The Reasoner

Volume 16, Number 1
January 2022

thereasoner.org
ISSN 1757-0522

Contents

Guest Editorial 1

Features 1

Dissemination corner 3

What's Hot in . . . 5

Events 7

Courses and Programmes 7

Jobs and Studentships 8

Guest Editorial

Dear Reasoners,

I am delighted to introduce Tomasz Jarmużek to you all! Tomasz is the chair of the Department of Logic at the Nicolaus Copernicus University in Toruń, Poland. His contributions span tableaux methods, positional logic, and relating semantics. Tomasz' work combines formal rigour with philosophical research on logic, its philosophy and its applications in the sciences.

This interview will focus on three main aspects of Tomasz' work: (a) his logical and intellectual background, (b) his interest on relating semantics, and (c) the work he is currently coordinating, as the chair of the Department of Logic at NCU-Toruń, to build new bridges between logic and philosophy of science.

Without further ado, here is Tomasz Jarmużek!

María Del Rosario Martínez-Ordaz
Philosophy, Federal University of Rio de Janeiro

Features

Interview with Tomasz Jarmużek

María Del Rosario Martínez-Ordaz: Tomasz, can you tell us a bit about your background?

Tomasz Jarmużek: I studied philosophy and sociology. Later on, with Jacek Malinowski as my supervisor, I got a PhD. in philosophy with emphasis on logic from the Nicolaus Copernicus University in Toruń, Poland. In my doctoral dissertation I applied temporal logic to analyze and reconstruct the master argument of Diodorus Cronus – a Greek philosopher of the Megarian school. The results of that research were lately published in the book *On the Sea Battle Tomorrow That May Not Happen: A Logical and Philosophical Analysis of the Master Argument*. Later on, I moved to research tableau methods and in 2021, I published *Tableau Methods for Propositional Logic and Term Logic*.

MM: And how did you get interested in logic?

TJ: Well, I always liked to consider possibilities from an early age. This interest in possibilities is what drove me first to philosophy and then to logic. I have always thought that the question of what is possible is both an ontological and logical question. In particular, from a logical point of view, you can answer this question very abstractly and look for limits to what is possible according to the different logical frameworks that you chose to work with.

In Poland, many people are interested in logic. Logic is a bit of a national discipline. In addition to music and the fight for freedom, logic is also part of our Polish identity. I became interested in logic because I had good logic teachers in college. They were the heirs of the Lwów-Warsaw School, and I felt that I would like to be part of this tradition. Despite the great *hekatomba* that hit Poland in the twentieth century, this tradition has not been lost. Its genesis is interesting, but that's a topic for another conversation.

MM: Your work has largely focused on a particular type of logics: *relating logics*. Which are the main features of these logics?

TJ: Yes, right now, I am mainly involved in relating logics. However, I prefer the name *relating semantics*, since semantics is the starting point for defining logical systems. The essence of relating semantics is that in addition to evaluating sentences in the model, we can also evaluate relationships between sentences. The latter can be determined recursive or non-recursive. For example, the logical value of the sentence: "Rain has fallen and grass has grown" cannot be reduced to the logical values of sentences "Rain has fallen" and "Grass has grown".

MM: Is there any particular phenomena that you think relating logics can capture better than any other logic?

TJ: Relating Logic (*RL*) is a logic of relating connectives – just as Modal Logic is a logic of modal operators. The basic idea behind a relating connectives is that the logical value of a given complex proposition is the result of two things: (i) the logical values of the main components of this complex proposition supplemented with (ii) a valuation of the relation between these components. The latter element is a formal representation of an intensional relation that emerges from the connection of several simpler propositions into one more complex proposition.

More formally, let $A_1, \ldots A_n$ be propositions with some fixed logical values and let $\mathtt{c}$ be an n-ary relating connective. Then the logical value of complex sentence $\mathtt{c}(A_1, \ldots, A_n)$ depends not only on the logical values of $A_1, \ldots A_n$, but additionally on the value of the connection between $A_1, \ldots A_n$. It therefore depends on an additional valuation of pairs (n-tuples) that is the part of the overall process of evaluation of the logical values of complex propositions built with relating connectives. This way we can form logical systems to deal with reasoning about non-logical relationships.

Often when we replace the parameters of classically valid arguments with real sentences and the classical connectives with certain natural language connectives, bizarre inferences result, such as the one below:

$$\frac{\begin{array}{c}\text{Ann has not died or Mark is in despair.}\\ \text{Mark is not in despair or Ann is calling for a doctor.}\end{array}}{\text{Ann has not died or Ann is calling for a doctor.}} \qquad \text{(a)}$$

The problem arises because when we construct everyday arguments, we consider not only the logical values of the sentences but also expect certain non-logical relationships to hold between them, such as a causal relationship in the case above. Further examples of such relationships conveyed by arguments expressed in natural language are analytic, temporal, content, preference and connexive relationships. A formal language needs more than the standard formal apparatus of disjunction and conjunction for handing extensional phenomena; it needs machinery to make sense of intensional phenomena too.

