CHAPTER 1

Introduction

COPYRIGHTED MATERIAL

1.1 THE MACHINE INTELLIGENCE RESEARCH

As the understanding of brain-like intelligence and developing self-adaptive systems to potentially replicate certain levels of natural intelligence remains one of the greatest unsolved scientific and engineering challenges, the brain itself provides strong evidence of learning, memory, prediction, and optimization capabilities within uncertain and unstructured environments to accomplish goals. Although the recent discoveries from neuroscience research have provided many critical insights about the fundamental mechanisms of brain intelligence, and the latest technology developments have enabled the possibility of building complex intelligent systems, there is still no clear picture about how to design truly general-purpose intelligent machines to mimic such a level of intelligence (Werbos, 2004, 2009; Brooks, 1991; Hawkins & Blakeslee, 2004, 2007; Grossberg, 1988; Sutton & Barto, 1998). The challenges of accomplishing this long-term objective arise from many disciplines of science and engineering research, including, but not limited to:

- Understanding the fundamental principles and mechanisms of neural information processing in the biological brain organism.
- Advancement of principled methodologies of learning, memory, prediction, and optimization for general-purpose machine intelligence.
- Development of adaptive models and architectures to transform vast amounts of raw data into knowledge and information representation to support decision-making processes with uncertainty.
- Embodiment of machine intelligence hardware within systems that learn through interaction with the environment for goal-oriented behaviors.
- Design of robust, scalable, and fault-tolerant systems with massively parallel processing hardware for complex, integrated, and networked systems.

To find potential solutions to address all of these challenges, extensive efforts have been devoted to this field from many disciplines, including neuroscience, artificial

Self-Adaptive Systems for Machine Intelligence, First Edition. Haibo He.
© 2011 John Wiley & Sons, Inc. Published 2011 by John Wiley & Sons, Inc.

intelligence, cognitive science, computational theory, statistics, computer science, and engineering design, among others. For instance, artificial neural networks have played an important role in the efforts of modeling functions of brain-like learning (Grossberg, 1988). Backpropagation theory has provided a powerful methodology for building intelligent systems and has demonstrated great success across many domains, including pattern recognition, adaptive control and modeling, and sensitivity analysis, among others (Werbos, 1988a, 1988b, 1990, 2005). There are many other representative works in this field as well, including the memory-prediction theory (Hawkins & Blakeslee, 2004, 2007), reinforcement learning (RL) (Sutton & Barto, 1998), embodied intelligence (Brooks, 1991, 2002), adaptive dynamic programming (ADP) (Werbos, 1997, 1998, 2004, 2009; Si, Barto, Powell, & Wunsch, 2004; Powell, 2007), the "new artificial intelligence" theory (Pfeifer & Scheier, 1999), and others. For instance, recently, a new theoretical framework based on hierarchical memory organization was proposed for designing intelligent machines (Hawkins & Blakeslee, 2004, 2007). This theoretical framework provides potential new solutions for how to understand memory and the prediction mechanism based on the neocortex. Because biological intelligent systems can learn through active interaction with the external environment, reinforcement learning has attracted much attention in the community and demonstrated great success in a wide range of applications (Sutton & Barto, 1998). The key idea of reinforcement learning is to learn how to map situations to actions to maximize the expected reward signal. One of the essential aspects of reinforcement learning is the value function, which specifies "good" from "bad" to guide the goal-oriented behaviors of the intelligent system. For instance, in biological systems, it could be a way of measuring happiness or pain (Starzyk, Liu, & He, 2006). The ideas for embodied intelligence originate from the observation that biological intelligent systems have biological bodies and are situated in a set of realistic environments (Brooks, 1991, 2002). The major research efforts for embodied intelligence are focused on understanding biological intelligent systems, discovering fundamental principles for intelligent behavior, and designing real intelligent systems, including living machines and humanoid robotics. Recently, it is recognized that *optimization* and *prediction* play a critical role to bring the brain-like general-purpose intelligence closer to reality (Werbos, 2009). For instance, the recently launched Cognitive Optimization and Prediction (COPN) program from the National Science Foundation (NSF) is a good indication to raise the attention to this critical area by bringing cross-disciplinary teams together to address the essential question of how the brain learns to solve complex optimization and resilient control problems (NSF, 2007). While optimization has a long-standing research foundation in control theory, decision theory, risk analysis, and many other fields, it has specific meanings in terms of machine intelligence research: learning to make better choices to maximize some kind of utility function over time to achieve goals. Extensive research efforts have suggested that ADP is the core methodology, or "the only general-purpose way to learn to approximate the optimal strategy of action in the general case" (Werbos, 2004, 2009). Of course, I would also like to note that many of the aforementioned fields are strongly connected with each other. For instance, ADP/RL approaches can be "embodied" (e.g., coupled with sensory-motor coordination with active interaction with the external environment) or built in a hierarchical way for effective goal-oriented multistage learning, prediction, and optimization (Werbos, 2009).

