Proceedings Second International Symposium on New Frontiers in Human-Robot Interaction, K. Dautenhahn, J. Saunders (Eds.), at the AISB 2010 convention, 29 March – 1 April 2010, De Montfort University, Leicester, UK

Smart Action Selection Architecture Taking into account both Goal-orientedness and Proactivity

Woo Young Kwon[1] and **Il Hong Suh**[2]

Abstract. Intelligent interaction with humans plays an important role in robotic applications. For such interaction, a robot has to select an action by considering not only the current situation but also the predicted future situation. Predictive abilities facilitate anticipation and smart decision-making by allowing the robot to decide what action has to be performed. In this paper, we propose smart action selection architecture taking into account both goal-orientedness and proactivity. To select actions with respect to a current situation, we use a reactive plan-based action selection mechanism that can deal with sequential goal-oriented behaviors within the built-in plan. To select actions appropriate to a future situation, we propose a probabilistic temporal prediction method, and integrate the prediction method with the reactive plan-based action selection mechanism. To show the validity of our proposed architecture, we present some experimental results.

1 Introduction

To accomplish a given task, a robot has to select an action that is appropriate for the current situation as well as a given goal. This requirement can be summarized as "what to do next" and is one of the most fundamental problems for a robot. To solve this problem, various approaches have been proposed including classical Artificial Intelligence (AI) techniques [1], ethology and biology-based techniques [2] [3] [4], a fuzzy-based technique [5], and a reinforcement learning based technique [6]. The architecture of early systems was based on traditional AI planning methods that consists of a sense-plan-act sequential cycle. These traditional AI planning methods have some limitations because they assume that sensors of a system provide accurate knowledge about the state of the world. This assumption is not valid because of a number of factors such as the changing state of the world, limited processing resources, and noisy unreliable sensory information [7]. As alternatives to classical AI techniques, behavior-based approaches 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 for a robot to accomplish a goal in a complex environment by selecting situation-adequate actions.

Although many researchers have addressed situation-adequate action selection, there has been little attention to proactivity; proactivity is acting in advance of a future situation, rather than just reacting. For a robot to be smart, it has to make a proactive action with respect to the predicted future situation. Proactive and prediction ability permits anticipation and smart decision-making by allowing a robot to decide which action to perform to obtain or avoid a predicted future situation [9]. By the 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.

On the basis of psychological and biological backgrounds, researchers have proposed computational models for the temporal prediction of future events; these models are essential for achieving proactivity. One of the most widely known approaches to prediction is classical conditioning. This is a learning theory that explains how animals learn to predict the emergence of an unconditioned stimulus from the presentation of a conditioned stimulus [10] [11]. This method is useful for explaining how an animal and/or a robot predicts the presentation of an unconditioned stimulus on the basis of conditioned stimuli. However, it is too simple to apply to real robotic applications because it usually consider only the static and temporal relations between two events. In [12], Bruke has integrated a temporal prediction model into a behavior-based architecture for autonomous virtual creatures based on scalar expectancy theory and a rate estimation theory by Gallistel and Gibbon [13] that is a temporal prediction model based on psychological theory. However, the scalar expectancy theory is too simple and the rate estimation theory is computationally ineffective. In [14], a proactive and goal-oriented action selection mechanism was proposed based on temporal prediction, but a deterministic temporal method was used in the paper. Therefore, it may not be suitable for real robotic applications.

In general, robots are required to learn complex relationships among many stimuli; moreover, these relationships can change dynamically and non-deterministically. In real robotic applications, therefore, an effective probabilistic method is required for the prediction of future events. Moreover, to select proactively an action, the temporal prediction method should be incorporated into an action selection mechanism. In this paper, we propose novel architecture for proactive action selection for smart robots. For situation-adequate and goal-oriented action selection, we use a reactive-plan based action selection mechanism [15] that is based on ethology. For proactive action selection, we propose a probabilistic temporal prediction method, and integrate this method with the action selection mechanism. By using the temporal prediction method, a robot can select an action with respect to the current situation as well as the predicted future situation. By the execution of a proactive action using the prediction of future events, the robot can reduce not only the task execution time for the action but also the human's waiting time.

This paper is organized as follows. In Section II, a proactive based action selection mechanism is presented. In Section III, a probabilistic temporal prediction method is discussed. Section IV presents the

[1] Hanyang University, Korea, email: wykwon@incorl.hanyang.ac.kr
[2] Hanyang University, Korea, email: ihsuh@hanyang.ac.kr (Corresponding author)

experimental results and concluding remarks follow in Section V.

