A Motivation-Based Action-Selection-Mechanism Involving Temporal Prediction

Woo Young Kwon and Il Hong Suh

Abstract— **Selecting the appropriate action to accomplish a given task is one of the most important functions for a service robot. In a dynamically changing environment a robot should select actions suitable for future as well as current situations using temporal prediction method. In this paper, a novel action-selection-mechanism(ASM) using a temporal prediction method is proposed. To select actions with respect to a current situation, our proposed ASM works as a fully connected finite state machine. It can moreover deal with sequential behaviors and can allow a state in the task program to migrate to any state in the task, in which a primitive node is associated with a state. To select actions appropriate to the future situation, a predictor using a temporal prediction method is added to the ASM. To show the validity of our proposed ASM, several experimental results will be illustrated.**

I. Introduction

A robot must select an action that is appropriate for the situation in which it operates in order to survive. Thus, a robot with sensors and actuators is usually equipped with an action selection mechanism (ASM) that relates its perceptions to its actions and makes it possible for it to adapt to its environment [1].

One of the most fundamental problems for a robot is deciding "what to do next" which is known as the action selection problem [2] . Before deciding what to do next, the following have to be taken into consideration: (1) Is the action to be decided goal-directed? In other words, does the action have to be selected in such a way that it presents the best way to attain a goal under the current situation, in which case a nominal sequence of state action pairs called a task program can be employed as a measure of distance between the current situation and the goal situation. (2) If there are multiple goals to be achieved, which goal-directed action has to be decided from among actions for different goals? The action selection problem has proven to be "a hard nut to crack" due to (a) incomplete knowledge, (b) unpredictable environments and surroundings, (c) imperfect sensors and actuators, and (d) limited resources [3]. As for technologies to deal with action selection problem classical Artificial Intelligence(AI) techniques have been proposed [4] as well as several ethology-based techniques [5]–[7].

This work was supported by the Technology Innovation Program of MKE/KEIT [2008-F-038-01, Development of Context Adaptive Cognition Technology]

W. Y. Kwon is with the Department of Electronics and Computer Engineering, Hanyang University, Korea(e-mail: wykwon@incorl.hanyang.ac.kr).

I. H. Suh is with the College of Information and Communications, Hanyang University, Korea(Corresponding author, e-mail: ihsuh@hanyang.ac.kr).

The architecture of the earlier systems, was based on traditional AI planning methods, and consisted of a sense-plan-act sequential cycle and interaction between sensing, planning, and action components. These traditional AI planning methods have some limitations because they assume that accurate knowledge about the state of the world state is provided by the system's sensors. This assumption is not valid due to a number of factors such as a changing world state, limited processing resources, and noisy unreliable sensory information [3]. As for alternatives to classical AI techniques, ethologically-inspired models of action selection have been proposed. Brooks [8] has suggested an architecture called "subsumption architecture," which is composed of competence modules with fixed priorities. This approach gives an advantage in that it can fulfill a goal in a complex environment. As menthoned, several researchers have proposed ethologically-inspired models of action selection [5]–[7]. Moreover, Sheutz [9] proposed a novel method to enable dynamic switching among different action selection strategies. All of these model have shown good performance in imitating real-life behavior, since action selection in these models has been based on competence modules with changing priorities. However, as shown in Fig. 1, the actions are selected not for achieving a long-term goal, but for satisfying a short-term drive generated by motivation flow. Thus, these models of action selection will not be adequate for a robot to achieve a task or equivalent long-term goal. Furthermore, most of those works involved "fixed" pre-designed state action behavior systems and did not incorporate learning. They may not, therefore, be appropriate in dynamic environments.

In action selections of robots, prediction is one of the key challenge. A predictive ability permits anticipation and smart decision-making by allowing a robot to decide which action to perform to obtain or avoid a predicted situation [10]. By proactive execution of actions with respect to predicted events, a robot can maximize the utility of task execution in many robotic applications such as human robot cooperation, navigation in mobile robots, and so on. In [11] Bruke has integrated a temporal prediction model into a behavior-based architecture for autonomous virtual creatures.

