Giovanni Sommaruga (ed): Formal Theories of Information: From Shannon to Semantic Information Theory and General Concepts of Information

Lecture Notes in Computer Science 5363, Springer-Verlag, Berlin, 2009, 267 pp, $64.95, ISBN: 9783-642-00658-6

Sebastian Sequoiah-Grayson

This new anthology on formal theories of information is based upon research presented at the May 2006 Muenchenwiler seminar of the Information and Knowledge research groups of the computer science departments of the universities of Bern, Fribourg, and Neuchatel. The collection is split into five sections: Philosophical Reflections; The Syntactical Approach; The Semantical Approach; Beyond the Semantical Approach; and Philosophical Conclusions. Giovanni Sommaruga's introduction is spot on, combining a brief but erudite introduction to the conceptual motivations underpinning research into formal theories of information with a thorough summary of the volume's contributions.

The first section, Philosophical Reflections, consists of Luciano Floridi's "Philosophical Conceptions of Information." Floridi's piece is a good choice for the opening contribution both for its content and its delivery. He takes the reader on a clear and accessible journey through every day uses of information, Shannon's mathematical theory of communication, Bar-Hillel and Carnap's theory of semantic information, Floridi's own theory of strongly semantic information, and a range of philosophical issues.

Accessibility is not achieved at rigour's expense; Floridi expects some philosophical acumen on the reader's part, covering the issues with care. Some time is taken to ensure that the reader understands that Shannon's communication theory is concerned with information measures based on the surprise value of the signals in question, and not concerned at all with the *content carried by* such signals. It was Bar-Hillel and Carnap's later work on semantic information that outlined an analysis (both qualitative and quantitative) of informational *content*. Although, both

S. Sequoiah-Grayson (✉)
Department of Theoretical Philosophy, Faculty of Philosophy, University of Groningen, Groningen, The Netherlands
e-mail: s.sequoiah-grayson@rug.nl

S. Sequoiah-Grayson
IEG—Computing Laboratory, University of Oxford, Oxford, UK

theories are similar at the formal level, their objects of analysis could not be more distinct.

Floridi covers sub-informational concepts such as *data* in detail, as well as applications of informationally-informed philosophical method to notoriously thorny philosophical issues surrounding *truth*, *hyperintensionality*, and *mental representations*. Perhaps the best aspect of Floridi's piece is the balance that it finds between expositing formal theories of information on the one hand, and demonstrating applications of such theories to traditional philosophical issues on the other.

There is also some dangerous philosophical terrain here. Floridi moves quickly between information itself, which is a commodity, various concepts *of* information, various formal analyses of the commodity of information and conceptual analyses of the concept of information, as well as pragmatist-style reflections on the correct use of the word 'information'. Such distinctions are not always marked. Marking them will reveal still more avenues for research, and Floridi has done a stellar job of identifying enough of these avenues to keep us busy for some time.

The second section, The Syntactical Approach, consists of two contributions, the first being Francois Bavaud's "Information Theory, Relative Entropy and Statistics." Bavaud's piece is a demonstration of the conceptual significance of applications of Shannon's mathematical theory of communication. The motivating idea is that we take the two arguments of Shannon's relative entropy function to correspond to probability measures of observational data and our expectations. In this way, Shannon's relative entropy function is understood as an asymmetric measure of the *dissimilarity* between empirical and theoretical probability distributions; hence relative entropy becomes "epistemicised".

Within this motivating framework, Bavaud covers notions such as hard falsifications, maximum entropy, factor analysis, alternating minimumisation, Markov chain models, and the applications of such notions to tasks such as hypothesis testing (for both single and multiple hypothesis cases), Bayesian selections, and textual analysis.

The main issue here is one of accessibility. Bavaud's contribution is the most demanding entry in the collection by a margin. This is so much the case that it is unlikely that those without a strong background in statistics will get very far at all. This is a shame, as the philosophical issues that Bavaud is dealing with are important.

The second contribution to the second section is Cristian S. Calude's "Information: The Algorithmic Paradigm," which focuses on the exposition of *algorithmic information theory* and its conceptual implications. Calude has a large knowledge of other work on formal theories of information (from semantic theories and logics of information to quantum and ecological information) and situates his entry in relation to these.

This is probably the clearest account of algorithmic information theory that one will come across. The sections on the halting problem and data compression in particular are especially transparent. By far the most fun is had in section four, which concerns itself with the issue of whether or not computers can create information. Calude never loses sight of the fact that the issue is being dealt with in

terms of information in the strict sense delineated by algorithmic information theory, and he is aware of the connections with the general issue of hyperintensionality as it applies to logical truths and deductions. With computational processes being canonical examples of deductive structures, it would be interesting to follow these connections through.

