Hecke, Oswald Teichmüller, Ernst Witt, Richard Courant, Edmund Landau, Felix Hausdorff, Ernst Peschl, Paul Riebesell, Helmut Ulm, Alfred Stohr, Ernst Zermelo, Gerhard Gentzen, Hans Petersson, Erich Kähler, and Wilhelm Süss. The names on this list range from the committed Nazis, through non-Nazi right wing nationalists, to naive, otherworldly men who seemingly didn't know much of what was going on, to those who stumbled into situations that were beyond their control. The stories are gripping. In assessing guilt the reader is hard pressed to decide with any certainty which people fall into which categories. Teichmüller comes off very badly, for example. Segal suggests that he in fact came up with the theory that Aryan mathematics was different and superior to "Jewish and French" mathematics, a notion usually credited to Bieberbach. Teichmüller claimed the student revolt against courses taught by Landau was not anti-Semitic but pro-German. He was a brilliant mathematician but a Nazi fanatic. He volunteered to go fight on the Russian front and was killed there at the age of 30. One should probably keep in mind that because of his early death, he, unlike many others, never had a chance to try to redeem his reputation during the denazification period after the war.

In this last chapter, the reader cannot help regretting that certain other mathematicians were not included for this more expansive treatment: Hans Zassenhaus, Gustav Doetsch, Georg Hamel, Helmuth Kneser, or Erhard Schmidt, for example. Perhaps Segal felt that their involvement was adequately described elsewhere in the text.

Of those treated, some behaved well, some badly. All were competent mathematicians; some were giants. Their mathematics, however, did not save some of them from being monsters.

Department of Mathematics and
Computer Science
Santa Clara University
Santa Clara, CA 95053-0290
USA
e-mail: galexand@math.scu.edu

The Mathematical Theory of Information

by Jan Kåhre

BOSTON, KLUWER. 2002. 520 PP., US$50.00
ISBN: 1-4020-7064-0

REVIEWED BY CRISTIAN S. CALUDE

Undoubtedly, the title of the book was well chosen: it is provocative, promising, and full of information. Syntactically, the title can be viewed as a variation on the titles of both the seminal paper [11] ("a" is replaced by "the") and the book [12] ("Communication" is replaced by "Information"). It provocatively questions Shannon's theory; according to [1] (page 215), "no prophet remains unchallenged for ever". And it promises "a new mathematical theory of information, built on a single powerful postulate: The Law of Diminishing Information."

The book was praised—"a bold new approach to classical information theory"—by von Baeyer [1], who dedicated a special chapter of his book to it; details about Kåhre and the fascinating autonomous islands Åland on which he lives are presented too.

Do we need a new information theory? Unsurprisingly, there is no one single theory of information, but several theories: semantic theories [2], algorithmic information theory [5,4], logic of information [7], information algebra [9], philosophy of information [8], information flow [3], quantum information theory [10], evolutionary information [13], to name just a few (a workshop devoted to various theories of information was recently held in Münchenwiler). Each theory focuses on some specific aspect of information, and overlaps are minimal. There is little evidence that the existing theories will converge towards a single, unified theory of information, so, indeed, there is ample room for (even a partial) unification.

The book discusses information from various angles, with interesting ideas and many examples. Bits and entropy are used for quantitative problems, while hits (the number of correct classifications), reliability, and nuts (von Neumann's utility) appear in more qualitative analyses.

Although the author's ambition is to develop a (if not *the*) "mathematical theory of information," the embodiment is *pre-mathematical*. It is neither a naive mathematical theory (as in naive set theory) nor is it abused mathematics (in the sense of mathematics applied in meaningless ways). However, the mathematical formalism is too rudimentary for a theory; I illustrate this point with two examples, the definitions of probability and algorithmic complexity. The probability $P(a)$ is a real number that satisfies the following three axioms (pages 25–26): probability cannot be a negative number, the probability of something that must occur is 1, and the probability that a or b will occur is the sum of their probabilities provided that a and b cannot both occur.[1] Kolmogorov complexity is defined (page 234) as the length $l(a_i)$ of the shortest algorithm generating a given b_j. In both cases the intuition is correct; even if some facts can be deduced from those definitions, there is still a long way to a satisfactory mathematical presentation.

The book is rather firmly based on Shannon's probabilistic view of information and entropy; the standard books [12, 6] are frequently used and cited. The information measure used in the book is defined by inf(B@A) = the **inf**ormation **B** gives about **A** (author's notation). Here inf(B@A) is a real function satisfying the Law of Diminishing Information (or, simply, the Law, as it is referred in the book): *Compared to direct reception, an intermediary can only decrease the amount of information.* If $A \rightarrow B \rightarrow C$ denotes a transmission chain, then the Law reads: inf(C@A) $\leq$ inf(B@A). The "theory of information" developed in the book is based on probability (as defined below) plus the Law (page 14):

> [The Law] will be used as the fundamental axiom of the mathematical theory of information. The Law is the pruning knife of information theory: we will argue that the Law is the necessary and sufficient con-

[1] It takes no fewer than 235 pages to realise that probabilities, defined in this way, apply only to finite sets (see section 8.4).

dition for a mathematical function to be accepted as an information measure, i.e. qualify as inf(B@A).