In this paper, we suggest a novel architecture based on ideas taken from ethology that allow proactive action-selection by temporal prediction of stimuli. Furthermore, we improve current ethology-based architectures to deal with sequential behaviors. Most typical tree structures organize actions into a hierarchy ranging from high-level "nodes" or activities to detailed primitive nodes via mid-level composite

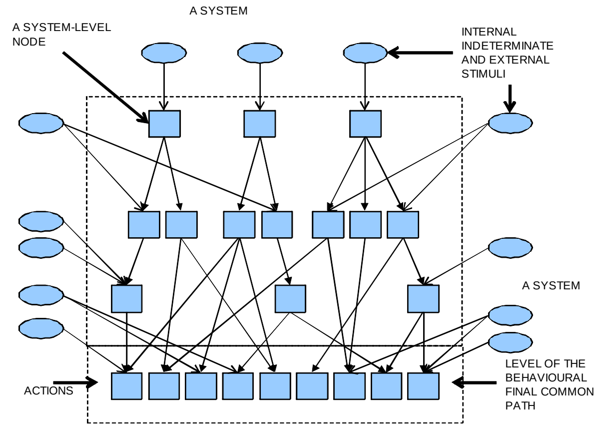


Fig. 1. Behavior selection process by motivation flows for generic ethology-based ASM [3].

actions. In these structures motivation flows are well described but behavior flows are, however, out of focus(see also Fig. 1). Thus, only primitive actions are actually executable. Our proposed ASM, however, is designed to select the most appropriate motivation in a given situation. In addition, our ASM can allow an appropriate action to be executed for every node within that motivation. As a result, our ASM is designed to choose correct sequential behaviors to satisfy a motivation and thus enable the system to learn necessary sequential behaviors. In addition to sequential behavior selections as shown in [12], we integrate a temporal prediction method into our proposed action selection mechanism. By the temporal prediction method, a robot equipped with our action selection mechanism selects preparatory actions. By the execution of these preparatory actions, the robot can reduce task execution time.

This paper is organized as follows. In Section II, a motivation based action selection mechanism to deal with sequential behavior selection will be presented. In Section III, a temporal prediction-based action selection method will be shown. Section IV presents the experimental results and concluding remarks follow in Section V.

II. Motivation based Action Selection

A. Design objectives of ASM

A robot task is usually described as a nominal sequence of state transitions from initial state s_i to goal state s_g. State transition is what constitutes a robot behavior. However, it should be noted that state can be changed beyond the control of any robot behaviors. For example, while a robot is approaching a person in a room to say "Hi", the person may leave the room. In this case, the robot does not have to say "Hi" any more, since the state that the robot is expecting has been changed to an unexpected state. That is, a state can be accidentally changed without consideration of the intended behaviors of robots. It is very difficult to take all such possible accidental state changes into account, when programming a robot task by classical programming languages. Classical programming assumes that all possible

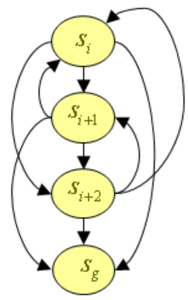


Fig. 2. An exemplar task program described as a fully connected finite state machine.

states can be defined for a robot and its environment. This assumption can be made to hold by defining any unexpected state as an exceptional state. Then, the problem becomes one of: (1) to program a robot task as an irreducible Markov process, in which a state in a robot task program can be migrated to any state in the robot task program. This problem requires ASM to be designed as a fully connected finite state machine as shown in Fig. 2. In addition, it is necessary to know how close the current state is to the goal state, since closeness to the goal state is a critical motivation to select actions and/or a task. (2) To fulfill such an additional requirement, a fully connected FSM such as a task program needs to be described as an ordered sequence of state action pairs. ASM has to support the programming of such an ordered sequence of state action pairs. (3) Furthermore, ASM has to be structured in such a way that some portions of such a sequence of state action pairs for a robot task need to be learned to adapt to new situations, and/or have to be reused another tasks of the robot.