Calude takes us all the way to algorithmic randomness and its connections to incompleteness and halting, as well as to incompleteness and uncertainty. Throughout, Calude remains careful of the distinction between formal theories of information on the one hand, and their objects of analysis on the other.

Section three, The Semantical Approach, consists of four articles, with the first being Jurg Kohlas and Cesar Schneuwly's dense but clear "Information Algebra." The important thing to note is that in this context 'information algebra' is a proper name, not a descriptive one. For example, there is nothing here on Heyting algebras and constructive logics, nor on Dunn's work. Instead, Kohlas and Schneuwly develop a detailed algebraic system based upon the idea that the domains of various algebraic relations correspond to pieces of information that give partial *answers* to the following question: Which of the elements of a Cartesian space represent the true values of the variables.

Conceptually, Kohlas and Schneuwly have developed their information algebra in order to model aspects of information processing, rather than identification of symbols acted on by such processes (as was Shannon's concern). They develop their algebra in both labelled and domain-free versions, introduce the operation of information-transport, and take propositional and predicate logic, as well as various contexts of informational transfer, as the objects of their analysis.

It is worth emphasising that Kohlas and Schneuwly do all of this with some pedagogical rigour. For example, section four is as good a demonstration of the relationship between qualitative and quantitative theories of information as one will find anywhere (although their terminology and notation departs from common algebraic usage in places, using "neutral element" for "unit" and so on).

The main issue with Kohlas and Schneuwly's piece is a philosophical one. They claim, variously, that their information algebra "covers important aspects of a general theory of information" (p. 98), or that "[i]nformation algebras represent a structure which captures essential features of any concept of 'information'". That information and the concept of information are distinct things notwithstanding, an information algebra (in Kohlas and Schneuwly's sense) possesses properties such as association, commutation, idempotency, the presence of unit, and so forth, that are violated by a range of information-processing scenarios (lexical databases and the resource-sensitive structures captured by variations of Hopf Algebras for just two examples). Relatedly, under an information algebra analysis, logically equivalent formulas in propositional logic describe the same information.

The second article in section three is Jurg Kohlas and Christian Eichenberger's "Uncertain Information." Their task is to expound algebras of basic probability assignments, and to then connect these algebras with Bayesian approaches. Kohlas and Eichenberger explain the notions of decoding and lossy channels from Shannon's communication theory in terms of uncertain information. This allows them to connect the information algebra framework with Shannon's communication

theory. Their central operation is a *hint*, defined as a quadruple consisting of a set of assumptions, a probability measure on this set, and a function from the set of assumptions to the power set of parameters. Hints represent the (uncertain) information drawn from experiment.

Kohlas and Eichenberger spend some time demonstrating the connections between hypothesis (as uncertain information) and assumptions in propositional logic. This is done neatly, as is their section five on the connections between information order and information measure (as was the case in Kohlas' previous piece on information algebra). Kohlas and Eichenberger prove that the algebra corresponding to assumption based inference satisfies all of the axioms of an information algebra, with the exception of idempotency. This makes perfect sense from a conceptual standpoint, as the same assumption may need to be made more than once in experimental settings.

Robert van Rooij's piece "Comparing Questions and Answers: A Bit of Logic, a Bit of Language, and some Bits of Information" is the third article in this section. This entry is getting on a little now, with most of the material having been written in 2000. Since then he has written several follow-up pieces, but the piece here remains foundational. van Rooij uses the notions of entropy and value-expectation from Shannon's communication theory in order to give a quantitative theory of the goodness of contextualised questions and answers. In short, van Rooij starts with Groenendijk and Stokhof's qualitative partitions theory, before moving on to Shannon's theory of communication and statistical decision theory. van Rooij's exposition of intensional truth-conditional semantics is *very* careful, and certainly one of the clearest that one will encounter in the literature. From here we are taken through several increasingly fine-grained approaches to providing a semantics for the goodness of answers to questions.

There is an initially appealing idea that a second answer is more informative than the first if the first answer entails the second, but in this case overinformative answers (with irrelevant conjuncts) would be preferable to less-informative but highly relevant answers, which they are not. From here we move to Groenendijk and Stokhof's partition theory, from which the basic idea is that the goodness of an answer may be measured in terms of the partition induced by the question; the more cells of the partition that a correct answer is incompatible with, the better the answer. Problems arise here when we incorporate the beliefs of the questioner as imposing constraints (as they do) on what counts as more or less informative answers to particular questions.

The expressive limitations of partition theory and its constituent partial orderings lead van Rooij to move to the total orderings and quantitative framework provided by Shannon's theory of communication. Ultimately, he suggests that a game theoretic framework will be needed in order to account for the full semantic behaviour of questions and answers. This is more than a little prescient of van Rooij, as the contemporary game-theoretic research environment attests.

