EUROPEAN JOURNAL OF COGNITIVE PSYCHOLOGY, 2003, *15* (3), 471–480

Book reviews

Gattis, M. (Ed.) (2001). *Spatial schemas and abstract thought*. Cambridge, MA: MIT Press. Pp. 352. ISBN 0–262–07213–0, £33.50/$50.00 (hardback).

In *Spatial Schemas and Abstract Thought*, Meredith Gattis brings together an expert group of cognitive psychologists, developmental psychologists, computer scientists, linguists, and anthropologists, to answer questions such as ''Which aspects of spatial cognition may be relevant to the use of space in abstract thought?'', ''Do spatial structures affect performance on abstract tasks?'', and ''How can spatial structures be adapted for nonspatial cognition?''. Humans, just as any other living being, depend heavily on their ability to perceive and represent spatial relations in the world they inhabit. Early in our life, we learn to observe the locations and movements of people and objects in our environment, as well as their spatial configurations. From these experiences, we develop *spatial schemas*, representing prototypical locations, movements, and configurations. The main topic of the book edited by Meredith Gattis is whether and how these spatial schemas may be used in abstract, nonspatial tasks. The contributors to the book present evidence that spatial schemas are indeed adapted to a variety of nonspatial tasks, supporting memory, communication, and reasoning. A critical question that recurs in many chapters of the book is related to the ''metaphor or mechanism'' alternative: Are spatial schemas only expressive tools for understanding abstract cognition, or are they actual internal representations and mechanisms? For instance, when we use spatial terms to talk about temporal relations, as in ''We'll meet *at* noon, and I'm looking *forward* to it'', are we using mere metaphors, or are we mapping spatial structures and temporal structures onto each other? And if the latter were true, how does our understanding of temporal relations benefit from this mapping?

Answers to these questions are provided in the 11 chapters of the book. After an introductory chapter by Meredith Gattis, the remaining 10 chapters are divided into three parts, each one addressing a different aspect of the question whether spatial schemas play a significant role in abstract cognition. The three chapters in Part 1 deal with the questions how humans and other animals represent space, and how spatial representations might affect reasoning and memory for nonspatial relations. The three chapters of Part 2 are related to cultural influences, and they give several interesting examples of spatial representations in specific cultural contexts. Finally, the four chapters in Part 3 propose mechanisms which might be used to adapt spatial structures for non-

http://www.tandf.co.uk/journals/pp/09541446.html DOI:10.1080/09541440244000201

spatial purposes, demonstrating how spatial schemas may be more than mere metaphors in nonspatial tasks.

The fact that spatial cognition is of vital importance to every animal is the basis of Chapter 2. In this chapter, William Roberts reviews the encoding of spatial relations in animals such as rats, pigeons, bees, and dogs. He also discusses a number of internal mechanisms available for navigating through space. Roberts also challenges the widespread assumption that animals use cognitive maps. He proposes several alternative explanations and notes some methodological challenges to showing that cognitive maps actually exist in animals. In Chapter 3, Lynn Liben takes on the topic of cognitive maps and compares them to cartographic maps. She proposes three principles of cartographic map understanding which determine the interpretation and use of maps: the purpose principle (maps are produced for specific purposes), the duality principle (maps have an existence in their own right, in addition to their standing for something else), and the spatialisation principle (maps do not represent just anything, they represent something intrinsically spatial). A large part of this chapter is concerned with developmental aspects, and Liben reviews research on children's developing comprehension of maps. She also gives some very illuminating examples of real uses of cartographic maps, and argues for expanding their use in education and psychological research. Part 1 of the book closes with a chapter by Barbara Tversky, in which she reviews many different types of spatial representations, surveying a variety of graphics. She focuses on graphs, charts, and diagrams, which use spatial relations to convey qualitative and quantitative information on varying levels (nominal, ordinal, interval). Her chapter closes with a short review of modern graphics using 3D or animation, a list of functions that graphic displays may serve, and an overview of other approaches to the study of graphics.

In Chapter 5, Sotaro Kita, Eve Danziger, and Christel Stolz report a study of gestures produced spontaneously during story telling. They compared gestures from individuals belonging to two different Mayan language groups. Despite many cultural features shared by the two groups, reliable language-related differences between their typical gestures were found, supporting the notion that culture plays an important role in providing spatial structures for thought. A similar point is made by Karen Emmorey in Chapter 6. She describes how signers of American Sign Language (ASL) use the three-dimensional signing space in front of them to represent not only spatial relations, but also abstract conceptual structures. She also argues that the signing space is particularly useful for representing time and order as well as relational aspects of conceptual structure. A developmental view is taken again in Chapter 7 by Peter Bryant and Sarah Squire. They discuss the relation between space and mathematics by examining how children's use of space can facilitate mathematical problem solving. They argue that children are particularly good at encoding spatial relations, and that these relations should be used as a starting point in teaching mathematics.