To solve design problems (1), (2), and (3) described above, we design Behavioral Motivation (BM) as a basic unit to control behavior flow as well as motivation flow. For this, BM is designed as a valued connector associated with Perception Filters[1] (PF) and a behavior (or Action Pattern[2] - AP), where PF is associated with promotion of motivation. A sequence of BM is represented as a task program. The BM is designed to be dynamically inserted or deleted together with its associated PF and behavior in a task program. In this sense, such a relocatable BM can be be called a "Dynamic BM"(DBM). On the other hand, a sequence of BM will be denoted as a task BM to represent a task program and to distinguish such a task BM from DBM, we will call such a task BM a "Static BM"(SBM).

Now, order of sequence of DBM is designed in such a way that goal is located at a leaf node. A DBM is designed such that value of a DBM flows into its child DBM, if the DBM is allowed to release its value by PF of one of its descendant DBMs. And then, a behavior is computationally selected among behaviors in a task program by comparing values of DBMs in the task program. By doing so, a sequence

[1]PF is a function to compute degree of matching for a given context and sensory information.

[2]An action pattern is a sequence of primitive actions

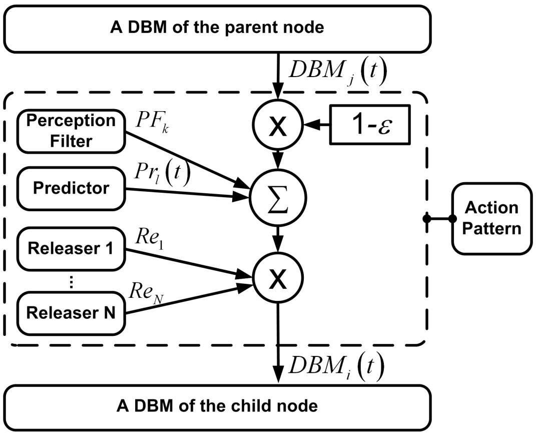


Fig. 3. A diagram of activation-value in a DBM.

of DBMs acts like a fully connected finite state machine. Moreover, nearness to the goal can be easily monitored by referring to the index of order of DBM.

B. Dynamic Behavioral Motivation (DBM)

A robot has to generate a series of behaviors and select the most appropriate one. To accomplish this, DBMs are organized into a flexible hierarchical network that can be changed by the learning process. A DBM has its own activation value that depends on the values from a PF, a predictor, a parent node, and releasers. A releaser is a special type of PF which has a value of only 0 or 1. It is used to pass or to block the activation value of a DBM. A DBM outputs its value to its child node while the relevant stimulus is incoming through the corresponding PF. The diagram of activation-value in a DBM is illustrated in Fig. 3. The activation value is accumulated through the path while the relevant stimulus is presented by the PF. Releasers play the role of blocking the flow of activation values. The value of a DBM is given as

$$DBM_i(t) = ((1-\varepsilon)DBM_j(t) + PF_k + Pr_l(t)) \prod_{m=1}^{N} Re_m, \quad (1)$$

where $DBM_i(t)$ is value of DBM i at time t, ε is small positive value, $DBM_j(t)$ is value of DBM j which is a parent node of DBM i, PF_k is value of PF connected to DBM i, $Pr_l(t)$ is value of a predictor l at time t connected to DBM i, and Re_m is value of a releaser m at time t.

A predictor outputs value to select an action in advanced of the action-related stimulus. A more detailed explanation will be given later. The sequence of value-updating in DBMs is from a parent node to a child node.

The most appropriate DBM to accomplish a task will be selected by choosing the maximum-valued DBM.

C. Static Behavioral Motivation (SBM)

An SBM implies a task, and contains a group of a sequence of DBMs to accomplish the task. Additionally, several SBMs (tasks) are organized as a flat network as shown in Fig. 4. Each SBM will take a value to be used for task selection process. An SBM is activated for given external stimuli and internal needs in a flat network, when its value is the largest for SBMs in the flat network. Here, internal needs will come out from the corresponding internal state (IS) to represent drives such as hunger or thirst. The AP, which reduces a certain IS or satisfies drives, is called a consummatory action, and other APs are called appetitive actions. The value of an IS is changed after an AP is executed. The relation of a AP to an IS can be defined based on Hull's theory [13]. The value of an SBM is computed by combining values of ISs with values of PFs. In addition, an SBM receives a feedback effect from DBM groups under

