Preface

This book is about neural computation and applied reasoning, and about how they are related and can be integrated within a machine-learning framework. Logic has a long and solid tradition of principled research, and has been influential in virtually all fields of scientifically-based human knowledge. Neural computation has now achieved maturity. Practitioners have been able to use artificial connectionist models in a number of real-world problems, including machine learning, pattern recognition, vision, cognitive modelling, and artificial agents. However, the area still lacks sound computational foundations, as several features of neural-computation models are still under investigation.

The book presents a number of principled, sound neural-computation models. It makes use of a number of applied (nonclassical) logical systems in order to do so. Logics lie at the foundation of computer science and artificial intelligence. For instance, logic research has provided the foundations of the theories of computability, recursion, and complexity. In artificial intelligence, the construction of automated reasoning systems, knowledge representation languages and formalisms, and theorem proving have been pursued since at least the 1950s. All these are based on logics.

More recently, applications in planning, multiagent systems, expert systems, the Semantic Web, learning theory, cognitive modelling and robotics, constraint satisfaction, searching, and the traditional fields of knowledge representation and reasoning have benefited from research on nonclassical logics. In particular, research on temporal, modal, description, intuitionistic, fuzzy, and nonmonotonic logics has been influential in artificial intelligence, both in relation to theory and principled languages for building models and systems.

Recently, nonclassical logics[1] have been found relevant in relation to the formalisation of integrated learning and reasoning [66, 72, 77, 148, 222]. They have contributed to the construction of sound models of cognitive reasoning. As is well known, human cognition successfully integrates the connectionist (brain-inspired)

---

[1] Nonclassical logics depart from the Aristotelian tradition in the sense that these logics lack some properties of classical logic, such as the law of the excluded middle, or extend them with additional features. For instance, the development of nonmonotonic logics, which has contributed to

and symbolic (mind-inspired) paradigms of artificial intelligence, language being a compelling case in point. Yet the modelling of cognition develops these separately in the fields of neural computation (and statistical artificial intelligence) and symbolic logic/artificial intelligence. There is now a movement towards a fruitful middle way in between these extremes, in which the study of logic is combined with connectionism. It is essential that these be integrated, thereby enabling not only a technology for building better intelligent systems, but also the understanding of fully fledged computational cognitive models.

The aim of neural-symbolic computation is to explore the advantages that each paradigm presents. Among the advantages of artificial neural networks are massive parallelism, fault tolerance (robust learning), efficient inductive learning, and effective generalisation capabilities. On the other hand, symbolic systems provide descriptions (as opposed to only discriminations); can explain their inference process, for example through automatic theorem proving; and use powerful declarative languages for knowledge representation and reasoning.

In this book, we explore the synergies of neural-symbolic integration from the following perspective. We use a neural network to simulate a given task. The network is obtained by being programmed (set up) or by adapting to and generalising over well-known situations (learning). The network is the mechanism to execute the task, while symbolic logic enables the necessary interaction between the network and the outside world. Differently from and complementing [66], which set the scence for neural-symbolic integration with the use of standard networks and logic programming, we are concerned here with applied (and nonclassical) logical systems. With this consideration at the forefront, *the book presents a rich cognitive model for reasoning and learning based on neural-network ensembles*. The book also illustrates the effectiveness of the model by experimentation, and shows that the connectionist model can compute a number of combined nonclassical logical systems [42], including modal, temporal, epistemic, and intuitionist logics. Finally, the book investigates the issue of relational learning and the representation of first-order logic, the combination (fibring) of models/networks, qualitative reasoning under uncertainty, and how neural networks can offer an effective approach to learning in argumentation frameworks [39, 212]. We conclude by summarising our case for a logic-based, cognitive connectionist model which we call *fibred network ensembles*. Overall, it offers a rich model from a symbolic computation/reasoning viewpoint, yet it is relatively simple and efficient as a learning model. We believe it strikes the right balance between expressiveness and computational feasibility. And, as a matter of principle, we believe that this is no coincidence, but a direct result of

---

the formalisation of practical or commonsense reasoning, led to the discovery of several new logical systems in which the principle of monotonicity may not hold. Further, several authors have defended a position that intuitionistic logics are more appropriate as a foundational basis of computation than classical logic [258]. Classical logic does not capture all of the nature of human practical reasoning. For instance, in order to formalise and automate computational (human) reasoning about time, knowledge, beliefs, or uncertainty several nonclassical logics have been found more appropriate than classical logic; see [42, 87, 100, 102]. Our point will be made clearer as the reader proceeds to the coming chapters.

our methodology, which seeks to unify systematically the most advanced concepts of symbolic computation with the physical constraints of a realistic connectionist machine.