From the practical application point of view, recent technology developments have enabled the growth and availability of raw data to occur at an explosive rate, such as sensor networks, security and defense applications, Internet, geographic information systems, transportation systems, weather prediction, biomedical industry, and financial engineering, to name a few. In many of such applications, the challenge is not the lack of the availability of raw data. Instead, information processing is failing to keep pace with the explosive increase of the collected raw data to transform them to a usable form. Therefore, this has created immense opportunities as well as challenges for the machine intelligence community to develop self-adaptive systems to process such vast amounts of raw data for information representation and knowledge accumulation to support the decision-making processes.

To this end, this book focuses on the computational foundations of machine intelligence research toward the "computational thinking" (Wing, 2006) capability for self-adaptive intelligent systems design. For instance, although the traditional artificial intelligence methods have made significant progresses and demonstrated great success across different specific application tasks, many such techniques lack the robustness, scalability, and adaptability across different knowledge domains. On the other hand, biological intelligent systems are able to adaptively learn and accumulate knowledge for goal-oriented behaviors. For instance, although today's computers can solve very complicated problems, they use fundamentally different ways of information processing than does the human brain (Hawkins & Blakeslee, 2004, 2007; Hedberg, 2007; Sutton & Barto, 1998). That is why a 3-year-old baby can easily watch, listen, learn, and remember various external environment information and adjust his or her behavior, while the most sophisticated computers cannot. In this sense, one may argue that modern computers are just computational machines without intelligence. This raises critical questions such as "What can humans do better than computers, and vice versa?" or, more fundamentally, "What is computable?" from the computational thinking point of view (Wing, 2006). We believe an in-depth understanding of such fundamental problems is critical for machine intelligence research, and ultimately provide practical techniques and solutions to hopefully bring such a level of intelligence closer to reality across different domains.

To give a brief overview of the major differences between traditional computation and brain-like intelligence, Figure 1.1 compares the major characteristics of these two levels of intelligence. One can clearly see that brain-like intelligence is fundamentally different to that of traditional computation in all of these critical tasks. Therefore, from the computational thinking point of view, new understandings, foundations, principles, and methodologies are needed for the development of brain-like intelligence. This book tries to provide the recent advancements in this field to address such critical needs in the community.

Traditional computation	Tasks	Brain-like intelligence
Sequential	Information processing	Parallel
Fixed	Complexity	Scalable
Centralized	Control mechanism	Distributed
Global	Interactions	Local
Programmed	Source of behaviors	Self-organizing/ conceptualizing
Limited	Fault tolerant	High
Custom designed	Architecture	Evolved
Some	Adaptability	High
Problem specific	Application domains	Robust

Figure 1.1: Comparison of traditional computation and brain-like intelligence.