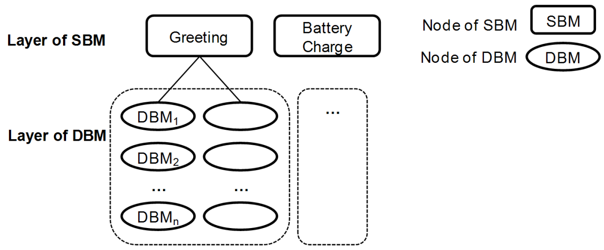


Fig. 4. Two SBMs in a flat network configuration.

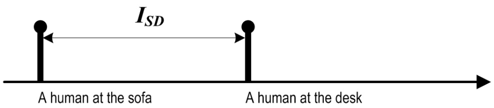
$$SBM_i = \sum_{all\ j} IS_j + \sum_{all\ k} PF_k + effect_i^{DBM}, \quad (2)$$

i, j, and k, respectively, imply the index of the ith SBM, the index of a related IS, and the index of a related PF.

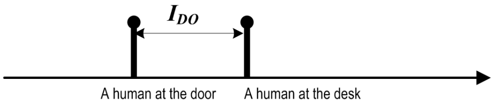
In (2), the term $effect_i^{DBM}$ indicates the strength and show how easily the goal can be achieved for a given current state of the environment. Thus, the value of an SBM may be high not only when the needs of the SBM become more important than those of other SBMs, but also when its goal is believed to be easily achieveable in the current state of the environment.

III. Proactive Action Selection by a Predictor

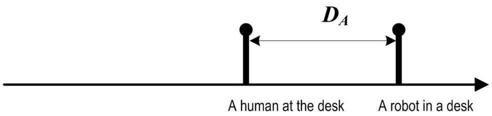
In this chapter, we will explain a proactive action-selection method by temporal prediction of an action-related stimulus to observe preceding stimuli. By preparing for an action, a robot can reduce the overall time taken to accomplish a given task. A predictor can play the role of simulating a value of the action-related stimulus beforehand by temporally predicting an action-related stimulus from many preceding neutral stimuli. To explain a proactive action selection, a greeting task in an office using a mobile robot is used in this paper. In a greeting task, a human will pass through the door and move to the front of a desk. Finally, he will sit on a chair in the front of the desk. From this sequence of human behaviors, the stimulus of "a human in the door" can be a preceding stimulus with respect to "a human in front of the desk". If a robot knows the temporal relation between "a human in the door" stimulus and "a human in front of the desk" stimulus, the robot can move to the front of the desk in advance using the predictive stimulus of "a human in the door".



(a) Prediction of "a human at the desk" from "a human at the sofa"



(b) Prediction of "a human at the desk" from "a human at the door"



(c) Action selection by using a perception of "a human at the desk"

Fig. 5. Predictions of perceptual events from preceding sensory events.

In many robotic applications, however, a robot is in a multi-stimulus situation. Therefore, a robot should learn causal and temporal relations among multiple stimuli and actions. Finding these relations among multiple stimuli and actions is more difficult than when there is a single stimulus and action. Thus, we will introduce a generalized proactive action selection method under multiple stimuli and actions will be presented.

The value of a predictor in a DBM i is obtained by

$$Pr_i(t) = m_i \sum_{\mathrm{all}\, j} w_{ij} \times f_{ij}(t - \tau_j) \tag{3}$$

, where t is time variable, i is an index of action-related stimulus, j is an index of any neutral stimulus, $Pr_i(t)$ is the value of a predictor for stimulus i at time t, m_i is the maximum value of a predictor i, w_{ij} is the weight between stimulus i and j. This weight means simultaneously frequency of presentation and temporal proximity between stimulus i and j. The weight will be explained the end of this chapter. τ_j is starting time of stimulus j, and $f_{ij}(t - \tau_j)$ is a temporal function of proactive action-selection. $f_{ij}(t - \tau_j)$ can be obtained by

$$f_{ij}(t - \tau_t) = \begin{cases} 1 & \text{if} \quad \max(0, I_{ij} - D_i) \leq t_j < I_{ij} \\ 0 & \text{else.} \end{cases} \tag{4}$$

, where I_{ij} is the averaged temporal interval between an ith stimulus and a jth stimulus, and D_i is the duration of an action pattern when DBM i is selected.

