BOOK REVIEW

TOWARD AN INFORMATION THEORETICAL IMPLEMENTATION OF CONTEXTUAL CONDITIONS FOR CONSCIOUSNESS

Wallace, R. (2005). Consciousness: A Mathematical Treatment of the Global Neuronal Workspace Model. Springer, Berlin. ISBN 0-387-25242-8, Euro 46.95; hbk).

A major driving force behind the attention that cognitive neuroscience has received in recent decades is the deep mystery of how consciousness is related to brain activity. Many scientists have been fascinated by the wealth of empirical data for individual neurons, neural assemblies, brain areas, and related psychological and behavioral features, and by progressively powerful computational tools to simulate corresponding cortical networks. At the same time, the interested public has been attracted by fancy illustrations of brain activity (e.g., from imaging techniques) and by pretentious claims of neural solutions to basic philosophical problems (e.g., free will versus determinism) in popular magazines and newspapers.

However, heaps of data, extensive simulations, pretty pictures and bold statements cannot replace the insight that is inevitable to relate the available facts to one another in an intelligible manner. I am talking about the old-fashioned stance that understanding is the ultimate goal of scientific effort. In this respect, the need for new conceptual and theoretical ideas in cognitive neuroscience begins to be recognized by prominent representatives of the field. Theory in this sense must not be confused with models fitting data, e.g. by regression algorithms, cluster analyses, etc. It is uncontroversial that experimental and numerical work is and will remain mandatory for scientific progress. But it can only unfold its full value if it is embedded within a profound theoretical framework.

To formulate a serious, clear-cut and transparent formal framework for cognitive neuroscience is a challenge comparable to the early stage of physics four centuries ago. Only very few approaches worth mentioning are visible in contemporary literature. I think that Wallace's book presents an appreciable step in the right direction. However, it is not the ultimate breakthrough yet. Much is left for future work before a full-fledged theory of consciousness will be established.

In two central conceptual respects, Wallace builds on earlier work: Dretske's (1981) usage of informational terms in consciousness research and Baars' (1988) global neuronal workspace model. As compared to Dretske, he restricts his account to syntactic information for which a detailed formalism is available. Semantics (or pragmatics), which is significant in Dretske's work, is left aside. As compared to Baars, his account intends a major extension from "brain only" to somatic, social, and even cultural domains. The resulting "bio-psycho-sociocultural" scope includes areas such as immunology, endocrinology, oncology and

psychosocial stress, or epigenetics and cultural inheritance, each with timely topics of current research.

In order to develop a formal framework for the conceptual ideas, advanced statistical physics (large deviations statistics, renormalization theory) and nonlinear dynamics (symbolic dynamics, complex networks) are utilized together with central ideas of information theory. Some basic terms and formalities are introduced in Chap. 2 (information theory) and in Chap. 4 (renormalization techniques). In a brief excursion in Sec. 5.6 Wallace addresses some, partly speculative, quantum aspects of information and delineates his approach from the popular quantum ideas of Penrose and Hameroff. (To be fair, his reference to Tegmark's (2000) paper against quantum coherence à la Penrose and Hameroff should be complemented by the response of Hagan et al. (2002), who give detailed arguments against Tegmark's results.)

Wallace is aware that the relation between information and physics is nontrivial (pp. 4/5), and that this applies even more so to the relation between information and cognition (Sec. 2.7). He postulates a "duality" of information and cognition (e.g., pp. X, 31) which, however, remains largely unexplained. I may mention that Miranker (2005) has recently proposed a fairly straightforward option to account for such a duality, considering a renormalized hierarchy of increasingly sophisticated kinds of cognition as processes of measurement that provide information.

It is correct that specific descriptions (observables, laws, etc.) of physical (and other) systems can be consistently reformulated in terms of information. But the price to be paid is the loss of much of the formal structure of physical theories and much of the specificity of their relevant observables. The details of a specific description can never be derived from a general information theoretical approach. This point applies also to alternative frameworks using variants of information theory for interdisciplinary purposes. Early examples are the non-equilibrium thermodynamics of Glansdorff and Prigogine (1971) or Haken's (1983) synergetics which Wallace does not refer to. The more recent work by Frieden and colleagues (Frieden 2004) with applications similar to Wallace's, and the formally sophisticated approach of computational mechanics (Shalizi and Crutchfield 2001) would additionally be interesting to compare.

A typical feature of many information theoretical frameworks of thinking is that they refer to a hierarchical organization of systems whose interlevel relations can then be formulated irrespective of the different level-specific descriptive terms. Such a relation between the neurobiology of the brain and the phenomenal content of mental states is a key issue of the hard problem of consciousness. In this respect, Wallace argues that the neurobiological level provides necessary (but not sufficient) conditions for the mental level. This differs from, but need not contradict, the current mainstream view (Chalmers 2000) that neurobiological brain states offer sufficient (but not necessary) conditions for conscious mental states, e.g. in the sense of multiply realized conscious states as in supervenience (Kim 1993).

