

Reinforcement Learning

R. S. Sutton and A. G. Barto

Cambridge, MA: MIT Press, 1998.

Hardbound, \$40.00. ISBN 0-262-19398-1

Reviewed by C. R. Gallistel

Reinforcement learning, as understood by Sutton and Barto, is a fusion of the trial-and-error “law-of-effect” tradition in psychology, optimal control theory in engineering, the secondary reinforcement tradition in learning, and the use of decaying stimulus traces in, for example, Hull’s (1952) concept of a goal gradient and, more recently, Wagner’s (1981) model of conditioning. This fusion has given researchers in artificial intelligence a number of ideas for computer algorithms that learn a policy that maximizes the agent’s long term return (amount of reward) from the performance of a task.

Although many of the ideas behind reinforcement learning originated in psychological theorizing, in recent years these ideas have been most extensively developed within the artificial learning community, particularly by the authors of this important summary, their students, and colleagues. The book is intended for use in a one-semester course for students interested in machine learning and artificial intelligence. It would probably not be suitable for a course intended for psychology and neuroscience students, because it does not present models of experimentally established behavioral or neuroscientific phenomena, and the problems given to illustrate how reinforcement learning algorithms may be applied are not necessarily problems that animals (even human animals) are notably good at solving (e.g., efficient scheduling problems). However, reinforcement learning, incentive, and utility remain central concepts in contemporary work on the neurobiological basis of learned, goal-directed behavior (Schultz, Dayan, & Montague, 1997; Shizgal, 1997), and this book is the place to look for the latest ideas on how these concepts may be developed into effective models for the direction of action.

The preface says that the only mathematical background assumed is familiarity with elementary concepts of probability, such as expectations of random variables. Most students will feel that rather more than that is in fact assumed. Nonetheless, the material is presented in an intuitively understandable form, emphasizing the basic ideas and giving helpful illustrations of their application, rather than elaborating proofs. It can be read with profit by motivated neuroscience and psychology students interested in a more rigorous development of these psychologically important ideas. The biographical

and historical sections at the end of each chapter are useful for the perspective they give. The many suggested exercises are challenging, open ended, and thought provoking.

Basic Concepts in Reinforcement Learning

The reward function is the objective feedback from the environment. Rewards are integer or scalar variables associated with some states or state-action pairs. This association (the reward function) defines the goal of the agent in a given situation. For example, a reward of 1 might be associated with the state of having moved all one’s pieces off the board in a backgammon game, while a reward of 0 is associated with all other states of the game (all the states leading up to the winning state). The agent’s sole objective is to maximize net long-term reward (e.g., in a backgammon playing agent, number of games won). The reward function defines what is objectively good and bad for the agent. The reward function is unalterable by the agent. In traditional psychological terms, the reward function specifies the states associated with primary reinforcement.

The concepts of *state* and *action* are very general. An action is any decision an agent might need to learn how to make, and a state is any factor that the agent might take into consideration in making that decision. The state on which an action is predicated may include a model of the environment. This model might represent past states of the actor’s environment. This representation would be considered part of the agent’s state at the time it decides on an action. Thus, reinforcement learning, unlike the more extreme (or pure) forms of neural net learning, is not necessarily asymbolic. It seems to be at least in principle, neutral on the issue of what knowledge is and how it determines action.

The agent’s *policy* (the policy function) is the mapping from possible states to possible actions. The mapping may be a simple look-up table, what is sometimes called a stored policy. A psychologist would call it a set of stimulus-response associations. Alternatively, the agent may rely on a computed policy. A computed policy may involve a search through an ever-changing tree of values,

using a model of the environment to look for action sequences that produce the greatest value.

The *value function* is the agent's current mapping from the set of possible states (or state-action pairs) to its *estimates* of the net long term reward to be expected after visiting a state (or a state-action pair) and continuing to act according to its current policy thereafter. In psychological terms, the value function is the secondary reinforcement values for the many states that are not associated with primary reinforcement. The value function is continually updated by the agent based on the experienced consequences of its actions. These experienced consequences are ultimately the rewards it receives, but the consequences that most immediately matter to the updating process are the values of the states to which the agent gets in the next few actions. In an episodically structured task, the current value of a state is the current estimate of the probability of getting to a reward state before the end of an episode. In a continuous task, the current value of a state is the estimate of the time-discounted amount of reward to be expected starting from that state and continuing to act according to its current policy. With time discounting, the contribution from a reward to the value of a state diminishes with the delay of that reward, so that rewards obtained after long delays contribute negligibly to value.