In Fig. 5, an illustration to explain a temporal function of proactive action-selection is shown. For an explanation of proactive action-selection, we will consider two cases for the greeting task.

In the first case, a human moves to the desk starting from the sofa. From the viewpoint of a robot, "a human at the sofa" stimulus was received first. After the time interval of I_{SD}, "a human at the desk" stimulus is received. In I_{SD}, a subscript of 'S' means a sofa and a subscript of 'D' means a desk. In the second case, a human move to the desk by passing the door. From the viewpoint of a robot, "a human at the door" stimulus was received firstly. After the time interval of I_{DO}, "a human at the desk" stimulus is received. In I_{DO}, a subscript of 'O' means a door. In both case, a robot will approach the desk to recognize the human after "a human at the desk" stimulus received. Averaged temporal duration of the approaching action is D_A, where a subscript of 'A' means approach. These are typical scenario of the greeting task.

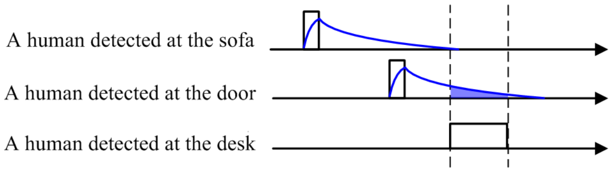


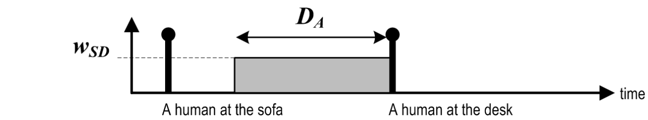
Fig. 6. An example of stimulus trace.

To accomplish the task quickly, it is better for a robot to be near the desk when a robot receive "a human at the desk" stimulus. By temporal prediction, onset time of "a human at the desk" stimulus with respect to "a human at the sofa" is I_{SD} as shown in Fig. 5-(a). If a robot start moving to the desk at the time of $I_{SD} - D_A$, a robot will be near the desk when a "a human at the desk" stimulus received without waiting time. This movement will ends at the time of I_{SD}. For the case of "a human at the door" stimulus, a robot may start moving to the desk at the time of $I_{DO} - D_A$.

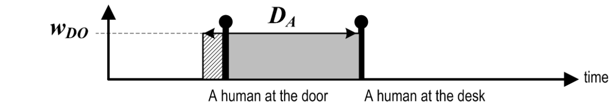
For the causal relation among many stimuli, we should know the degree of proximity among stimuli. In animal timing theories, many researches had proposed causal relation among many stimuli. In rate estimation theory, Gibbon [14] used a ratio of a overlapped time to all observation time of stimulus. If two stimuli observed at the same time, the ratio will be closed to 1. Otherwise, if overlapped time of two stimuli is small, the ratio will be closed to 0. In real robotic applications, however, there are many relations among stimuli without overlapped time. These relations may have obvious causal relations. Therefore, we use concept of stimulus trace as shown in Fig. 6.

The static relation between a neutral stimulus and an action-related stimulus can be obtained from the degree of temporal contiguity between two stimuli. If the neutral stimulus disappears before the an action-related stimulus starting, two stimuli would not be present at the same time. Thus, some form of internal stimulus trace of a preceding stimulus is required as shown in Fig. 6, where, overlapped area between preceding stimulus and a target stimulus is large if the difference in temporal contiguity between preceding stimulus and a target stimulus is small. We can say that the relation between "A human detected at the door " and "A human detected at the desk " is stronger than the relation between "A human detected at the sofa " and "A human detected at the desk".