1.2 THE TWO-FOLD OBJECTIVES: DATA-DRIVEN AND BIOLOGICALLY INSPIRED APPROACHES

Figure 1.2 illustrates a high-level view of the machine intelligence framework that we focus on in this book. Here, there are two important components: the intelligent core such as neural network organizations and learning principles, and the interaction between the intelligent core and the external environment through sensorimotor pathways (embodiment). To this end, this book includes

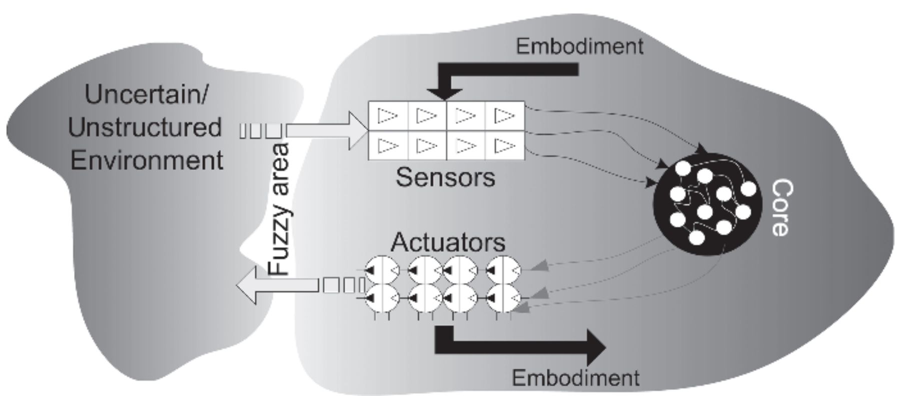

Figure 1.2: A high-level view of machine intelligence.

two major parts to address the two-fold objectives: data-driven approaches and biologically inspired approaches for machine intelligence research. This will not only allow us to understand the foundations and principles of the neural network organizations and learning within the intelligent core, but it also allows us to advance the principled methodologies with a focus on the data processing path (sensing, acquisition, processing, and action). The key is to understand how a brain-like system can adaptively interact with unstructured and uncertain environments to process vast amounts of raw data to develop its internal structures, build associations and predictions, accumulate knowledge over time, and utilize self-control to achieve goals.

The underlying motivation of data-driven approaches is quite straightforward: Data provide the original sources for any kind of information processing, knowledge transformation, and decision-making processes. From the computational intelligence point of view, data are almost involved in every aspect of "intelligence": reasoning, planning, and thinking, among others. Therefore, data can be a vital role for machine intelligence development in different formats, such as sensing, acquisition, processing, transformation, and utilization. You can think about many examples in real-world applications from this perspective, ranging from picking up a pen from your office desk, to driving a car in the metropolitan area of New York City, to scheduling your calendar for the next month. All of these tasks involve data analysis at different levels. If one would like to design an intelligent machine to possibly replicate certain levels of brain-like intelligence, many critical questions are raised from the data computational point of view, such as: What kind of data are necessary to support the decision-making processes? How can an intelligent machine continuously learn from non stationary and noisy data? How do you effectively combine multiple votes from different hypotheses based on different data spaces for optimal decisions?

Specifically, in this book we will discuss the following data-driven approaches for machine intelligence research:

- Incremental Learning. Incremental learning is critical to understand brain-like intelligence and potentially bringing such a level of intelligence closer to reality in at least two aspects. First, intelligent systems should be able to learn information incrementally throughout their lifetimes, accumulate experience, and use such knowledge to benefit future learning and decision-making processes. Second, the raw data that come from the environment with which the intelligent system interacts becomes incrementally available over an indefinitely long (possibly infinite) learning lifetime. Therefore, the learning process in such scenarios is fundamentally different from that of traditional static learning tasks, where a representative data distribution is available during the training time to develop the decision boundaries used to predict future unseen data. Furthermore, how to achieve global generalization through incremental learning is a crucial component in the correct understanding of such problems. Therefore, it is critical to go beyond the conventional "compute–store–retrieve" paradigm for the development of natural intelligent systems for such large-scale and complicated data processing systems.

