様　式　C－１９、F－１９－１、Z－１９　（共通）

科学研究費助成事業　　研究成果報告書

令和　５ 年　 ６ 月 １５ 日現在

機関番号：３８００５
研究種目：若手研究
研究期間：2020～2022
課題番号：２０K１９９０１

研究課題名（和文）Self-Organized Multi-Level Working Memories Facilitate Predictive Coding Based Action Panning

研究課題名（英文）Self-Organized Multi-Level Working Memories Facilitate Predictive Coding Based Action Panning

研究代表者
ＱＵＥＩＳＳＥＲ　Ｊｅｆｆｒｅｙ（Queisser, Jeffrey）

沖縄科学技術大学院大学・認知脳ロボティクス研究ユニット・スタッフサイエンティスト

研究者番号：８０８６９２０６

交付決定額（研究期間全体）：（直接経費）　　1,600,000円

研究成果の概要（和文）：本研究ではAIシステムへのワーキングメモリについて調査を行った
物体操作の内容の直接学習ではなく、視覚的ワーキングメモリで表現される内容から学習することで、未学習の状況における汎化性能の向上を達成した。
モデルは、目標指向型の計画が必要なブロックの積上タスクにおいて、ロボットが未知の物体を操作できることを示した。また、開発した「内容にとらわれない情報処理」の概念を、固有感覚と言語生成の文脈に拡張した。
結果、導入したワーキングメモリモジュールが、明示的な制約なく学習したモデルの内部表現を階層的に整理できること、特定の記憶を結合し活用するモデルの作業成果が反復/計数タスクにおいて向上することを示した。

研究成果の学術的意義や社会的意義
This research increased sample efficiency and lowering the computational complexity for online robotic behavior generation, where “big-data” is not available and systems need efficient generalization from few observations.
Further, the models can be valuable for understanding brain (dis-)functions.

研究成果の概要（英文）：The conducted work explored working memories (WM) for AI systems.
By learning to manipulate content in a visual WM instead of learning to represent the content directly, an improved generalization performance to unlearned situations could be achieved: The developed model is able to control a robot to manipulate previously unseen objects in a block stacking scenario that requires goal‐directed planning. Generalization was tested for new colors of objects and applied textures. Further, the concept of content-agnostic information processing was extended to the context of proprioception and language generation.
Results show that the introduced working memory modules result in hierarchical organization of internal representations of learned models without imposed explicit supporting constraints. As a further result, supporting evidence for an improved task performance of models utilizing a certain class of memory connectivity for repetition/counting tasks could be found.

研究分野：AI, Machine learning, Robotics

キーワード：Generalization　Free Energy Minimization　Predictive Model　Robot　Learning　Planning

１．研究開始当初の背景

The aim of the project is the implementation of brain like intelligence for robotic systems based on two kinds of motivation: firstly, the extension of the state of the art of current robotic applications in real world environments (RG1), and secondly, to reveal the evolutionary developed strategies of the brain to cope with complex and variable environments (RG2).
Remarkable developments in the machine learning community, i.e. in the field of Deep Learning (DL), allow the application of learning methods to a wide range of complex regression and classification tasks. Nevertheless, for "Big Data" - a large amount of training data is necessary to allow for a good generalization performance. Prominent example tasks are sequence-to-sequence learning, image classification, or feature space compression. Particularly the application of those learning architectures in interactive systems and for online learning is limited as the availability of training samples is restricted.
From the perspective of the gain of knowledge for the neuroscience community, insights from Deep Learning are limited as well:

(1) Current models employ many layered processing levels that are difficult to relate to the brain substrate with respect to the processing speed.
(2) In many cases the designed machine learning models a treated as a black box as the brain itself.

From the opposite point of view, models based on findings in neuroscience are hard to implement on a big scale and on real-world scenarios due to their complexity, like spiking neuronal networks or detailed simulations on the neuronal information transmission.
For the conducted project we referred to recurrent neuronal networks (RNNs) [1,2] as a compromise in the abstraction between complexity and modeling details. RNNs have been widely used in neurorobotics studies to investigate temporal organization in temporal prediction of sensorimotor tasks. RNNs form attractors that store previously perceived / presented patterns but pose the problem that attractors are slow to learn, knowledge transfer over longer periods of time is difficult and generalization is challenging as the predictions tend to converge to the previously seen patterns.
Prior works (e.g. [3, 4]) show the remarkable ability to allow goal-directed planning and to self-organize the utilization of a simple buffer as visual working memory, in the course of end-to-end learning. For the experiments, the models were trained on sequences consisting of visual and proprioceptive data of goal directed actions of a robot.
Without imposing specific constraints on the functionality of the attention and buffer modules of the model, end-to-end learning was applied. During learning, the model was trained to predict the visual and proprioceptive signals of the respective next time-step, such that after the training phase, it can be used as a predictive sequence generator. In the case of a closed-loop configuration, the initial state of the model determines following predictions of the successive time steps, additionally, probabilistic sampling of the initial state and/or hidden states can be introduced. Goal-directed planning can be conducted by presentation of a desired goal state at a later time-step in combination with a back-propagation of error

signals that relate to the mismatch between the predicted and desired goal state at the final time step. It was shown that as a result of learning, the robot's dynamic visual attention follows the manipulated objects in the representation of the visual stream and the visual buffer of the model is utilized for temporary storage of unattended content in the visual images.