Time discounting is another venerable idea in psychological theorizing. In reinforcement learning models, this discounting is generally exponential (a constant percentage discount per unit of delay). Empirically, however, animals of all kinds, including humans, appear to use hyperbolic discounting (Fantino, 1969; Green, Fristoe, & Myerson, 1994; Killeen, 1982; Mazur, 1984; McDiarmid & Rilling, 1965). For each reward, they compute a rate of reward by dividing the magnitude of the reward by the delay to obtain it; then, they average these rates to determine the value of an option. It would be nice if reinforcement learning simulations could shed light on why agents formed by natural selection use this peculiar form of discounting.

What reinforcement learning is most centrally concerned with are different methods for computing and updating the value function, the agent's current estimate of the return to be expected once one has arrived at a given state if one continues to pursue the current policy. The general approach is to back up value in accord with the Bellman equation, which sets the value of a state equal to the discounted value of its successor states, weighted by the probabilities of visiting each of those successor states. Value (secondary reinforcement) propagates backwards from the states associated with primary reinforcement to the many intrinsically neutral states or state-action pairs that precede them. These intrinsically neutral states acquire secondary reinforcing value insofar as they lead sooner or later to primary reinforcement.

In general, the policies considered here are rules for choosing actions given the values of the states to which those actions will immediately lead (the values of the possible next states). A *greedy policy* always selects the action that leads to the state with the greatest value. The problem with a completely greedy policy is that it does not enable the agent to learn where little-trying and so-far low value actions might lead if pursued further. That is, it does not allow exploration of the consequences of so-far untried action sequences. If the agent is to acquire an accurate value function (see below), greedy policies need to be modified to allow exploratory actions, at least occasionally. Greedy policies thus modified are called ϵ -greedy.

A central concern in the design of reinforcement learning algorithms is finding the optimal trade-off between exploratory actions, which lead to more accurate knowledge of the value function, and goal-oriented actions, which use the current value function to generate actions that produce reward. Policy improvement and the determination of the value function are interdependent processes, because the secondary reinforcing value of a state depends on the average amount of primary reinforcement obtained after visiting that state, which depends in turn on what the agent has done when visiting that state and subsequent states. Thus, the value of a state depends on the policy being followed, while the actions a given policy produces themselves depend on the values of the next possible states. The path to the best policy is traced out by a dance between a changing value function and a changing policy, each dancer responding to the other.

A *model* in reinforcement learning is a representation of the environment, which is used for planning, that is, for determining what values a contemplated sequence of actions will lead to. Insofar as they make use of models as part of the agent's state space, the reinforcement learning approaches in artificial intelligence depart considerably from the reinforcement learning tradition in psychological theorizing. In psychological theorizing, reinforcement learning has often been offered as an alternative to model building. In practice, however, the reinforcement learning approach in artificial intelligence, like the reinforcement learning tradition in psychology, has little to say about how models of the environment are acquired. ("We do not address the issues of constructing, changing, or learning the state signal" [p. 61].) Moreover, learned models of the world are rarely incorporated into the state space of the agent, because doing so is thought to lead to computational intractabilities. I suspect that these intractabilities arise from deep and not carefully examined difficulties with the kinds of representational schemes that are considered, in particular with the tendency to represent all structure in discrete and probabilistic form.

Reinforcement Learning Methods

The authors analyze three types of reinforcement learning methods:

The *dynamic programming approach* is a collection of algorithms that may be used to compute the optimal policy given a complete model of the environment. The key idea in these algorithms is to use the incremental (action-by-action) updating of value functions to organize and structure the search for good policies, policies that maximize the net long term reward. The process of finding a good policy is accomplished by executing an action in accord with that policy and then backing up the value of the state preceding the action on the basis of the values of all the possible states attainable from the successor state, weighted by the probabilities of attaining them.

There are two problems with dynamic programming as a general solution to the reinforcement learning problem: 1) It assumes a complete model of the environment, which is often not in fact realistically obtainable; 2) It rapidly becomes computationally intractable as the number of state variables increases, and, hence, the size of the state space for which the value function must be iteratively computed. (The difficulty of obtaining exact solutions when states and actions are continuous rather than discrete variables is also a problem, although it is not entirely clear to what extent this is less of a problem for the other methods considered.)

Monte Carlo methods average complete returns from sampled action sequences rather than from all possible action sequences, as in dynamic programming. They do not require a model of the environment, only sample experience. However, they can be applied only to tasks that have an episodic structure, because the computation of the return must come to an end (at the end of the episode) before the algorithm can compute the return for that sample action sequence and back up the values of the intermediate states or intermediate state-action pairs (the states or state-action pairs leading to the episode terminating state). Thus, Monte Carlo methods are not incremental; that is, they improve their policy and their value function only on an episode by episode basis rather than on an action by action basis. They do not profit as rapidly as they might from the knowledge gained from acting.