The Law is easy to understand and informally seems correct (for example, to anyone who has played the Telephone game in which one person chooses a sentence, whispers it into the ear of her left neighbour, who in turn whispers it into the ear of her left neighbour, and so on down the line). It ties in well with other principles such as the second law of thermodynamics, the data processing inequality [6], and the invariance of algorithmic complexity under computable transformations [4]. Moreover, it is not difficult to see that there are infinitely many functions satisfying the Law (trivially, each constant function satisfies the inequality). Finally, a weaker form of the Law has already been discussed in [6] (page 32) as a consequence of the data processing inequality.

According to the book, the Law is ubiquitous. It makes physics possible: "systems are forgetting their past as they reach equilibrium, or rather, the initial conditions can be eliminated from their description. Otherwise, physics would be complicated beyond comprehension". It also explains evolution. It applies to information technology, game theory, legislation, logic of research, algorithmic information, chaos theory, control engineering, medical tests. It can even be used as a legitimacy test: any acceptable information measure must satisfy the Law.

Is the Law true and should it be adopted? First, there are exceptions. The author himself discusses one: the Chinese paper. Assume that the channel $A \rightarrow B \rightarrow C$ consists of A = an Englishman tries to read an article in Chinese, B = an interpreter translates the article into English, C = the English translation of the article. Clearly, inf(C@A) > inf(B@A), hence the Law fails. Second, similar but less subjective violations of the Law can be easily constructed using algorithmic complexity. This signals a problem: which restrictions should be imposed?

The book, which covers more than 500 pages, discusses a wealth of topics grouped in 14 chapters, from specific information measures, statistical information and algorithmic information to control and communication, information physics and quantum information and applications. Some topics are better presented than others. The chapter on algorithmic information, which is close to my expertise, is far from satisfactory, as one can see by browsing the paragraphs 2, 3, and 4 on page 238. The main aim, a grand unification theory of information, is certainly not achieved. Despite this, the book, written by an original thinker, contains a number of interesting ideas which may inspire mathematically oriented readers to continue the project.

REFERENCES

1. H. C. von Baeyer. *Information: The New Language of Science*, Harvard University Press, Cambridge, Mass., 2003 (paperback edition 2004).
2. Y. Bar-Hillel (ed.). Language and Information: Selected Essays on Their Theory and Application, Addison-Wesley, Reading, Mass, 1964.
3. J. Barwise and J. Seligman. Information Flow: The Logic of Distributed Systems, Cambridge University Press, Cambridge, 1997.
4. C. S. Calude. Information and Randomness: An Algorithmic Perspective, 2nd Edition, Springer-Verlag, Berlin, 2002.
5. G. J. Chaitin. Algorithmic Information Theory, Cambridge University Press, Cambridge, 1987.
6. T. M. Cover and J. A. Thomas. *Elements of Information Theory*, Wiley, New York, 1991.
7. K. J. Devlin. Logic and Information, Cambridge University Press, Cambridge, 1991.
8. L. Floridi. What is the philosophy of information? Metaphilosophy 33(1-2) (2002), 123–145.
9. J. Kohlas. Information Algebras: Generic Structures for Interference, Springer-Verlag, London, 2003.
10. M. A. Nielsen and I. L. Chuang. Quantum Computation and Quantum Information, Cambridge University Press, Cambridge, 2000.
11. C. E. Shannon. A mathematical theory of communication, *Bell System Technical Journal* **27**(1948), 379–423, 623–656.
12. C. E. Shannon and W. Weaver. *The Mathematical Theory of Communication*, University of Illinois Press, Urbana, Illinois, 1949 (paperback edition 1963; special fiftieth anniversary edition in 1999).
13. T. Stonier. Information and Meaning: An Evolutionary Perspective, Springer, Heidelberg, 1997.

Department of Computer Science
The University of Auckland
Private Bag 92019
Auckland
New Zealand
e-mail: cristian@cs.auckland.ac.nz

Negative Math: How Mathematical Rules Can Be Positively Bent

by Alberto A. Martínez

PRINCETON UNIVERSITY PRESS, 2006, 267 PP., ISBN-13: 978-0-691-12309-7

REVIEWED BY ERIC GRUNWALD

Extricating the book from its packaging, I was greeted by a picture of a large spoon on the dust wrapper. What could this mean? Some sort of reference to spoon-bending in the last word of the subtitle? Is the reader going to be spoon fed? Surely the book isn't written in spoonerisms. Immediately after the title page, came the following:

> You can use a spoon to drive a screw into a wall. With practice, you can become skillful at it. You can also learn many juggling tricks with the spoon, and thus impress and bewilder people who don't juggle spoons. And you can make all of this more puzzling by calling the spoon a 'fork'. And you can write books about it and form societies with other people who also juggle spoons called forks. And even then, sure, you can use a spoon to drive screws into a wall.
>
> But a screwdriver is better. And even if you've never seen a screwdriver, you can just as well invent one. It might resemble the spoon in some ways though not in others. So you can keep your spoon as well; for eating soup, for juggling, or even, occasionally, for driving screws into walls. At least until you have more skill with a better tool.