Introduction

This book contains the presentations given during the Duke Symposium on Computational Models of Conditioning, which took place between May 15th and May 17th of 2009 at the Duke Campus in Durham, N.C. The meeting was sponsored by the Duke Department of Psychology and Neuroscience, the Duke Office of the Vice Provost for International Affairs, and the Duke Arts and Sciences Research Council. All the participants and I are indebted for their generous support.

The meeting was organized with the assistance of my friend and former Ph.D. advisor Professor John Moore (University of Massachusetts at Amherst). I am particularly thankful to John for helping me in finding a group of participants who contributed both well-established and novel theories of classical conditioning. I am also grateful to Munir Gunes Kutlu for his help in running many aspects of the meeting.

The models

John Kruschke and Rick Hullinger (Indiana University, USA) prepared the chapter on "The evolution of learned attention." In this chapter, the authors use simulated evolution, with adaptive fitness measured as overall accuracy during a lifetime of learning, and show that evolution converges to architectures that incorporate attentional learning. They also describe the specific training environments that encourage this evolutionary trajectory, and how we assess attentional learning in the evolved learners. Interestingly, the resulting attentional mechanism is similar to that proposed by Mackintosh (1975).

A question regarding the evolution of complex systems is whether a simple, basic associative system first evolves, and then attentional (and also configural) mechanisms would appear in later stages of the simulated evolution. The plausibility of such a process is partly supported by the observation that most vertebrates share a basic brain structure that is supplemented by more complex processing brain areas (like the neocortex) in certain species (Romer, 1970).

Justin A. Harris (University of Sydney, Australia) contributed the chapter on "The arguments of associations." In this chapter, Harris demonstrates that, although it is commonly assumed that the solution to patterning and biconditional discriminations requires "configural" elements, these discriminations can be solved by assuming that some elements of each stimulus are inhibited when two stimuli are presented in compound.

Harris's approach is similar, in some respects, to that introduced by Grossberg (1975) and further developed by Schmajuk and Di Carlo (1989), although neither of these authors applied this "normalization" mechanism to discrimination problems. It also interesting that when a configural model is used (e.g., Schmajuk & Di Carlo, 1992), the configurations produced by the model in its "hidden layer" look very much like those suggested by Harris. In addition, this normalization mechanism might have an important role as a "front end" for a competitive network (e.g., Rescorla & Wagner, 1972) when a large number of stimuli are used which might overwhelm the network, thereby destabilizing the algorithm. Despite all its positive properties, the fact that Harris's (2006) model cannot describe some occasion-setting paradigms (e.g., like those in which the feature is weaker than the target, such as positive-feature, simultaneous discriminations with a weak feature, motivational learning or contextual learning) suggests that configural mechanisms cannot be completely discarded.

Michael Le Pelley (Cardiff University, United Kingdom) wrote the chapter on "The hybrid modeling approach to conditioning." In this chapter, Le Pelley describes his "hybrid" model of associative learning which incorporates two associabilities; α defined as in Mackintosh's (1975) theory, and σ defined as in the Pearce–Hall model, with the overall learning rate for a stimulus being determined by α x σ. The chapter discusses evidence from recent studies of animal conditioning and discrimination learning that provides support for the hybrid modeling approach. Common to both Le Pelley's (2004) hybrid model and the Schmajuk, Lam and Gray (1996) SLG model, is that blocking is the result of two mechanisms simultaneously at work (Schmajuk & Larrauri, 2006).

Ralph Miller and James Witnauer (Binghamton University, USA) provided the chapter on "The role of within-compound associations in cue interactions: models and data." In this chapter, the authors use computational modeling to review and simulate experiments related to the role of within-compound associations in negative mediation, positive mediation, and counteraction.

A mathematical model that attributes all cue interactions to within-compound associations was shown to provide a better fit to some negative mediation phenomena than a model that attributes negative mediation effects to variations in outcome processing. Overall, the results of this analysis suggest that within-compound associations are important for cue competition, conditioned inhibition, counteraction effects, retrospective revaluation, and second-order conditioning.