Third, they consider *temporal difference learning*, which they argue is the idea that is most central and novel in reinforcement learning. Temporal difference learning is like dynamic programming in that it is fully incremental; the value function and policy function are updated after each action. Each action leads to a state whose value is used to back up (change) the value(s) of the state(s) that preceded it. The value of the preceding state moves some fraction of the distance toward the value of the succeeding state. In psychological terms, the secondary reinforcing value of the post-action state is

used as if it were a reward value. The secondary reinforcement value of the pre-action state is updated (“backed up”) by adding to its previous value some fraction of the difference between that previous value and the value of the post-action state. Temporal difference learning is unlike dynamic programming algorithms in that it does not require a complete model of the environment; indeed, it does not require any model, although it can profit from a model if one is available. It is unlike Monte Carlo methods in that it does not require that the task be organized into episodes and that the backing up of the value function be postponed until an episode is completed. By introducing decaying eligibility traces, which determine how far back in the sequence of preceding states the backing up process goes, temporal difference methods can be made to range from pure dynamic programming methods to pure Monte Carlo methods, sometimes substantially improving on both.

Psychologists and neuroscientists will recognize that the temporal difference algorithm is closely related to the principle for associative updating proposed by Rescorla and Wagner (1972). The well known Rescorla-Wagner formula updates the strength of an association by a fraction of the difference between the reinforcement actually received on a trial and the predicted reinforcement, where the predicted reinforcement is determined by summing over all the pretrial associative strengths. The temporal difference formula updates the value of a predecessor state by a fraction of the difference between the value of that state and the value of a successor state. It might be regarded as the Rescorla-Wagner formula applied to the computation of secondary reinforcing value, rather than to the updating of associative strengths.

Caveats and Questions

The authors are careful to point out in several places that “particular states and actions vary greatly from application to application [of a reinforcement learning algorithm], *and how they are represented can strongly affect performance*” (p. 54, italics mine). To which they add that “. . . representational choices [by the modeler] are at present more art than science.” This caveat is related to a question that I believe merits more extensive discussion, namely, for what kinds of problem is reinforcement learning useful?

Reinforcement learning seems to be valuable for dealing with problems in which, even when you know the rules and constraints, it is not at all clear what is the best action to follow. The rules of backgammon are, for example, easily learned, but how to play backgammon effectively is another matter. One of the greatest successes of reinforcement learning has been Tesauro’s (1994) use of reinforcement learning to create a program that rapidly learned to play backgammon at a level equal to the very best human players. The program discovered strategies

that have subsequently been widely adopted by champion human players.

There are other problems for which determining the best policy is trivial given the requisite knowledge *and* an appropriate representation of that knowledge. For these latter problems, one does not need reinforcement learning, either because the agent's problem is trivial or because it has been solved in advance, yielding an innate or preprogrammed policy (an unconditioned response). In those cases, the agent needs only an efficient exploratory strategy, which rapidly builds up the knowledge base on which optimal action depends. Most interestingly, it also needs an efficient representation of the knowledge gained from its explorations, a representation that does not take up too much memory and that minimizes the amount of computation that must be done to determine an optimal action for achieving a specified goal. It probably also needs a set of shrewdly chosen default assumptions that fill in the gaps in its empirically based knowledge with useful surmises.

This question merits discussion because it is not always clear whether a problem belongs to the category of problems for which reinforcement learning is going to be useful or to the category for which it is not going to be relevant. Skinner (1957) thought that language learning could reasonably be treated as a problem in reinforcement learning, a problem of learning what utterances to make in which situations. Chomsky (1959) famously argued that this was nonsense, that the problem of language learning was the problem of learning the grammar (building up the agent's knowledge base, i.e., the agent's model of its linguistic environment). The universal grammar proposals for learning language, now favored by many linguists, assume that the human brain contains an already evolved machine that solves the agent's problem (the problem of generating acceptable utterances) once the agent has learned the grammatical parameters of its language environment.

More prosaically, take the example of matching behavior, in which animals appear to learn to match the relative amounts of time (and/or responses) they allocate to different options to the relative incomes (amounts of reward per unit time) that they receive from those options. This strategy equates the returns realized from the different options being sampled, which is an optimal or nearly optimal policy under many (but by no means all) circumstances. This would appear to be a reinforcement learning problem par excellence. Yet, Heyman (1982) and others (Gibbon, 1995; Mark & Gallistel, 1994) have argued on the basis of experimental findings that matching behavior is elicited (unconditioned) behavior rather than learned (conditioned) behavior. They argue that matching is simply what animals are programmed to do given knowledge of the relative amounts of income to be expected from the different options. Here, too, it may be the case that the agent's role (deciding on the policy to pursue) is predetermined rather than learned; the

behavior we see may depend only on the knowledge acquisition process, the process that determines the agent's state, rather than its policy.

Another question that I would like to have seen addressed is the time-scale invariance or lack thereof in reinforcement learning algorithms, particularly time-difference algorithms. An algorithm is time-scale invariant if the time scale of the events to which it is applied has no effect on the results, that is, if the only relevant aspect of the temporal intervals in the problem space is their relative durations. All time-series analyses in conventional statistics are time scale invariant; the results of the analysis depend only on the numbers, not the units attached to the numbers.

