PRESERVING KNOWLEDGE IN SIMULATED BEHAVIORAL ACTION LOOPS

by

Rachel St.Clair

PREVIEW

A Dissertation Submitted to the Faculty of

Charles E. Schmidt College of Science

In Partial Fulfilment of the Requirements for the Degree of

Doctor of Philosophy

Florida Atlantic University

Boca Raton, FL

May 2022

Copyright 2022 by Rachel St.Clair

PREVIEW

PRESERVING KNOWLEDGE IN SIMULATED BEHAVIORAL ACTION LOOPS

by

Rachel St.Clair

This dissertation was prepared under the direction of the candidate's dissertation advisor, Dr. Elan Barenholtz and Dr. William Hahn, Center for Complex Systems and Brain Sciences, and has been approved by all members of the supervisory committee. It was submitted to the faculty of the Charles E. Schmidt College of Scienceand was accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

SUPERVISORY COMMITTEE:

Elan Barenholtz
Elan Barenholtz (Mar 29, 2022 11:10 EDT)

Elan Barenholtz, Ph.D.
Dissertation Co-Advisor

William Hahn (Apr 5, 2022 11:08 EDT)

William Hahn, Ph.D.
Dissertation Co-Advisor

Susan Schneider
Susan Schneider (Apr 5, 2022 11:37 EDT)

Susan Schneider, Ph.D.

Gary W Perry
Gary W Perry (Apr 5, 2022 11:53 EDT)

Gary Perry, Ph.D.
Chair, Center for Complex Systems and Brain Sciences

Teresa Wilcox

Teresa Wilcox, Ph.D.
Interim Dean, Charles E. Schmidt College of Science

Robert W. Stackman Jr., Ph.D.
Dean, Graduate College

April 8, 2022

Date

PREVIEW

ACKNOWLEDGEMENTS

The author wishes to express sincere gratitude to her committee members for all of their guidance and support, and a very special thanks to my advisors for their persistence, patience, encouragement, countless pep-talks, and endless support during the preparation and typing of this manuscript. The author is grateful to the Rubin Gruber Sandbox for providing the research equipment and space to conduct the study. Last but not least, the author wishes to thank her lab mates for all of their support.

PREVIEW

ABSTRACT

Author: Rachel St.Clair

Title: Preserving Knowledge in Simulated Behavior Action Loops

Institution: Florida Atlantic University

Dissertation Advisor: Dr. Elan Barenholtz and Dr. William Hahn

Degree: Doctor of Philosophy

Year: 2022

One basic goal of artificial learning systems is the ability to continually learn throughout that system's lifetime. Transitioning between tasks and re-deploying prior knowledge is thus a desired feature of artificial learning. However, in the deep-learning approaches, the problem of catastrophic forgetting of prior knowledge persists. As a field, we want to solve the catastrophic forgetting problem without requiring exponential computations or time, while demonstrating real-world relevance. This work proposes a novel model which uses an evolutionary algorithm similar to a meta-learning objective, that is fitted with a resource constraint metrics. Four reinforcement learning environments are considered with the shared concept of depth although the collection of environments is multi-modal. This system shows preservation of some knowledge in sequential task learning and protection of catastrophic forgetting in deep neural networks.

PREVIEW

DEDICATION

This manuscript is dedicated to my family, particularly my supportful and patient mother, Kimberly, and father, Joe, who have encouraged me through these many years of research, and especially to my cat, Boopi, who is the joy of my life. I also dedicate this work to my friends who have allowed me to allot an extraordinary amount of my cognitive processing time to this work by offloading some basic human functioning to their own accord, and my grandparents, Phyllis and Don, both of whom believed in the pursuit of my dreams unequivocally.

PREVIEW

PRESERVING KNOWLEDGE IN SIMULATED BEHAVIORAL ACTION LOOPS

List of Tables ix

List of Figures x

1 Introduction 1

Significance 6

Theoretical Basis 7

Relevant Literature 12

Dilemma 15

Research Questions and Hypotheses 16

Research Question One 16

Research Question Two 16

2 Literature Review 17

History of Computing 17

Backpropagation 19

Genetic Algorithms 20

Artificial General Intelligence 21

Knowledge Preservation 22

Theory Relevant Literature 23

Connectionism 23

Symbolism 23

Artificial General Intelligence 24

Neuro-Symbolic Architectures 25

PREVIEW

Backpropagation as a Plausible Brain Mechanism 25
Contending Frameworks and Theories of Brain Activity 29
Current Empirical Literature 31
Catastrophic Forgetting 31
Transferring Knowledge 33
Neuroevolution 33
3 Method 35
Environment 35
Computational Framework 39
Measures 41
Research Design 43
Procedures 46
Data Analysis 50
Research Question One 50
Research Question Two 51
4 Results 53
Model and Environment Analysis 53
Statistical Analysis 58
Research Question One 58
Research Question Two 62
5 Discussion 65
Summary 65
Conclusions 65
Importance of Layer Learning 65
Environment and Model Viability 67
Training with Resource Constraints in Evolutionary Loops . . 68
Mixed Sequences and Depth Sequences 68

