID: 9TX7Z3

Open Access: https://doi.org/10.32388/9TX7Z3

Author: Karl Javorszky

The submission is peculiar to say the least. I have read it several times, but the reason why the author produced this text —escaped from my mind.

Finally, the manuscript can be considered as an example of Gödel's incompleteness theorems, nothing more.

Many of the statements are apodeictical, i.e., necessarily true without thinking the issue over.

On the other hand, there are a plethora of deficiencies, which limits the understanding and fair evaluation on a large scale.

Any scientific contribution should conform several basic rules and attitudes, e.g.,

- The meaning of the terms should be unequivocal and should conform with the earlier usage.
- Any work should enumerate the earlier findings and place the new achievements into the context of the present status of knowledge, as well as discuss the similarities and differences from previous approaches just to see the novelty, if any. One can ignore other people's findings, but it is not a scientific attitude even if it can be observed frequently in nowadays "science".
- A new term should be defined (and abbreviated) properly at the first mentioning.
- An "elaborate explanation model which assumes the world to be of a basic duality" is necessarily damned to failure. The World or Nature is not of duality, or not multitudinous —we use an n-dimensional (Euclidean) space to its description (and understanding).

Why just a two-dimensional grid was selected? Why not three or more? The course of dimensionality is well known [2].

Two characteristic examples show the unavoidable misunderstanding while reading the manuscript.

The title is not understandable. I looked for vocabulary and thesaurus to make the meaning clear without success. The

meaning of incongruence is "lack of consistency or appropriateness" [3]; why should information be inconsistent or inappropriate? And why is it "indwelling" or "inherent" [4]? Anyway, the information is a vague (cloudy) term. There are many forms of it; which one is considered here?

"Two combinatorial functions cross twice in the region 1.n..140"—Apart from the lack of any explanation about background or motivation, the functions are arbitrarily defined and obviously the study is limited to the first 140 numbers of natural numbers. In fact, it is an enormously large reduction of the meaning for the word "information".

Perhaps the author thinks that some form of information is inherently not suitable (consistent)? For what?

How can one speak about information without mentioning the Shannon's entropy? Why should Shannon's formula for entropy being inconsistent?

…

"Everyday observations show, (SIC!) that Nature uses a conflict management system"—i) everyday observations show the opposite, as well; ii) Nature does not use anything, we use some mathematical formulae to describe and understand observations (and using assumptions and neglects).

"Gay-Lussac law"—The author probably thinks on the principle of least constraint: a principle formulated by the French physicochemist Henry Louis Le Châtelier and the German physicist Karl Ferdinand Braun: if we apply some kind of constraint to an external effect on a system in chemical equilibrium, the equilibrium shifts in such a way that it balances this constraint or at least reduces it—explained in ref. [5] in detail.

"Information is the extent of deviation of the observed value to the expected value"—Well, it is called widely and unambiguously as residua (residuals).

There are many other dichotomies apart from expected (predicted) and observed values: foreground vs. background, proximity vs. distantness, etc.

Let us consider similarity and diversity. These terms have contradictory, diffuse, and inconsistent usage in the scientific literature, though their indices have a well-defined meaning. We should quantify them if we do not want to classify our knowledge as unsatisfactory [6]. Table 4 of ref. [7] lists 25 similarity indices classified by seven basically different ways of calculating the similarities between pairs of attributes. Any of them seems to be suitable to quantify similarity. A more comprehensive list of binary similarity coefficients can be found in ref. [8]. The antonym of similarity is diversity or distance. One should pay attention that similarity is defined between 0 and 1, but distance between 0 and + endless. If the distance is scaled between 0 and 1 it is generally called dissimilarity. One may start at Wikipedia [9] and return to Shannon index or have a look at part: 3.14.6 Similarity-Dissimilarity Measures, Distance Metrics in ref. [10]. Recent developments solve the problem of comparisons n-times [11,12].

The manuscript overlooks the inherent features of ranking: i) rankings has direction (top-down, down-top); ii) equal numbers (ties) can be ranked differently (average, minimum, maximum); iii) the "rank difference" should not be one

necessarily partial ranking [13], multi label ranking [14] are also known. Rank aggregation is also bound to ranking [15]. Hasse diagram is particularly suitable for visualization [16].

Naturally, a short contribution cannot cover all aspects of similarities, diversities, ranking, sequencing, etc. However, many factors influence the problem at hand. The author has properly realized that these factors, indices, cohorts, partitioning, etc. are conflicting: "logical conflicts", "contradictions among sorting orders", etc. One step further is missing desperately: only multicriteria decision making (MCDM) algorithms (analyses) are suitable. At least a comparison with MCDM techniques should have been provided.

Minor errors

"5,4EE+92" – Come on!

"We propose a world view" – no one can believe it.

"all possible states of the world" (or World?) - it sounds like nonsense!

"a description of the world" – with 140 natural numbers?