The third part of the book is opened by Dedre Gentner. In her Chapter 8, she examines how spatial structures are used to process temporal information. She argues that in general, the process of mapping from spatial structures to abstract cognition may occur at four possible levels, called lexical relations, structural parallelism, cognitive archaeology, and system mapping. She also gives a detailed presentation of experiments, examining which of these mappings explains temporal reasoning with two kinds of space–time metaphors (the time-moving metaphor vs. the ego-moving metaphor). The results of these studies show quite nicely that spatial metaphors do indeed affect the processing of temporal relations. In Chapter 9, Meredith Gattis first suggests four general constraints on mapping any type of abstracts concepts onto spatial representations, namely iconicity, associations, polarity, and structural similarity. In support of her arguments, she presents recent studies with graphs and artificial sign languages. In Chapter 10, Brendan McGonigle and Margaret Chalmers examine the role of spatial schemas in reasoning. They review reasoning and seriation studies, concluding that both spatial coding and temporal relations play an important role in relational encoding. Finally, in Chapter 11, John Hummel and Keith Holyoak present a process model of human transitive inference. Their LISA model uses a spatial array representation of relations between elements. They argue that this spatial array representation successfully models human performance on transitive reasoning problems, thereby supporting the role of spatial schemas in reasoning.

To summarise, I found the chapters contained in this book very interesting and stimulating. Although not all of the details contained in all of the chapters may be easily understood by every reader, the writing is mostly clear and concise. Up-to-date information is presented by leading experts in the field, and the collection of chapters presented in this book makes for a comprehensive treatment of the common topic. Probably, most readers will not find all of the chapters close to their personal interests, but all of them are valuable reading. My list of favourites is based solely on my personal preferences, not on quality of the chapters. Naturally, the chapter on graphs by Barbara Tversky and the chapter on temporal relations by Dedre Gentner are among my favourites because they discuss topics closely related to my own research interests. In addition, the chapter on sign languages by Karen Emmorey made it to the top of my list. I did expect an inherently spatial language such as ASL to use spatial schemas in order to facilitate abstract thinking. However, I was quite surprised to learn just how extensive and differentiated this process is. For instance, who would have thought that ASL uses three different spatial time lines (a deictic time line, a sequence time line, and an anaphoric time line) to represent distinct types of temporal information? There is just one thing that the book leaves to be desired: more intensive integration of the chapters and their topics. As it stands, the book is a collection of fine chapters, but each one is somewhat solitary, and I would have appreciated an integrative discussion of the chapters. I admit that this is a challenging task, given the diversity of topics, but it would have been extremely valuable. It could have been attempted in a final chapter, in which either the editor or the authors jointly try to point out the general conclusions that may be drawn from their chapters. Moreover, given that the book is based on a symposium at which the authors were present to discuss the topics of the book, it is a pity that we do not learn anything about these discussions, including points of agreement and disagreement. De Vega, Intons-Peterson, Johnson-Laird, Denis, and Marschark (1996) have recently published a book containing this type of discussion, and *Spatial Schemas and Abstract Thought* would have profited from it as well.

REFERENCES

De Vega, M., Intons-Peterson, M. J., Johnson-Laird, P. N., Denis, M., & Marschark, M. (1996). *Models of visuospatial cognition*. New York: Oxford University Press.

MIKE RINCK
Dresden University of Technology, Germany

O'Reilly, R. C., & Munakata, Y. (2000). *Computational explorations in cognitive neuroscience: Understanding the mind by simulating the brain*. Cambridge, MA: MIT Press. Pp. 512. ISBN 0–262–65054–1. £36.95/$55.00 (paperback).

Harder, better, faster, stronger: A review of *Computational Explorations in Cognitive Neuroscience*

Just like the sequel to a successful movie, O'Reilly and Munakata's *Computational Explorations in Cognitive Neuroscience* aims to follow up and expand on the original 1986 *Parallel Distributed Processing* volumes edited by James McClelland, David Rumelhart, and the PDP research group (McClelland & Rumelhart, 1986; Rumelhart & McClelland, 1986). This kinship, which is explicitly recognised by the authors as the book is prefaced by Jay McClelland, is perceptible throughout *Computational Explorations*: Not only does this volume visit many of the problems and paradigms that the original books were focused on (so making *Computational Explorations* sometimes feel more like a remake than like a sequel), but there also is an instantly recognisable, and clearly ''psychological'' approach to the role of computational modelling in the cognitive neurosciences, much like in Elman et al.'s (1996) *Rethinking Innateness*. The result is a highly effective, wonderful introduction to the ideas, methods, and problems that characterise this still burgeoning domain.