There are some conceptual stumbles in section four, where van Rooij conflates Shannon's communication theory with Bar-Hillel and Carnap's theory of semantic information, but anyone familiar with the area (or who has read Floridi's piece in the volume) should make it through safely enough.

The fourth and final contribution to the section on The Semantical Approach is Jeremy Seligman's "Channels: From Logic to Probability." Seligman's entry takes at the outset information to be a commodity which reduces uncertainty, and then takes the reduction of uncertainty to be modelled by both logic and probability theory. The point of this is to extend the analysis of information-via-logic from the channel theoretic framework that he developed with Jon Barwise to a probabilistic approach. This is no easy task, but Seligman brings the reader along with him at every turn.

Seligman's theory is *dynamic*; concerned with modelling information flow rather than simply individuating information structures. The important difference between the approaches to information flow via dynamic epistemic logic and that taken via channel theory, is that the former preserves the use of expressions in a formal language to express content. The goal of Seligman's (and Barwise's) theory, is for such dynamics to be captured in a way that is independent of any underlying formal system.

The first task undertaken by Seligman is to expound the channel-theoretic notions of *classifications*, *infomorphisms*, and *channels*. A classification is essentially a *typing* of tokens. An infomorphism is a pair of inverse functions between classifications. A (binary) channel between two classifications S (source) and R (receiver) consists of a third classification C, and a pair of infomorphisms such that one is a pair of inverse functions between C and S, and the other a pair of inverse functions between C and R. C is called the *core* of the channel. Seligman then moves through a channel-theoretic analysis of information structures from the logic side (Tarski-channels) and probability side (Shannon-channels).

The next channel-theoretic notion Seligman introduces is that of a *link*. A link is basically a pair on a classification such that the pair consists of a subset of the tokens of the classification, and a binary relation between types of the classification such that if this relation holds, then it is respected by every token in the subset. Seligman then dedicates section three to assigning links with both logic and probability theory. Restating Seligman's ultimate goal in more detail, it is to develop a theory of links independently of any underlying formal system.

There is a great deal of contrasted conceptual work in Seligman's final sections, with connections made to various aspects of category theory, Dretske's theories of information flow, and the Barwise-Seligman models of information flow as the movement of local logics (logics that model parts of the world) around a network of classifications. This is distinct from the model of information flow in the present entry, as a relation between types or tokens in the classifications S or R. A technical appendix covering the details of certain proofs finishes things off.

Seligman's entry is perhaps the most conceptually adventurous of the collection, as it is breaking ground that readers are not likely to have encountered elsewhere, whilst also situating itself within contemporary research on logics of information flow.

The fourth section of the collection, Beyond the Semantical Approach, consists of Keith Devlin's "Modelling Real Reasoning." Devlin's entry is an extremely accessible introduction to the conceptual motivations behind the use of situation

semantics to model human reasoning in real-world situated scenarios. Devlin introduces the relevant concepts (of which there are many) with dialectical ease.

Devlin is clear that he is modelling information itself (not the concept of it), and he notes its status as a commodity. He goes to considerable lengths to lay out the properties of real-world reasoning (the fact that it is holistic and not always linear etc.), before deploying situation semantics towards modelling it.

One of Devlin's central insights, that is as obvious as it is frequently overlooked, is that humans notice regularities in the world, and that it is this ability to perceive regularities that allows humans to acquire information about their environment. Once gotten, such information can be stored and processed and reasoned with in various ways, and it is this reasoning that Devlin's framework is designed to model. The catch here is that Devlin's theory is not actually a theory of *reasoning* at all. Rather, it is a theory of the results of such reasoning procedures. That is, reasoning is a dynamic procedure, and Devlin's theory is static.

The final section, Philosophical Conclusions, is made up of Giovanni Sommaruga's "One or Many Concepts of Information." Sommaruga's entry on informational pluralism is both retrospective and speculative in the best way possible. Using the volume's contributions as a platform, Sommaruga goes to considerable lengths in order to indicate their overlapping points of view, as well as their points of departure. This is quite a task, and Sommaruga accomplishes it with both clarity and brevity. That Sommaruga argues in favour of an informational pluralism was to be both expected and welcomed, as such a state of affairs was predicted by Shannon, for excellent reasons, at the beginning of research into formal information theories (as articulated in Floridi's entry to this volume).

Sommaruga's entry prompts further questions of a foundational nature. For one, to what extent are formal theories of information *itself* constrained by results from the conceptual analysis of the *concept* of information? A certain lack of constraint seems likely, as the operational irrelevance of philosophy of mathematics on the ontological status of numbers has been to much of progress in mathematics suggests. But the existence of something is not the same thing as the sanction of it, and such debates are likely to continue for some time.

This is just one of the open questions that Sommaruga's volume leaves us with, and there are more. This is both what we should expect and what we should desire. Formal theories of information and their philosophical analysis are being developed right now, and this is what makes a volume of this quality so welcome.