The question of time-scale invariance of reinforcement learning arises because many of the algorithms use eligibility traces and discount decay rates that decay by some fixed fraction for each fixed time step. These decay rates impose a time scale. Does this mean that the algorithms are not time scale invariant? Does how well they work for a given task depend on the time scale of the task? If so, an algorithm that did a good job of discovering a scheduling policy for elevators that moved at one speed would not do a good job of discovering a scheduling policy for elevators that moved more slowly, prolonging all the intervals by a scaling factor. (Scheduling elevators is one of several applications of reinforcement learning discussed at some length.)

Perhaps time-scale invariance can be achieved by adjusting the time steps in these discrete-time algorithms to the scale of the task. But is this something that the modeler—a *deus ex machina*—must do, something that the algorithm could not accomplish by itself? If reinforcement learning algorithms cannot be made to adjust automatically to the time scale of the task they confront, then I would have reservations about them as general purpose solutions to the problem of determining the policy that maximizes reward in continuous time tasks. Similar problems might arise in spatial tasks, if the algorithms could not adjust automatically to the spatial scale.

Acknowledgments

I thank Peter Dayan and Rochel Gelman, for their critical comments and suggestions, from which this review benefited greatly.

Reprint requests should be sent to C. R. Gallistel, Department of Psychology, UCLA, 405 Hilgard Avenue, Los Angeles, CA 90095-1563.

REFERENCES

- Chomsky, N. (1959). Review of *Verbal Behavior* (by B. F. Skinner). *Language*, 35, 26–58.
- Fantino, E. (1969). Conditioned reinforcement, choice, and the psychological distance to reward. In D. P. Hendry (Ed.), *Conditioned reinforcement* (pp. 163–191). Homewood, IL: Dorsey Press.

- Gibbon, J. (1995). Dynamics of time matching: Arousal makes better seem worse. *Psychonomic Bulletin and Review*, 2, 208–215.
- Green, L., Fristoe, N., & Myerson, J. (1994). Temporal discounting and preference reversals in choice between delayed outcomes. *Psychonomic Bulletin and Review*, 3, 383–389.
- Heyman, G. M. (1982). Is time allocation unconditioned behavior? In M. Commons, R. Herrnstein, & H. Rachlin (Eds.), *Quantitative analyses of behavior*, vol. 2: *Matching and maximizing accounts* (pp. 459–490). Cambridge, MA: Ballinger Press.
- Hull, C. L. (1952). *A behavior system*. New Haven, CT: Yale University Press.
- Killeen, P. R. (1982). Incentive theory: II. Models for choice. *Journal of the Experimental Analysis of Behavior*, 38, 217–232.
- Mark, T. A., & Gallistel, C. R. (1994). The kinetics of matching. *Journal of Experimental Psychology: Animal Behavior Processes*, 20, 79–95.
- Mazur, J. E. (1984). Tests of an equivalence rule for fixed and variable delays. *Journal of Experimental Psychology: Animal Behavior Processes*, 10, 426–436.
- McDiarmid, C. G., & Rilling, M. E. (1965). Reinforcement delay and reinforcement rate as determinants of schedule preference. *Psychonomic Science*, 2, 195–196.
- Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II* (pp. 64–99). New York: Appleton-Century-Crofts.
- Schultz, W., Dayan, & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science*, 275, 1593–1599.
- Shizgal, P. (1997). Neural basis of utility estimation. *Current Opinion in Neurobiology*, 7, 198–208.
- Skinner, B. F. (1957). *Verbal behavior*. New York: Appleton-Century-Crofts.
- Tesauro, G. J. (1994). TD-Gammon, a self-teaching backgammon program, achieves master level play. *Neural Computation*, 6, 215–219.
- Wagner, A. R. (1981). SOP: A model of automatic memory processing in animal behavior. In N. E. Spear & R. R. Miller (Eds.), *Information processing in animals: Memory mechanisms* (pp. 5–47). Hillsdale, NJ: Erlbaum.

The Two Sides Of Perception

Richard B. Ivry and Lynn C. Robertson
Cambridge MA: MIT Press, 1998
Hardbound, \$55.00. ISBN 0-262-09034-1

Reviewed by David Poeppel

With *The Two Sides of Perception*, Ivry and Robertson have done the field of cognitive neuroscience a major service. The provocative new theory of perceptual asymmetries in vision and audition that they develop is simple and elegant, unifies a variety of otherwise unrelated observations, and is presented in a manner so explicit that it can be evaluated and refined in a straightforward way. It doesn't hurt that the book is also extremely well

written, and has, for the most part, very appealing graphics.