Our aim here is to offer a principled way in which robust learning and effective reasoning can be realised. In line with one of the recent grand challenges in computer science, according to the British Computer Society – nonclassical computation – we seek to unify the classical and nonclassical (natural) computational paradigms, with the nonclassical computational models presented here showing how alternative formulations can deal with cognitive dimensions usually studied under separate formalisms.

London, UK, and Porto Alegre, Brazil, February 2005 to June 2008.

Chapter 1
Introduction

1.1 Motivation

The construction of robust computational models integrating reasoning and learning is a key research challenge for artificial intelligence. Recently, this challenge was also put forward as a fundamental problem in computer science [255]. Such a challenge intersects with another long-standing entry in the research agenda of artificial intelligence: the integration of its symbolic and connectionist paradigms. Such integration has long been a standing enterprise, with implications for and applications in cognitive science and neuroscience [51, 66, 130, 178, 179, 238, 240, 247, 248, 250]. Further, the importance of efforts to bridge the gap between the connectionist and symbolic paradigms of artificial intelligence has also been widely recognised (see e.g. [51, 66, 229, 242, 243]).

Valiant [255] has pointed out that the construction of rich computational cognitive models is one of the challenges computer science faces over the next few decades. A positive answer to this challenge would provide a characterisation of a semantics for cognitive computation,[1] as follows:

> The aim here is to identify a way of looking at and manipulating commonsense knowledge that is consistent with and can support what we consider to be the two most fundamental aspects of intelligent cognitive behaviour: the ability to learn from experience, and the ability to reason from what has been learned. We are therefore seeking a semantics of knowledge that can computationally support the basic phenomena of intelligent behaviour [255].

Valiant also described the characteristics of the semantic formalisation needed for supporting learning and reasoning:

> One set of requirements [for a semantics to be adequate for commonsense reasoning] is that it should support integrated algorithms for learning and reasoning that are computationally

---

[1] The article [255] was published in 2003 in the Journal of the ACM, in a special issue celebrating its 50th anniversary. In that issue, the editor-in-chief at the time (J.Y. Halpern) invited winners of the Turing Award and Nevanlinna Prize to discuss up to three problems that these prominent researchers thought would be major challenges for computer science in the next 50 years.

tractable and have some nontrivial scope. Another requirement is that it has a principled way of ensuring that the knowledge-base from which reasoning is done is robust, in the sense that errors in the deductions are at least controllable [255].

Aiming at building integrated reasoning and learning methods, our approach provides a unified computational foundation for neural networks and nonclassical reasoning. Knowledge is expressed by a symbolic language, whereas deduction and learning are carried out by a robust connectionist engine. This book also seeks to contribute to the long-term aim of representing expressive symbolic formalisms in learning systems [253], by means of neural-symbolic integration [66]. Ultimately, the goal is to produce biologically motivated models with integrated reasoning and learning capabilities, in which neural networks provide the inspiration and the machinery necessary for cognitive computation and learning, while several nonclassical logics provide practical reasoning and explanation capabilities to the models, facilitating the interaction between the models and the outside world. This book contributes to the integration of both research programmes into a unified foundation; both of these programmes are now widely but separately used in many areas of computer science and artificial intelligence [42, 66, 87, 125].

A historical criticism of neural networks was raised by McCarthy back in 1988 [176]. McCarthy referred to neural networks as having a "propositional fixation", in the sense that they were not able to represent first-order logic. This, per se, has remained a challenge for a decade, but several approaches have now dealt with first-order reasoning in neural networks (see e.g. [43] and Chap. 10). Perhaps in an attempt to address McCarthy's criticism, many researchers in the area have focused attention only on first-order logic. This has suppressed developments in other important fronts, mainly in nonclassical, practical reasoning, which also should be at the centre of neural-symbolic integration research owing to the practical nature of neural-network research. We have shown recently that nonclassical reasoning can be used in a number of applications in neural-symbolic systems [33, 68–72, 74, 76–79, 157]. This has been possible through the integration of nonclassical logics and neural networks.