PREVIEW

Model Weight Change 69
Evolutionary Approach to Learning 70
Hypothesis Analysis 70
Limitations 72
Recommendations for Future Research 74
Appendix. 75
References 86

PREVIEW

LIST OF TABLES

3.1 Set-Up for All Environments 39

3.2 Set-up for Computational Framework 41

3.3 Accuracy Measures per Model and Environment 43

4.1 P-values for Depth and Mixed Test DQN+GA and DQN-GA. 61

PREVIEW

LIST OF FIGURES

3.1 Basic Environment Set-up 37
3.2 Match Environment 38
3.3 DQN Model Flow. 40
3.4 Novel Model Flow 45

4.1 DQN Hyperparameter Importance 54
4.2 DQN Training Results 55
4.3 Base Model Training Accuracy 56
4.4 DQN+GA Hyperparameter Importance 57
4.5 DQN+GA Hyperparameter Search 57
4.6 DQN+GA, DAN+GA-RC, DQN-GA Depth Sequence Training Resuts. 59
4.7 DQN+GA and DQN-GA Depth Sequence Training Results. 60
4.8 DQN+GA and DQN-GA Depth Sequence Test Results. 62
4.9 Model Weight Change 63
4.10 DQN+GA and DQN-GA Test Results. 64

PREVIEW

Chapter 1

INTRODUCTION

One basic goal of artificial learning systems is the ability to continually learn throughout that system's lifetime. The ability for continual learning, or life-long learning implies that the artificial learning system, or agent, can learn without forgetting prior learning. The potential utility of learning without forgetting, or 'knowledge preservation', is twofold: first, the system can return to previously learned tasks after learning something new. Secondly, bootstrapping previous knowledge may allow for faster learning of a novel task. Both of these potential benefits of preserving learned information may serve to conserve resources. Allowing the same computational hardware to be 'multiplexed' for another task, rather than requiring allocation of distinct resources, reduces the required energy resources involved in the learning process. Multiplexed information is stored in such a way that the system can switch between previously learned tasks while retaining performance. Information can then be accessed for learning new tasks in a contextually relevant manner. Multiplexed information is thus stored and retrieved in a way that facilitates the conservation and use of prior learning, broadening the ability of a learning system to multiple task domains by encoding high-level 'multi-use' concepts.

Although many metrics exist for characterizing desirable properties of artificial general intelligence (AGI), the term is commonly used to describe a software program that is highly adaptable to a variety of environments while displaying a range of functionality similar to, or beyond, human cognition. The term is distinct from traditional artificial intelligence research, which aims to make domain-specific prediction programs. AGI, or rather superintelligent AGI, which has the potential to

PREVIEW

drastically improve the quality of human life, being able to solve problems at computational speeds that far exceed the capacity of several human lives combined and with perspectives on systemic societal issues currently unavailable to human bias. However, it is not without ethical consideration that the creation of such a program could be adversarial to human life if the sueprintelligent AGI does not hold the same values as it's biological counterparts [1]. Nevertheless, several approaches for AGI have been created and play an important role in understanding artificial learning and cognition.

Both approaches, AGI and AI, to creating artificial learning systems encompass a wide variety of approaches, all attempting to create an intelligence in which being capable of learning without forgetting to is desirable to some degree. Transitioning between tasks and re-deploying prior knowledge is thus a desired feature of artificial learning. Each of these approaches has a unique perspective to offer upon our understanding of intelligence as a whole, as it relates to human, animal, and machine 'life'. However, no approach thus far has been able to demonstrate the degree of knowledge preservation necessary to achieve life-long learning to the degree of biological counterparts.

