Advanced Robotics, Vol. 20, No. 10, pp. 1127–1145 (2006)

Also available online - www.brill.nl/ar

Full paper

Lexicon acquisition based on object-oriented behavior learning

SHINYA TAKAMUKU [1], YASUTAKE TAKAHASHI [2] and MINORU ASADA [2,*]

[1] *HANDAI Frontier Research Center, Osaka University, Osaka, Japan*
[2] *Department of Adaptive Machine Systems, Graduate School of Engineering, Osaka University, Yamadagaoka 2-1, Suita, Osaka, 565-0871, Japan*

Received 7 April 2006; accepted 31 May 2006

Abstract—Studies on lexicon acquisition systems are gaining attention in the search for a natural human–robot interface and a test environment to model the infant lexicon acquisition process. Although various lexicon acquisition systems that ground words to sensory experience have been developed, existing systems have clear limitations on the ability to autonomously associate words to objects. This limitation is due to the fact that categories for words are formed in a passive manner, either by teaching of caregivers or finding similarities in visual features. This paper presents a system for lexicon acquisition through behavior learning. Based on a modified multi-module reinforcement learning system, the robot is able to automatically associate words to objects with various visual features based on similarities in affordances or in functions. The system was implemented on a mobile robot acquiring a lexicon related to different rolling preferences. The experimental results are given and future issues are discussed.

Keywords: Active sensing; affordance; categorization; lexicon acquisition; multi-module reinforcement learning system.

1. INTRODUCTION

Various robots intended to coexist with human users in everyday life have been developed [1, 2]. Users expect rich communication with these robots. However, current domestic robots can just refer to fixed types of objects or behaviors in their language communication. In order to enrich their communication, the design of the ability to acquire and share a lexicon with users is a fundamental issue to be developed. Building robots that acquire a lexicon also gives insights on language acquisition in human children. The idea to consider words directly grounded in a physical embodiment as the basis for acquiring more abstract words and syntax, that

*To whom correspondence should be addressed. E-mail: asada@ams.eng.osaka-u.ac.jp

has been suggested in cognitive linguistics [3], is now supported by several robot experiments. For example, Harnad *et al.* showed how symbol manipulation can be constrained by directly grounding basic symbols in cognitive representations such as categories [4]. Sugita and Tani showed how compositional syntax could emerge from an attempt to acquire a generalized correspondence between word sequences and sensorimotor flows [5].

Several robotics research projects have investigates the problem of acquiring a new lexicon from sensory experience. Steels and Kaplan [6] investigated lexicon acquisition in pet robots, and showed that intense interaction between a learner and a caregiver (e.g., the feedback from the caregiver about the utterance by the learner is effective for lexicon acquisition) is essential for acquiring object categories for communication. Roy [7] developed a system that learns words from untranscribed acoustic and video inputs, and showed that the speech segmentation process can be accelerated by considering mutual information between speech segments and visual categories of objects that co-occur. Although these studies have done much to advance our understanding of the lexicon acquisition problem from the learner's point of view, their systems have clear limitations on the ability to autonomously associate words to objects. In their system, categories for word acquisition are mostly formed in a passive manner, either by finding similarities in visual features or the teaching of caregivers. Research on computer vision has shown that such passive approaches have many difficulties in object recognition [8] that seems essential for lexicon acquisition. On the other hand, some psychological studies on word association of human children indicate the importance of learners' actions. For example, Nelson *et al.* [9] showed that even 2-year-old children can associate names to new objects in accordance with the functions when the function of the object is apparent to them. This kind of word association seems effective, since our environment is full of artifacts created for particular functional purposes. Another study by Kobayashi [10] showed that a caregiver's action on objects influences children's inferences about word meanings. The result of these studies can be interpreted as that those children associate words to new objects in accordance to the affordances [11] of the objects. Acquiring words related to affordance is important for children to know the appropriate behaviors for the objects. Despite the expected effect on lexicon acquisition, existing robot systems of lexicon acquisition are not capable of such word associations based on the learner's behaviors and object functions.

This paper focuses on the issue of how lexicon learners can acquire words and associate them to different objects based on similarities in functions or in affordances. Following the discussion of symbol grounding [12], we formulate this issue as a task to acquire object categories related to function or affordance and associate appropriate labels to them (we assume one-to-one correspondence between a label and a category when they are related). With regard to the issue of obtaining object categories, several researchers have indicated how effective the learner's actions are. Fitzpatrick and Metta [13] showed how object affordance can be explored by rolling the objects and Ogata *et al.* [14] adopted RNNPB with multi-modal sensory input to improve the object recognition performance. Their system succeeded in forming representations suitable for object recognition. However, the variety of categories obtained by their approaches seems limited due to the fact that actions performed by their system are fixed *a priori*. Actions for different functions or affordances vary and it seems almost impossible for the designers to prepare all kinds of behaviors in advance. Thus, we propose a system that forms categories by learning and identifying specific behaviors that can be performed with the objects. In order to acquire the word 'cylinder', for example, the learner acquires the behavior to face the lateral aspect of the object and roll it. The behaviors with the objects are learned and identified by a multi-module learning system [15] modified for reinforcement learning. We show a method to associate sensorimotor concepts represented in the multi-module learning system with labels without assuming the labels to be given simultaneously with the activation of those learning modules.

The paper is organized as follows. First, we explain our approach of lexicon acquisition based on behavior learning. Then, experimental results show the validity of our approach. Finally, future issues are discussed and conclusions are given.

2. LEXICON ACQUISITION BASED ON BEHAVIOR LEARNING

2.1. Basic idea

In our study, we address the task of learning labels for objects given by a caregiver (Fig. 1) and associating appropriate labels to new objects based on similarities in functions or in affordances. The learner is considered successful when it answers appropriate labels for given objects in view. The actual lexicon acquisition task involves various difficult tasks such as extracting individual words from continuous utterances and attending to the object intended by the caregiver. However, in order to focus on the label association issue, we do not deal with these tasks here.