Notwithstanding this evidence, little attention has been given to nonclassical reasoning and its integration with neural networks. We believe that for neural computation to achieve its promise, connectionist models must be able to cater for nonclassical reasoning. Research on nonclassical logics, including new results on modal, temporal, intuitionistic, and nonmonotonic logics and their combinations, has been relevant not only to computer science and artificial intelligence, but also to economics and the physical sciences. We believe that neural-symbolic systems can benefit from the results and achievements that nonclassical logics have had in all these areas.

In summary, we shall argue in this book that nonclassical reasoning is fundamental in the construction of computational connectionist models. If one assumes that neural networks can represent rich models of human reasoning and learning, and can offer an alternative solution to the challenges confronted by intelligent computation, it is undeniable that nonclassical logics should play a fundamental role at the centre of this enterprise.

1.2 Methodology and Related Work

Several approaches have been proposed for integrating the connectionist and symbolic paradigms of artificial intelligence. Most provide a solution to the learning of classical propositional logic or fragments of first-order logic by means of neural networks or related methods (see e.g. [12, 43, 148, 228, 229, 250, 254]). Our book [66] surveys the work on neural-symbolic integration done until 2002, and proposes a methodology for dealing with nonmonotonicity in artificial neural networks, including knowledge extraction. In [12], a survey of recent developments in classical-logic learning in neural networks is presented. Further, [66] showed that neural-symbolic systems are also appropriate for tackling learning in real-world problems. In particular, the analysis presented in [66] shows that neural-symbolic systems can be used effectively in a number of applications, ranging from the detection of large-scale power system failures to DNA sequence analysis.

Despite the significant contributions of the developments in first-order logic to knowledge representation, learning, and reasoning in artificial intelligence, a truly intelligent agent or multiagent system, in the sense defined in [271], has several dimensions that cannot be appropriately managed solely by the use of first-order classical logic.

There are several extensions and alternatives to classical logic. Nonclassical logics have become useful in computer science and artificial intelligence over the last few decades. Such logics have been shown to be adequate for expressing several features of reasoning, allowing for the representation of temporal, epistemic, and probabilistic abstractions in computer science and artificial intelligence, as shown for example, in [42, 87, 106, 121].

For instance, temporal, modal, and intuitionistic logics are now amongst the most successful logical languages used in computing. Born in philosophy and mathematics, they have benefited from research efforts in applications of computing. Several semantic models and (automated) proof systems have been designed for nonclassical logics [42, 88, 104]. Temporal logic has had a successful history in computer science and artificial intelligence since the pioneering work of Pnueli, back in 1977 [207], as it allows an accurate and elegant formalism for reasoning about the dynamics of computing systems. Temporal logic has had a large impact in both academia and industry [89, 103]. Modal logic, in turn, has also become a lingua franca in the areas of formalisation, specification, verification, theorem proving, and model checking in multiagent and distributed computing systems [42, 50, 87, 106, 143, 154]. Nonmonotonic reasoning dominated research in artificial intelligence in the 1980s and 1990s, and intuitionistic logic is considered by many to be an adequate logical foundation in several core areas of theoretical computer science, including type theory and functional programming [258]. Other applications of nonclassical logics include the characterisation of timing analysis in combinatorial circuits [180] and in spatial reasoning [23], with possible use in geographical information systems. For instance, Bennett's propositional intuitionistic approach provided for tractable yet expressive reasoning about topological and spatial relations. In [106], several applications of many-dimensional modal logic are illustrated.

Automated reasoning and learning theory have been the subject of intensive investigation since the early developments in computer science and artificial intelligence [81, 174, 251]. However, while machine learning has been developed mainly by the use of statistical and connectionist approaches (see e.g. [125, 173, 184, 262]), the reasoning component of intelligent systems has been developed using classical and nonclassical logics (see e.g. [42, 87, 100, 104]). The acceptance of the need for systems that integrate reasoning and learning into the same foundation, and the evolution of the fields of cognitive and neural computation, has led to a number of proposals integrating logic and machine learning [43, 51, 66, 77, 79, 118, 148, 164, 229, 250, 254, 255].