It is easy to observe that if we interpret the expression *or* present in (a) in models $\langle v, R\rangle$ (where v is a binary valuation of variables and R is a binary relation defined on a set of formulas) in the following way: $\langle v, R\rangle \models A \vee B$ iff $\langle v, R\rangle \models A$ or $\langle v, R\rangle \models B$, and $R(A, B)$, then inference (a) is not valid (preserving, of course, the classical meaning of negation and writing $\vee$ instead of *or*). However, if we assume that R is transitive, then (a) is valid.

Although the simplest model for a relating logic is a pair: $\langle v, R\rangle$, the situation may get more complicated. We can use multi-relating models to represent more types of non-logical relations between sentences. In addition, the valuation of relationships between sentences may not be binary but may be many-valued or more subtly graded. Furthermore, we can mix relating semantics with possible world semantics, equipping all worlds with additional valuations of complex sentences. Last, but not least, any semantics may be treated as relating one, when we assume that in case of complex sentences a relationship is represented by a universal relation.

The solution that relating logics offers seems to be quite natural, since when two (or more) propositions in natural language are connected by a connective, some sort of emergence occurs. In fact, the key feature of intensionality is that adding a new connective results in the emergence of a new quality, which itself does not belong to the components of a given complex proposition built by means of the same connective. An additional valuation function determines precisely this quality. Talk of emergence is justified here, because the quality that arises as a result of the connections between the constituent propositions is not reducible to the properties of those propositions. Consequently, if the phenomenon of emergence is to be properly captured, we need additional valuations in a model. The key feature of relating semantics is that it enables us to treat non-logical relations between sentences seriously.

MM: What are you working on right now?

TJ: After we have demonstrated the effectiveness of the *alpha* algorithm, which enables the axiomatization of various relating logics, I am working on a *beta* algorithm. Its task is to move from formulas to the properties of relations in the model. We are constantly working on the application of relating semantics to various philosophical and non-classical logics, e.g. connexive logic, deontic logic or causal logic.

MM: How do you see the future of relating logics? What are the future challenges?

TJ: The community interested in relating semantics is increasing, which makes the future more promising. It is not only the Toruń Logic Group (at Nicolaus Copernicus University in Toruń) working on relating semantics, but a number of people from all over the world. After the success of the first Workshop on Relating Logic, we are organizing another one, in July 2022 in Cagliari, Sardinia, Italy. We started publishing papers in collaboration with international colleagues, we are organizing more conferences and publications, and we want to keep conducting research in this area. This future of relating logic is also connected with my students, especially Mateusz Klonowski, who is a talented and promising logician. Mateusz defended his Ph.D., entitled *Axiomatization of Monorelational Relating Logics*. It was the first doctoral dissertation devoted to this new field. I believe that the next generation will make relating logic a separate and important branch of logic.

MM: So, Tomasz, I heard that you are now working on applications of logic to the philosophy of science. How do you think these two disciplines can help each other?

TJ: I am involved in the project of formalization of mechanistic philosophy. The idea is to give an axiomatic theory that would allow to precisely lay out the assumptions and consequences of mechanistic philosophy. I think that logic can always be helpful in formulating complex theoretical concepts.

MM: What can we expect to see in the near future from your involvement in the philosophy of science?

TJ: In October 2022, on the occasion of the 550th anniversary of the birth of our compatriot, Nicolaus Copernicus, we are preparing a large philosophy of science conference titled *Hypothesis in science: A permanent feature or a temporary expedient to be tolerated?* We expect it to be a great opportunity to have leading philosophers of science interacting with leading logicians of science here in Toruń. This is one of the collective projects that I am most excited about.

MM: Actually, this event is the result of the newly "emerging field of Logic and Philosophy of Science" at the Nicolaus Copernicus University. Which are the future plans for this "emerging field"?

TJ: The research field "Logic and Philosophy of Science" is created by scientists associated primarily with the Department of Philosophy and Social Sciences of the Nicolaus University; however, an important part of the field's composition are employees of the Department of Logic. The field's staff is expected to grow with more people who, regardless of their affiliation, work on philosophical problems of logic and its application in other disciplines. The issues we deal with within the field are: non-classical logics (paraconsistent, relating, positional, deontic, epistemic, non-monotonic, etc.), applications of logic to the construction of formal theories (mereology, pointless geometries, etc.) construction and application of logical systems in the formalization of theories from the social sciences and the humanities, as well as the formal analysis of related problems in the philosophy of language and pragmatics.