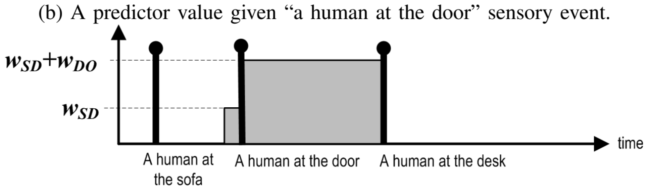
In previous computational models of classical conditioning, many method to model a trace of the stimulus have been introduced [15]. We use this equation



(a) predictor value given “a human at the sofa” sensory event.



(b) A predictor value given “a human at the door” sensory event.



(c) A predictor value of “a human at the desk” sensory event given “a human at the door” and “a human at the sofa” sensory events.

Fig. 7. An example of a predictor value.

$$w_{ij} = \frac{\int_{t_j=0}^{\infty} Tr_j(t_j) S_i(t_j) dt_j}{\sum_{all\, j} \int_{t_j=0}^{\infty} Tr_j(t_j) S_i(t_j) dt_j} \quad (5)$$

as the stimulus trace, where $Tr_i(t)$ is defined as

$$Tr_i(t+\varepsilon) = (1-\delta)Tr_i(t) + \delta S_i(t)$$

$$S_i(t) = \begin{cases} 1 & \text{stimulus i present} \\ 0 & \text{stimulus i absent} \end{cases}, \quad (6)$$

where δ is a real value between 0 and 1 related to a slope of the trace, and ε is an interval of sampling time. And, S_i is stimulus of a preceding stimulus S_j is an action related stimulus.

In Fig 7, an illustration of computing predictor values are shown. If the frequency of the sequential stimuli of “A human at the door” and “A human at the desk” stimulus is almost the same as sequential stimuli of “A human at the sofa” and “A human at the desk”, a weight of w_{DO} is higher than w_{SD} because the temporal proximity of ”A human at the door” and “A human at the desk” is smaller. In Fig 7-(a) and Fig 7-(b), a value of a predictor given “a human at the sofa” stimulus using $w_{SD} \times f_{SD}(t-\tau_S)$ and a value of a predictor given “a human at the door” stimulus using $w_{DO} \times f_{DO}(t-\tau_D)$ are shown, respectively. In Fig 7-(b), $I_{DO} - D_A$ is less than zero. However, an elapsed time from the onset of “a human at the sofa” stistimulus cannot precede “a human at the sofa” stimulus. Therefore, a temporal region with negative time is ignored. An example of predictor value using (4) under considerations of both “a human at the sofa” and “a human at the door” stimuli is shown as Fig 7-(c). This predictor value has an influence on DBM value to select a proactive action.

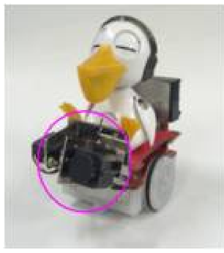


Fig. 8. The robot platform used in the experiment.

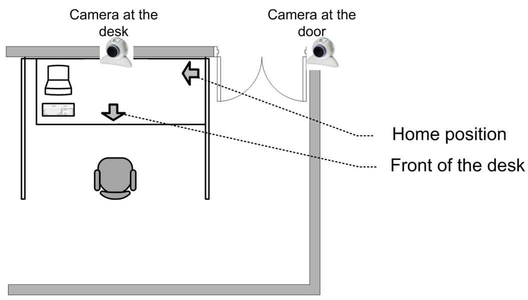


Fig. 9. An experimental setup.

IV. Experiment

In this section, we report on an evaluation of the usefulness of the proposed ASM using a mobile robot which can move on top of a desk. Our own mobile robot is designed and employed for experiments as shown in Fig. 8. The robot has been designed to have a single CCD camera with a tilt guide, and to be controlled by a two-wheeled differential drive system. The robot is equipped with speaker and a microphone to interact with a human, and actuators to express emotions such as eyes, wings, and a lip.

An illustration of an experimental environment is shown in Fig. 9. To show the validity of our proposed ASM integrating a temporal prediction method, an experimental set-up for a greeting task of a mobile robot was organized as shown in Fig. 9. The experimental environment is assumed to a room with a desk and a door. There are two cameras. One is for detecting a human at the door. Another is for detecting a human at the desk.

