氏　　　　名	チン ウェイ ホン CHIN WEI HONG
所　　　　属	システムデザイン研究科　システムデザイン専攻
学 位 の 種 類	博士（工学）
学 位 記 番 号	シス博　第 122 号
学位授与の日付	令和元年 9 月 30 日
課程・論文の別	学位規則第 4 条第 1 項該当
学位論文題名	A Declarative Memory Recurrent Neural Model for Lifelong Learning of Intelligent Agents （知的エージェントのライフロング学習のための陳述記憶に基づくリカレントニューラルモデル）
論文審査委員	主査　教　授　久保田　直行
	委員　准教授　武居　直行
	委員　准教授　福井　隆雄
	委員　准教授　和田　正義　（東京農工大学大学院）

【論文の内容の要旨】

Intelligent learning agents and autonomous robots interacting in complex environments must constantly acquire and refine knowledge over long periods of time. For example, a neural network is often used for knowledge acquisition, but incremental learning remains a challenge because of catastrophic forgetting in which new sensory data impedes with existing learned knowledge and significantly degrades the performance of neural networks. This limitation represents a main disadvantage for state-of-the-art neural networks, which commonly learn from uniform training data batches without considering data that is sequentially available over time.

Declarative memory is a human long-term memory that store facts and previous experiences over time while integrating new data. Declarative memory is further divided into episodic memory and semantic memory. The episodic memory represents memory of experiences and specific events in time. The semantic memory is formed by the episodic memory to generate a more generalized knowledge. Complementary Learning Systems (CLS) theory claims that the hippocampus promptly encodes episodic events that can be reactivated during resting (memory recall). In this way, knowledge is consolidated by reactivating encoded experiences in terms of multiple internal replays. The CLS theory provides the basis for the development of cognitive neural models to

generalize over experiences while maintaining particular memories in a lifelong learning manner. For example, dual-memory framework with probabilistic models has been proposed in previous works related with CLS theory. However, there is no empirical evidence showing that these approaches can be applied to real world implementation.

In this thesis, I propose a novel recurrent neural model that mimics human declarative memory system for lifelong learning. The aim of the proposed method is to overcome the catastrophic forgetting dilemma. The proposed neural model consists of three hierarchical memory layers: i) Working Memory, ii) Episodic Memory and iii) Semantic Memory layer. Each memory layer comprises a self-organizing adaptive recurrent incremental network (SOARIN) that generates a topological network for a particular learning task. The Working Memory quickly learns object instances from input sources (clustering); while the Episodic Memory learns fine-grained spatiotemporal relationships of object instances (temporal encoding). The Semantic Memory utilizes task-relevant cues to adjust the level of network plasticity (categorization). To alleviate the catastrophic forgetting, the Episodic Memory regularly reactivates previously learned temporal neurons activations and replays to itself and the Semantic Memory (memory replay) without requiring external sensory data.

The effectiveness of the proposed recurrent neural model is evaluated through experiments specifically in both spatial and temporal domains. I implemented and validated my proposed model on the tasks of robot spatial mapping and human skeleton actions recognition. Robot spatial mapping is one of the incremental learning tasks in the spatial domain as the robot must understand and update its knowledge in a place where environmental conditions are progressively changing. Another challenging incremental learning task in the temporal domain is human action recognition in which robots must observe sequences of human motions and categorizes these motions into different actions.

This thesis introduces a novel recurrent neural model that generates topological networks incrementally for adapting incoming sensory data. In addition, the memory replay alleviates the catastrophic forgetting during incremental learning. Thus, the proposed neural model can constitute a basis for intelligent learning agents to acquire a higher level of cognitive capabilities for accomplishing real-world learning tasks.

In Chapter 1, I provide the social background that leading to this thesis research objectives. Contributions of the research work are highlighted in this chapter as well.

Chapter 2 provides an introduction in biological aspects of incremental learning and catastrophic forgetting in neural networks for a better understanding of the challenges reflected in this thesis. In addition, I review well-established findings regarding human memory in the brain along with a background on computational architectures for state-of-the-art human memory modeling.

In Chapter 3, I propose a cognitive neural model with long-term memory that generates a topological network to quickly learn to categorize sensory data in the real-world and remember it, without forgetting previously-learned knowledge. The model has been validated in real robot for generating the topological network as a spatial map. The generated spatial map is then integrated with path planning for navigating robots to move from one place to another. However, due to the learning nature of the model, the number of neurons is increasing monotonically which requires large memory to store the network in the long run.

In Chapter 4, I develop a long-short-term memory neural model to overcome the limitation of the model in Chapter 3. The neural model contains a working memory that generates short-term memory and equipped with a mechanism to remove neurons that inactivate for long time. The neural model successfully implemented in robot spatial mapping that dynamically adjust the number of neurons in the topological network. However, the neural model is incapable of encoding sensory temporal data. Therefore, path planning is still required in order to perform robot navigation.

In Chapter 5, I further extend the neural model in Chapter 4 with the episodic memory for encoding spatiotemporal connections of neurons as episodes while generating topological networks incrementally. As such, the generated topological network contains both spatial and temporal sensory data. Naturally, the episodic memory navigates robots to move from one place to another without path planning. The neural model is successfully implemented in autonomous goal navigation. However, regulating structural plasticity levels and developing more compact network from episodic memory remain a challenge for this model.

In Chapter 6, by incorporating a semantic memory layer hierarchically, the model in Chapter 5 is expanded. The semantic memory layer generates a more compact topology network from episodic memory. To this end, episodic neuron activation and labeling signals are received by the semantic memory layer to modulate the plasticity of the network. In addition, sequences of previously-activated neurons replay to both networks for consolidation of knowledge to overcome the catastrophic forgetting. The model is implemented in continuous spatial mapping with place labeling. Furthermore, the model is successfully applied in continuous human action recognition for temporal domain validation. Experimental results showed that the extended model scales up to learning novel data incrementally. In addition to encoding spatiotemporal sensory data, the generated networks constitute an interface for human-robot interaction in both the spatial and temporal domains of lifelong learning.

Concluding remarks are in Chapter 7, the experimental results are discussed from the viewpoint of the research objectives, limitations with the proposed recurrent neural model. Finally, I provide several future research paths.