PREVIEW

In recent years, practical implementations in the field of artificial intelligence (AI) have largely converged on deep learning methodologies. The underlying mechanics of the majority of these approaches involve artificial neural networks constructed with weights which are tuned by backpropagating an error signal in accordance with some update rule (i.e. stochastic gradient descent). While deep learning has shown state-of-the-art performance in many domains such as computer vision, natural language processing, time-series forecasting, etc., one major problem is that of catastrophic forgetting, in which new learning overwrites prior learning. Challenges in overcoming catastrophic forgetting are bound by practical considerations such as resource constraints and model applicability. As a field, we want to solve the catastrophic forget-

ting problem without requiring exponential computations or time, while demonstrating real-world relevance. Learning without catastrophic forgetting would potentially allow an AI model to transfer its knowledge base between tasks. The study of transfer learning is still in its infancy in neural network models. More research is needed on how to create robust algorithms that employ prior knowledge without requiring excessive training time and previous learning rehearsal. Thus, new learning without overwriting prior learning provides a model with a broadened intelligence, rather than the narrow, domain-specific intelligence seen in current AI approaches. I suspect a broad intelligence basis, when paired with other adequate algorithms, might allow for the emergence of previously elusive cognitive behaviors such as causal inference, common sense, rational choice, etc.

The field of AGI encompasses many techniques and metrics. Most notable in recent years are described in [2], which describes symbolic, emergentist, artificial life, developmental robotics, universalist, and hybrid approaches to AGI. In symbolic architectures, which have seen the majority of adoption in the AGI community, have the challenge of designing memory components capable of strong learning. Other approaches exhibit strong learning, such as associative memory capabilities, but lack the ability for high-order concepts, i.e. complex language and reasoning. While still, others lack the ability to model the brain to the degree of which larger artificial learning systems can be constructed, failing to result in algorithms which exhibit multiple cognitive behaviors. Hybrid architectures attempt to leverage the various symbolic and connectionist approaches but lack in their ability to be understandable in terms of design and offer little insight into human intelligence. That, these approaches are often a hodge-podge of design elements taken from computational experiment, and not lack systematic design based on properties of human intelligence as we understand it from the perspective of neuroscience. Interoperability for understanding the resulting artificial learning system is important as it protects against adverse interpretations of

PREVIEW

reality as we understand it, compared with reality as a computer may understand it; thus, protecting ethical considerations and real-world applicability [3]. While many of these techniques have resulted in useful, intelligent algorithms, we have yet to see algorithms that result in knowledge preservation close to the level in which vertebrate brains can perform. Even though our understanding of neuroscience is still incomplete, core features from foundational mechanisms of action can be recruited to better design AGI systems. Thus, it is foreseeable that AGI approaches would benefit from better understanding how knowledge preservation occurs in biological systems.

It can be inferred from biological intelligence that one core property of how biological brains learn is their ability to preserve knowledge within resource constraints. That the number of neurons for an animal is mostly fixed and animals with advanced nervous systems can learn multiple distinct tasks without forgetting prior tasks. Evolution designed vertebrate brains in a highly conserved manner in over about 3.5 million years [4]. This information is stored molecularly, mostly by way of genetic encoding. The modern human brain sequesters this information during early prenatal development before any online learning starts to take place. Advantageous features of the brain to be repairable, easily constructible, conserve resources, modifiable, and synchronous favored the highly conserved architectural design of biological brains [5]. In vertebrates, these evolutionary pressures resulted in nervous system topology structures with distinct condition detection devices (neurons) arranged in segregated populations (brain areas) with ambiguous roles which coordinate with the larger specific tasks of their respective area. From such architecture, emergent fundamental features arise allowing for higher cognitive processes, according to the recommendation architecture theory of cognition [6].

When neuron populations are segregated based on functional tasks (eg. visual vs auditory streams), they have to share information via some third party mechanisms of integration to signal to the overall system the appropriate interpretation of the current

PREVIEW

condition. This area to area signaling is not exactly information compression, it's more like communication or control and may be a foundational principle of emergent cognitive phenomena. It is plausible that this communication between brain areas lends itself to the self-generated, "subjective", experience. The modular paradigm found in the human brain uses several tools to manage information flows through the system, including (but not limited to) inhibition, feedback, behavior recommendation and selection, indirect activation, receptive field expansion, and reward cascades [7].