If there are no humans in the room, the robot spends time charging up its battery. After some time, a user enters to the room by opening the door, and he sits on the chair in front of the desk. When a camera in the front of the desk detects a human, the robot will move to the front of the desk. The next step is recognizing the user and greeting him. In Fig. 10, the BM tree of a greeting task and a charging task are shown. Without temporal prediction, a robot can move to the desk only after detecting a human in front of the desk. With temporal prediction, a robot can move to the desk before a human is detected at the desk using predictive information.

A time-line expression of all PFs and actions when a predictor doesn’t apply to the ASM are shown in Fig. 11-(a). In Fig. 11-(a), “ a human detected at the door” does not have an effect on selecting an action selection. Fig. 11-(b) is a

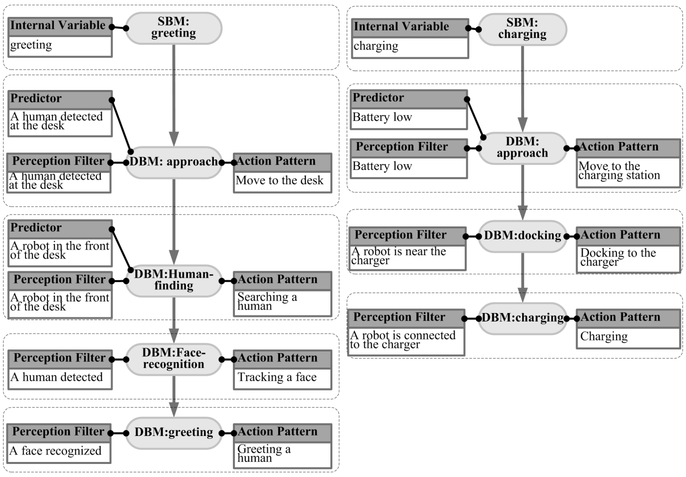
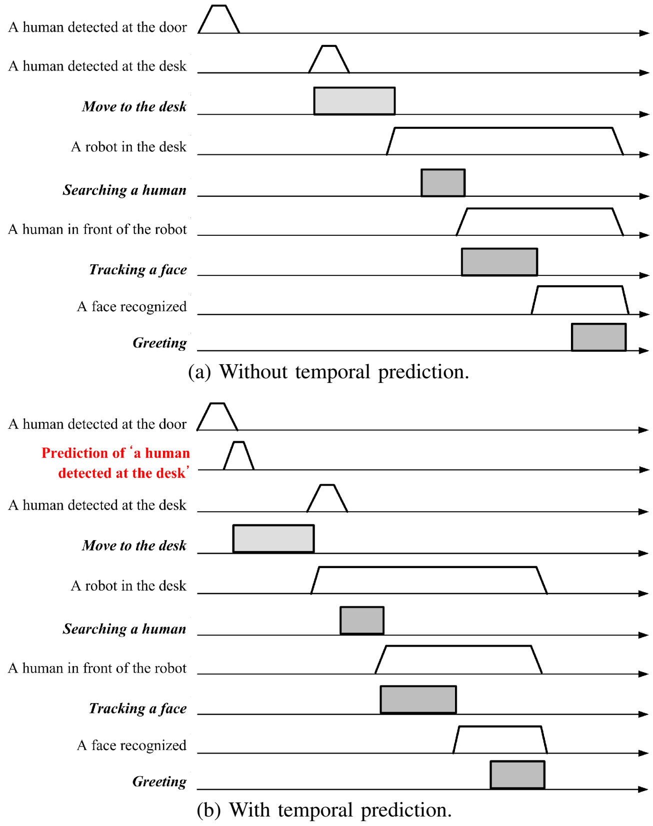


Fig. 10. BM tree for a greeting task using predictors.



(a) Without temporal prediction.

(b) With temporal prediction.

Fig. 11. Temporal expression of PFs, predictors, and APs consisting of a greeting task.

time-line expression of all PFs, predictors, and actions when predictors do apply to the ASM. After the onset of the "A human detection at the door" stimulus, the predictor of "A human detection at the desk" outputs a value. Receiving the value from the predictor of "A human detection at the desk", a BM having an action of "Move to the desk" as an action pattern will be selected. Fig. 11-(b), verifies that a predictor has a role in reducing the time taken to accomplish a task.