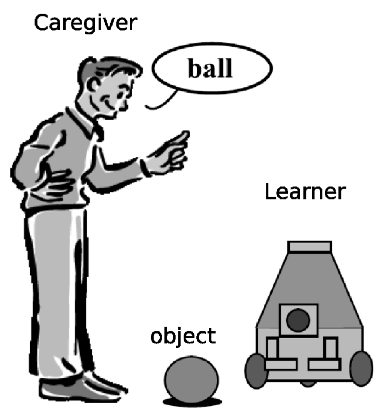


Figure 1. Lexicon acquisition environment.

The lexicon acquisition process of the system consists of four sub-processes:

A Learn object-oriented behaviors.

B Categorize (photometric) visual feature space based on object-oriented behaviors.

C Categorize (photometric) visual feature space based on labels given by the caregiver.

D Learn the correspondence between object-oriented behaviors and labels.

A is an process where the learner learns the behavior for given objects. The behaviors specified for a certain affordance or function of objects such as cutting behavior for the label 'scissors' is referred to as object-oriented behaviors. During this behavior learning process, the learner builds models of object behavior caused by the robot's behavior which can later be utilized to identify behaviors that the objects afford. **B** is a process of categorizing the visual feature space based on the object behavior model acquired in process **A**. This process enables the learner to identify the behavior for the object in view without physical interaction. As for visual features, we adopt photometric features such as color histograms or edge histograms. Although visual categories related to object-oriented behavior are acquired by **A** and **B**, the learner still needs to associate labels to the object-oriented behaviors in order to communicate about them. This association is not an easy task since categories of behaviors are gradually obtained and not all labels are related to those categories. In order to solve the association task, the learner categorizes the visual feature space based on the labels (**C**) and learns the association by matching the categories of behaviors and of labels (**D**).

2.2. System overview

The overview of our system for lexicon acquisition based on behavior learning is shown in Fig. 2. Inputs to the system are classified into three types, i.e., photometric features for object identification, state variables for controlling the objects (such as current position and orientation of objects) and labels given by the caregiver. The idea to separate visual information of objects into two main categories, i.e., photometric features and state variables, comes from the study by Milner and Goodale [16], which indicates that visual information for identification and control are handled by different processes in the human brain. Similar separation of visual information is also adopted in the system of Fitzpatrick and Metta [13] to learn the affordance of objects. As previously mentioned, process **A** is realized by a multi-module reinforcement learning system taking state variables as input. Each module of the multi-module reinforcement learning system corresponds to a particular object-oriented behavior. As **A** performs, **B** easily performs by categorizing the photometric feature space according to the identification of object-oriented behaviors. On the other hand, the system also categorizes the photometric feature space according to the labels given by the caregiver (**C**). This means that two different categorizations are performed on the same photometric feature space. Adap-

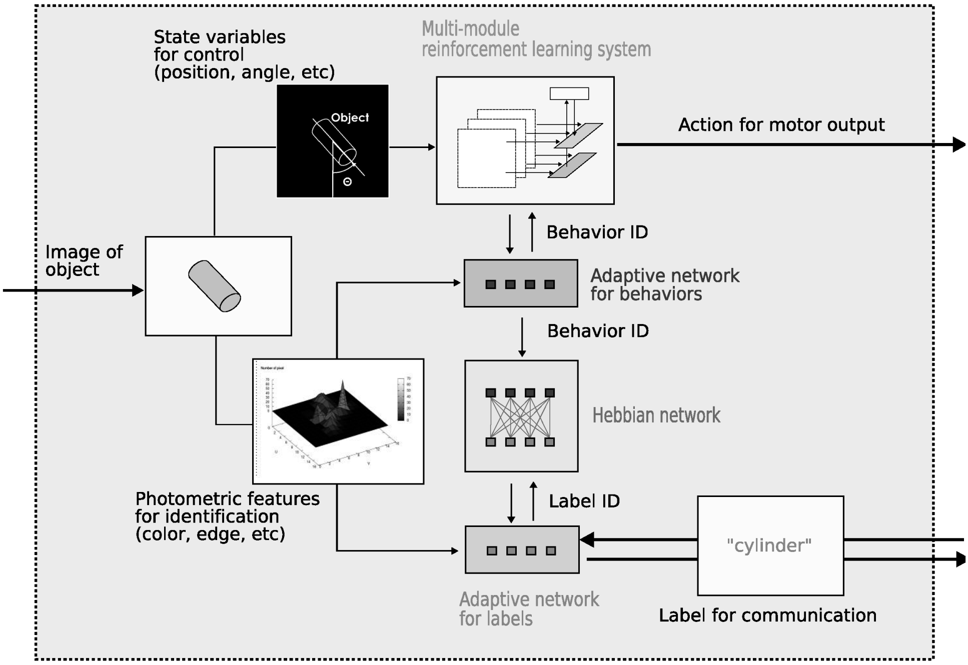


Figure 2. Sketch of the system for lexicon acquirition based on behavior learning.

tive networks are adopted for these categorizations. Finally, process **D** is realized by learning the correspondence between object-oriented behaviors and labels by a Hebbian network, connecting a behavior and a label whose category is selected for same photometric features. Details on each learning system are explained in the following sections. Although the different learning processes are explained one at a time, the system is designed to run all the learning processes in parallel. Scheduling of lexicon acquisition is introduced only for the simplicity of showing the results.

2.3. Learning and identifying object-oriented behaviors

We adopt the multi-module reinforcement learning system shown in Fig. 3 for learning and identifying object-oriented behaviors. The system consists of multiple learning modules, each of which consists of a predictor and a planner, and a gate to select the appropriate module based on reliability representing the accuracy of the goal-directed state transition prediction by each learning module. In the current system, one-to-one correspondence between learning modules and object-oriented behaviors is assumed. The system learns object-oriented behaviors as state-action mappings and identifies the object-oriented behaviors based on the reliabilities of the learning modules. If no learning module with sufficient reliability is found, a new learning module is assigned to learn a new object-oriented behavior. We chose Q-learning [17] associated with state transition models and reward prediction models as the reinforcement learning method. The method can acquire the behaviors with relatively little prior knowledge of the task or the environment.