Research on cerebral lateralization has produced a rich literature full of interesting empirical observations. Unfortunately, the literature very much lacks coherence. There is, for example, little agreement to what extent asymmetry is driven by perceptual or motor phenomena, or both (although recent contributions explicitly take a stand on this issue, e.g., Corballis, 1998). More problematically, much of the work is still couched in terms of task-based specializations (e.g., language is a left-hemisphere function, spatial cognition is a right-hemisphere function, interpretation is left, evaluation right, and so on) and is not sufficiently sensitive to the computational requirements underlying perceptual operations. One of the most serious shortcomings, however, is that many generalizations about cerebral lateralization are based on a single class of evidence, be it neuropsychological, dichotic, or visual-half-field data.

In the context of the large but conceptually disparate literature, Ivry and Robertson's contribution illustrates how cognitive neuroscience can drive progress by providing a more integrative and computationally explicit perspective that is equally sensitive to data deriving from neuroscience and cognitive science.

The book has a brief—but particularly clear—opening chapter laying out the core ideas from neuropsychological and cognitive psychological research that form the basis of our current thinking about cerebral lateralization (e.g., aphasia, spatial neglect, split-brains, dichotic-listening, etc.). In chapter 2, Ivry and Robertson develop their “double filtering by frequency” (DFF) theory of hemispheric asymmetries in perception. Two subsequent chapters discuss the DFF theory in the context of asymmetries in visual perception, two more cover auditory and speech perception. A computational model implementing the DFF theory is shown (a excellent aspect of the book), and a brief chapter of how DFF might account for neuropsychological observations about alcoholism, schizophrenia, and dementia follows. This section (chapter 8) is the weakest part of the book, interrupting the otherwise smooth flow of argumentation and adding little to the discussion of the theory.

Three assumptions form the basis of the DFF theory. First, the frequency content of a visual or auditory input signal is used to build a perceptual representation. Frequency-based perceptual representations are, of course, a natural way to characterize visual and auditory signals, and there is ample evidence that the spectral structure of stimuli is in fact extracted in the visual and auditory systems. Second, frequency information is argued to be asymmetrically represented. In particular, higher frequency spectral information is preferentially processed in the left hemisphere, lower frequency information in the right hemisphere. Crucially, Ivry and Robertson assume that “the difference in how the two hemispheres amplify information is one of relative scale rather than

absolute scale” (p. 36). How is the reference point provided that establishes the relative scale? That is the consequence of their third assumption: a selective attention process provides a mechanism by which the frequency-based perceptual representation is further specified. Depending on the perceptual requirements or task demands, the selective-attention mechanism highlights the part of the spectrum necessary to execute the required operation.

An effective way to conceptualize the DFF is to think of it as a three-stage process with the selective attention and asymmetric amplification processes as filters—hence “double filtering by frequency.” The initial construction of the perceptual representation is bilateral and symmetric, providing the same input signal to both hemispheres and thus allowing for some degree of redundancy in the representations that are ultimately constructed. Next, selective attention picks out a spectral range. In frequency terms, this stage acts as a band-pass filter, preferentially letting through those frequencies necessary to execute the task. This stage is “intended to capture the fact that selective attention, whether it operates on restricted spatial regions or on objects [presumably visual or auditory] determines which region of the sensory spectrum will be elaborated for further processing” (p. 60). In the second filtering stage, the range selected by the selective attention mechanism is asymmetrically amplified in the two hemispheres. This stage can be thought of as low-pass and high-pass filtering. The left hemisphere acts as a high-pass filter, the right as a low-pass filter. It is subsequent to this stage that non-identical representations emerge. Prior to the second filtering operation, the two hemispheres have access to the same representations; after it, the left and right hemispheres are dealing with representations that, while partially overlapping, emphasize different aspects of the frequency spectrum.

The intuitive appeal of the theory is that it provides a simple way for information to be simultaneously represented at multiple scales. When presented with a complex (visual or auditory) spectral input, the representations constructed by the two hemispheres emphasize different aspects of the input, permitting the system to simultaneously extract whatever the lower- or higher-frequency content best represents (e.g., local detail versus global shape, or formant versus fundamental frequency). In this framework, task-related perceptual asymmetries are not caused by the fact that certain tasks are exclusively performed by one or the other hemisphere (e.g., language/left versus spatial/right), but are rather the consequence of the fact that executing a particular task requires access to the part of the spectrum that contains the relevant information.

Ivry and Robertson go on to apply the theory to a range of phenomena in vision and audition. The model is presented in action by examining the responses to simple and complex visual stimuli, and a lot of data from

neuropsychological patients are reinterpreted. Although there are not sufficient data available from the DFF perspective to make a clear case, the authors discuss some very nice predictions for the case of face perception and recognition—again rooting potential asymmetries in the specific requirement imposed by the perceptual task. The alternative model for hemispheric asymmetries that is considered most seriously (in chapters 3 and 7) is the categorical-coordinate distinction pushed by Kosslyn and colleagues. That proposal, like the DFF theory, accounts for asymmetries by appealing to the computational requirements underlying the execution of tasks rather than in assumptions about which hemisphere which psychological function lives in. For adherents (or opponents) of the categorical-coordinate distinction, this book is, unsurprisingly, a must.