How has connectionism changed between then and now, one might ask? The answer is straightforward: The central difference between early neural network models and the models described in *Computational Explorations* is a far greater concern for what we know of the biological underpinnings of information processing. In other words, O'Reilly and Munakata spend a great deal of time making sure that the algorithms they introduce are constrained by biologically realistic principles, and the resulting models thus incorporate detailed assumptions about such things as membrane potentials, leak currents, or spiking rates. Most importantly, the relevance of each assumption (i.e., why should psychologists care about such details?) is motivated in the context of current cognitive neuroscience so as to make it clear which computational principles they make it possible to capture.

The 500-page book is extremely well organised, which makes it an ideal foundation as the main text around which to build an introductory course on modern neural network modelling of information processing in the brain. The volume is divided in two main sections, with a third one dedicated to providing technical details about a simulation environment called PDP++.

The first section, titled ''Basic Neural Computational Mechanisms'', comprises 5 chapters, and is essentially dedicated to developing a novel, biologically motivated learning algorithm called Leabra (pronounced Libra), which, rather intricately, stands for ''local, error-driven and associative, biologically realistic algorithm''. Leabra, in contrast to the now almost *passé* backpropagation algorithm, expands on the ideas put forth by McClelland in his GRAIN framework (McClelland, 1993), and takes it as a starting point that networks of real neurons exhibit several properties that are incompatible with the assumptions of vanilla backpropagation, including the facts (1) that they are bi-directionally connected (and hence fully interactive), (2) that they almost systematically involve competition (implemented in the form of ''virtual'' inhibitory interneurons), and (3) that connection weights never change sign during learning, and that it is never the case that connections are used in reverse to backpropagate error information during learning.

Consistently with the ''reconstructivist'' strategy that is a hallmark of O'Reilly and Munakata's approach, these ideas are introduced and motivated progressively over the five chapters of this section, starting with known computational principles of information processing by individual neurons (Chapter 2), moving up to networks of neurons (Chapter 3), and continuing by characterising the computational problems that such networks face when learning (Chapters 4 and 5), to finally arrive at a detailed understanding of how the mathematical underpinnings of Leabra offer the best way to capture the relevant computational principles (Chapter 6). The real strength of this section, and in fact of the book as a whole, lies in how effectively it manages to combine an introduction to the biology of information processing in the brain with an understanding of the computational principles involved, all in the context of developing a detailed, implemented, mathematical characterisation of the problems.

In this respect, the most rewarding aspect of the book is its focus on the broad computational principles that characterize information processing in the brain. For instance, one cannot help but be convinced by the arguments that motivate a central feature of Leabra, namely that the brain is generally faced with two different problems—that of developing accurate, predictive models of the environment with which it interacts (''model learning''), and that of developing the highly idiosyncratic input–ouptut mappings involved in learning how to perform specific tasks (''task learning''). O'Reilly and Munakata suggest that each of these computational objectives is fulfilled by different learning algorithms—Hebbian learning in one case, and error-driven learning in the other. In Leabra, this latter component is implemented in the form of biologically plausible ''Contrastive Hebbian Learning'', which, unlike backpropagation, never involves passing around actual error signals, but assumes instead that each unit computes its own error based on the differences in its activation under different input–output conditions.

The final chapter in this section (Chapter 6) also introduces several other algorithms that are not part of Leabra itself (but that are implemented in PDP++ environment), such as the Temporal Differences algorithm or Reinforcement Learning.

In the second section, titled ''Large-Scale Brain Organization and Cognitive Phenomena'', O'Reilly and Munakata aim to demonstrate the strength of their approach by applying the ideas introduced in the first section to a wide variety of cognitive phenomena in domains such as Perception and Attention, Memory, Language, and Higher-Level Cognition (Chapters 8–11). Chapter 7, which opens this second section, is dedicated to an overview of overall brain organisation, and offers a brilliant introduction to basic anatomical and functional facts about its different systems, as well as a very useful discussion of the general computational problems that cognitive agents are faced with (such as binding, abstraction, or generalisation). This section, with its broad focus, gives O'Reilly and Munakata a chance to showcase both Leabra's ability to offer new insights into old problems and to address novel ones. Here again, there is a compelling focus on computational principles, such as for instance the idea that memory in the brain is organised in separable systems because of the incompatible computational objectives memory has to address, namely storing information about specific instances versus storing information that is shared by many instances. This entire section offers interesting speculations on the role of language in cognition and on the nature of consciousness.