- Imbalanced Learning. In many real-world applications, an intelligent system needs to learn from skewed data distributions to support decision-making processes. Such skewed distribution with underrepresented data may significantly compromise learning capability and performance. For instance, many of the existing learning algorithms assume or expect balanced data distributions to develop the decision boundary. Therefore, when presented with the imbalanced data, such learning algorithms fail to properly represent the distributive characteristics of the data and resultantly provide worse learning performance. Due to the inherent complex characteristics of imbalanced data and its wide occurrence in many real systems, the imbalanced learning problem has presented a significant new challenge to society with wide-ranging and far-reaching application domains.
- Ensemble Learning. Generally speaking, ensemble learning approaches have the advantage of improved accuracy and robustness compared to the single model–based learning methods. In the ensemble learning scenario, multiple hypotheses are developed and their decisions are combined by a voting method for prediction. Since different hypotheses can provide different views of the target function, the combined decision will hopefully provide more robust and accurate decisions compared to single model–based learning methods. There are two critical aspects related to ensemble learning. First, how can one develop multiple hypotheses given the training data? For instance, to obtain the diversified hypotheses, many techniques such as bootstrap aggregating (bagging), adaptive boosting, random subspace, stacked generalization, mixture of experts, and others have been proposed in the community. Second, how can one effectively integrate multiple hypotheses outputs for improved final decisions? This mainly includes different types of combinational voting strategies, which will also be discussed in this book.

In addition to the data-driven approaches, we have a keen interest to understand and develop biologically inspired approaches for machine intelligence. Recent brain science research has provided strong evidence that the biological brain uses fundamentally different ways in handling various tasks than today's computers (Hawkins & Blakeslee, 2004, 2007; Hedberg, 2007). For instance, although IBM's Deep-Blue can win a chess game against a world champion over a decade ago, it did not tell us too much about the development of general-purpose brain-like intelligent machines as it uses completely different ways of information processing as in the human brain. On the other hand, the evolutionary algorithm has recently showed great potential to develop the self-learning capabilities for a master-level chess program, which could provide us important insights to understanding the essence of machine intelligence (Fogel, Hays, Han, & Quon, 2004). From this perspective, the important question is how to develop biologically inspired system-level models and architectures that are able to mimic certain levels of brain intelligence. In this book we will discuss three major components on this.

- Adaptive Dynamic Programming (ADP). ADP has been widely recognized as the key methodology to understand and replicate general-purpose brain-like intelligence in the community (Werbos, 1994, 1997, 2004, 2009; Si et al., 2004; Powell, 2007). There are two major goals of the ADP research that can contribute to the machine intelligence research: *optimization* and *prediction*. Specifically, optimization in this case can be defined as learning to make better choices to maximize some kind of utility function over time to achieve goals (Werbos, 2009). To this end, the foundation for optimization over time in stochastic processes is the Bellman equation (Bellman, 1957), closely tied with the Cardinal utility function concept by Von Neumann. In addition to optimization, recently strong evidence from neurobiology research suggested that prediction is another equally important component to provide a level of adaptive general-purpose intelligence (Werbos, 2009). Prediction in the ADP design can be considered in a more general way to include much important information, such as the future sensory inputs from observed data as well as modeling and reconstructing of the unobserved state variable, with the objective of facilitating action selection toward optimization (Werbos, 2009). In this book, we propose a hierarchical ADP architecture with multiple-goal representations to effectively integrate optimization and prediction together for machine intelligence research.
- Associative Learning. Associative memory plays a critical role for natural intelligence based on information association and anticipation. Generally speaking, there are two types of associative memories: hetero-associative and auto-associative memory. Hetero-associative memory makes associations between paired patterns, such as words and pictures, while auto-associative memory associates a pattern with itself, recalling stored patterns from fractional parts of the pattern. It is believed that the human brain employs both hetero-associative and auto-associative memory for learning, action planning, and anticipation (Rizzuto & Kahana, 2001; Brown, Dalloz, & Hulme, 1995; Murdock, 1997). The memory evolved in the human brain is self-organized, hierarchically distributed, and data driven. For instance, self-organization is responsible for formation of hierarchically organized structures not only in the human brain but also in the nervous systems of lower vertebrates (Malsburg, 1995). In this book, I will focus on the essential characteristics of associative learning, including self-organization, sparse and local connection, and hierarchical structure.
- Sequence Learning. Sequence learning is widely considered among one of the most important components of human intelligence, as most human behaviors are in the sequential format, including, but not limited to, natural language processing, speech recognition, reasoning and planning, and others. Therefore, the understanding of fundamental problems of sequence learning could provide us critical insights for machine intelligence research.