Evolution is a slow process of optimization for which we can only observe the final products and few intermediate states. By contrast, researchers can design and implement computational architectures speedily and flexibly, creating the potential for new models that predict and test hypotheses about the evolution of the nervous system. Currently, most artificial network architectures are arbitrarily designed due to lack of systematic evidence for what types of connectivity topology best facilitates what type of learning. Here, connectivity refers to how information units are linked to other information units, whether it be nodes in a neural network, subject and predicate links in graphs, etc.. Vertebrate brains are highly conserved in that they have continuously evolved without discarding much of the previously adapted architecture. If neuron connectivity lends itself to learning environmental solutions, then some previously learned solutions are already stored in the earlier evolved parts of our brains.

PREVIEW

The niche field of neuroevolution tries to replicate what occured naturally by allowing network topology to evolve using various genetic algorithms given input. Most of the best neuroevolutionary models investigate emergent architectures with specific inputs for task optimization. To date, no published approach uses knowledge preservation as an optimization metric. Likewise, there is little information on how connectivity architectures emerge from an evolving environment subject to similar pressures in which the human brain developed.

Significance

The importance of investigating what types of architectures are well suited for preserving knowledge lies in constructing artificial learning systems that can not only provide utility in our daily lives, but also give insight into our understanding of human learning and intelligence.

Currently, topologies of artificial learning systems are somewhat haphazardly designed. Informing our architecture designs with systematic research, will lead to better insight into how models of artificial cognition can be better constructed. The key idea is modeled from the human brain where it is likely that network topology is at least, if not more, important than later learning algorithms in achieving knowledge preservation. This allows us to avoid ending up with a system that is doing the bulk of the learning process after it has been configured. In a system that prioritizes initial connectivity, this manifests as understanding what is the appropriate topology to facilitate continual learning, over the model's lifetime. Our current artificial learning algorithms drastically reduce their biological counterparts to imitative mathematics that do not yet withstand the demands of real-world problem solving within resource conservative systems.

The major problem of catastrophic forgetting prevents artificial neural networks from learning in a similar fashion to the human brain. An abundance of evidence suggests that biological brains are capable of learning without forgetting and that they do so within resource constraints. Brains do not create a new neuron for every piece of information they need to learn. Instead, they multiplex existing learning in novel contexts. If we better understand how connectivity between neurons facilitates knowledge preservation, we can not only create more generally intelligent artificial agents, but also learn more about how the human brain is facilitating learning.

In a broader sense, knowledge preservation lends itself to general learning which

PREVIEW

is useful for designing artificial intelligences with less computational resources and in better understanding how the organizations of neurons in the human brain may facilitate different types of life-long learning. The significance of having an artificial general intelligence is that if AGI is created with benevolent,ethical consideration, it is speculated to be useful in solving a range of everyday tasks as well as tackling some of humanity's toughest challenges. We will have created something that is capable of solving problems, just as a human would but with much more speed and processing power. The reason for these exponential returns is that silicon chips are inherently much faster than biological neurons. As semiconductor research progresses, it is foreseeable that any model of cognition that runs in a computer will be much faster than any biological counterpart. The difficulties in creating a computational model that has the ability to help humanity solve our toughest challenges, eve those we cannot foresee, is that the model must be ethical (sharing humanitie' values), capable of general intelligence (that it can learn in a exceedingly comparable fashion to human intelligence), and that it can understand the world as we understand it, or rather, better than we understand it.

PREVIEW

Theoretical Basis

Lifelong learning, or continual learning, is the idea that a computational model should learn across tasks without being reset to an initialization state. In lifelong learning, a model is trained upon inception on a task and retains it's training so long as it is in use [8]. This means that the model weights are not reset, as they typically are in current AI models, between training tasks. Knowledge preservation concerns itself with lifelong learning models. Those that re-initialize the weights are inherently destroying all previous learning. Being able to switch between tasks, using prior knowledge, is a deserada of general intelligence. Thus, lifelong learning models play a critical role in understanding the functions of intelligence in biological brains.