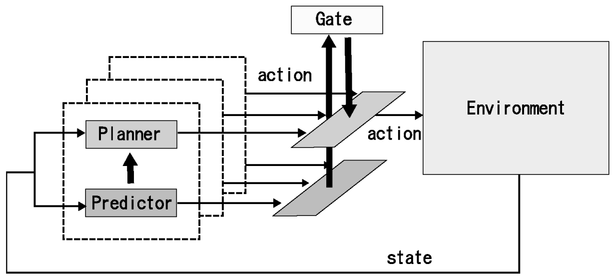


Figure 3. Multi-module reinforcement learning system.

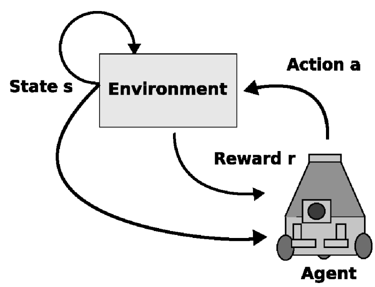


Figure 4. Basic model of learner–environment interaction in reinforcement learning.

2.3.1. Behavior learning. In general reinforcement learning, interaction between the learner and environment is modeled as shown in Fig. 4. In every time step, the learner obtains a discrete representation of the current state $s_t \in \mathcal{S}$ (S is the set of possible states) and selects an action $a_t \in \mathcal{A}(s_t)$ ($\mathcal{A}(s_t)$ is the set of possible actions at state s_t). Then the next state $s_{t+1} \in \mathcal{S}$ and reward $r_{t+1} \in \mathcal{R}$ are determined, depending only on the state and action selected by the learner. The task of reinforcement learning is to choose a policy $a = f(s)$ which maximizes the decaying sum of reward:

$$\sum_{n=0}^{\infty} \gamma^n r_{t+n}, \tag{1}$$

where γ is the decay factor ($0 < \gamma < 1$).

The predictor in each learning module builds two models of the environment through the interactions with the objects. One is a state transition model, which is a set of probabilities of all state transitions:

$$\hat{\mathcal{P}}^a_{ss'} = Pr\{s_{t+1} = s' | s_t = s, a_t = a\}. \tag{2}$$

Another is a reward prediction model, which is a set of expected reward values for all state-action sets:

$$\hat{\mathcal{R}}^a_s = \sum_{s'} E\{r_{t+1} | s_t = s, a_t = a\}. \tag{3}$$

As the state transition model and reward prediction model are built, the planner for each learning module calculates the action value $Q(s, a)$ (a set of expected decaying reward sums for every state-action set) by simulating the learning process offline. The offline learning process takes place at the end of each trial of interaction with the environment. All the state transition data from a continuing interaction are assumed to come from the same object:

$$Q(s, a) = \sum_{s'} \hat{\mathcal{P}}^a_{ss'} \big(\hat{\mathcal{R}}^a_s + \gamma \max_{a'} Q(s', a') \big). \tag{4}$$

When $Q(s, a)$ converges, the rational policy is given as:

$$f(s) = arg \max_{a \in \mathcal{A}(s)} Q(s, a). \tag{5}$$

In order to collect data for building the models, the learner selects actions for the objects by first specifying the best-matching learning module and then selecting the action within the learning module. A learning module is usually selected based on its reliability (defined in the next section). However, in the initial phase of interaction where no state transition is observed, the learner selects the learning module corresponding to the best-matching object-oriented behavior for the given photometric feature of the object. The adaptive network is used for this selection and discussed in detail in Section 2.4. Once a learning module is selected, the learner selects the action based on the ϵ-greedy method. For more effective identification, the learner can choose the action which gives the most different state value transition among the learning modules.

2.3.2. Behavior identification. We calculated the reliability for each learning module as:

$$rel = DQ_{\text{threshold}} - \Delta Q(s_t, a_t), \tag{6}$$

in which $\Delta Q(s_t, a_t)$ is the action value error defined as:

$$\Delta Q(s_t, a_t) = \Big| r_{t+1} + \gamma \max_{a_{t+1}} Q(s_{t+1}, a_{t+1}) - Q(s_t, a_t) \Big|. \tag{7}$$

This means that when the learner encounters an object, it identifies the object-oriented behavior by choosing the learning module with the most accurate prediction of the state value transition. We adopted this criterion instead of common state transition predictions in order to put more importance on task-oriented state changes. When action value errors of all existing behavior modules exceed a predefined threshold $DQ_{\text{threshold}}$, all reliabilities have negative values and a new learning module is assigned to learn the new behavior. The threshold was balanced to avoid redundant computation, but to acquire enough categories for object handling.

2.4. Categorization of photometric feature space

We adopted adaptive networks, modified radial basis function neural networks, for the categorization of the photometric feature space. The method is more tolerant to noise than simple nearest-neighbor methods and considered as a model for categorical perception of biological systems [18]. The method is also adopted by Steels and Belpaeme [19] to acquire categories for words.

An adaptive network consists of a set of networks, each of which corresponds to a single category. The network of each category outputs a scalar value $y(\boldsymbol{x})$ for an input of photometric feature vector $\boldsymbol{x}$ and the category with the largest network output is selected as the best-matching category. Each category network consists of locally reactive units whose responses are greatest at a central value $\boldsymbol{m}$ and decay exponentially around this central value. The response of local unit j is:

$$z_j(\boldsymbol{x}) = \mathrm{e}^{-\frac{1}{2}\left(\frac{d(\boldsymbol{x},\boldsymbol{m}_j)}{\sigma}\right)^2}, \tag{8}$$

where $d(\boldsymbol{x}, \boldsymbol{m}_j)$ is the distance between central value $\boldsymbol{m}_j$ and input vector $\boldsymbol{x}$. Constant values are set for variance σ. The output of the network for category k is:

$$y_k(\boldsymbol{x}) = \sum_{j=1}^{J} w_{kj} z_j(\boldsymbol{x}), \tag{9}$$

where J is the number of locally reactive units. A schematic of the calculation of the output for category networks is shown in Fig. 5.