The main body of the book concludes with a thorough and insightful discussion of the pros and cons of computational modelling in general. A final section of the volume is dedicated to an introduction to the PDP++ simulation environment. The software, available free of charge on the internet, is massively complex and under continuous development. This makes it a genuine challenge to learn. While the numerous detailed examples and exercises offered throughout the book are easy to present and understand in the context of a dedicated course, ''Exploring on Your Own'' requires a truly dedicated mind and is generally not for the faint of heart. It takes real understanding, both of the often arcane details of the simulator itself and of the underlying principles, to progress in developing your own simulations. In this sense the PDP++ simulator illustrates a principle dear to early connectionists, namely that the ''rich get richer'': Reaping the rewards of developing a successful original model requires a very substantial initial investment that undergraduate students might simply not have the time for, and that seasoned researchers might think twice about. The appeal of the book as a general introduction to neural network modelling is also somewhat mitigated by its focus on a single set of algorithms. Important alternative approaches, such as generative models, are thus simply not represented in the volume.

Despite our slight concerns about the complexity of the accompanying software, ''Computational Explorations'' is truly a magnificent volume that does wonders at demonstrating how combining neuroscience, cognitive psychology, and mathematical methods in an profoundly integrated manner helps to bring out the overarching computational principles that characterise the problems of information processing. This integrated approach holds the promise of resulting in the development of genuinely explanatory theories of cognition—the holy grail that we are all collectively engaged in a quest for.

REFERENCES

Elman, J. L., Bates, E. A., Johnson, M. H., Karmiloff-Smith, A., Parisi, D., & Plunkett, K. (1996). *Rethinking innateness: A connectionist perspective on development*. Cambridge, MA: MIT Press.

McClelland, J. L. (1993). The GRAIN model: A framework for modeling the dynamics of information processing. In D. E. Meyer & S. Kornblum (Eds.), *Attention and performance XIV: Synergies in experimental psychology, artificial intelligence, and cognitive neuroscience* (pp. 655–688). Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

McClelland, J. L., & Rumelhart, D. E. (1986). *Parallel distributed processing: Explorations in the microstructure of cognition: Vol. 2. Psychological and biological models*. Cambridge, MA: MIT Press.

Rumelhart, D. E., & McClelland, J. L. (1986). *Parallel distributed processing: Explorations in the microstructure of cognition: Vol. 1. Foundations*. Cambridge, MA: MIT Press.

AXEL CLEEREMANS AND ARNAUD DESTREBECQZ
Université Libre de Bruxelles, Belgium

Marcus, G. F. (2001). *The algebraic mind: Integrating connectionism and cognitive science*. Cambridge, MA: MIT Press. Pp. 208. ISBN 0–262–13379–2. £18.50/$27.95 (hardback).

In a critique of a nascent connectionist cognitive science, Fodor and Pylyshyn (1988) issued a dilemma to researchers developing neural network models of cognition. They argued that connectionist models either merely implement symbolic systems, or fail to capture essential properties of human cognition. This dilemma depends upon properties that they deemed requisite to cognition. Gary Marcus adapts the same style of argumentation in his recent book *The Algebraic Mind*, without the associated pessimism for symbolic implementation. Marcus' book is organised according to three important features of our cognitive architecture as he sees it. Any model of human cognition, he suggests, must allow abstract relations between variables and recursive representations, along with an ability to distinguish between kinds and individuals. Much of the book is devoted to demonstrating that certain implementations of popular types of connectionist models, multilayer perceptrons and simple recurrent networks, cannot account for these aspects of cognition.

The first two properties Marcus proposes, abstract relations and recursive representation, can be traced back at least to Fodor's (1975) ''language of thought'' hypothesis. By abstract relations, Marcus means open-ended schemas or rules that hold between whole classes of entities, much like algebraic equations in mathematics or computer programming (p. 35). He asserts that the mind represents these relations and other facts recursively in that new knowledge can be constructed by combining simpler elements into more complex elements. Both of these properties underlie a kind of language of thought or ''mentalese'' described by Fodor long ago. In fact, Fodor and Pylyshyn (1988) maintained that the systematicity inherent in mentalese cannot be captured by connectionist models, at least not without implementing a symbolic cognitive system. Marcus appears to adapt a similar approach in this book.