It is crucial to note that the aforementioned functionality of the modules is not predefined, but rather organizes itself in the course of learning the regeneration of exemplar sequences. Based on the results of the discussed studies, the conducted work continued this line of research and tried to explore and extend the limits of the model in terms of generalization capabilities.

２．研究の目的

The core idea of the project involves a learning system that does not predict future states of the environment and its own actions directly, it rather learns a manipulation sequence of an initial state of the environment. By preference for the manipulation of content over the sensory predictions directly, an improved generalization of the system to new situations is expected. Figure 1 conceptualizes the implemented system in a bio-inspired context. The prediction of sensory sensations and generalization of motor signals is based on a combined prediction by the central pattern generator and the current representations in the working memory (WM). The working memory is based on two separate modules, first a static memory that lives in the sensory space, and a second memory that is able to manipulate the content of the working memory. Manipulation can be initiated by a gating process that attends some specific features in the WM and an optional update of the PFC. The PFC itself is limited to the manipulation of its own data through transformations. The transformed data can then be transferred back to the WM for later use and used by the system to generate sensory predictions.

As an example case, assume the system is trained with complex object manipulations. The system can represent manipulations like rotations of objects as transformation functions of the memory.

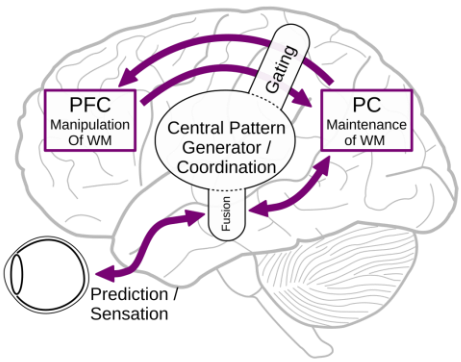

Figure 1: Illustration of the visual pathway of the proposed model for sensorimotor prediction. The working memory of the PC allows a long term storage of information. The PFC module allows the manipulation of the WM content. The central pattern generator performs sensory prediction and the coordination / fusion with the WM.

This is beneficial as the parameterization of the transformation (i.e. object becomes larger) remains the same independent of the object identity. Thereby a form of variable binding emerges that allows the separation of identities and actions.

In comparison to classical network architectures, the central pattern generator learns to control the memory modules, the transformation function, and the data fusion instead of the prediction of the direct pixel values for a rotated object. E.g. the system learns that an object becomes larger in case the robot moves the object closer to itself.

３．研究の方法

The implementation of the model is based on the predictive coding type deep visuomotor recurrent neural network model (P-DVMRNN) as proposed in [3,4]. The model is extended by an attention mechanism that introduced a unified representation of focal (high-resolution, i.e. attended) and peripheral (low-resolution, i.e. unattended) areas in the visual stream. Additionally, a second visual memory in the space of attended vision is introduced as the previously described dynamic memory structure. The model is differentiable and can be trained by state-of-the-art gradient descent methods like the Adam optimization technique. The functionality of the WM is supposed to emerge during training due to restrictions of the connectivity. The prediction of the sensorimotor signals and control signals for gating and fusion operations is performed by stacked (convolutionary) LSTM cells. The memory modules are implemented as stationary buffers as in [4]. For the conducted research, the transformation of the dynamic memory is based on a predefined transformation, i.e. affine transformations. Open future work includes the aim for the extension of these transformations to learned parameterized functions and self-organization.

４．研究成果

(1) Core Results

The conducted research revealed that even though earlier models show good generalization capabilities in their learned domain, they fail in case it is applied for planning tasks in which strong deviations in the visual appearance are present.

When testing these models on previously unseen objects, they show the behavior to fall back to their internal learned attractor dynamics, with manifests in morphing unfamiliar visual appearances of novel objects into previously seen ones.

My research showed the developed concept of *content-agnostic information processing* is able to overcome this limitation and to allow the model to perform generalization in an extended domain (details in [5]).

This project demonstrated that it is possible to apply a learned manipulation sequence to objects with arbitrary visual appearances. The model developed a way to create two spatially distributed representations, including ①, the visual appearances of objects, and ②, the executed sequence of manipulations on them.

It is important to note, that the representations of the visual appearances of objects are manipulated in a modality-specific manner, and by content-independent transformation functions. The executive control of ②, on the other hand, is performed by a generation of task-specific but modality-independent modulations of the internal pathways of the model and the parameterization of the image transformations of ①.

This novel representation of goal-directed actions for planning has drastic consequences on the efficiency of the representation of learned tasks, as manipulation of objects can be learned independently from the representation of content.