Also unsurprisingly, there are some things that are not so satisfying. Three of the shortcomings are briefly raised. First, the theory places a major burden on the selective attention system, and the discussion of how, precisely, the attentional filtering process works is not presented in sufficient detail. In the case of the visual system, there is a large literature that Ivry and Robertson can draw on, but the discussion of selective attention in audition is quite thin. Because of the crucial role this process plays, the authors at least owe us some speculation about when, where, and how the system selects the critical information.

Second, there is room for considerably more detail on the possible neurophysiologic mechanisms that underlie the implementation of the DFF. The lack of attention to neurophysiologic detail is particularly frustrating with respect to the asymmetric filtering stage of the model. One wants to know, for example, what type of receptive field properties would yield the asymmetric representations? One could imagine models that incorporate a varying temporal integration time for neurons in the two hemispheres (e.g., left hemisphere neurons in the relevant cortical areas integrate over, say, 25 msec, right hemisphere neurons over, say, 250 msec) or models that implement a delay-line architecture that provides varying temporal integration constants. Whatever mechanism the authors envision, more words about this are necessary.

Finally, the discussion of audition and speech is provocative, but leaves one with too many questions. In these two chapters, Ivry and Robertson discuss DFF in the context of pitch perception, music, prosody, and phoneme perception. But how things fit together is not as clear. Probably many of the questions raised by these chapters will be answered when the issue of how selective attention operates in this domain are specified. What is confusing is that sometimes it appears as if the attentional filter selects a very small spectral range to execute a task (e.g., in the Ivry-Lebby tasks), other times a much broader range appears necessary (e.g., in the missing fundamental tasks, if the missing fundamental corre-

sponds to a reasonably high frequency). Some of the explicit predictions are also not supported. For example, DFF predicts that the harmonic structure of music is preferentially processed in the left hemisphere because harmonic structure is in the upper part of the spectrum (e.g., p. 52), whereas there is evidence that suggests the opposite result (Tramo & Bharucha, 1991). Presumably these topics and a few others that deserve more attention (including neuroimaging data that is relevant to testing some of the ideas) will be addressed with more enthusiasm in subsequent editions.

Importantly, the theory developed by Ivry and Robertson does not try to be a theory of everything (about lateralization). The authors concentrate on hemispheric asymmetries in visual and auditory perception, although the theory may of course generalize to other perceptual or motor phenomena. As is probably true for any broad new proposal, it is likely to be wrong about many details and maybe even about some of the major assumptions. Nevertheless, it feels like it is on the right track.

Reprint requests should be sent to David Poeppel, Department of Linguistics and Department of Biology, Neuroscience and Cognitive Science Program, University of Maryland, College Park, 1405 Marie Mt. Hall, College Park, MD 20742-7505.

REFERENCES

- Corballis, M. (1988). Cerebral asymmetry: Motoring on. *Trends in Cognitive Sciences*, 2, 152-157.
- Tramo, M., & Bharucha, J. J. (1991). Musical priming by the right hemisphere post-callosotomy. *Neuropsychologia*, 29, 313-325.

Toward a Science of Consciousness II. The Second Tucson Discussions and Debates

Edited by S. R. Hameroff, A. W. Kaszniak, and A. C. Scott
Cambridge, MA: MIT Press, 1998.

Hardbound, 764 pp., \$70.00. ISBN 0-262-08262-4

Reviewed by Diana S. Woodruff-Pak

The Tao of Consciousness

For someone looking for a volume addressing scientific perspectives of consciousness, *Toward a Science of Consciousness II* provides some scientific approaches and a lot more. Included in this collection of 64 chapters are articles invited from among the 42 plenary talks and close to 500 concurrent oral presentations and posters presented at the second "Toward a Science of Consciousness 1996" held in Tucson in April 1996 (and affectionately called, "Tucson II"). The organizers' classification of the major themes included in the conference and the book yielded the following categories: Neuroscience,

Cognitive Science and Psychology, Physics and Mathematics, Biology, Philosophy, Experiential Approaches, Anomalies of Consciousness, Humanities, and Culture and Society. The number of presentations in the first four categories that represent a traditional scientific approach to consciousness was 247, and the number of presentations in the last five categories that do not follow a scientific tradition was 260. The book reflects this bias, including more perspectives that do not fall in the realm of scientific than do. Thus, this volume might be considered as the "Tao" of consciousness as it imparts *the Way of Consciousness* in humanistic, philosophical, and experiential terms with somewhat less content exploring the topic scientifically.