The question then becomes, how do we structure models to facilitate learning across a vast range of tasks, storing knowledge in such a way that it can return to previous tasks later and use prior learning to learn new tasks efficiently. Connectionism proposes that the brain is composed of units, or neurons, which act on other units to create a functional dynamical network of total units active at one time [9]. Activity denotes excitation or inhibition. This idea has been extensively extended to computer science where simulations of nodes activating other nodes within a network occur. Thus, neural networks rely extensively on the connectionist paradigm by incorporating connected nodes updated with backpropagation. In backpropagation, node weights are learned by propagating and error backwards through the hierarchy of connections [10]. In these approaches, knowledge is built by iterating over environmental input which is forced through hierarchical encodings [11]. Here, hierarchy enforces decomposition of higher concepts into parts, which relies on the combinatorial expression of the network to make sense of. The 'mind' results from the environmental input's impression on the neural substrate. Parallel processing enables the network to create a dynamical system which in turn, takes on new functions for the network itself, interacting with other networks as a complex system [12]. The main point of connectionism is that the connectome, or connections between neurons, construct the mind, or rather, cognition. The architecture of the brain results in learning which transfers input to output in a series of parallel computations of units activating other units in a network. In the field of artificial intelligence, recent research has only just begun to look at what types of architectures facilitate which types of learning. While many deep learning approaches already use inductive biases of biological brains, such as visual cortex hierarchies in the convolutional neural network, there is much that remains to be understood on the impact of architectures on inductive biases of neural networks. The topology of various networks considers how computational units, or neurons, are connected in such a way that facilitates

PREVIEW

learning appropriate information and transcoding that information into appropriate output, or behavior. Connectionism in artificial neural networks measures learning by stimulus-response pairs, or rather how well the model can predict an output from an input. Artificial learning models, especially those using deep learning methods, suffer in the ability to learn new tasks without forgetting prior-learning. In the artificial intelligence community, this problem is known as catastrophic forgetting (CF), or the plasticity/stability trade-off. This is in part due to the fundamental nature of back-propagation in deep learning models to rewrite learning; where weights are updated most heavily towards current learning at the expense of weights learned in more distant tasks [13]. The result of backpropagation, even when assuaged by CF mitigation, is narrow feature learning that will eventually reach a resource constraint asymptote as learning increases throughout the model's lifetime. In recent years, practical implementations in the field of artificial intelligence (AI) have largely converged on deep learning methodologies. The underlying mechanics of these approaches involve artificial neural networks constructed with weights which are tuned by backpropagating an error signal in accordance with some update rule (i.e. stochastic gradient descent). While deep learning has shown state-of-the-art performance in many domains such as computer vision, natural language processing, time-series forecasting, etc., one major problem is that of catastrophic forgetting (CF), in which new learning overwrites prior learning. Challenges in overcoming CF are bound by practical considerations such as resource constraints and model applicability. As a field, we want to solve the catastrophic forgetting problem without requiring exponential computations or time, while demonstrating practical relevance. Learning without CF would potentially allow an AI model to transfer its knowledge base between tasks. Thus, new learning without overwriting prior learning provides a model with a broadened intelligence, rather than the narrow, domain-specific intelligence seen in current AI approaches. We suspect a broad intelligence basis, when paired with other adequate algorithms,

PREVIEW

might allow for the emergence of previously elusive cognitive behaviors such as causal inference, common sense, rational choice, etc. Embodied cognition describes a theory of cognition in which the brain needs the ability to execute actions in an environment in a cyclical fashion. Behavioral outputs react with the environment, producing the next set of inputs, aiding in higher cognitive functions, such as learning and agency. In embodied cognition, representations of the mind are manipulated to interact with nature in a way that produces a dynamical system [14]. The being, or agent of the brain, is directly entwined with the interaction of it's situation. Physiology cannot be distinctly separated from goals of the agent, and thus, must be viewed as part of the overarching nature of cognition [15]. Embodied cognition describes a system in which an agent takes input from the environment, computes that information, and provides output which in turn creates it's next input. Reinforcement learning (RL), a subset of artificial intelligence and dynamic programming, models embodied cognition in the form of state, action, and reward scenarios. The core foundation of RL is that most behaviors are sequential, requiring active engagement in sequestering the next input to the sequence [16]. A balance between exploration and exploitation is critical in successful RL algorithms as the first results in new solutions and the former in reliance on previous learning. Therefore, retaining prior learning and being able to use it in new situations is imperative for RL architectures. The mathematical foundations of RL rely upon markov decision making [17]. Excessive trial-and-error is needed to iterate through effective policies for decision making before an optimal solution is converged upon. This results in the need for an abundance of computational resources; unlike biological brains which are bound by the number of neurons and biochemical processes available to the system.

