Advances in Experimental Medicine and Biology 1194

Panayiotis Vlamos *Editor*

GeNeDis 2018

Computational Biology and Bioinformatics

EXTRAS ONLINE

Springer

Logical Analysis and Validation of Publications in Bioinformatics

Konstantinos G. Papageorgiou

1 Introduction

Statistics is not a proving method in mathematics, nor can it be used to prove anything in the strict sense; the only proving method in mathematics is logic (Stoll 1963). Logic is *a priori*, in the sense that it is set up even before the first word is written, as, for example, it is used in the logical setup of a scientific paper (arguments, i.e., premises leading to conclusions, logical inferences), not *a posteriori*, as is the case with statistics. Statistics per se may only be as good (i.e., credible) as the level of logical consistency and logical analysis used. Essentially, logical analysis per se is the scientific method (Papageorgiou and Lekkas 2018). It only follows naturally that logic concerns all of scientific and academic writing, in any form and for any purpose, whereas statistics is of concern to a great deal of scientific papers, but not of all.

Despite occasional problems in regard to the quality of the statistical analysis used – or even the manipulation of results towards making them more appealing in medical research (Ioannidis 2005, 2012) – statistical analysis is the golden standard of the contemporary scientific method. The problem has, therefore, already been identified; however, no solution has ever been proposed.

We theorize about a novel contribution in scientific writing that would greatly impact future publications in bioinformatics, and even other fields make the evaluation of scientific papers "measurable," and provide the researchers and academics with an important proving tool for structuring their core arguments properly: a logical tool, an expert system, that would help academics to apply core logical principles and to avoid logical fallacies. Finally, the theoretical work required for such an endeavor would create a reference point for both academic writing and training in logic. No matter how complicated they seem, all logical arguments are reducible to

K. G. Papageorgiou (✉)
Philosophy of Science, Ionian University, Department of Informatics, Kerkira, Greece

P. Vlamos (ed.), *GeNeDis 2018*, Advances in Experimental Medicine and Biology 1194, https://doi.org/10.1007/978-3-030-32622-7_47

basic logical structures. Already in classical antiquity, Aristotle proposed 14 such structures, "syllogisms" (easily extendable to at least 20), classified by Latin scholars and known to us by coded names, e.g., "Barbara," "Celarent," etc. (Hodgkin 2005; Netz 2003). In the twentieth century, the so-called linguistic turn occurred within the context of the newly developed symbolic logic (starting with Frege, Russel, etc.). Logical positivism (the Vienna School) and the still popular analytic philosophy were just some results of these developments in mathematics and linguistics, propagating the need of both scientists and philosophers to express themselves more accurately, more systematically, and more in a direction toward "proving" their ideas (Psillos 2007). Despite recent advancements in information technology and computational linguistics, no research in this direction has ever been conducted – in spite of the enormous existing bibliography about logic or in spite of the educational material related to logic. Problems in relation to fallacies and problematic argumentation have already been presented (Bencivenga 1979; Massey 1981), but no tool has ever been developed to address such issues.

The expected functions of such a computer program (an expert system) would be to:

- Identify the structure of the arguments used and, after translating them to symbolic language, be able to process them in various ways: perform logical inversions, predict and analyze conclusions, evaluate given conclusions based on given premises, pinpoint ambiguities, and convert arguments to all possible equivalent forms
- Connect to a database of saved arguments and present to the researcher identical or similar arguments

Therefore, researchers and academics may expect the following:

- To better plan their protocols by automating the process of analysis of the premises given.
- Be able to locate and analyze arguments from the text of the scientific paper, or from certain parts of it (e.g., the "conclusions" section), or even from the whole text, based on advanced AI algorithms.
- Evaluate papers, giving a score of logical consistency and completeness (i.e., how many conclusions have been investigated in comparison to all possible conclusions based on the set of premises used by the researchers; how ambiguous premises are).
- Provide a new standard for future publications in any form and in any scientific field.
- Contribute to a publicly accessible database that would become a search engine for arguments.