A third essential property Marcus discusses, the ability to distinguish between kinds and individuals, is a novel approach to criticising connectionist models. Marcus' coverage of the relevant literature, however, leaves much to be desired. It is unclear at this time that such a skill is as widespread in our cognitive system as he suggests. Indeed, research in social cognition, for example, reveals a blurring of kinds and individuals at some levels of processing (Banaji, Hardin, & Rothman, 1993). A wider canvassing of the relevant empirical literature is needed before claiming that this skill is as much of a desideratum for models as Marcus argues.

As case examples, Marcus considers several neural network implementations that fail to satisfy the first two properties of abstract relations and recursive representation. These limitations stem from well-known properties of network dynamics, and are usually overcome in practice by choosing input representations and training regimes suitable to the task at hand. For example, Marcus describes a multilayer perceptron of his own that copies or inverts a binary number (p. 49). His model failed to copy sequences not seen in training. However, our own exploratory simulations revealed that sensible changes to Marcus' design (e.g., total connectivity and some additional copy training) will result in the expected performance. He also demonstrates that some prominent models fail to make the distinction between kinds and individuals, though it should be noted that these models were actually not originally developed for making such distinctions. Ironically, Marcus seems to mirror the failings of his own networks on a theoretical level when he argues from the failure of a few individual models that the whole kind of ''multilayer perceptrons do not offer an adequate basis for cognition'' (p. 7). But if there is one area of human endeavour in which distinguishing between kinds and individuals is crucial, it is in arguments regarding the relative merits of scientific theories. Thus, in modern post-Popperian philosophy of science, disconfirming individual instantiations of a theory does not necessarily falsify the theory as a whole. We therefore urge the reader not to confuse problems with individual network implementations for problems with kinds of networks as a whole.

The book's vision of our cognitive architecture and its critical approach to connectionism are heavily dependent upon a Fodorian foundation. The three properties of cognition that organise the core of the book have a long pedigree in symbolic theory. In a later chapter on evolutionary psychology, he seems to embrace the modularity hypothesis that often marks similar perspectives (pp. 146, 150). The overall enthusiasm for symbol manipulation and modularity surfaces clearly in the final pages of the book: ''To understand human cognition, we need to understand how basic computational components are integrated into more complex devices—such as parsers, language acquisition devices, modules for recognizing objects, and so forth'' (p. 172). However, there may be reason to question that this framework can offer as deep an understanding as Marcus hopes. Fodor, the framework's forefather, has himself recently issued arguments that should temper such optimism. Discussing classical computation, modularity, and adaptationism, Fodor writes: ''The three together constitute not an utterly implausible account of some aspects of cognition. As the reader will no doubt have noticed, it's the part of cognition that *doesn't* work that way that I'm worried about, the indications being that it's quite a big part, and that much of what's special about our kinds of minds lives there'' (Fodor, 2000, p. 80). Fodor argues that there are vast regions of our cognitive architecture about which Marcus' approach would have nothing to say, and that it is ''light years from being satisfactory'' (Fodor, 2000, p. 5). *The Algebraic Mind* therefore seems unbalanced in its criticisms, offering only a unidirectional critique of eliminative connectionism. Perhaps more revealing, the book offers little in terms of what actual symbolic implementations could replace the ones he criticises (for example, even though he offers fairly detailed suggestions for a theory of ''treelets'' [p. 108], they are not accompanied by any implementation). One cannot help wonder why Marcus devotes so much effort to connectionist implementations that do not work rather than offering up algebraic ones that do.

Marcus' book incorporates lines of criticism against neural network models that are instructional, and lead to important discussion and debate. In his own demonstrations of network failings, he offers lessons on what to avoid and overcome when designing connectionist models. The book does suffer from only considering potential problems of eliminative connectionism, without equal weight dedicated to discussing the well-known shortcomings of the symbolic approach; and this despite a subtitle that promises to *integrate* connectionism and cognitive science. A more suitable subtitle, one that would at least lead to appropriate expectations for the reader, might have been ''Assimilating Connectionism into Symbolic Cognitive Science''.

REFERENCES

Banaji, M. R., Hardin, C., & Rothman, A. J. (1993). Implicit stereotyping in person judgment. *Journal of Personality and Social Psychology*, *65*, 272–281.

Fodor, J. (1975). *The language of thought*. New York: Crowell.

Fodor, J. (2000). *The mind doesn't work that way*. Cambridge, MA: MIT Press.

Fodor, J. A., & Pylyshyn, Z. W. (1988). Connectionism and cognitive architecture: A critical analysis. *Cognition*, *78*, 3–71.

RICK DALE AND MORTEN H. CHRISTIANSEN
Cornell University, Ithaca, USA