Because one of the aims of the "emerging field" is to widen our academic community, we will host 15 visiting researchers – all of whom will be selected through an open competition. In addition, publication wise, we have prepared a few interesting special issues journals: a volume devoted to the history of positional logic, *The Heritage of Jerzy Łoś's Philosophical Logic*. We are also working on special issue for the journal *Logic and Logical Philosophy* which is devoted to relating semantics and it contains the products of the *1st Workshop on Relating Logic*. Furthermore, other collective publications in outstanding philosophical venues are planned. Finally, we have organized the *1st Workshop on Relating Logic* (2020), the workshop *Eclectic approaches to causation and explanation* (2021). And, in 2022, we expect to host the *World Congress of Paraconsistent Logic* (by the way, paraconsistent logic was invented in Toruń by Stanisław Jaśkowski in 1948), the conference *Bridges between Logic, Ethics and Social Sciences*, the conference series *Philosophical Perspectives on Sciences*, the *2st Workshop on Relating Logic* (in Cagliari University, Sardinia), and finally, the conference *Non-classical Logics: Theory and Applications 2022*. As you can see, we really working hard here on connecting logic and philosophy of science.

MM: Indeed!The present and near future look really promising. Congratulations, and please, do not forget to keep us, the Reasoners, updated on the activities of this emerging field.

TJ: Will do, María.

MM: Thanks for the interview, Tomasz!

TJ: Thank you for having me, Reasoners! (And also, thanks to you, Maria for nice and kind words!)

Dissemination corner

One True Logic

Logic is about *logical consequence*, the relation that holds between the premises and conclusion of a logically valid argument. It is usually thought that there are facts about the validity of arguments in natural language (here, English and its extensions), whether they're about the law, economics or metaphysics. The *one true logic* is then the formal system that matches these facts. If a natural-language argument is valid, the one true logic respects this, and similarly if it's invalid.

Nevertheless, this view has been seriously threatened since the publication of Jc Beall and Greg Restall's *Logical Pluralism* (2006), though pluralism has historical precedents in the work of Rudolf Carnap and others. Logical pluralists hold that there can be many true logics. That is, many distinct formal systems can do equally well in matching the validity facts of English. The logical pluralist can cheerfully use classical logic at one time, intuitionistic at another and relevant at yet another. Since Beall and Restall's work, a plurality of logical pluralist views have sprung up, entertaining different logics and for different reasons.

Our monograph *One True Logic*, to be published by OUP in 2022, is in one sense traditional and in another radical. As the name suggests, we believe that logical pluralism is false. There are different problems for different formulations, but we believe that there is also a general recipe for knocking them down. If a pluralism entertains a large number of logics—such as Stewart Shapiro's *eclectic* pluralism—then it faces severe metalogical difficulties. Any pluralist will want to carry out some metalogical reasoning. They may want to argue that a particular logic is appropriate for a particular domain. They may want to argue for the soundness or completeness of a logic. And they will certainly want to argue for the truth of their view. These are all examples of metalogical reasoning.

A question naturally arises: which logic should the pluralist use for metalogical reasoning? If they use only rules licensed by some of their logics and not others, they're not meeting their own standard: that's an area of deductive reasoning where one logic is less successful. So they should aim to use reasoning that's allowed in all of the logics canvassed (there can be different arguments in different logics, of course). But here's the problem: the eclectic pluralist entertains a huge range of logics and trying to work in them all leaves the pluralist with an unworkably weak metatheory, or even logical *nihilism*: the view that there are no exceptionless inferences.

Perhaps the pluralist's eclecticism was their mistake. Perhaps they let too many flowers bloom and should instead endorse just some small number of logics. This is how Beall and Restall's *modest* pluralism works. They only have to work in a few logics and the metatheoretic challenge then looks easier to answer. But why do they stop where they do? What justifies a concern with these logics and not others? Beall and Restall's stopping place is, we argue, unmotivated. And we're not convinced that any other can work either. So we're not logical pluralists.

One True Logic's endorsement of logical monism is its more traditional component. The next step is more radical. If there is one true logic, which is it? Our answer is that the one true logic must be highly infinitary. This claim is the L∞G∞S Hypothesis (pronounced 'logos hypothesis'). The two infinity symbols indicate that the one true logic is as infinitary as can be.

We offer two kinds of argument for the L∞G∞S Hypothesis: 'bottom-up', based on examples and fairly low-level principles, and 'top-down', based on logic's nature. Consider the argument $\mathcal{A}$: 'There is at least one planet, there are at least two planets, ..., there are at least n planets, ..., *so* there are infinitely many planets'. This English argument is valid: if its premises are true, then its conclusion must be true. And the one true logic, remember, is in the business of capturing such arguments. So the one true logic had better capture the validity of $\mathcal{A}$. Suitably elaborated, this shows that the one true logic cannot be first-order, the logic we teach budding logicians.

But we can't stop there. If $\mathcal{A}$ is valid, then so $\mathcal{B}$: 'There is at least one planet, there are at least two planets, ..., there are at least $\aleph_0$ planets, there are at least $\aleph_1$ planets, ..., *so* there are at least κ planets'. Here, κ is an uncountable limit cardinal. The one true logic had better capture the validity of $\mathcal{B}$. So the one true logic had better be highly infinitary. That's a sketch of one bottom-up argument for the L∞G∞S Hypothesis.

Top-down arguments, based on the nature of logic, also lead to the L∞G∞S Hypothesis. Any account of logical consequence needs a selection of logical constants. What makes, say, '$\wedge$', '$\neg$', '$\exists$' and '$=$' distinctively logical but not expressions such as 'because' or 'is red'? A popular thought is that the former are *topic neutral*: they are expressions without any particular subject matter needed to carry out inference in any domain.