The system categorizes the feature space by modifying the network weights and adding new local units. When a training data set (photometric features with a label

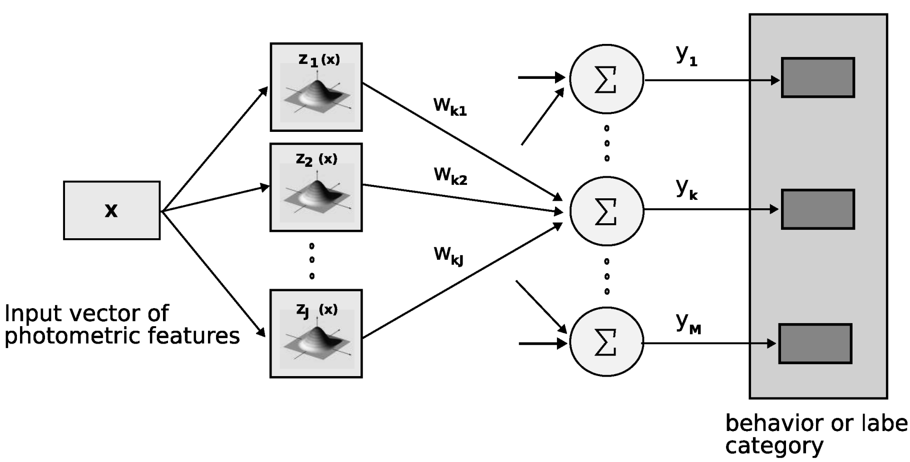


Figure 5. Schematic of the adaptive network for behavior or label. The category with the largest network output is selected.

or photometric features with a behavior) is given and categorization is successful, the network weight of the matching category k is increased as:

$$w_{kj} \longleftarrow w_{kj} + \beta z_j(\boldsymbol{x}). \tag{10}$$

If the categorization is not successful, a new local unit with a central value of the unsuccessfully categorized input is assigned and the weight to the correct category is set to a predefined value of w_0 (other weights are set to 0). When new categories are added, such as the case when new behavior modules are assigned or new labels are taught, a new category network is assigned with local units capable of classifying inputs for those new categories. The weights of all local units of all categories are decreased whenever the network is modified:

$$w_{kj} \longleftarrow \alpha w_{kj} \quad (\text{for } k = 1, \ldots, K, \text{ and } j = 1, \ldots, J). \tag{11}$$

This process enables the system to forget unused categories and keep adapted to changes of the environment. $\beta \in [0, 1]$ is a learning rate and $\alpha \in [0, 1]$ is a decay factor.

2.5. Learning the relation between object-oriented behaviors and labels

We adopt a Hebbian network for learning the relation between object-oriented behaviors and labels. The Hebbian network connects the nodes of object-oriented behaviors and the nodes of labels, and modifies the weight of the network so that corresponding nodes are strongly connected. Whenever the learner captures an object, the learner extracts the photometric features of the object and selects the best-matching behavior and the label utilizing both adaptive networks. In the case where behavior m and label n are selected, the weight between those two nodes W_{mn} is increased as:

$$W_{mn} \longleftarrow W_{mn} + \delta_{\text{inc}}, \tag{12}$$

and weights of other connections to node m or n are decreased with δ_{inh} to disregard labels unrelated to object-oriented behaviors:

$$W_{st} \longleftarrow W_{st} - \delta_{\text{inh}} \quad (s = m \text{ or } t = n), \tag{13}$$

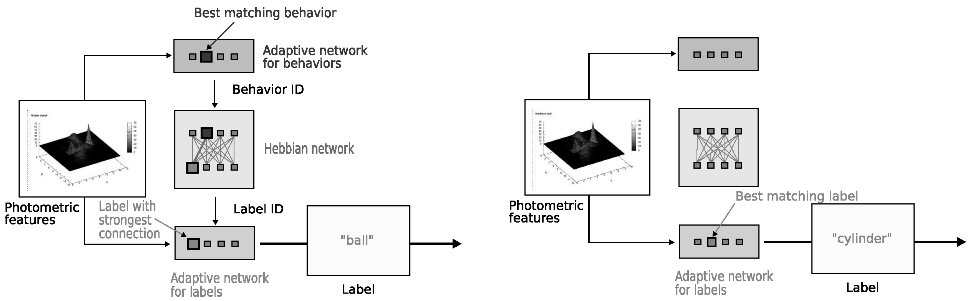
δ_{inc} and δ_{dec} have positive values smaller than 1. When new nodes are added, the weights are initialized to 0s.

2.6. Label association policy

The process of selecting labels for objects in view is shown in the following. The process includes label selection based on the similarity of behaviors that the object affords.

(i) Extract photometric features of objects.

(ii.A) Select the best-matching behavior utilizing the adaptive network for behaviors.



(a) label guessing by behavior category (b) label guessing by label category

Figure 6. Label guessing process.

(ii.B) Select the best-matching label utilizing the adaptive network for labels.

(iii) Compare the outputs of the selected behavior category network and the selected label category network.

(iv) If the output of the behavior category network is larger than that of the label category network and there exists a label unit in the Hebbian network which is connected to the unit of the selected behavior with a weight higher than a predefined value W_{limit}, go to step (v). Else, go to step (vi).

(v) Output the label with the strongest connection to the behavior selected in step (ii.A).