An effective model of integrated reasoning and learning has been shown to be attainable by means of neural-symbolic learning systems [66, 69–72, 74, 79]. This book advocates the use of nonclassical logics as a foundation for knowledge representation and learning in neural-symbolic systems. We propose a new approach for representing, reasoning with, and learning nonclassical logics in a connectionist framework, which leads, in a principled way, to a powerful but computationally light cognitive model combining expressive nonclassical reasoning and robust learning; we call it *fibred network ensembles*.

In contrast to symbolic learning systems, neural networks' learning implicitly encodes patterns and their generalisations in the networks' weights, so reflecting the statistical properties of the trained data [35]. The merging of theory (background knowledge) and data learning (learning from examples) into neural networks has been shown to provide a learning system that is more effective than purely symbolic or purely connectionist systems, especially when the data are noisy [246, 250]. This result has contributed to the growing interest in developing neural-symbolic learning systems. By integrating logic and neural networks, neural-symbolic systems may provide (i) a logical characterisation of a connectionist system, (ii) a connectionist (parallel) implementation of a logic, or (iii) a hybrid learning system that brings together features from connectionism and symbolic artificial intelligence.

Until recently, neural-symbolic systems were not able to fully represent, compute, and learn expressive languages other than propositional logic and fragments of first-order, classical logic [12, 43, 51, 238]. To the best of our knowledge, research efforts towards representing nonclassical logical formalisms in connectionist systems were scant until the early 2000s. However, in [67, 70, 73, 74, 76–78], a new approach to knowledge representation and reasoning in neural-symbolic systems based on neural-network ensembles was proposed, namely *connectionist nonclassical logics*. In [75], connectionist modal logic (CML) was introduced, showing that modalities can be represented effectively in neural networks. In [70, 72, 73], the language of the Connectionist Temporal Logic of Knowledge (CTLK) was introduced, and in [76–78] the computation of intuitionistic reasoning was shown to be learnable within neural networks. This new approach shows that a variety of nonclassical logics can be effectively represented in artificial neural networks. To the best of our knowledge, this was the first approach to combining nonclassical logics and neural networks.

Recently, it has also been shown that value-based argumentation frameworks can be integrated with neural networks, offering a unified model for learning and reasoning about arguments, including the computation of circular and accrual argumentation [68, 69]. The study of formal models of argumentation has also been a subject of intensive investigation in several areas, notably in logic, philosophy, decision theory, artificial intelligence, and law [25, 31, 39, 48, 83, 210, 212]. In artificial intelligence, models of argumentation have been one of the approaches used in the representation of commonsense, nonmonotonic reasoning. They have been particularly successful when modelling chains of *defeasible arguments* so as to reach a conclusion [194, 209]. Although logic-based models have been the standard for the representation of argumentative reasoning [31, 108], such models are intrinsically related to artificial neural networks, as we shall show in Chap. 11. This relationship between neural networks and argumentation networks provides a model in which the learning of arguments can be combined with argument computation.

This book also presents a new neural-network architecture based on the idea of fibring logical systems introduced by Gabbay [101]. Fibring allows one to combine different systems (here, neural networks) in a principled way. Fibred neural networks may be composed not only of interconnected neurons but also of other networks in a recursive architecture. A fibring function then defines how this recursive architecture must behave, by defining how the networks should relate to each other (typically by allowing the activation of one network to influence the changes of the weights of another). We show that, in addition to being universal approximators, fibred networks can approximate any polynomial function to any desired degree of accuracy, and are thus more expressive than standard feedforward neural networks.

Neural-symbolic systems that use simple neural networks, such as single-hidden-layer feedforward or recurrent networks [125], typically only manage to represent and reason about propositional symbolic knowledge or *if then else* rules [36, 66, 95, 205, 250]. On the other hand, neural-symbolic systems that are capable of representing and reasoning about (fragments of) first-order logic are normally less capable of learning new concepts efficiently [136, 149, 229, 243]. There is clearly a need to strike a balance between the reasoning and learning capabilities of such systems, and between expressiveness and computational complexity.

As argued in [43], if connectionism is an alternative paradigm to artificial intelligence, neural networks must be able to compute symbolic reasoning in an efficient and effective way. It is also argued that connectionist systems are usually fault-tolerant, whereas symbolic systems may be 'brittle and rigid'. We seek to tackle these problems by offering a principled way of computing, representing, and learning nonclassical logics within connectionist models.