Building on the work of Alfred Tarski and Gila Sher, we endorse *isomorphism invariance* as an explication of topic neutrality. Very roughly, the logical operations are those which behave the same way on any domain and however the objects in that domain are arranged. More precisely: the logical operations have extensions which are invariant under isomorphism. Intuitively, the extension of 'is red' will be sensitive to presence of red objects in the domain, whereas conjunction has no such sensitivity.

If you take this approach to the logical constants seriously, what logical constants must the one true logic contain? Reworking a result of Vann McGee's, we show that the one true logic must include maximally infinitary resources. That's a sketch of a top-down argument for the L∞G∞S Hypothesis.

Overall, we see the remarkable convergence of bottom-up and top-down arguments. They converge on a highly infinitary logic, the one true logic of our title.

Owen Griffiths
Philosophy, University of Cambridge
A.C. Paseau
Philosophy, University of Oxford

BRIO

Con brio is a term indicating a musical direction, meaning "with spirit" or "with vigor". With a wink to this original meaning, BRIO is also the acronym of a soon to be started 3-years project funded by the Italian Ministry of Research: BRIO here stands for "Bias, RIsk and Opacity: design, verification and development of Trustworthy AI".

In the wake of the extensive deployment of AIs in autonomous decision making systems, the issue of what does it mean and how it is possible to trust such systems is becoming increasingly urgent. With safety a major concern, problems of opacity, bias and risk are pressing. Creating Trustworthy AI (TAI) is thus of paramount importance, as testified by the issue making the headlines almost daily. Nonetheless, technical advances in AI design struggle to offer implementations driven by conceptual knowledge and qualitative approaches helping in this respect. The BRIO project aims at addressing these limitations, by developing design criteria for TAI based on philosophical analyses of transparency, bias and risk combined with their symbolic formalization and technical implementation for a range of platforms, including both supervised and unsupervised learning. We argue that this can be obtained through the explicit formulation of epistemic and normative principles for TAI, their development in formal design procedures and translation into computational implementations.

BRIO is articulated around 4 main objectives and will be jointly developed by 5 research units.

The first objective is epistemological and ethical. We need to offer a normative analysis of TAI as undermined by bias and risk, not only with respect to its reliability, but also to social acceptance. Our objective is to analyse the epistemological and ethical elements in the components of trust, and to provide a characterization which accounts for both the epistemic and the non-epistemic levels. We believe increasing the level of trust in AI technologies is needed but that technical reliability is not sufficient. We also must focus on how such technologies actually work in interaction with human and other artificial agents. This research task will be jointly developed by members of the META Group (Social Sciences and Humanities for Science and Technology) at Politecnico di Milano and members of the Department of Electrical Engineering and Information Technologies at the Università degli Studi di Napoli.

The output of the first objective is essential to define a comprehensive formal ontology, including a taxonomy of biases and risks and their mutual relations for autonomous decision systems. Formal ontologies as artefacts are built to make explicit the hidden assumptions behind the use of terms and notions in the model or representation of a system. Hence, modeling a system with a formal ontology helps in making its working mechanisms more transparent, and its performances as human-understandable as possible. The next task is to offer a systematic characterization of bias types and make them viable for formal and automatic identification, as well as to help the identification of possible sources of risk in the design phase to avoid harm for humans interacting with AI. This research task will be jointly developed by the Institute for Cognitive Sciences and Technologies of the National Research Council in Genova and the Department of Philosophy at Università degli Studi di Genova.

A formal ontology of bias and risk in AI is to be crucially integrated within the design of (sub)-symbolic formal models to reason about safe TAI, and produce associated verification tools. We will articulate the cognitive representation of classification biases in ML by extensions of Description Logics (DL) in terms of concepts with thresholds, defined by a list of weighted features that represent their importance for the classification, and are tuned to reflect the likelihood of the properties attached by the ML algorithm. On this basis, the uncertain behaviour of complex AI systems can be modelled by inferential probabilistic processes, including their expected outputs and the agent's beliefs about them. We will develop proof-theories and semantics to model output divergence from the intended and expected outcomes and the margins of risk in obtaining unintended side-effects. These models are apt for implementation aimed at formal verification through proof- and model-checkers. This research task will be jointly developed by members of the Logic, Uncertainty, Computation and Information Group at the University of Milan (luci.unimi.it/) and the Department of Philosophy at Università degli Studi di Genova.

Finally, BRIO aims at developing a novel computational framework for TAI systems explanation capabilities, to mitigate the opacity of Machine Learning (ML) models. The proposed framework will go beyond widespread deficiencies of XAI approaches based on low-level or high-level features, explaining the ML model responses in terms of hierarchical structures and compositional properties of middle-level features. In view of their identified compositionality properties, a roadmap will be outlined towards a hybrid neural-symbolic framework for explanation in TAI, integrating the proposed computational framework for the explanation of ML model behaviours with other symbolic approaches (formal ontologies, logical inference, formal specification and verification methods). This research task will be led by members of the Department of Electrical Engineering and Information Technologies at the Università degli Studi di Napoli.