"a picture of the world, the concept of mass comes to mind." – Really? To whom?

"two interrelated aspects of all things that make up the world" – What? The presented case is an example of interrelation of two things and has nothing to do with World.

"basic design properties of the world depicted. Our world being subject to periodic changes" and "Adaptation to periodic changes can be pictured by ranking" – Oh No!

"space" "spatial", etc. In fact, an integer sequence and a 2D grid are considered.

"an elementary moment consists of 3 phases" – fairly the author uses an example with 3 phases (aspects).

"logical sentences that describe a state of the world" – by no means. Does the author think about "system" but says world?

…

"Nature makes use of accounting procedures" – I cannot believe my eyes. The Nature does not use any procedures, we use some procedures to describe and understand phenomena.

…

There is no use continuing the enumeration of such misleading and bombastic statements.

The author thinks on a level of ideas I cannot follow. Maybe this is my fault.

Still, I insist on placing present findings into the up to date achievements.

References

[1] https://en.wikipedia.org/wiki/G%C3%B6del%27s_incompleteness_theorems Access date: May 23 / 2023

[2] https://en.wikipedia.org/wiki/Curse_of_dimensionality Access date: May 23 / 2023

[3] https://dictionary.apa.org/incongruence Access date: May 23 / 2023

[4] https://www.dictionary.com/browse/immanent Access date: May 23 / 2023

[5] Michal Pavelka, Miroslav Grmela, Braun-Le Chatelier principle in dissipative thermodynamics, AAPP, Physical, Mathematical, and Natural Sciences, pdf file available at https://doi.org/10.1478/AAPP.97S1A22

[6] "When you can measure what you are speaking about, and express it in numbers, you know something about it; but when you cannot measure it, when you cannot express it in numbers, your knowledge is of a meagre and unsatisfactory kind." —Lord Kelvin, physicist.

[7] J.W. Johnston, Similarity Indices I: What Do They Measure? Battelle Pacific Northwest Laboratories, Richland, Washington 99352, November 1976, BNWL-2152 RC-1, Addendum https://www.osti.gov/servlets/purl/7256702

[8] Roberto Todeschini, Viviana Consonni, Hua Xiang, John Holliday, Massimo Buscema, and Peter Willett, Similarity Coefficients for Binary Chemoinformatics Data: Overview and Extended Comparison Using Simulated and Real Data Sets Journal of Chemical Information and Modeling 52 (2012) 2884-2901. https://doi.org/10.1021/ci300261r

[9] https://en.wikipedia.org/wiki/Diversity_index Access date: June 08 / 2023

[10] Dávid Bajusz, Anita Rácz, Karoly Héberger, (2017) Chemical Data Formats, Fingerprints, and Other Molecular Descriptions for Database Analysis and Searching. In: Comprehensive Medicinal Chemistry III vol. 3, pp. 329–378. Samuel Chackalamannil, David P Rotella and Simon E Ward, (eds.) Oxford: Elsevier.

http://dx.doi.org/10.1016/B978-0-12-409547-2.12345-5

available via Research Gate: https://www.researchgate.net/publication/315513438

[11] Ramón Alain Miranda-Quintana, Dávid Bajusz, Anita Rácz, Károly Héberger, Extended similarity indices: the benefits of comparing more than two objects simultaneously. Part 1: Theory and characteristics Journal of Cheminformatics, 13; Article No: 32 (2021) https://doi.org/10.1186/s13321-021-00505-3

[12] Ramón Alain Miranda-Quintana, Dávid Bajusz, Anita Rácz, Károly Héberger, Extended similarity indices: the benefits of comparing more than two objects simultaneously. Part 2: speed, consistency, diversity selection, Journal of Cheminformatics 13; Article No: 33 (2021)

https://doi.org/10.1186/s13321-021-00504-4

[13] Juan C. Alfaro, Juan A. Aledo, José A. Gámez, Pairwise learning for the partial label ranking problem, Pattern

Recognition 140 (2023) Article No: 109590.

[14]   Guoqiang Wu, Chongxuan Li, Kun Xu, Jun Zhu Rethinking and Reweighting the Univariate Losses for Multi-Label Ranking: Consistency and Generalization, 35th Conference on Neural Information Processing Systems (NeurIPS 2021) https://openreview.net/forum?id=42yEyjooGSC

[15]   Xue Li, Xinlei Wang, and Guanghua Xiao, A comparative study of rank aggregation methods for partial and top ranked lists in genomic applications, Briefings in Bioinformatics, 20 (2019) 178-189. https://doi.org/10.1093/bib/bbx101

[16]   Martin Charles Golumbic, Algorithmic Graph Theory and Perfect Graphs Annals of Discrete Mathematics, Chapter 5 - Comparability graphs 57 (2004) 105-148. https://www.sciencedirect.com/topics/mathematics/hasse-diagram

June 09 /2023