The combination of nonclassical reasoning and connectionism is achieved by means of algorithms that translate logical clauses into neural-network ensembles. Such algorithms can be proved correct in the sense that the ensembles compute a semantics of the original theory. An immediate consequence of our approach is the ability to perform learning from examples efficiently, by applying, for example, the backpropagation learning algorithm [224] to each network of the ensemble. We also

show the effectiveness of our approach as a (distributed) knowledge representation, reasoning, argumentation, and learning mechanism by applying it to the well-known test beds found in the literature [87, 121]. Our approach paves the way for modular, integrated computation and learning of distributed, nonclassical knowledge, with a broad range of applications from practical reasoning to evolving multiagent systems.

Technical aspects of this work will be presented throughout the book as the need arises. No assumption is made that the reader has prior knowledge of nonclassical logic or neural networks. Connectionist modal logic, a (one-dimensional) ensemble of neural networks [66], is used to represent modalities such as *necessity* and *possibility*. In CTLK, a two-dimensional network ensemble is used to represent the evolution of knowledge through time. In both cases, each network ensemble can be seen as representing a possible world that contains information about the knowledge held by an agent in a distributed system. Learning in the CML system is achieved by training each network in the ensemble independently, corresponding to the evolution of an agent's knowledge within a possible world. It is important that these logics are investigated within the neural-computation paradigm. For instance, applications in artificial intelligence and computer science have made extensive use of decidable modal logics, including the analysis and model checking of distributed and multiagent systems, program verification and specification, and hardware model checking.[2] In the case of temporal and epistemic logics, these logics have found a large number of applications, notably in game theory and in models of knowledge and interaction in multiagent systems [87, 89, 103, 207].

From a machine-learning perspective, the merging of theory (background knowledge) and data learning (learning from examples) in neural networks has provided learning systems that are more effective than purely symbolic or purely connectionist systems [246, 250]. In order to achieve this merging, first one translates the background knowledge into a neural network's initial architecture, and then one trains the network with examples using, for example, backpropagation. In the case of CML, for instance, learning is achieved by training each individual network, each of which is a standard network.

Another long-term aim is to contribute to the challenge of representing expressive symbolic formalisms within learning systems. We are thus proposing a methodology for the representation of several nonclassical logics in artificial neural networks. We believe that connectionist approaches should take these logics into account by means of adequate computational models catering for reasoning, knowledge representation, and learning. This is necessary because real-world applications, such as failure diagnosis, fraud prevention, and bioinformatics applications, will require the use of languages more expressive than propositional logic. Bioinformatics, in particular, requires very much the ability to represent and reason about relations such as those used in predicate logic [6]. In summary, knowledge is represented by a symbolic language, whilst deduction and learning are carried out by a connectionist engine.

---

[2] It is well known that modal logic corresponds, in terms of expressive power, to the two-variable fragment of first-order logic [264]. Further, as the two-variable fragment of predicate logic is decidable, this explains why modal logic is so robustly decidable and amenable to multiple applications.

1.3 Structure of the Book

This research monograph is divided into the following chapters. Chapter 1 (the present chapter) introduces the subject and overviews developments in the area of connectionist models for integrated reasoning and learning. Chapter 2 introduces the basic concepts of logic and knowledge representation. Chapter 3 introduces the concepts of neural networks. Chapter 4 introduces the foundations of neural-symbolic integration. Chapter 5 introduces connectionist modal logic, covering some foundational results and introductory examples. Chapter 6 presents CTLK and its applications in distributed temporal knowledge representation and learning. Chapter 7 introduces intuitionistic reasoning and learning in neural networks. Chapter 8 describes some applications of connectionist nonclassical reasoning. Chapter 9 introduces the idea of combining (fibring) networks, for example CTLK and intuitionism. Chapter 10 describes the combination networks to represent and learn relations in a first-order setting with variables. Chapter 11 establishes a close relatioship between connectionist models and argumentation frameworks, and uses argumentation as an application of fibring. Chapter 12 introduces symbolic reasoning under uncertainty in neural networks and illustrates its feasibility using well-known test beds, including the Monty Hall puzzle [121]. Chapter 13 concludes the book and indicates directions for further research.

An extensive list of references cited is provided at the end of the book. The list is by no means complete, as the literature in this field is vast.