Integration of Multimodal Categorization and Reinforcement Learning for Robot Decision-Making

Kazuki Miyazawa, Tatsuya Aoki, Chie Hieida, Kensuke Iwata, Tomoaki Nakamura, and Takayuki Nagai

***Abstract*— It is very interesting to contemplate question how humans learn, plan, and decide on behaviors. In addition, the determination of the correlation between high level functions, such as action planning/decision, language understanding, and thinking is an open problem. To solve these problems, we propose a framework to simultaneously learn concepts, actions, and language. This can be achieved by integrating multimodal categorization by using multilayered multimodal latent Dirichlet allocation (mMLDA) and reinforcement learning.**

I. Introduction

The understanding of the real-world is important for robots to perform actions there. Researchers have been studying the understanding of the real-world by robots, including language, from the viewpoint of symbol emergence in robotics [1]. In some studies, the core idea is that the categorization of multimodal information acquired through robotic experiences leads to a bottom-up concept formation [2],[3]. By using this concept acquired in the bottom-up manner, a robot can predict unobserved information across the modality from the actual observations. We strongly agree that the prediction is the basis of understanding. In addition, we assume that language can be understood and generated in the same framework. How do the robots experience and collect the data in the first place? Furthermore, how do the robots use the acquired knowledge to decide their own actions? We contemplate that answering these questions is important for the framework to integrate concepts and language acquisition, real-world understanding, action planning, etc., from motion learning by trial and error.

The proposal of this paper is based on the multilayered multimodal latent Dirichlet allocation (mMLDA), which is an extension of the MLDA, and expresses multiple concepts and the relationship between the concepts stochastically [4]. In the current study, we propose a framework to realize concept learning, knowledge acquisition, language learning, and action decision by linking mMLDA and reinforcement learning. Next, we show that it is possible to learn appropriate actions through the proposed framework based on the mMLDA and reinforcement learning.

*This work was supported by JST, CREST (JPMJCR15E3).

[1]Kazuki Miyazawa, Tatsuya Aoki, Chie Hieida, Kensuke Iwata, Tomoaki Nakamura, and Takayuki Nagai are with the Department of Mechanical Engineering and Intelligent Systems, University of Electro-Communications, 1-5-1 Chofugaoka, Chofushi, Tokyo 182-8585, Japan {miyazawa, aoki, hchie, iwata, naka_t}@apple.ee.uec.ac.jp, tnagai@ee.uec.ac.jp.

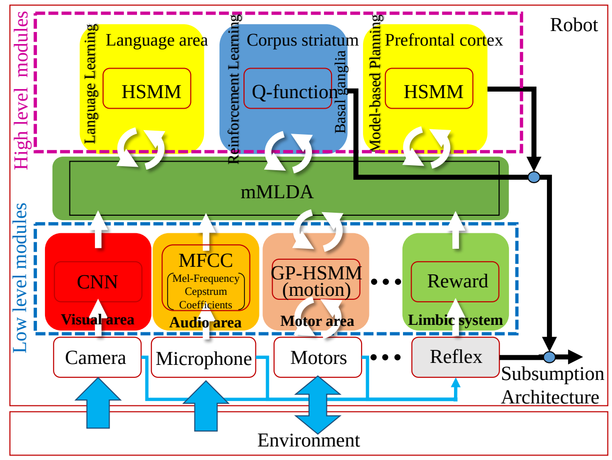

Fig. 1. Overview of the proposed integrated model.

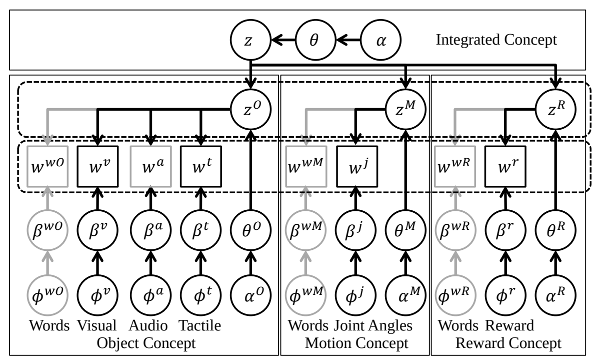

Fig. 2. Graphical model of the mMLDA used in this paper.

Until now, there are few cases, in which behavior learning through reinforcement learning, concept acquisition, and even language learning are handled in a unified manner. The main contribution of this research is that a unified framework for realizing such a behavior and knowledge acquisition loop is assembled around mMLDA.

II. Proposed Model

Figure 1 shows an overview of the proposed integrated model consisting mainly of mMLDA and a combination of several modules. In the following, we describe mMLDA which is the core of this model followed by the details of the proposed integrated model.

A. multilayerd multimodal LDA

mMLDA has a hierarchical structure with multiple MLDAs that express subordinate concepts, such as objects, motions, and places, in the lower layer and an MLDA that integrates them in the upper layer. The mMLDA allows the categorization of each sensor-motor signal, such as movement, place, object,and person, and simultaneous learning of the relationship between these concepts unsupervised [5]. The graphical model of mMLDA used in the experiment is illustrated in Fig. 2, in which z is a category representing an integrated concept, and z^O, z^M, and z^R are objects, motion, and reward categories corresponding to subordinate concepts, respectively.