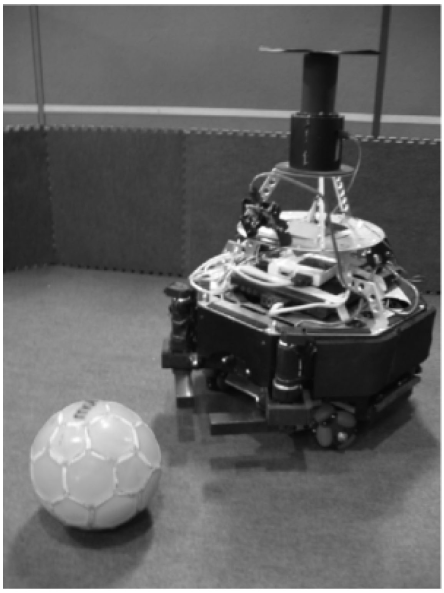
(vi) Output the label selected in step (ii.B).

Figure 6a and 6b shows the information flow for selecting the label in steps (v) and (vi). The learner outputs labels according to behavior when it knows the behavior for the object and has acquired the label for it. A learner with a low W_{limit} associates labels more often with behavior categories, whereas a learner with a high W_{limit} is more conservative about association.

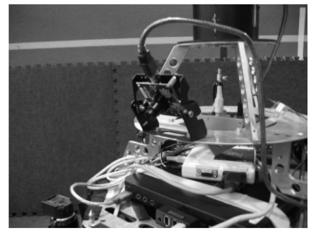
3. EXPERIMENTS

3.1. Task

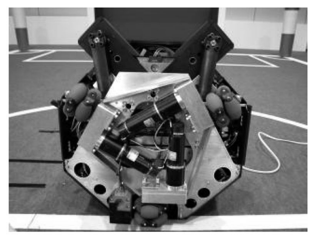
To show the validity of our system, we implemented the system on a mobile robot (Fig. 7) that learns a lexicon about objects with different rolling preferences. We used the objects shown in Fig. 8, and gave the labels ‘ball’, ‘box’, ‘cylinder’ and ‘car’. After presenting the objects shown in Fig. 8a–d paired with the labels corresponding to them, we introduced the new objects shown in Fig. 8e–h without the corresponding labels. Note that some components of an object category are very similar to components in other object categories (e.g., the new ball is more similar to the new cylinder captured from the cap side than to the old ball). Categorizing a unseen object based only on similarities of visual feature is useless in this case.



(a) robot

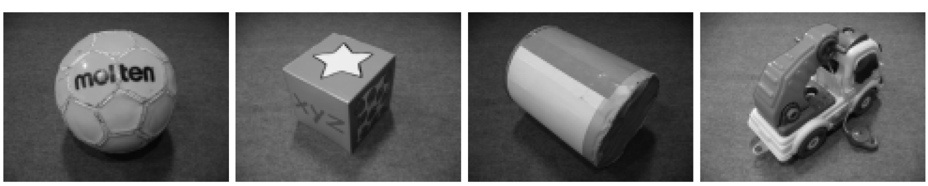


(b) CCD camera

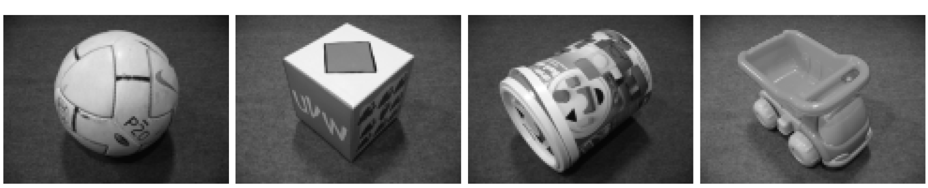


(c) omni-directional wheels and kicking device

Figure 7. Mobile robot used in the experiment.



(a) ball (b) box (c) cylinder (d) car



(e) new ball (f) new box (g) new cylinder (h) new car

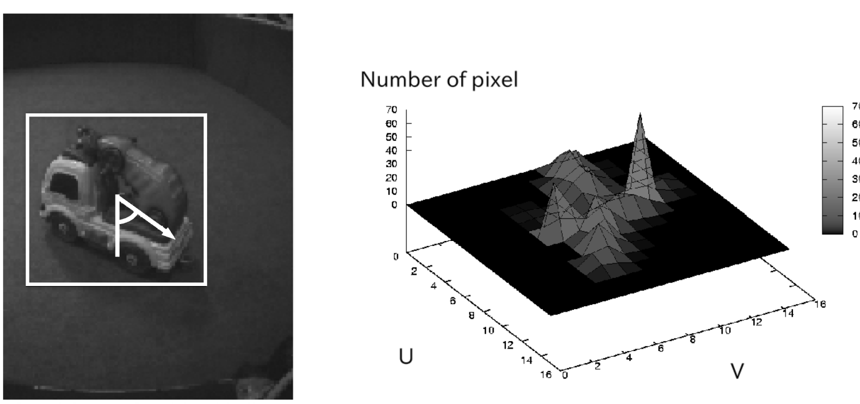
Figure 8. Objects used in the experiment.

However, if the system successfully learned the correspondence between object-oriented behaviors and labels, it should be able to associate the labels to the new objects.

3.2. Learner setup

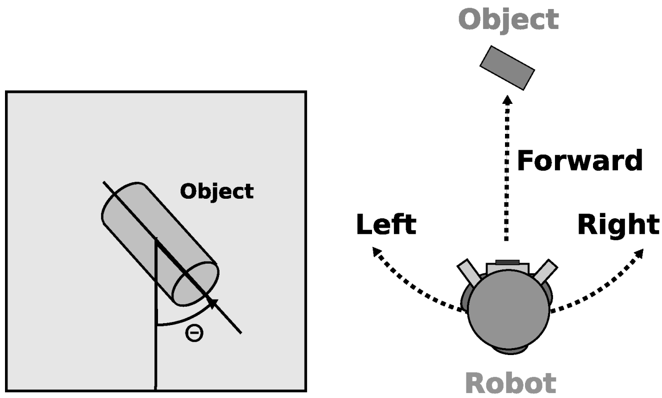
3.2.1. Hardware and environment. The mobile robot utilized in the experiment is shown in Fig. 7. The robot is equipped with a CCD camera (FireFly2 provided from Point Grey Reseach) with a fixed view angle (Fig. 7b) and a kicking device in front with two arms that rotate independently in the horizontal plane (Fig. 7a and 7c). The robot is able to move in any direction in the horizontal plane by the use of omni-directional wheels (Fig. 7c). The robot interacted with the objects in a 3 m $\times$ 3 m space covered with a green color.