The work would also result in a number of publications about logic, truth values, Venn diagrams, the nature and analytic features of negation, logical inversions, comparing and contrasting concepts vs. non-concepts and their systemic treatment, logical complementarity, issues of conciseness, consistency, completeness, decidability and inference in relation to set theory and Boolean algebras, Gödelian considerations, and issues connected to Turing machines, all for the theoretical support of the program.

2 Development of the Expert System

The functionality of such a computer program would extend on many levels.

On the level of premises, it would:

- Convert the premises to equivalent ones
- Perform logical inversions of the premises
- Present all possible negations of the premises (contrarieties and contradictions)
- Discover ambiguous expressions

On the level of conclusions:

- Check the level of consistency between the premises and the conclusions, attributing a score based on this evaluation.
- Predict all possible conclusions based on a given set of premises.
- List all possible variations of the conclusion (logical inversions, contrarieties, contradictions, and equivalences).

On the level of the complete argument (premises and conclusions):

- Identify the form of the argument.
- Convert it to equivalent forms, whether simpler or more complicated.
- Rate and evaluate the argument as a whole.
- Access a public database and present the researcher with identical or similar arguments (based on the search filter) – and with the scientific publications that these arguments belong to.

Such a process would afford a series of benefits to academics:

- A better way of planning their research by exhaustively analyzing their ideas.
- A more objective way of evaluating the logical aspect of their arguments.
- An insight into similar arguments.
- A tool toward making the application of important theoretical knowledge easy and a manageable part of the everyday writing routine

3 Conclusions

Such a tool will ultimately provide a series of practical guidelines for writing research papers which are based on robust logical foundations and contain strong and sound proofs employing more tools in addition to, or instead of, statistical analysis. The key make-or-break factors behind scientific papers of high impact will be identified. Moreover, a global perspective of what really is a scientific contribution as well as what are the components of scientific output will be gained. Such a tool would provide a significant solution, or answer, to the fierce debate that is going on about peer reviewing (Resnik et al. 2008; Teixeira da Silva and Dobránszki 2015). The face of many papers has evolved; more and more journals, e.g., require their hosted papers to begin with bullet points summing up key points and contributions. By being able to assess the quality of the core arguments of a publication independently, objectively, and in a standardized way, the "inadequate reviewing" problem will, at least, be minimized. The innovation of this expert system, therefore, lies not only in the multidisciplinary approach to a theoretical problem about the evaluation of arguments but also in the combined know-what and know-how of mathematics (logic), computational linguistics, and information science, in order to create a practical application in the form of an easy-to-use computer program.

References

Bencivenga E (1979) On good and bad arguments. J Philos Log 8(1):247–259

Hodgkin L (2005) A history of mathematics: from Mesopotamia to modernity: from Mesopotamia to modernity. Oxford University Press, Oxford/New York

Ioannidis JPA (2005) Why most published research findings are false. PLoS Med 2(8):e124

Ioannidis JPA (2012) Why science is not necessarily self-correcting. Perspect Psychol Sci 7(6):645–654

Massey G (1981) The fallacy behind fallacies. Midwest Stud Philos 6(1):489–500. Mikolov T, Chen K, Corrado G, Dean J (2013) Efficient estimation of word representations in vector space, CoRR, abs/1301.3781,2013

Netz R (2003) The shaping of deduction in Greek mathematics: a study in cognitive history. Cambridge University Press, Cambridge

Papageorgiou KG, Lekkas D (2018) On the methodology of the analytic method: historical account, epistemological suggestions, stages. Epistēmēs Metron Logos (1)

Psillos S (2007) Philosophy of science A-Z. Edinburgh University Press, Edinburgh

Resnik DB, Gutierrez-Ford C, Peddada S (2008) Perceptions of ethical problems with scientific journal peer review: an exploratory study. Sci Eng Ethics 14(3):305–310

Stoll RR (1963) Set theory and logic. Freeman, San Francisco

Teixeira da Silva JA, Dobránszki J (2015) Problems with traditional science publishing and finding a wider niche for post-publication peer review. Account Res 22(1):22–40