Some similarities between the role of CX–CS and CS–CX associations in Miller's comparator model (Miller & Schachtman, 1985), Wagner's (1981) Sometimes Opponent Process (SOP), and Schmajuk *et al.*'s (1996) SLG model should be noted. For instance, in all three models these associations decrease responding. In the comparator model, the CS–CX association activates the CX–US association, which subtracts from the CS–US association, thereby decreasing the output of the comparator. In the SOP and SLG models the CR is attenuated by decreasing the activation of the CS–US association when the CS is predicted by the CX.

Allan R. Wagner (Yale University, USA) and Edgar Vogel (Universidad de Talca, Chile) contributed the chapter on "Associative modulation of US processing: implications for understanding of habituation." In this chapter, the authors analyze, in the light of data from their and other laboratories, Wagner's (1976, 1979) suggestion that long-term habituation might be the result of an associative process by which stimuli come to be "expected" in the context in which they have been exposed. The authors indicate that one complication is that extended contexts (as well as discrete cues) can control response-potentiating, conditioned-emotional tendencies, in addition to the presumed decremental effects. They describe experiments that separate these effects and illustrate how the approach offered by SOP and AESOP (Wagner, 1981; Wagner & Brandon, 1989) describes those results.

The Wagner and Vogel chapter has some commonalities with Stout and Miller's (2007) modeling of the extended comparator hypothesis. In their chapter, Miller and Witnauer indicate that the comparator sometimes subtracts and at other times adds the CS–US associations, in order to explain why sensory preconditioning training results in responding to the non-reinforced CS with a few CS1–CS2 and CS1–US alternated trials (a potentiating effect), but results in the CS being inhibitory with an increasing number of trials (a decremental effect).

Together with Munir Gunes Kutlu, Joey Dunsmoor, and Jose Larrauri (Duke University, USA), we wrote the chapter on "Attention, associations, and configurations in conditioning." This chapter describes a number of computational mechanisms (associations, attention, configuration, and timing) that first seemed necessary to explain a small number of conditioning results, and then proved able to account for a large part of the extensive body of conditioning data. The chapter first presents Schmajuk *et al.*'s (1996) SLG model, a neural-network theory that includes attentional and associative mechanisms, and applies it to the description of compound conditioning with different initial associative values, extinction of the conditioned excitor decreasing the retardation of conditioning of its conditioned inhibitor (conditioned inhibition as a slave process), facilitation of conditioning by context preexposure, recovery and absence of recovery from blocking, latent inhibition–overshadowing synergism and antagonism, summation tests in the context of extinction, and spontaneous recovery. Although other models can increase processing when the predictor of a CS disappears and the CS is presented by itself (change of CX in latent inhibition), only the SLG model can increase attention to a CS upon presentation of another, novel stimulus (e.g., presenting a novel stimulus preceding an extinguished CS to produce disinhibition). Also, the SLG model is unique in yielding mediated attentional changes, whenever a CS_A predicted by a CS_B is absent when the CS_B is presented by itself, thereby increasing Novelty and attention to the predicted, absent CS_A.

The chapter also introduces the Schmajuk, Lamoureux and Holland (1998) SLH model, a neural network that incorporates configural mechanisms, and applies it to the description of response form in occasion setting. Finally, it is shown how the combination of configural and timing mechanisms describes timing of occasion setting, and how the combination of attentional, associative, and configural mechanisms describes causal learning. Because these computational mechanisms were implemented by artificial neural networks, which can be mapped onto different brain structures, the approach permits the establishment of brain–behavior relationships, like the Burgos and Mauk models described in the next two chapters.

Michael Mauk (University of Texas, USA) contributed the chapter on "Computer simulation analysis of cerebellar mechanisms of eyelid conditioning." This chapter first outlines work that identifies the essential principles of cerebellar learning and lays down the foundation for simulating cerebellar function. Then it describes a number of new computational findings that offer a relatively accurate description of the essential computational unit of the cerebellum and how it works in a noisy input background.