3.2.2. Visual information processing. A 240 (H) $\times$ 320 (V) size image of objects was captured by a CCD camera and sent to a laptop PC through an IEEE1394 interface. The object region is extracted assuming that the green color regions are background regions. The direction of the principal axis is then calculated from the object region by principal component analysis as shown in Fig. 9a and utilized as a state variable to run the object rolling behavior. On the other hand, we adopt YUV color space and the UV color histogram as photometric features. UV space is quantized into 16 $\times$ 16, so the color histogram is a 256-dimensional vector representing the frequency of quantized colors in the UV space. Figure 9b shows an example of the UV space histogram. The system uses χ^2-divergence as the distance metric for categorization. The χ^2-divergence of two photometric feature vectors $\boldsymbol{a}$



(a) extracted image (b) color histogram

Figure 9. Color histogram of an object.



(a) state (b) action

Figure 10. State and action for the task.

and $\boldsymbol{b}$ is calculated as follows:

$$d(\boldsymbol{a}, \boldsymbol{b}) = \chi^2(\boldsymbol{a}, \boldsymbol{b}) = \sum_i \frac{(a_i - b_i)^2}{a_i + b_i}. \tag{14}$$

3.2.3. Learning system. The state space for the learning modules of the multi-module reinforcement learning system consists of the direction of the principal axis of the object $\theta \in [-90, 90]$ as shown in Fig. 10a. The state space is quantized into 7 and another state is added to represent a case when the principal axis is uncertain. The learner is able to choose from three actions (Fig. 10b), i.e., kicking the object forward, and moving clockwise and anti-clockwise around the object. Finally, a reward whose value is proportional to the moving distance of the object is given to the learner.

Parameters for the multi-module reinforcement learning system, adaptive networks and Hebbian network are shown in Table 1.

3.3. Experiment on object-oriented behavior learning and identification

The learner acquired the rolling behavior for the objects shown in Fig. 8a–d. The assignments of new learning modules and changes of reliability for each module are shown in Fig. 11, where curves with different line types correspond to different learning modules. The horizontal axis indicates the number of trials of interaction with the objects. One trial finished when the learner kicked the object 5 times or when the learner lost the object. This indicates that if the object was to roll, approximately one successful kick is included in each trial. The learner experienced the objects in the fixed order of ball, box, cylinder and car. They interacted with the objects for five trials in the first 20 trials, for three trials in the next 12 trials, etc., the object was switched after each trial. We introduced the scheduling to the behavior learning process, since if the learning modules were too immature, they

Table 1.
Parameters for the learning system

Multi-module reinforcement learning system	
decay factor γ	0.3
threshold for action value error $DQ_{\text{threshold}}$	1.0
Adaptive network for behavior	
learning rate β	1.0
decay factor α	0.99
initial value of network weight w_0	0.5
Adaptive network for label	
learning rate β	1.0
decay factor α	0.99
initial value of network weight w_0	0.5
Hebbian network	
increasing value of weight δ_{inc}	0.1
decreasing value of weight δ_{dec}	0.1
limit value of weight considered W_{limit}	0.0

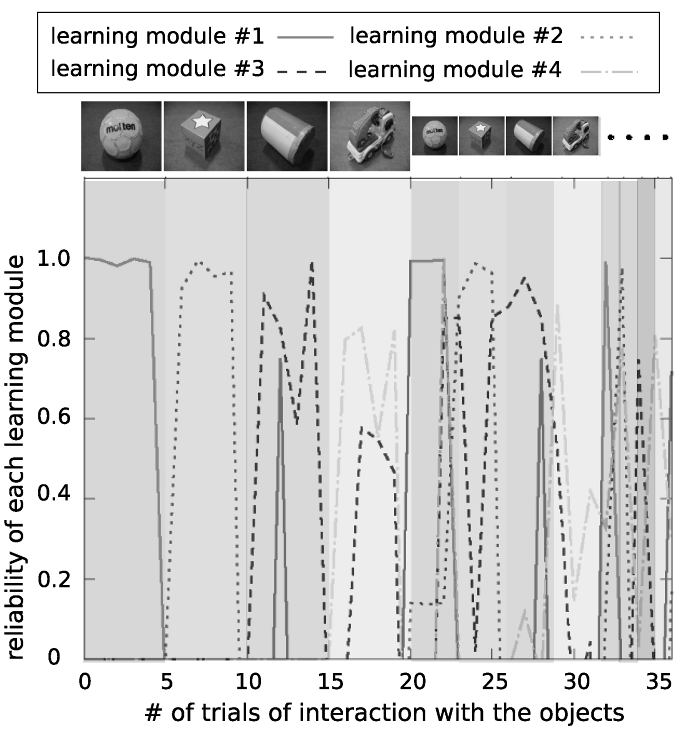


Figure 11. Action value error of each learning module while learning object-oriented behaviors.