In the following months members of the BRIO project will regularly post their findings and thoughts on this outlet. For more news, updates and job openings associated with the project follow us at sites.unimi.it/brio/

With spirit and vigor, the philosophical and logical underpinnings of this project will make every effort to contribute to the development of AIs with less bias, more transparent and where associated risks are known.

Giuseppe Primiero
Philosophy, University of Milan

What's Hot in . . .

Statistical Relational AI

Over the past decade, the annual workshop on statistical relational AI (branded StaRAI) was a popular venue for work in progress as well as substantial technical contributions to the field. While it was previously colocated with one of the broad flagship artificial intelligence conferences AAAI, UAI or IJCAI, this year for the first time it was combined with conferences and workshops on Inductive Logic Programming (ILP), Neural Symbolic Integration (NeSy) and Program induction (AAIP) in a brandnew format, the *1st International Conference on Learning and Reasoning (IJCLR)*. Originally intended to be held in September 2020 in Athens, it was postponed for a year and eventually held as an online meeting in October 2021. In addition to the traditional formats of the participating events, the IJCLR featured joint sessions of both talks and posters as well as the option to submit papers to a designated journal track.

To give a flavour of the synergies between statistical relational AI and neural-symbolic reasoning, I report on the first joint session of the IJCLR, devoted to a statistical relational perspective on Graph Neural Networks (GNNs) GNNs, first proposed more than a decade ago, adapt the concepts of deep learning to input data carrying a graph structure. By dealing natively with structured input, GNNs avoid the need for preprocessing the input data into a vectorised form, a step that leads to intransparency and a loss of information. Since their introduction, hundreds of variants of GNNs have been proposed, and GNNs have started to have practical impact in areas such as chemoinformatics and natural language processing. By working with graph-structured input to answer queries both about the graph as a whole and about individual nodes, GNNs are carrying out very similar tasks to statistical relational approaches, suggesting both integration of and comparison between those areas as urgent research issues.

The joint session of the IJCLR featured three papers covering different aspects of this issue:

Sourek, Zelezny and Kuzelka discuss lifted relational neural networks as a declarative differential programming environment for the high-level representation of GNN arcitectures. Syntactically a paramterised version of the well-known Datalog language, lifted relational neural networks encode GNN architectures as rules whose literals carry tensor parameters. Learning examples are then encoded as parametrised facts, where the parameter takes unit value for logical facts and tensor values otherwise. The authors then demonstrate how various GNNs and extensions thereof can be expressed in the proposed formalism, and demonstrate experimentally that lifted relational neural networks can actually outperform dedicated contemporary GNN implementations on a wide variety of molecular classification tasks.

Dash, Srinivasan and Vig discuss the problem of how best to incorporate domain-knowledge into GNNs. This is particularly relevant in highly structured domains and where learning examples are scarce, such as the drug evaluation scenario underlying the real-world datasets used for the empirical evaluation of the paper. The authors introduce *vertex-enrichment* to include polyadic relational background knowledge onto a binary graph; in this way, for instance, knowledge of functional groups and rings can be added to the molecular structure encoded as input for a GNN model. In addition to adding the background knowledge taken from the literature, the authors also investigate the usefulness of adding an explicit preprocessing step in which the inductive logic programming system Aleph is used to infer higher predicates from the background knowledge, and then to add them to the vertex-enriched GNN. The experimental evidence provided suggests that both steps improve the accuracy of the learned GNN model.

Embar, Srinivasan and Getoor compare GNNs and statistical relational approaches on their performance in answering aggregate graph queries, more complex queries that aggregate information about multiple nodes to elucidate properties of the graph as a whole. Aggregate graph queries can be used to assess the community structure of a coloured graph, for instance by querying the numbers of intra- and inter-community edges between graph nodes. The authors discuss various techniques for approximating the answers to such queries, necessitated by the intractability of any direct exact computation. As aggregate graph queries reference the entire graph at once, statistical relational formalisms which explicitly model the joint distribution allow for a straightforward approach through sampling from the joint distribution, while GNN models that do not model a joint distribution have to rely on extrapolating from individual predictions. Empirical results reflect this, and statistical relational methods outperform GNNs on several citation datasets.

Overall, the research presented here has made a good case for the continuing relevance of symbolic relational approaches in the era of deep learning, and in particular underlined that statistical relational representations have a role to play even in settings where accuracy is more important than explainability. On the other hand, the comparisons made here pitted GNNs and symbolic methods in application domains that are considered traditional strongholds of relational techniques. Indeed, predicting molecular properties from their structure was one of the first breakthrough domains of inductive logic programming in the early 90s, and highly complex, aggregate queries seem tailor-made for the expressive and explicit representation of joint probabilities that statistical relational models provide.