Wallace attempts to combine Dretske's (1981) emphasis on the significance of

necessary conditions for consciousness with Baars' (1988) emphasis on contexts that constrain conscious activity. Such contexts can be unconscious networks shaping conscious contents, or motives and emotions, so-called "goal contexts". Wallace proposes (i) to formulate such contexts in information theoretical terms, (ii) to implement them as necessary conditions for consciousness using asymptotic limit theorems, and (iii) to define conscious mental states by appropriate coarse grainings of a mental state space. In simpler words, the key idea is to sort out kinds of conscious activity that are most likely under particular contextual constraints.

The spirit of Wallace's framework bears some similarity with a recently introduced, formally elaborated approach denoted as "contextual emergence" (Primas 1988, Atmanspacher and Bishop 2006). Here, stability criteria rather than informational constraints are utilized to implement contexts. This implies an altered (usually coarsened) topology of the considered state space, permitting the definition of contextually emergent features. An application of this procedure to the relation between mental and neural descriptions has been worked out by Atmanspacher and beim Graben (2006).

What does or can concretely follow from Wallace's proposal? I think its primary substance is a sketch of a general formal framework to deal with consciousness as constrained by higher-level and lower-level contexts. A central idea, also expressed in Wallace (2006), is to treat the transition from (lower-level) unconscious activity to (higher-level) consciousness as a non-equilibrium phase transition in a complex network (Chapter 4). Other topics of application, focusing on higher-level constraints of consciousness, are speculatively indicated in Chapter 6. They range from sociocultural context as evolutionary selection pressure to an impressive list of disorders (immune dysfunction, tumor, coronar heart disease, anxiety, schizophrenia) which are related to psychosocial stress. Detailed elaborations in terms of the proposed formalism are not contained in the book.

Beyond references that are necessary for basic elements of his work, the author has put together additional interesting reading material, of course not comprehensive, from numerous areas of research. On the other hand, some significant literature, as mentioned above, is missing. Unfortunately the book does neither have a subject index nor an author index. Also, a concluding summary of results and achievements would have been helpful for readers, including myself, to judge the relevance of the approach more clearly and to clarify which parts of the formalism serve which purpose.

At the end of his preface Wallace lists some familiarity with Baars' model (as well as with current debate) and advanced undergraduate mathematical knowledge as desirable for readers of his book. However, an additional issue is important: skills for and experience with formal thinking. Readers without this capacity will have severe difficulties to make proper sense of his proposals. Therefore, I am afraid that many cognitive neuroscientists will have a hard time to relate the formalism to their everyday business. On the other hand, it might not be easy for mathematicians and physicists who are unfamiliar with consciousness research to

estimate the problems that arise when Wallace's ideas are to be implemented in concrete detail.

Wallace's book is surely not easy reading. It is not a textbook for beginners or a pedagogical introduction to formal models in cognitive neuroscience or, more specifically, consciousness research. It is a book that can unfold its potential to well-educated readers with genuinely interdisciplinary ambitions, who are able to use it as an inspiring source of ideas to work with. Hopefully Wallace's account will find many such readers to develop, refine and apply its author's proposals.

REFERENCES

Atmanspacher, H. and P. beim Graben (2006): Contextual emergence of mental states from neurodynamics. *Chaos and Complexity Letters*, in press. Manuscript preprint available at http://xxx.lanl.gov under q-bio.NC/0512034.

Atmanspacher, H. and R.C. Bishop (2006): Stability conditions in contextual emergence. *Chaos and Complexity Letters*, in press. Manuscript preprint available at http://www.igpp.de/english/tda/pdf/potsdamb5.pdf.

Baars, B. (1988): *A Cognitive Theory of Consciousness*, Cambridge University Press, Cambridge.

Chalmers, D. (2000): What is a neural correlate of consciousness? In *Neural Correlates of Consciousness*, ed. by T. Metzinger, MIT Press, Cambridge, pp. 17–39.

Dretske, F. (1981): *Knowledge and the Flow of Information*, MIT Press, Cambridge.

Frieden, B.R. (2004): *Science from Fisher Information: A Unification*. Cambridge University Press, Cambridge.

Glansdorff, P. and I. Prigogine (1971): *Thermodynamic Theory of Structure, Stability and Fluctuations*, Wiley, New York.

Hagan, S., S.R. Hameroff, and J.A. Tuszynski (2002). Quantum computation in brain microtubules: decoherence and biological feasibility. *Physical Review E* **65**, 061901-1 to -11.

Haken, H. (1983): *Synergetics. An Introduction*, Springer, Berlin.

Kim, J. (1993): *Supervenience and Mind*. Cambridge University Press, Cambridge.

Miranker, W. (2005): The Hebbian synapse: Progenitor of consciousness. *Mind and Matter* **3**(2), 87–102.

Primas, H. (1998): Emergence in exact natural sciences. *Acta Polytechnica Scandinavica* **91**, 83–98.

Shalizi, C.R. and J.P. Crutchfield (2001): Computational mechanics: pattern and prediction, structure and simplicity. *Journal of Statistical Physics* **104**, 817–879.

Tegmark, M. (2000): Importance of quantum decoherence in brain processes. *Physical Review E* **61**, 4194-4206.

Wallace, R. (2006): A modular network treatment of Baars' global workspace consciousness model. Manuscript available at http://cogprints.org/4528/.

Harald Atmanspacher, email: haa@igpp.de
Department of Theory and Data Analysis
Institute for Frontier Areas of Psychology
Wilhelmstr. 3a, D–79098 Freiburg, Germany