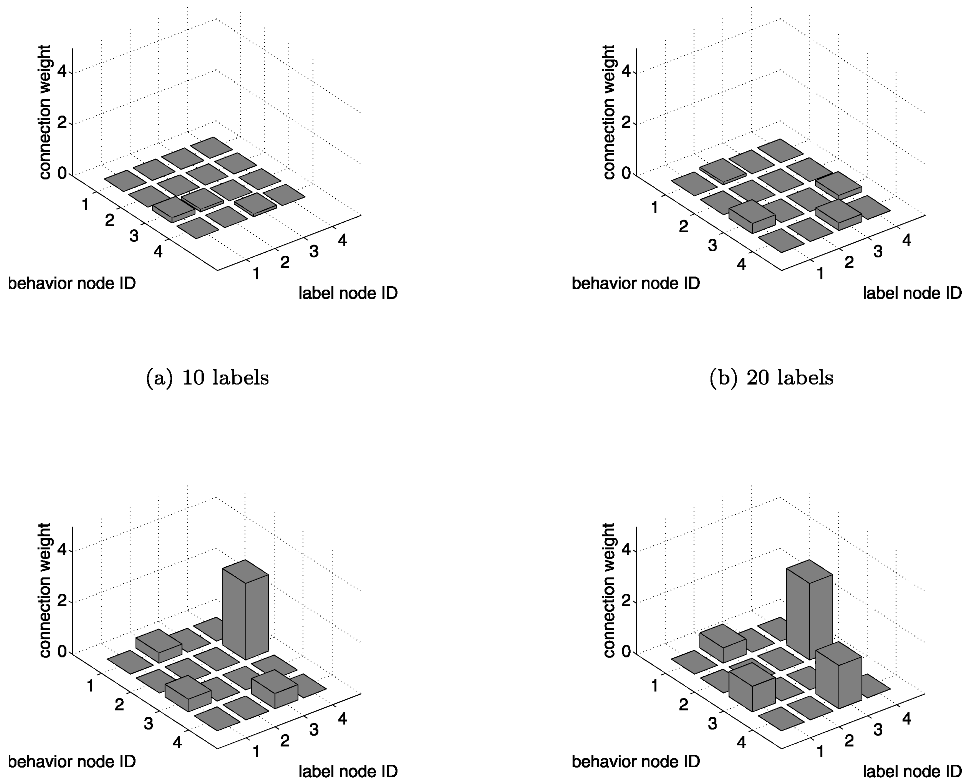
would give large action value errors, even though they were interacting with the corresponding objects. Taniguchi and Sawaragi [20] discuss this topic in detail. As shown in Fig. 11, every time an unfamiliar object was introduced, the action value error exceeded the predefined limit $DQ_{\text{threshold}}(s, a) = 1.0$ (which makes the reliability less than 0) and a new learning module was assigned. Figure 11 also shows that the same learning module is selected for each object. This shows that the system successfully acquired the set of learning modules that can be utilized to identify the objects based on object-oriented behaviors.

3.4. Experiments on learning the relation between object-oriented behaviors and labels

After the learner acquired the object-oriented behaviors and categorized the photometric feature space based on the behaviors, the caregiver taught the labels of the objects which the learner learned to roll (Fig. 8a–d) in random order. When the label was given, the system categorized the photometric feature space by assigning simultaneously given photometric features to the category of the given label. As the category of the label grew, the relationship between object-oriented behaviors and labels is learned by modifying the weight of the Hebbian network. Figure 12 shows the Hebbian network's weight transition recorded from a real robot experiment in accordance with the number of times the label was given. We can observe that the system successfully learned the one-to-one correspondence of object-oriented behaviors and labels.

3.5. Experiment on label association

After learning the relation between object-oriented behaviors and labels, the learner was presented with the new objects shown in Fig. 8e–h without the labels. Since all the object-oriented behaviors of the new objects were known, the learner was expected to be able to associate the learned labels to the new objects. Figure 13 shows the transition of the success rate in labeling new objects as the learner interacted with the new objects and identified their object-oriented behaviors. The horizontal axis shows the number of interaction trials the learner experienced with the new objects. During this interaction with the new objects, the multi-module reinforcement learning system continued the behavior learning process to adapt the learning modules to the new objects. The success rate of labeling shown on the vertical axis was calculated based on 400 sets of object images and label (100 sets of testing data were prepared for each object) produced in the environment of the experiment. The success rate of labeling is only about 20% at the start, which shows that existing methods are helpless for associating labels to new objects autonomously. As the learner interacted with the new objects, it expanded the categories of object-oriented behavior to the photometric features of new objects. Through this growth of the behavior-based category, the learner was able to assign appropriate labels to new objects following the process discussed in Section 2.6.



(a) 10 labels (b) 20 labels

(c) 30 labels (d) 60 labels

Figure 12. Weight transition of the Hebbian network.

After about 12 trials of interaction with the new objects, the mean success rate of labeling for all new objects reached 80%. The success rate of labeling for the new car and the new box was not as high as the other two objects. This was due to the fact that the new car had a very similar color to the boxes. We can overcome this problem by introducing other photometric features such as edge histograms to form a better space for categorization.

4. DISCUSSION AND CONCLUSIONS

We proposed a lexicon acquisition system which can associate labels to new objects with very different visual features according to the behavior that can be performed with the objects. The system was implemented on a robot learning labels about objects with different rolling preference. The robot acquired the rolling behaviors for each object, formed photometric categories of those behaviors and successfully associated the labels to newly introduced objects. An interesting point of our

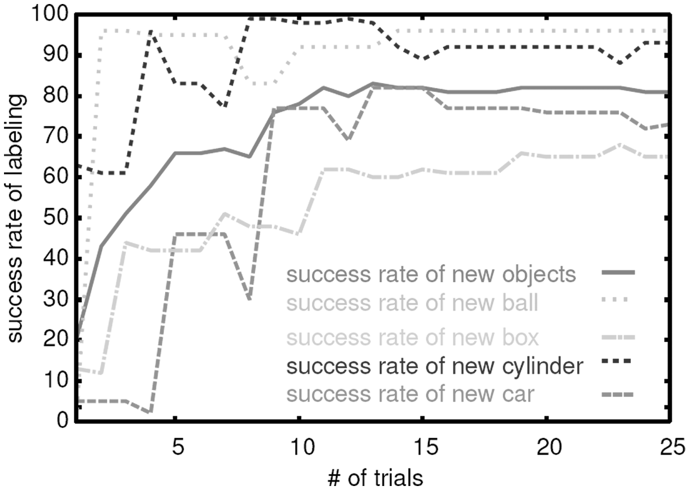


Figure 13. Association of labels to new objects.