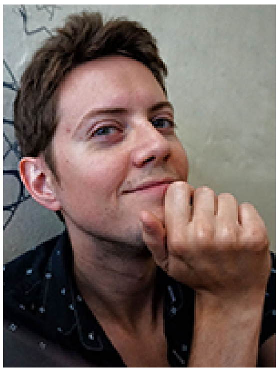
The second edition of the IJCLR is already planned for September 2022 as an in-person meeting in Windsor, UK. However, the composition of associated events will change, and the StarAI workshop will no longer be colocated with the IJCLR. Instead, the HLC workshop on human-like computing will be included, alongside ILP, NeSy and AAIP.

The website of the IJCLR 2021 includes a list of all accepted papers, including those of the associated conferences, and the website of IJCLR 2022 will be updated with calls for papers and further details in due course.

Felix Weitkämper
Computer Science, LMU Munich

Mathematical Philosophy

Understanding has lately taken center stage in many branches of philosophy, both as an object of independent interest and a key ingredient in the analysis of other important concepts. We've seen, for instance, understanding-based accounts of rationality (Schafer, "Rationality as the capacity for understanding", *Noûs* 2019) scientific progress (Dellsén, "Understanding scientific progress: The noetic account", *Synthese* 2021) and explanation (Wilkenfeld, "Functional explaining: A new approach to the philosophy of explanation", *Synthese* 2014).



All fine and dandy. But where's the mathematics angle here, you ask? Well, following Wilkenfeld's lead, Inglis and Mejía-Ramos have recently defended an account of explanatory proof in which mathematical understanding plays a starring role ("Functional explanation in mathematics", *Synthese* 2019). IMR's work is notable because, unlike that of many understanding-fans, it offers specific, psychologically informed details about the nature of understanding and the kinds of reasoning that promote it.

A basic component of IMR's account is the concept of a *schema*, introduced by Piaget and still widely used. A schema is, roughly speaking, a structured mental representation of a subject matter S that facilitates remembering, recognizing, predicting and reasoning about information related to S.

Schemas are thought to play a role in many cognitive tasks. For instance, you have a *house* schema, which represents your understanding of the important general attributes of houses (e.g. that they have parts, are made of materials, and have functions) as well as some default specific values for those attributes (e.g. that the parts of a house are rooms, that the materials are wood or brick, and that the function is human habitation). Accessing this schema can provide various benefits. Since its components are strongly cognitively interlinked, the schema allows you to quickly classify something as a house on the basis of a few key characteristics. And once classified, the schema helps you recognize other expected house-features, make inferences about features you didn't directly perceive and remember what you saw later.

We also have schemas for more abstract subjects, like chess games and mathematical problems, which play cognitive roles similar to schemas for everyday objects. A skilled chess player will have a detailed *Sicilian Defense* schema, for instance, encoding the situations in which this opening is likely to be played, its most important variations and the counterplays for each, typical middle-game positions to which the Sicilian tends to lead, and so on. A player with many such schemas at her disposal will be able to choose moves quickly and accurately, whereas one without them will have to painstakingly note the positions of each piece on the board, call to mind their possible moves and calculate the consequences of each.

On IMR's view, the creation and consolidation of appropriate schemas is constitutive of mathematical understanding. As they write, "one can be said to have understood something when a sufficiently well-organised schema... has been encoded into long-term memory" (S6381). For noeticists like IMR, the task of a theory of explanatory proof is therefore to say what kinds of proofs let us generate and store high-quality mathematical schemas.

To fill in these details, IMR draw on a simple model of human cognitive architecture. The key components of this "modal model" are sensory memory, working memory and long-term memory. The job of sensory memory is to briefly store sense impressions before passing them on for further processing. Long-term memory is the storage site for cognitive schemas. Working memory receives inputs from both sensory memory and long-term memory, and it consists of independent verbal and visual information channels.

In order for a new schema to be created or an old one modified, information has to pass from working memory to long-term memory. There are several constraints on the efficiency of this process. One is the limited bandwidth of working memory. Another is the inability of working memory to hold information for more than a short time. To get around these limitations, working memory needs to be presented with efficiently formatted content. For instance, things go better if working memory receives input from long-term memory in the form of well-organized and easily manipulable schemas. And if a given content is too complex for either the verbal or visual channel to handle on its own, it should be presented in a way that splits the informational load between the two channels.

Putting these considerations together, IMR arrive at the following account of explanatory proof:

> Our account suggests that the archetypal explanatory proof would have at least three properties. First, it would have features that make it easy, or at least as easy as possible, to select the information from sensory memory into working memory that is neces-

> sary for a successful processing stage. ...Second, it would have features that make it easier to coordinate the new knowledge contained in the proof with existing schemas retrieved from long-term memory, and therefore to reorganise the new and existing information into coherent new schemas. Finally, it would be likely to split the working memory load it gives to its readers between their visual and verbal/auditory channels so that the chances of their working memory capacity being exceeded during the schema-organisation process is minimised. (S6381)

These are certainly pleasant and useful things for a proof to do. And many explanatory proofs do indeed rate highly on IMR's criteria. But I wonder whether this is the whole story about mathematical explanation and understanding.