(2) Further outcomes that are not published yet (in preparation)

The concept of content-agnostic information processing has been further extended to the context of proprioception and language generation.

Results show that the added working memory modules to network architectures result in hierarchical organization of internal representations of learned models without imposed explicit supporting constraints. As a further result, supporting evidence for an improved task performance of models utilizing a certain class of working memory connectivity for repetition/counting tasks could be found.

(3) Ongoing studies and future work

The project supported the development of tactile sensor devices, Figure 2(a-b), that are used in an experimental setup involving a humanoid robot. Integration and utilization are currently being explored in a collaborative project at OIST. The aim is to use tactile feedback in the predictive framework as elaborated above to facilitate inference of environmental states and guidance of adaptation of generation of motor-primitives.

Further, a quadruped robot was designed and implemented (Figure 2c) and is currently prepared for experiments involving action planning through the developed active-inference framework involving explicit working memory. The ongoing studies involve application specific scenarios that highlight and stress the key features of the developed methods, namely fast adaptation to novel and uncertain scenarios as well as data efficiency through superior generalization due to content-agnostic information processing.

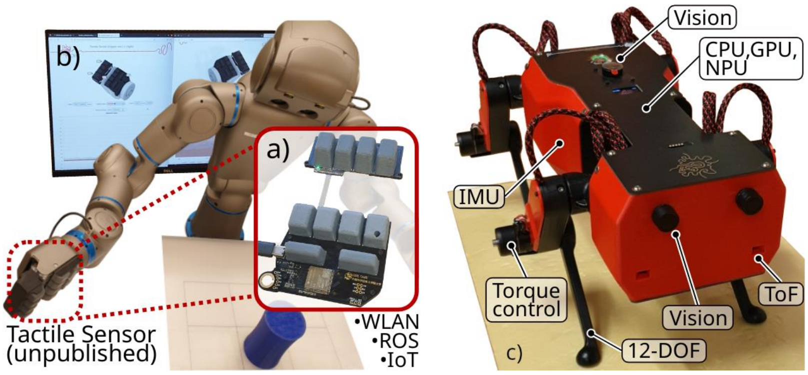

Figure 2: Tactile sensor device (a) of right gripper and its application (b). Monitor displays web-UI including live data (via websocket); Quadruped Robot (c). Torque control, and on-board NPU for online-learning experiments.

REFERENCES:

[1] M. Z. Alom, T. M. Taha, C. Yakopcic, S. Westberg, P. Sidike, M.S. Nasrin & V.K. Asari, "A state-of-the-art survey on deep learning theory and architectures", Electronics, 8(3), 292, 2019

[2] T. M. Yong Yu, Xiaosheng Si, Changhua Hu and Jianxun Zhang, "A Review of Recurrent Neural Networks: LSTM Cells and Network Architectures, ", Neural Computation, 2019, 31:7.

[3] M. Choi, T. Matsumoto, M. Jung, and J. Tani. Generating goal-directed visuomotor plans based on learning using a predictive coding type deep visuomotor recurrent neural network model. arXiv abs/1803.02578, 2018. URL http://arxiv.org/abs/1803.02578.

[4] M. Jung, T. Matsumoto, J. Tani, "Goal-Directed Behavior under Variational Predictive Coding: Dynamic Organization of Visual Attention and Working", IEEE International Workshop on Intelligent Robots and Systems, 2019

[5] J. F. Queißer, M. Jung, T. Matsumoto, J. Tani, "Emergence of Content-Agnostic Information Processing by a Robot Using Active Inference, Visual Attention, Working Memory, and Planning". Neural Computation 2021; 33 (9): 2353–2407.

[6] N. Kau. Stanford pupper, "A low-cost agile quadruped robot for benchmarking and education", 2022, arXiv:2110.00736 [cs.RO].

５．主な発表論文等

〔雑誌論文〕　計1件（うち査読付論文　1件／うち国際共著　1件／うちオープンアクセス　1件）

１．著者名	４．巻
Queisser Jeffrey Frederic、Jung Minju、Matsumoto Takazumi、Tani Jun	33
２．論文標題	**５．発行年**
Emergence of Content-Agnostic Information Processing by a Robot Using Active Inference, Visual Attention, Working Memory, and Planning	2021年
３．雑誌名	**６．最初と最後の頁**
Neural Computation	2353～2407

掲載論文のDOI（デジタルオブジェクト識別子）	査読の有無
10.1162/neco_a_01412	有
オープンアクセス	**国際共著**
オープンアクセスとしている（また、その予定である）	該当する

〔学会発表〕　計0件

〔図書〕　計0件

〔産業財産権〕

〔その他〕

Publication Webpage
https://direct.mit.edu/neco/article/33/9/2353/102624/Emergence-of-Content-Agnostic-Information
Animation, explanation of model
https://groups.oist.jp/cnru/movies

６．研究組織

	氏名 （ローマ字氏名） （研究者番号）	所属研究機関・部局・職 （機関番号）	備考

７．科研費を使用して開催した国際研究集会

〔国際研究集会〕　計0件

８．本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関