José E. Burgos (Universidad de Guadalajara, Mejico) wrote the chapter "The operant/respondent distinction: a computational neural-network analysis." This chapter outlines a neural-network interpretation of the distinctions between types of stimulus–response relations (operant versus classical), reinforcement contingencies (response-dependent versus stimulus-dependent), and their effects. The operant–classical distinction is interpreted in terms of the difference between two types of input–output relations that involve different types

Cambridge University Press
978-0-521-11364-9 — Computational Models of Conditioning
Nestor Schmajuk
Excerpt
More Information

of input units, as well as different types of output units. The distinction between different types of contingencies is interpreted in terms of the difference between two types of training protocols (S-dependent versus R-dependent). The distinction between the effects of the contingencies is interpreted in terms of the difference between changes in the different types of output units in the presence of S activations. Finally, the chapter suggests possible roles of hippocampal and dopaminergic systems in conditioning.

Using the same nomenclature used in the rest of the book, the Burgos model can be re-described as follows. In classical conditioning, presentation of the unconditioned stimulus (US, S*) activates the unconditioned response UR (R*) which strengthens its connection with the conditioned stimulus (CS, S). Subsequent presentations of the CS (S) will activate the CR (R*). In operant conditioning, generation of an arbitrary response R will result in the delivery of the US (S*), which will strengthen the connection with the S present at the time. Subsequent presentations of the CS (S) will activate the operant R. In this way, the network clarifies the differences and similarities between classical and operant mechanisms and how they interact.

Beyond parsimony: redundancy and reliability

Most chapters in this book seem to reinforce the notion that more than one mechanism is needed to account for the reported results on classical conditioning. Conditioning can be described by a rather complex combination of different mechanisms. At the Duke Meeting, Mike LePelley's talk title pointed out that "parsimony is overrated."

Overall, they suggest that simple notions like Mackintosh's (1975) attentional theory, Rescorla–Wagner's (1972) delta rule, Pearce and Hall's (1980) attentional model, Miller and Schachtman's (1985) first comparator hypothesis, or Grossberg's (1975) attentional competition model failed to explain important aspects of conditioning. Instead, the field has moved to increasingly complex models that have incorporated those ideas. Even these models have limitations and might require the incorporation of additional mechanisms to provide a complete account of associative learning. It is apparent that the complexity of these models puts them beyond the ability of our intuitive thinking and makes computer simulations irreplaceable.

The content of this book suggests that the different mechanisms required for a full description of the data can be analyzed separately and then combined into an integrated model. As indicated before (Schmajuk, 2010), the method is reminiscent of the Wright brothers' approach to airplane design, which consisted of the independent development and testing of the individual components of the plane before assembling them together into a flying machine

(Padfield & Lawrence, 2003). Interestingly, that approach was based on a study of bird flight by George Cayley (1773–1857), who realized that the lift function and the thrust function of bird wings were separate and distinct, and could be imitated by separate systems on a fixed-wing craft. Imitation of each separate, relatively simple function permits the explanation of how each function is achieved.

If more than one mechanism is required for a full description of classical conditioning, therefore making parsimony unrealistic, it is possible that some of the properties provided by these mechanisms might overlap. As mentioned above, both Le Pelley's (2004) hybrid model and the SLG model describe blocking as the result of two mechanisms. That is, multiple, complex mechanisms might provide redundancy and increased reliability, which is a much-desired property of both technologically designed products and biologically evolved systems. In sum, computational models might be becoming less parsimonious, but the added redundancy increases reliability. A clear demonstration of such increased reliability is that some functions (like blocking) can survive the effect of hippocampal lesions (Holland & Fox, 2003).

Evaluation of the models

In order to quantify the quality of a model's simulated results, some chapters (e.g., Miller and Witnauer's) have used correlations. Alternative methods have been used in the past, for instance, Schmajuk *et al.* (2001) used χ^2, and Schmajuk and Larrauri (2006) applied analysis of variance using the actual variance of the experimental subjects. Of these alternative methods we prefer to use correlations because, although they disregard the importance of the variance in the data, they indicate when to reject the null hypothesis (that simulated values and experimental data are not correlated). Instead, χ^2 indicates when to accept the null hypothesis (that simulated values and experimental data are equivalent). Finally, the analysis of variance used in Schmajuk and Larrauri (2006) requires knowledge of the values of the variance of the data, which is not always reported in the experimental studies.