To this end, we propose a biologically inspired model for sequence learning, storage, and retrieval within distributed hierarchical neural organizations. Prediction capability is an essential element of this sequence learning model.

1.3 HOW TO READ THIS BOOK

This book includes both introductory discussions of machine intelligence backgrounds, as well as the latest advanced theoretical and practical developments in the field. In addition to the presented mathematical foundations, learning principles, models, and algorithms, this book also provides a large number of case studies to demonstrate the applications of such methods to solve real-world problems. All of these could provide a valuable resource for researchers and practitioners interested in the machine intelligence field.

This book can also be used for graduate- and undergraduate-level courses in the area of machine learning, data mining, computational intelligence, and adaptive systems. For courses focused on the data processing perspective, it is suggested students use the materials in Part I of this book (Chapters 2, 3, and 4). For courses focused on biologically inspired learning, students can use the contents in Part II (Chapters 5, 6, and 7). Chapter 8 provides an interesting discussion regarding the hardware design of machine intelligence based on massive, parallel processing architecture, including both the dedicated very large-scale integration (VLSI) systems and reconfigurable field-programmable gate array (FPGA) technology. This chapter could be an interesting resource for anyone who would like to know how existing and emerging technologies (including *memristor*) could potentially change our society by building such complex intelligent systems in hardware. A one-semester course in the computational intelligence topic should be able to cover all the chapters discussed in this book. Furthermore, to improve the readability and flexibility of using this book for course materials at different levels, the chapters have been written to be readable in any sequence to the best effort. However, some interdependence is unavoidable for a coherent and systematic presentation of the materials.

Figure 1.3 shows the organization of the entire book, and a brief review of the chapters is summarized as follows.

1.3.1 Part I: Data-Driven Approaches for Machine Intelligence (Chapters 2, 3, and 4)

Chapter 2 presents basic concepts of incremental learning and its importance for machine intelligence research. The focus is to understand the principles and foundations of adaptive learning over time for knowledge accumulation to support decision-making processes. A specific learning framework is presented and various practical design considerations are given in this chapter. Application case

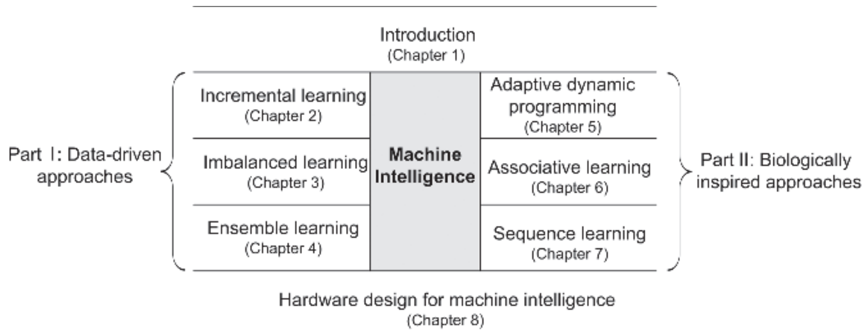

Figure 1.3: The organization of this book.

studies based on video stream data and e-mail data are used to demonstrate the effectiveness of this method.

Chapter 3 covers the imbalanced learning problem. As a relatively new field in the community, the focus for this chapter is to understand the nature of the imbalanced learning problem, and investigate the state-of-the-art techniques to address this problem. Four major categories of methodologies including sampling methods, cost-sensitive methods, kernel-based learning methods, and active learning methods are presented. Assessment metrics used to evaluate learning performance under imbalanced data and the major challenges as well as opportunities on this topic are also discussed in this chapter.