Consider the modern proof of Gauss's quadratic reciprocity theorem based on algebraic number theory. Originally due to Hilbert, the proof derives quadratic reciprocity through an analysis of the behavior of prime ideals in algebraic number fields. This style of proof, it seems, strikingly fails to satisfy IMR's criteria. Its argumentation is complex. It's full of sophisticated technical notions—ramification of prime ideals, discriminants of algebraic number fields, Galois groups, Frobenius automorphisms, and so on. It features no visual (or easily visualizable) elements to speak of. So the proof exhibits none of the cognitive facility represented by IMR's conditions. Nevertheless, many number theorists believe this approach best explains why quadratic reciprocity holds. It's been said of the proof, for instance, that it "goes... straight to the root of the phenomenon", that it "provides the correct explanation of Gauss's reciprocity law", and that "the content of the quadratic reciprocity law only becomes understandable" using its methods (D'Alessandro, "Proving quadratic reciprocity", *Synthese* 2021).

I've argued that number theorists value Hilbert's proof for its *depth*—an explanatory virtue having to do with exposing hidden reasons and situating a theorem in a broader network of concepts and results. It's not just that deep proofs are explanatory *in spite of* not offering familiar or easily consolidated schemas. Rather, their explanatory value is closely tied to the novel machinery they employ, with all the cognitive effort required to grasp it. The fact that such proofs lead our minds down rough and unfamiliar roads is an explanatory feature, not a bug.

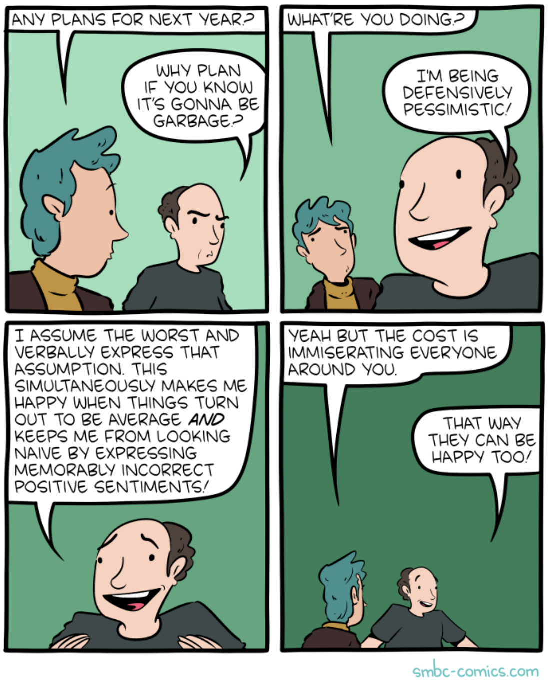
Whether or not this particular worry is on the right track, I think IMR deserve heaps of credit for proposing an account of understanding that makes specific, nontrivial, scientifically motivated predictions. And I have a lot of sympathy for understanding-based approaches to mathematical explanation in general. I hope we're only seeing the first surge of a wave of new work in this direction—long live the noeticist revolution!

William D'Alessandro
MCMP, Munich

Events

January

PoMaL: Cambridge Graduate Conference on the Philosophy of Mathematics and Logic, University of Cambridge, 22–23 January.



Courses and Programmes

Courses

CiE: Computability in Europe 2021: Connecting with Computability Tutorials, 5–9 July.

Programmes

MA in Reasoning, Analysis and Modelling: University of Milan, Italy.

APhil: MA/PhD in Analytic Philosophy, University of Barcelona.

Master Programme: MA in Pure and Applied Logic, University of Barcelona.

Doctoral Programme in Philosophy: Language, Mind and Practice, Department of Philosophy, University of Zurich, Switzerland.

Doctoral Programme in Philosophy: Department of Philosophy, University of Milan, Italy.

LogiCS: Joint doctoral program on Logical Methods in Computer Science, TU Wien, TU Graz, and JKU Linz, Austria.

HPSM: MA in the History and Philosophy of Science and Medicine, Durham University.

Master Programme: in Statistics, University College Dublin.

LoPhiSC: Master in Logic, Philosophy of Science and Epistemology, Pantheon-Sorbonne University (Paris 1) and Paris-Sorbonne University (Paris 4).

Master Programme: in Artificial Intelligence, Radboud University Nijmegen, the Netherlands.

Master Programme: Philosophy and Economics, Institute of Philosophy, University of Bayreuth.

MA in Cognitive Science: School of Politics, International Studies and Philosophy, Queen's University Belfast.

MA in Logic and the Philosophy of Mathematics: Department of Philosophy, University of Bristol.