Evaluation of the data

In addition to the question of the evaluation of the models, the issue of the robustness of the data was an important concern during our discussions. For example, the Schmajuk *et al.* chapter refers to the contradictory results regarding the combined effect of preexposure, which usually yields latent inhibition, and compound training, which usually results in overshadowing. As indicated in that chapter, it has been reported that preexposure and compound training can add and cancel each other. Because the SLG model (Schmajuk *et al.*, 1996) explains away these contradictions in terms of differences in experimental parameters used in the different reports, parametric studies are needed to test those explanations. Most important, in addition to testing the model, these parametric studies would serve to determine the range in which some reported results are valid and can be replicated.

Future challenges

A general problem of the models presented in this book is that most of them, with the exception of Mauk's, neither take into account any specific preparation (e.g., rabbit's eyeblink conditioning, rat's conditional emotional response, taste aversion, human ratings), nor the different experimental values (e.g., duration of the CS, salience of the CS, the duration and strength of the US, context salience, intertrial interval, trials to criterion) used in the experiments run with those preparations. Therefore, most models are "generic" models of classical conditioning. We expect that future models will (1) adopt parameters appropriate for specific preparations, and (2) use simulation values (e.g., stimulus duration and salience, trials to criterion) that are scaled to those used in the corresponding experiments. The resulting models will provide more accurate descriptions of the data.

Society for Computational Modeling of Associative Learning

One important achievement of the Duke meeting was the creation by all the participants of the Society for Computational Modeling of Associative Learning. The purpose of the Society is to (1) foster communication about computational models of associative learning among those who do computational modeling, between those who create models and those who might be instructed by them, and between those who do experiments on associative learning and those whose models might be instructed; and (2) to promote the use of computational models for addressing conceptual issues in associative learning. After the meeting, a number of researchers were invited to become members and at the time of this writing the society has more than 30 members.

References

Grossberg, S. (1975). A neural model of attention, reinforcement, and discrimination learning. *International Review of Neurobiology*, **18**, 263–327.

Harris, J. A. (2006). Elemental representations of stimuli in associative learning. *Psychological Review*, **113**, 584–605.

Holland, P. C. & Fox, G. D. (2003). Effects of hippocampal lesions in overshadowing and blocking procedures. *Behavioral Neuroscience*, **117**(3), 650–656.

Le Pelley, M. E. (2004). The role of associative history in models of associative learning: a selective review and a hybrid model. *The Quarterly Journal of Experimental Psychology*, **57B**, 193–243.

Mackintosh, N. J. (1975). A theory of attention: variations in the associability of stimuli with reinforcement. *Psychological Review*, **82**, 276–298.

Miller, R. R. & Schachtman, T. (1985). Conditioning context as an associative baseline: implications for response generation and the nature of conditioned inhibition. In R. R. Miller and N. E. Spear, eds., *Information Processing in Animals: Conditioned Inhibition*. Hillsdale, NJ: Erlbaum, pp. 51–88.

Padfield, G. D. & Lawrence, B. (2003). The birth of flight control: an engineering analysis of the Wright brothers' 1902 glider. *The Aeronautical Journal*, December, 697–718.

Pearce, J. M. & Hall, G. (1980). A model for Pavlovian conditioning: variations in the effectiveness of conditioned but not unconditioned stimuli. *Psychological Review*, **87**, 332–352.

Rescorla, R. A. & Wagner, A. (1972). A theory of Pavlovian conditioning: variations in the effectiveness of reinforcement and non-reinforcement. In A. H. Black and W. F. Prokasy, eds., *Classical Conditioning II: Current Research and Theory*. New York: Appleton–Century–Crofts, pp. 64–99.

Romer, A. S. (1970). *The Vertebrate Body*. Philadelphia: W. B. Saunders.