Chapter 4 covers the ensemble learning methods. The focus of this chapter is to discuss techniques used to develop multiple diversified learning hypotheses and to integrate such multiple hypotheses to support the final decision-making processes. Several major approaches including bootstrap aggregating (bagging), adaptive boosting (AdaBoost), and subspace method, among others, are discussed. Numerous combination voting strategies and detailed margin analysis are also presented in this chapter.

1.3.2 Part II: Biologically-Inspired Approaches for Machine Intelligence (Chapters 4, 5, and 6)

Chapter 5 discusses the ADP methods for machine intelligence research. Our focus here is to understand the fundamental principles of ADP design for optimization and prediction. A hierarchical learning architecture with three types of networks is proposed, and a specific learning algorithm based on backpropagation is also presented. We also demonstrate a case study of this architecture to the cart-pole balancing control problem.

Chapter 6 covers self-organizing memory in hierarchical neural network organization. This includes the association learning mechanism, neural network organization, and network operation. Application cases of such memory organization for both hetero-associative and auto-associative learning are also presented.

Chapter 7 presents a neural network structure for complex sequence learning, storing, and retrieving. This architecture features two essential characteristics for machine intelligence: hierarchical neural organization and distributed information processing. We present the detailed system-level architecture and its learning mechanism. A case study with four-level hierarchical structure for text analysis is used to demonstrate the application of this model.

Finally, Chapter 8 provides a discussion regarding hardware design for machine intelligence based on the latest hardware platforms, including dedicated VLSI technology as well as reconfigurable FPGA technology. The goal is to provide useful insights about critical hardware design considerations, such as power consumption, design density, memory requirement, and speed requirement, to potentially build a large-scale, integrated, and complex system into real hardware. We also point to some emerging technologies such as *memristor* for future brain-like intelligent systems design to conclude this book.

1.4 SUMMARY AND FURTHER READING

This book aims to advance the fundamental understanding of general-purpose brain-like intelligence research, and develop principled methodologies and practical techniques to bring such a level of intelligence closer to reality through a wide range of domains. The underlying methodology presented in this book is based on two paths of research: data-driven approaches and biologically inspired approaches for machine intelligence.

Machine intelligence research draws on theories and concepts from many disciplines. For the latest research development in this field, interested readers can find a number of good sources from many international journals, which include, but are not limited to, *IEEE Transactions on Neural Networks*, *Neural Networks*, *Neural Computation*, *IEEE Transactions on Evolutionary Computation*, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *IEEE Transactions on Knowledge and Data Engineering*, *Artificial Intelligence*, *Cognitive Brain Research*, *Machine Learning*, and others. Meanwhile, there are also a number of conferences that cover different aspects of machine intelligence research, including *International Joint Conference on Neural Networks (IJCNN)*, *National Conference on Artificial Intelligence (AAAI)*, *International Joint Conference on Artificial Intelligence (IJCAI)*, *Neural Information Processing Systems (NIPS)*, *International Conference on Machine Learning (ICML)*, *International Conference on Data Mining (ICDM)*, *Annual Meeting of the Cognitive Science Society (CogSci)*, and others.

REFERENCES

Bellman, R. E. (1957). *Dynamic programming*. Princeton, NJ: Princeton University Press.

Brooks, R. A. (1991). Intelligent without reason. *Proc. Int. Joint Conf. on Artificial Intelligence*, pp. 569–595.

Brooks, R. A. (2002). *Flesh and machines: how robots will change us*. New York: Pantheon.

Brown, G. D. A., Dalloz, P., & Hulme, C. (1995). Mathematical and connectionist models of human memory: a comparison. *Memory*, *3*(2), 113–145.

Fogel, D. B., Hays, T. J., Han, S. L., & Quon, J. (2004). A self-learning evolutionary chess program. *Proc. IEEE*, 92, 1947–1954.