The evolution of the human nervous system has been speculated to be a product of darwinian fitness. The central dogma of Darwinian evolution is that adaptations which are advantageous get passed down via genetic traits [18]. Here, adaptations

PREVIEW

are behaviors encoded in genes and advantage, or fitness, refers to the organism's increased survival and reproduction. The class of computational algorithms, genetic algorithms (GA), capitalize on this fundamental feature; that a fitness function is used as a metric to determine which genes should propagate to future individuals in a population. GAs are mathematical constructs to embody natural evolution [19]. While most GAs use the strategy of transferring genes, or rather computational instructions, to the next generation, or rather iteration, of the population (i.e. program), according to a fitness function. Together, the fitness function and the mathematical code creates an optimization tasks, which with enough computational resources and time, is generally successful at deriving solvable solutions. This approach, although human derived, is quite similar to natural selection in Darwinian evolution since the program's behavior results in advancing the system forward, by recombination and mutation of genetic material that provides an increase in fitness. The fitness function thus becomes instrumental in advancing the program's outcome and constitutes the evolutionary pressures of development.

Resource constraints are considered as a key indicator of multiplexing information in life-long learning systems here based on the theoretical framework of the Recommendation Architecture (RA) as theorized in [6]. Understanding recommendation architecture (RA) requires thinking about the brain as a system in which different anatomical structures like the cortex, basal ganglia, hippocampus, thalamus, amygdala and cerebellum are subsystems that perform different, distinctive types of information processes, each contributing in key ways to the brain's conscious activity. This framework for understanding higher cognition describes brain function in a similar mechanism to understanding functions of a computer; by relating levels of understanding to levels of anatomical description. The key feature of RA is that it describes foundational properties of cognition in terms of core system interactions. This means that properties of the brain can be somewhat abstracted from

PREVIEW

their biological properties of neurochemical interactions and dynamical systems into artificial systems which embody the system interactions by mechanisms outside of physiology and anatomy. Here, we rely on the RA's description of evolutionary pressure and constraints of biological brains to provide support for artificial cognition being constrained within resource conservative systems. This approach is very useful in understanding the problem of knowledge preservation. Thus, we can define the concept of knowledge preservation as the ability to learn information for later multiplexing under resource constraints. The constraint of limited resources enforces the multiplexing of information units. That is, in order to keep the number of neurons needed from growing exponentially, neurons need to be used in various ways without degrading prior learning. Once information is stored appropriately, it can then be bootstrapped to inform novel contexts. With this consideration, learning can then be measured, in part, in terms of using prior information in novel context and returning to previous context without information degradation, all within minimal requirements of the amount of computational resources available.

PREVIEW

Relevant Literature

Recent literature suggests that the intersection of continual deep reinforcement learning with genetic algorithms may be useful for preserving knowledge. State of the art learning algorithms in reinforcement learning rely on deep-learning based Deep Q Networks (DQNs) and Markov decision based processes, namely Proximal Policy Optimization (PPO) and Soft Actor-Critic (SAC) models. In both of these cases, an agent uses one of these decision making models to perform tasks in an environment. Some decisions in the tasks provide the agent's model with a reward value. The reward value is used to calculate the appropriateness of the decision making policy within the model. Over time, the single-task or multiple-task objective is completed, often in a continual learning process. In genetic algorithms, a continual learning

process spans over a period of time where populations of agents learn according to a fitness function, providing the 'reward' for agents individually and/or as a whole population of agents. Thus, both RL and GAs are similar in their overall design for continuous learning, but the mechanisms by which policies are constructed and rewards are given are specific to the model design.

In the works of [20], a single-task continual reinforcement learning objective is used to test how context-dependent relevance to tune the information in the models in such a way that the task performance is increased and memory requirements are decreased compared to other state-of-the-art RL models. The key insight to their work is that the sequential nature of learning is often dependent on prior contexts. They successfully showed that by incorporating some metric of context, the model is more robust to changes in input states because the distribution can adapt more smoothly by incorporating gradient information of recent context. Thus, by distributing prior knowledge into the current state, the dynamically leverages the plasticity/stability. However, their approach did not explicate how the information is managed in the weights of the network, or how long the continual learning tasks could be performed before the benefits of the model reach resource constraints. This work suggests that incorporating sequential contextual information across the continual learning process of the model has benefits for mitigating catastrophic forgetting.

The works of [21] further suggest that continual deep reinforcement learning can be modified to mitigate catastrophic forgetting using another technique called policy reuse, where task-specific policies are cached and switched between tasks. Their findings support this approach for lifetime learning. The key to their work is that multiple-task learning objectives can be learned independently over a continual learning process while minimizing catastrophic forgetting. Here the limitation is that the tasks are finite and pre-determined, leaving little room for adaptability and generalization to different categories of tasks and resource constraints for storing policies.

PREVIEW

Reproduced with permission of copyright owner. Further reproduction prohibited without permission.