Schmajuk, N. A. (2010). *Mechanisms in Classical Conditioning: A Computational Approach*. Cambridge: Cambridge University Press.

Schmajuk, N. A. & Di Carlo, J. J. (1989). A neural network approach to hippocampal function in classical conditioning. *Behavioral Neuroscience*, **105**, 82–110.

Schmajuk, N. A. & Di Carlo, J. J. (1992). Stimulus configuration, classical conditioning, and the hippocampus. *Psychological Review*, **99**, 268–305.

Schmajuk, N. A. & Larrauri, J. A. (2006). Experimental challenges to theories of classical conditioning: application of an attentional model of storage and retrieval. *Journal of Experimental Psychology: Animal Behavior Processes*, **32**, 1–20.

Schmajuk, N. A., Cox, L. & Gray, J. A. (2001). Nucleus accumbens, entorhinal cortex and latent inhibition: a neural network approach. *Behavioral Brain Research*, **118**, 123–141.

Schmajuk, N. A., Lam, Y. & Gray, J. A. (1996). Latent inhibition: a neural network approach. *Journal of Experimental Psychology: Animal Behavior Processes*, **22**, 321–349.

Schmajuk, N. A., Lamoureux, J. A. & Holland, P. C. (1998). Occasion setting: a neural network approach. *Psychological Review*, **105**, 3–32.

Stout, S. C. & Miller, R. R. (2007). Sometimes-competing retrieval (SOCR): a formalization of the comparator hypothesis. *Psychological Review*, **114**, 759–783.

Wagner, A. R. (1976). Priming in STM: an information-processing mechanism for self-generated or retrieval-generated depression in performance. In T. J. Tighe and R. N. Leaton, eds., *Habituation: Perspectives from Child Development, Animal Behavior, and Neurophysiology*. Hillsdale, NJ: Erlbaum, pp. 95–128.

Wagner, A. R. (1979). Habituation and memory. In A. Dickinson and R. A. Boakes, eds., *Mechanisms of Learning and Motivation*. Hillsdale, NJ: Lawrence Erlbaum.

Wagner, A. R. (1981). SOP: A model of automatic memory processing in animal behavior. In N. E. Spear and R. R. Miller, eds., *Information Processing in Animals: Memory Mechanisms*. Hillsdale, NJ: Erlbaum, pp. 5–47.

Wagner, A. R. & Brandon, S. E. (1989). Evolution of a structured connectionist model of Pavlovian conditioning (AESOP). In S. B. Klein and R. R. Mowrer, eds., *Contemporary Learning Theories: Pavlovian Conditioning and the Status of Traditional Learning Theory*. Hillsdale, NJ: Erlbaum, pp. 149–189.

1

Evolution of attention in learning

JOHN K. KRUSCHKE AND RICHARD A. HULLINGER

Abstract

A variety of phenomena in associative learning suggest that people and some animals are able to learn how to allocate attention across cues. Models of attentional learning are motivated by the need to account for these phenomena. We start with a different, more general motivation for learners, namely, the need to learn quickly. Using simulated evolution, with adaptive fitness measured as overall accuracy during a lifetime of learning, we show that evolution converges to architectures that incorporate attentional learning. We describe the specific training environments that encourage this evolutionary trajectory and we describe how we assess attentional learning in the evolved learners.

Birds do it, bees do it; maybe ordinary fleas do it. They all learn from experience. But why is learning so ubiquitous? Why not just be born already knowing how to behave? That would save a lot of time and a lot of error. Presumably, we are born ignorant either because evolution is unfinished or because what we need to know is too complex to be fully coded in the genome. Either way, it seems that evolution has cleverly found a mechanism for dealing with the birth of ignorance; a mechanism that we call learning.

Of course, it may be that learning is merely something that organisms do for fun in their spare time. Perhaps there is not much adaptive value in learning, and little cost, and therefore no selective pressure on the mechanisms of learning. To the contrary, there is good evidence that learning is metabolically costly (Mery & Kawecki, 2003) and, therefore, it is probably achieving something of reproductive value (Johnston, 1982).