Grossberg, S. (1988). *Neural networks and natural intelligence*. Cambridge, MA: MIT Press.

Hawkins, J., & Blakeslee, S. (2004). *On intelligence*. New York: Times Books.

Hawkins, J., & Blakeslee, S. (2007). Why can't a computer be more like a brain? *IEEE Spectrum*, *44*(4), 20–26.

Hedberg, S. R. (2007). Bridging the gap between neuroscience and AI. *IEEE Intel. Syst*., *22*(3), 4–7.

Malsburg, C. V. (1995). Self-organization and the brain. In M. Arbib (Ed.), *The Handbook of Brain Theory and Neural Networks* (pp. 840–843). Cambridge, MA: MIT Press.

Murdock, B. B. (1997). Context and mediators in a theory of distributed associative memory (todam2). *Psychological Review*, *104*, 839–862.

NSF. (2007). Emerging frontiers in research and innovation: Cognitive optimization and prediction: From neural systems to neurotechnology (copn). *[online], available: http://www.nsf.gov/pubs/2007/nsf07579/nsf07579.htm.*

Pfeifer, R., & Scheier, C. (1999). *Understanding intelligence*. Cambridge, MA: MIT Press.

Powell, W. B. (2007). *Approximate dynamic programming: Solving the curses of dimensionality*. Hoboken, NJ: Wiley.

Rizzuto, D. S., & Kahana, M. J. (2001). An autoassociative neural network model of paired-associate learning. *Neural Computation*, *13*, 2075–2092.

Si, J., Barto, A., Powell, W. B., & Wunsch, D. (2004). *Handbook of learning and approximate dynamic programming*. Piscataway, NJ: IEEE Press.

Starzyk, J. A., Liu, Y., & He, H. (2006). Challenges of embodied intelligence. *Proc. Int. Conf. on Signals and Electronic Systems*. Lodz, Poland.

Sutton, R. S., & Barto, A. G. (1998). *Reinforcement learning: An introduction*. Cambridge, MA: MIT Press.

Werbos, P. J. (1988a). Backpropagation: Past and future. *Proc. IEEE Int. Conf. Neural Netw*., I-343–353.

Werbos, P. J. (1988b). Generalization of backpropagation with application to a recurrent gas market model. *Neural Netw*., *1*, 339–356.

Werbos, P. J. (1990). Backpropagation through time: What it does and how to do it. *Proc. IEEE*, 78, 1550–1560.

Werbos, P. J. (1994). Approximate dynamic programming for real time control and neural modeling. In P. J. White & P. J. Sofge (Eds.), *Handbook of intelligent control* (pp. 493–525). New York: Van Nostrand.

Werbos, P. J. (1997). Brain-like design to learn optimal decision strategies in complex environments. *Proc. Conf. Decision and Control*, pp. 3902–3904.

Werbos, P. J. (1998). A brain-like design to learn optimal decision strategies in complex environments. In M. Karny, K. Warwick, & V. Kurkova (Eds.), *Dealing with complexity: A neural networks approach*. London: Springer.

Werbos, P. J. (2004). ADP: Goals, opportunities and principles. In J. Si, A. G. Barto, W. B. Powell, & D. Wunsch II (Eds.), *Handbook of learning and approximate dynamic programming* (pp. 3–44). Piscataway, NJ: IEEE Press.

Werbos, P. J. (2005). Backwards differentiation in AD and neural nets: Past links and new opportunities. In H. M. Bucker, G. Corliss, P. Hovland, U. Naumann, & B. Norris (Eds.), *Automatic differentiation: Applications, theory and implementations, lecture notes in computational science and engineering* (Vol. 50, pp. 15–34). Berlin, Germany: Springer-Verlag.

Werbos, P. J. (2009). Intelligence in the brain: A theory of how it works and how to build it. *Neural Networks*, *22*, 200–212.

Wing, J. M. (2006). Computational thinking. *Communications of the ACM*, *49*(3), 33–35.