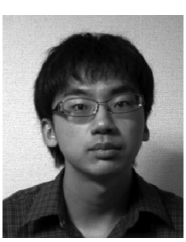
research is that the acquired lexicon can be interpreted as both nouns and verbs depending on the context. Our research indicates the possibility that verbs can be grounded to some behaviors indirectly through features of objects with certain affordances. For future work, we consider the possibility of symbol grounding by modifying our system. The reliabilities of the multi-module reinforcement learning system can be associated to basic symbols for symbol grounding [21, 22] as in the case of the parametric bias of Sugita and Tani's system [5]. Compared to the distributed expression adopted in their approach, local expressions adopted in our approach are advantageous in the sense that no memory interference occurs. Most symbol grounding systems developed so far require basic symbols to be given simultaneously with the corresponding sensorimotor experience. The idea of indirect association through features of objects with certain affordances may also be useful for finding correspondences between basic symbols and sensorimotor concepts.

REFERENCES

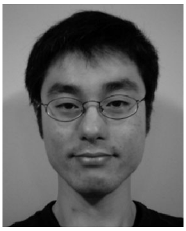
1. M. Fujita and H. Kitano, Development of a quadruped robot for robot entertainment, *Autonomous Robots* **5**, 7–18 (1998).
2. H. Ishiguro, T. Ono, M. Imai, T. Kanda and R. Nakatsu, Robovie: an interactive humanoid robot, *Industrial Robot* **29**, 498–503 (2001).
3. G. Lakeoff and M. Johnson, *Metaphors We Live By*. University of Chicago Press, Chicago, IL (1980).
4. A. Cangelosi, A. Greco and S. Harnad, Symbol grounding and the symbolic theft hypothesis, in: *Simulating the Evolution of Language*, A. Cangelosi and D. Parisi (Eds), pp. 191–210. Spring, London (2002).
5. Y. Sugita and J. Tani, Learning semantic combinatoriality from the interaction between linguistic and behavioral processes, *Adaptive Behav.* **13**, 33–52 (2005).

6. L. Steels and F. Kaplan, AIBO's first words: the social learning of language and meaning, *Evolution Commun.* **4**, 3–32 (2000).
7. D. Roy, Learning from sights and sounds: a computational model, PhD thesis, MIT Media Laboratory (1999).
8. S. Ullman, *High-Level Vision. Object Recognition and Visual Cognition*, MIT Press, Cambridge, MA (1966).
9. K. Nelson, R. Russell, N. Duke and K. Jones, Two-year-olds will name artifacts by their functions, *Child Dev.* **71**, 1271–1288 (2000).
10. H. Kobayashi, The influence of adults' actions on children's inferences about word meanings, *Jpn. Psychol. Res.* **41**, 35–49 (1999).
11. J. Gibson, *The Ecological Approach to Visual Perception*. Lawrence Erlbaum, Englewood Cliffs, NI (1979).
12. S. Harnad, The symbol grounding problem, *Physica D* **42**, 335–346 (1990).
13. P. Fitzpatrick and G. Metta, Early integration of vision and manipulation, *Adaptive Behavior* **11**, 109–128 (2003).
14. T. Ogata, H. Ohba, K. Komatani, J. Tani and H. G. Okuno, Extracting multi-modal dynamics of objects using RNNPB, in: *Proc. IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, Edmonton, pp. 160–165 (2005).
15. D. Wolpert and M. Kawato, Multiple paired forward and inverse models for motor control, *Neural Networks* **11**, 1317–1329 (1998).
16. A. D. Milner and M. A. Goodale, *The Visual Brain in Action*. Oxford University Press, Oxford (1995).
17. C. Watkins and P. Dayan, Technical note: Q-learning, *Machine Learn.* **8**, 279–292 (1992).
18. M. Hassoun, *Fundamentals of Artificial Neural Networks*, MIT Press, Cambridge, MA (1995).
19. L. Steels and T. Belpaeme, Coordinating perceptually grounded categories through language: a case study for colour, *Behav. Brain Sci.* **28**, 469–529 (2005).
20. T. Taniguchi and T. Sawaragi, Assimilation and accomodation for self-organizational learning of autonomous robots: proposal of dual-schemata model, in: *Proc. IEEE Int. Symp. on Computational Intelligence in Robotics and Automation*, Kobe, pp. 277–282 (2003).
21. K. Samejima, K. Katagiri, K. Doya and M. Kawato, Symbolization and imitation learning of motion sequence using competitive modules, *Trans. Inst. Electron. Inform. Commun. Eng.* **J85-D-II**(1), 90–100 (2002).
22. T. Taniguchi and T. Sawaragi, Self-organization of inner symbols for chase: symbol organization and embodiment, in: *Proc. Int. Conf. on Systems, Man and Cybernetics*, Hague (2004).

ABOUT THE AUTHORS



Shinya Takamuku received the BE and ME degrees from Osaka University in 2003 and 2005, respectively. He is currently a PhD candidate in the department of Adaptive Machine Systems, Osaka University. His current research interest is object category formation based on behavior learning.



Yasutake Takahashi received the ME and PhD degrees from Osaka University in 1996 and 2002, respectively. He is an assistant professor at the department of Adaptive Machine Systems, Graduate school of Engineering, Osaka University from 2000. He is currently a guest researcher at Fraunhofer AiS, Germany.

Minoru Asada received the BE, ME and PhD, degrees in control engineering from Osaka University, Osaka, Japan, in 1977, 1979 and 1982, respectively. From 1982 to 1988, he was a Research Associate of Control Engineering, Osaka University. On April 1989, he became an Associate Professor of Mechanical Engineering for Computer-Controlled Machinery, Osaka University. On April 1995 he became a Professor of the same department. Since April 1977, he has been a Professor of the department od Adaptive Machine Systems at the same university. From August 1986 to October 1987, he was a visiting researcher of Center for Automation Research, University of Maryland, College Park, MD. Since 2002, he has been the president of the International RoboCup Federation. In 2005, he has been elected to the grade of Fellow of IEEE for contributions to robot learning and applications. In the same year, he has also been elected to the research direction of 'ASADA Synergistic Intelligence Project' of ERATO (Exploratory Research for Advanced Technology by Japan Science and Technology Agency).