The editors of the volume and organizers of the conference have worked most effectively to blend and provide cohesion to this array of material. One valuable feature of the book is an Introduction that chronicles the planning that resulted in the conference's organizational structure along with a detailed narrative of the daily events that transpired. The "spirit" of the conference is communicated in this chronology written by Keith Sutherland, Jean Burns, and Anthony Freeman and reprinted from the *Journal of Consciousness Studies*. Each of the fifteen sections of the volume are introduced with an Overview, an additional feature that aids in integration.

If your aim is to read the most thoughtful of contemporary philosophical views of consciousness, this volume is most valuable. John Searle, David Chalmers, Daniel Dennett, and Patricia Churchland are among the luminaries who enlighten and debate issues in this domain. The degree to which they disagree highlights the preliminary state of understanding about consciousness. Dennett, the reductionist, maintains that consciousness arises solely from neurons, in contrast to Chalmers, who asserts that consciousness may be an irreducible, fundamental property of the universe. In different ways, Searle and Churchland are optimists. Searle addresses nine standard objections to the scientific study of consciousness, and more or less concludes that all of these objections can be overcome. Churchland considers that the problem of consciousness in neuroscience may be easier to solve than the problem of timing, and she wholeheartedly endorses further research in experimental psychology and neuroscience.

Anesthesiology offers a tantalizing perspective on consciousness. The chapter by Nicholas Franks and William Lieb addresses the molecular basis of general anesthesia and points out the neural proteins that are critically affected by anesthetics. Identification of these sites might be important in finding essential loci for consciousness. Postsynaptic receptors responsive to GABA, glycine, serotonin, and nicotinic acetylcholine receptors are the most sensitive to anesthetics. It is not yet understood how anesthetics activate inhibitory receptors such as GABA and glycine, yet inhibit the excitatory activation of serotonin and nicotinic cholinergic receptors. Hans

Flohr links intravenous dissociative drugs such as ketamine and the street hallucinogen phencyclidine (PCP) to NMDA receptors and discusses the unique properties of these receptors. Flohr asserts that NMDA receptors play a pivotal role in higher-order representational states. Thus, in Flohr's view, consciousness is a consequence of the brain's representational activity. Brain imaging and anesthesia are the tools employed by Michael Alkire, Richard Haier, and James Fallon to search for the site(s) of consciousness. Imaging participants during an awake state and under propofol anesthesia, Alkire and associates observed a global effect and specific regional effects. Additional testing suggested that it was when global neuronal functioning is reduced below some critical threshold that consciousness is lost. From this perspective, consciousness is associated with brain metabolism and global neuronal activation above some critical threshold level.

Among the most engaging chapters in the book are those in the section entitled, "Neural Correlates of Consciousness" whose authors defend hypotheses about specific brain sites or mechanisms for consciousness. In the Overview of this section, the editors point out that David Chalmers argues that we need a device that detects consciousness before we can determine its neural substrates. We are told that during his talk at Tucson II, Chalmers demonstrated a prototype consciousness meter that was said to have been developed by combining "neuromorphic engineering, transpersonal psychology, and quantum gravity, although it strikingly resembled a hair dryer" (p. 216).

Brain connectivity as an emergent property of nonspecialized, divergent neuronal networks is the substrate for consciousness postulated by Susan Greenfield in her chapter addressing, "A Rosetta Stone for Mind and Brain?" Greenfield reasons by analogy, using comparisons to phenomena such as infant development, animal consciousness, and electrical lights on Oxford Street to make her points. Joseph Bogen who localized creativity to the corpus callosum, suggests that the state he calls, "subjectivity" is engendered by neuronal activity in and immediately around the intralaminar nuclei of each thalamus. His position is based primarily on evidence with patients who have experienced small bilateral lesions of this region that cause immediate unresponsiveness.

Conscious awareness of emotion is localized to anterior cingulate cortex in the chapter by Richard Land and associates who also note that this region is activated in cognitive states engaging attention, perception, and mental effort. The main source of evidence presented in the Land chapter is brain imaging. Bernard Baars, James Newman, and John Taylor present neuronal mechanisms of consciousness in a chapter written in outline form and conclude that a brain mechanism for consciousness is like a global-workspace system. This system is best understood using the metaphor of the theater-of-consciousness in which consciousness corresponds to the bright

spotlight on the stage of a theater. This model relies on the construct of working memory and the interaction between input and memory.