MA Programmes: in Philosophy of Science, University of Leeds.
MA in Logic and Philosophy of Science: Faculty of Philosophy, Philosophy of Science and Study of Religion, LMU Munich.
MA in Logic and Theory of Science: Department of Logic of the Eotvos Lorand University, Budapest, Hungary.
MA in Metaphysics, Language, and Mind: Department of Philosophy, University of Liverpool.
MA in Mind, Brain and Learning: Westminster Institute of Education, Oxford Brookes University.
MA in Philosophy: by research, Tilburg University.
MA in Philosophy, Science and Society: TiLPS, Tilburg University.
MA in Philosophy of Biological and Cognitive Sciences: Department of Philosophy, University of Bristol.
MA in Rhetoric: School of Journalism, Media and Communication, University of Central Lancashire.
MA programmes: in Philosophy of Language and Linguistics, and Philosophy of Mind and Psychology, University of Birmingham.
MRes in Methods and Practices of Philosophical Research: Northern Institute of Philosophy, University of Aberdeen.
MSc in Applied Statistics: Department of Economics, Mathematics and Statistics, Birkbeck, University of London.
MSc in Applied Statistics and Datamining: School of Mathematics and Statistics, University of St Andrews.
MSc in Artificial Intelligence: Faculty of Engineering, University of Leeds.
MSc in Cognitive & Decision Sciences: Psychology, University College London.
MSc in Cognitive Systems: Language, Learning, and Reasoning, University of Potsdam.
MSc in Cognitive Science: University of Osnabrück, Germany.
MSc in Cognitive Psychology/Neuropsychology: School of Psychology, University of Kent.
MSc in Logic: Institute for Logic, Language and Computation, University of Amsterdam.
MSc in Mind, Language & Embodied Cognition: School of Philosophy, Psychology and Language Sciences, University of Edinburgh.
MSc in Philosophy of Science, Technology and Society: University of Twente, The Netherlands.
MRes in Cognitive Science and Humanities: Language, Communication and Organization: Institute for Logic, Cognition, Language, and Information, University of the Basque Country (Donostia San Sebastián).
Open Mind: International School of Advanced Studies in Cognitive Sciences, University of Bucharest.
Research Master in Philosophy and Economics: Erasmus University Rotterdam, The Netherlands.

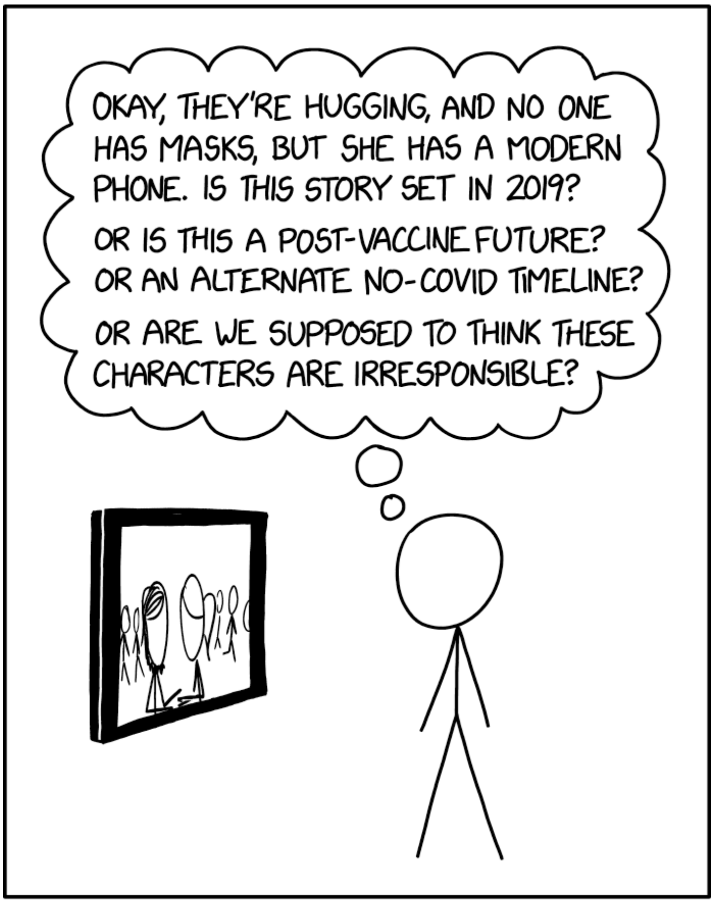
Jobs and Studentships

Studentships

Doctoral Programme in Philosophy: Language, Mind and Practice, Department of Philosophy, University of Zurich, Switzerland.
LogiCS: Joint doctoral program on Logical Methods in Computer Science, TU Wien, TU Graz, and JKU Linz, Austria.

Jobs

Assistant Professor: in Logic and Philosophy of Mathematics, University of Warwick, deadline 9 January.
PhD position: in Uncertainty Quantification for Precision Medicine, University of Exeter, deadline 10 January.
Post-doc: in History of Philosophy of Science, Tilburg University, deadline 14 January.
Post-doc: in Causality & Knowledge Graphs, University of Edinburgh, deadline 23 January.
Professor: in Philosophy of Medicine, University of Bordeaux, France, deadline to be determined.



MOVIES AND SHOWS THAT ARE VAGUELY SET IN "THE PRESENT" WILL BE AWKWARD FOR A WHILE.