B. Integrated Model

The proposed integrated model consists of three parts.

The first part is responsible for learning concepts. It is composed of low-level modules and the mMLDA. The low-level modules transform sensor-motor signals into feature vectors. Then, the robot learns concepts by categorizing these feature vectors using the mMLDA. In addition, primitive actions used in the proposed model are realized through motion learning by using the Gaussian process hidden semi-Markov model (GP-HSMM) [6].

The second part involves with high-level cognition. It is composed of three modules, which receive inputs from the mMLDA. These high-level modules include hidden semi-Markov model (HSMM) handling language (grammar), Q-learning (Q-function) for model-free action learning, and another HSMM involving action planning. Regarding language, words are grounded on real-world information (concepts) through mMLDA. By applying syntactic information representing by the HSMM to this word information, the proposed model is able to generate sentences according to the observations. In addition, by decomposing sentences using the HSMM (syntactic information), and by predicting real-world information using the mMLDA, the robot can understand the meaning of sentences. In terms of action selection, the reinforcement learning module treats concepts, which are learned by the mMLDA, as a state space and selects actions that maximize expected rewards. In the action planning part (model-based planning), long-term actions can be planned by using Viterbi algorithm based on the HSMM, which represents sequences of actions and states.

The third part integrates an action generated by the reflex module, Q-learning (model-free) module, and model-based planning by the HSMM. Therefore it consists of three behavioral modules connecting through the subsumption architecture. By using the subsumption architecture, the robot is able to take actions in various levels by connecting reflexes, model-free actions, and actions through the planning using the HSMM.

With this framework, robots can learn from motion concepts to language in a unified framework. In this paper, we aim to realize the above framework by considering integration of each part of this integrated model.

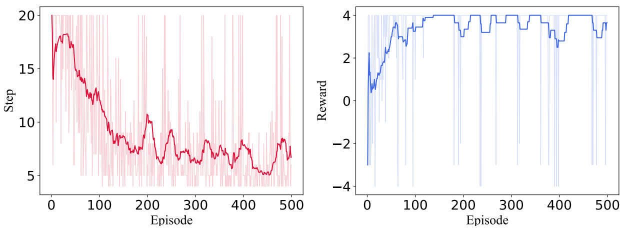

Fig. 3. Reward and number of steps for each episode.

III. Experiment

As the first step , the robot learned in a simple environment. In this environment, a state transition matrix and a reward function are defined. In one state, the robot acquires multimodal information of an object. One object is randomly selected from a certain category which is linked to a particular state. There are 134 items of 8 categories in total, which are acquired by a real robot. The robot can select one action from a set of four actions. In such an environment, the robot learns to transition state from the start to the goal. Figure 3 shows the learning result. As the episode progresses, the number of steps per episode decreases. Also, the reward has increased. These show that the robot can reach the goal by learning actions in a simple simulation environment.

IV. Conclusions

This paper proposed an integrated model for understanding real-world and selecting actions based on understanding by robots. Furthermore, a simulation environment was designed in order to verify the proposed model, and a simple experiment was conducted. In the future, we aim to complete the integrated model and to implement the model on a real robot.

References

[1] T. Taniguchia, T. Nagai, T. Nakamura, N. Iwahashi, T. Ogata, and H. Asoh, "Symbol Emergence in Robotics: a Survey, " *Advanced Robotics*, vol.30, 11-12, pp.706–728, 2016.

[2] T. Nakamura, T. Nagai, and N. Iwahashi, "Grounding of Word Meanings in Multimodal Concepts Using LDA, " *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, pp.3943–3948, 2009.

[3] J. Nishihara, T. Nakamura, T. Nagai, "Online Algorithm for Robots to Learn Object Concepts and Language Model, " *IEEE Transactions on Cognitive and Developmental Systems*, Issue 99, 2016.

[4] M.Fadlil, K.Ikeda, K.Abe, T.Nakamura, and T.Nagai, "Integrated Concept of Objects and Human Motions Based on Multi-layered Multimodal LDA," *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, pp. 2256–2263, 2013.

[5] M.Attamimi, Y.Ando, T.Nakamura, T.Nagai, D.Mochihashi, I.Kobayashi, and H.Asoh, "Learning Word Meanings and Grammar for Verbalization of Daily Life Activities Using Multilayered Multimodal Latent Dirichlet Allocation and Bayesian Hidden Markov Models," *Advanced Robotics*, 30(11-12), pp.806–824, 2016.

[6] T.Nakamura, K.Iwata, T.Nagai, D.Mochihashi, I.Kobayashi, A.Hideki, and M.Kaneko, "Continuous Motion Segmentation Based on Reference Point Dependent GP-HSMM," IROS 2016 Workshop on Machine Learning Methods for High-Level Cognitive Capabilities in Robotics, 2016.