For the most part, Jeffrey Gray is pessimistic about the possibility of locating brain substrates for consciousness, but he does present some possible approaches. He suggests that conscious content is similar to the output from a comparator that may be localized in the subiculum of the hippocampus. An approach to determine if consciousness is neurally generated or learned from experience would be to test individuals with word-color synesthesia (who experience visual sensations such as color to auditory input, or vice versa) with brain imaging techniques. "If synesthetes have color experiences when they hear words because unusual patterns of neuronal connectivity lead to activity in brain circuits, with no need for a life history of associative learning, this would imply that the features characterizing specific conscious experiences depend upon neuronal events, not upon information processing" (p. 289). The implication from this result would be that attempts to understand consciousness using nonneuronal systems, such as computers, would not be efficacious.

A review of this extensive coverage of the topic of consciousness cannot even touch upon all the perspectives that are presented. Language and consciousness are addressed primarily in chapters dealing with communication in animals. Paul Bloom does raise the interesting question of whether consciousness and mental life arose in the human species because we developed language rather than language being an emergent consequence of consciousness. His conclusion is that at least some unique aspects of human mental life exist independently of language and that the nature and origin of these aspects must be explained in another way.

Eight chapters are devoted to vision and consciousness, which seems quite excessive in the absence of sections devoted to other sensory modalities and cognitive processes like attention, working memory, declarative memory, and executive function. Blindsight, as discussed in chapters by Larry Weiskrantz and by Stanley Klein, is among the most directly relevant of the topics on visual perception to consciousness. With blindsight, a phenomenon observed in patients with circumscribed lesions of the visual system, it can be demonstrated that accurate performance on a visual task can be separated from conscious awareness. In addition to blindsight, there may be more effective strategies for assessing visual perception and consciousness than are presented in the section on Vision and Consciousness. For example, an approach recently published by Lumer, Friston, and Rees (1998) using experimental manipulations to produce binocular rivalry during fMRI recording revealed a role for frontoparietal areas in conscious perception.

Given the vast coverage of topics in this book ranging from quantum physics to extrasensory perception, it would seem likely that readers approaching conscious-

ness from a number of different perspectives would find a satisfying amount of material covering their interests. However, in the case of cognitive neuroscientists, this book may be disappointing. From a cognitive neuroscientific perspective, brain imaging techniques provide tools useful in the scientific investigation of consciousness. Brain imaging (PET, MRI, or fMRI) is mentioned briefly in four different chapters, but for the most part the potential of these techniques is ignored. Apparently Robert Tootell described functional MRI maps of activities related to consciousness in a presentation at the conference, but he did not contribute a chapter to the volume. We are left with the philosopher Michael Lockwood pointing out the relative futility of searching for insights about consciousness with brain imaging.

Indeed, it may not be utterly fanciful to suppose that one day, a supercomputer hooked up to a brain scanner may, with unerring accuracy and in minute detail, describe what is going on in the conscious mind of the person whose brain is being scanned. Perhaps such methods will also yield new insights into the minds of nonhuman animals.

But we must ask how far such work can be seen as a steady progress in the direction of understanding what consciousness really *is*, in physical terms, or of *integrating* it into our physical worldview. (pp. 84–85)

Another cognitive neuroscientist who presented a topic central to the issue of consciousness is Daniel Schacter, who discussed implicit brain memory systems. Alas, Schacter's chapter is not included in this volume, either. The topic of learning and memory in the absence of awareness is one that has been addressed by extensive research in the past decade, and Daniel Schacter has been a leading contributor in this area. His chapter would also likely have addressed consciousness using brain imaging techniques. In this sense, omitting a chapter by Schacter eliminates two dimensions of interest to

cognitive neuroscientists. The topic of unconscious learning and memory provides important insights about consciousness, yet the terms, "implicit" and "nondeclarative" do not even appear in the extensive index of this book.

Tucson II collected a diversity of scholars to discuss a topic that has been a pariah to science for most of the twentieth century. Cognitive neuroscience has been one of the disciplines to embrace the study of consciousness and to contribute significantly to progress in describing it with precision both in its presence and absence; Tucson III would benefit from more input from cognitive neuroscientists. An entire section devoted to brain imaging approaches to consciousness and another added section addressing implicit forms of learning and memory are warranted. More input including case study material from neuropsychology and abnormal psychology would provide better empirical evidence on consciousness, altered states of consciousness, and the absence of consciousness. Totally ignored (including no entry in the index) are motor systems and motor control. Motor learning is a phenomenon that makes the transition from conscious to unconscious. The atomization of behavior, and the role that brain structures such as the cerebellum play in this process is a central part of the conscious-unconscious continuum. The fact that one might propose additional topics and have the desire to read yet another huge volume on multidisciplinary perspectives of consciousness attests to the vitality and value of the present volume.

Reprint requests should be sent to Diana S. Woodruff-Pak, Department of Psychology, Temple University, 1701 N. 13th St., Philadelphia, PA 19122.

REFERENCES

- Lumer, E. D., Friston, K. J., & Rees, G. (1998). Neural correlates of perceptual rivalry in the human brain. *Science*, *280*, 1930–1934.