V. Conclusion

In this paper, we have proposed an ethology-based action selection method using a temporal prediction method. It is most important for a robot to be able to select and learn the best behaviors to survive in an environment. To accomplish this, we propose a hierarchical organization of competence modules called SBMs and DBMs. The SBM was used to select the most appropriate motivation in a given situation. The DBM is used to select a behavior that can satisfy its motivation. By utilizing releasers used to pass or block the activation-value of a DBM, a DBM tree can generate sequential behaviors. Thus, not only can our proposed ASM select the most appropriate behavior in a given situation, it can also deal with sequential behaviors. Furthermore, our proposed ASM included a temporal prediction module. By using the temporal prediction module, a robot can prepare some preliminary actions. Thus, a robot can save time when accomplishing given tasks.

References

[1] A. Guillot and J.-A. Meyer, "Computer simulations of adaptive behavior in animals," in *Proc. Computer Animation '94*, 25–28 May 1994, pp. 122–131.

[2] P. Maes, "Modeling adaptive autonomous agents," *Artifical Life*, vol. 1, no. 1-2, pp. 135–162, 1994.

[3] T. Tyrrell, "Computational mechanisms for action selection," Ph.D. dissertation, University of Edinburg, 1993.

[4] M. J. Huber, S. Kumar, and D. McGee, "Toward a suite of performatives based upon joint intention theory," *Lecture Notes in Computer Science*, vol. 3396, pp. 226–241, 2005.

[5] E. Spier and D. Mcfarl, "A finer-grained motivational model of behaviour sequencing," in *From Animals to Animats 4: Proceedings of the Fourth International Conference on Simulation of Adaptive Behavior*, 1997.

[6] B. M. Blumberg, P. M. Todd, and P. Maes, "No bad dogs: Ethological lessons for learning in hamsterdam," in *From Animals to Animats, Proceedings of the Fourth International Conference on the Simulation of Adaptive Behavior*, 1996.

[7] Y. Hoshino, T. Takagi, U. Di Profio, and M. Fujita, "Behavior description and control using behavior module for personal robot," in *Proc. IEEE International Conference on Robotics and Automation ICRA '04*, vol. 4, Apr 26–May 1, 2004, pp. 4165–4171.

[8] R. Brooks, "A robust layered control system for a mobile robot," *IEEE Transactions on Robotics and Automation*, vol. 2, no. 1, pp. 14–23, Mar. 1986.

[9] M. Scheutz and V. Andronache, "Architectural mechanisms for dynamic changes of behavior selection strategies in behavior-based systems," *IEEE Transactions on Systems, Man and Cybernetics, Part B*, vol. 34, no. 6, pp. 2377–2395, Dec. 2004.

[10] P. Capdepuy, D. Polani, and C. Nehaniv, "Construction of an internal predictive model by event anticipation," *Lecture Notes in Computer Science*, vol. 4520, p. 218, 2007.

[11] R. Burke, "It's about time: Temporal representations for synthetic characters," Ph.D. dissertation, Massachusetts Institute of Technology, 2001.

[12] S. Lee, I. Suh, and W. Kwon, "A Motivation-Based Action-Selection-Mechanism Involving Reinforcement Learning," *nternational Journal of Control, Automation, and Systems (IJCAS)*, vol. 6, no. 6, pp. 904–914, 2008.

[13] C. L. Hull, *Principle of behavior*. New York: Appleton-Century-Crofts., 1943.

[14] C. Gallistel and J. Gibbon, "Time, rate, and conditioning," *PSYCHOLOGICAL REVIEW-NEW YORK-*, vol. 107, no. 2, pp. 289–344, 2000.

[15] C. Balkenius, J. Morén, and L. Kungshuset, "Computational models of classical conditioning: a comparative study," in *From Animals to Animats 5: Proceedings of the Fifth International Conference on Simulation of Adaptive Behavior*. MIT Press, 1998, pp. 348–353.