2 Action Selection Architecture Taking into account both Goal-orientedness and Proactivity

A robot task is usually described as a nominal sequence of state transitions from an initial state s_i to a goal state s_g. However, a state can be accidentally changed without consideration of the initial task plan. It is difficult to consider all such accidental state changes when programming a robot task using classical programming languages. Classical programming assumes that all possible 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: (1) to program a robot task as an irreducible Markov process, in which a state in the robot task program can be migrated to any other state. This problem requires an action selection mechanism to be designed as a fully connected finite state machine as shown in Figure 1.

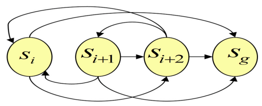


Figure 1. Example of task program described as fully connected finite state machine.

(2) In addition, it is necessary to know how close the current state is to the goal state, since the degree of closeness is an important motivation when selecting actions. To fulfill this requirement, a fully connected finite state machine such as a task program needs to be described as an ordered sequence of state-action pairs. The action selection mechanism has to support the programming of this ordered sequence of state-action pairs. (3) Furthermore, the action selection mechanism has to be structured in the 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 for other tasks.

To solve problems (1), (2), and (3) above, we designed Behavioral Motivation (BM) in [16] as a basic unit to control behavior flow and motivation flow. A BM is designed as a valued connector associated with Perception Filters[3] (PF) and Action Pattern (AP)[4], where a PF is associated with the promotion of motivation. BMs are organized into a flexible hierarchical network, and they have an activation value that depends on the values from a PF, a parent node, and releasers. A releaser is a special type of PF that has a value of 0 or 1. It is used to pass or to block the activation value of a BM. A BM outputs its value to its child node while the relevant stimulus is incoming through the corresponding PF.

However, a robot equipped with a motivation-based action selection mechanism in [16] cannot proactively select an action; it can select only situation-adequate and goal-oriented actions because a BM can receive values only from a PF and the BM on the parent node. For proactivity, we designed a new computational module called a predictor that proactively outputs a value to select an action in advance of an action-triggering event. Figure 2 illustrates an activation-value in a BM. The activation value is accumulated through the path while the relevant stimulus is presented by the PF. The releasers block the flow of activation values. The value of a BM is given as

$$V_i^{BM} = \left((1-\varepsilon) V_j^{BM} + V_{PF} + \sum_{k=1}^{N} V_k^{Pr} \right) \prod_{l=1}^{M} V_l^{Re}, \tag{1}$$

where V_i^{BM} is the value of *BM i*; ε is a small positive value to discount the value of the BM at the parent nodel; V_j^{BM} is the value of *BM j*, that is a parent node of BM i; V_{PF} is the value of a PF that filters an action-triggering stimulus in BM i; V_k^{Pr} is the value of a predictor to predicts an action-triggering stimulus from stimulus k; and V_l^{Re} is the value of releaser l. The most appropriate BM to accomplish a task will be selected by choosing the maximum-valued BM.

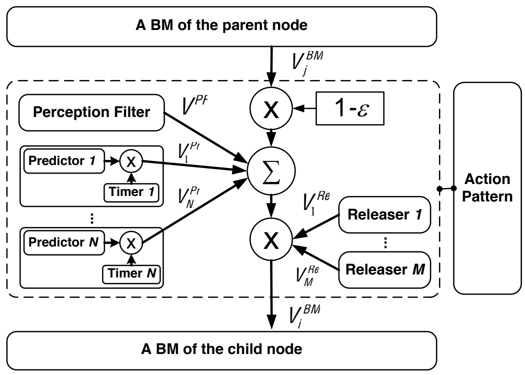


Figure 2. Activation value in BM.

For proactivity, we should consider not only the action-triggering stimulus that define a PF value but also neutral stimuli. A neutral stimulus is not an action-triggering event, but it has a causal and temporal relation with the action-triggering event. Without the predictor, however, the neutral stimulus does not affect the action selection because the PF does not output a value when observing a neutral stimulus. To predict the occurrence of a future action-triggering stimulus from any neutral stimuli, two predictions are necessary: when the stimulus will present and the probability of the stimulus. Two additional computational modules in Figure 2 perform these two predictions: a predictor and an internal timer. The predictor calculates the probability of an action-triggering stimulus from a neutral stimulus, and the internal timer predicts when the action-triggering stimulus will occur from observation of the neutral stimulus.

The value of the predictor V_k^{Pr} can be given by

$$V_k^{Pr} = P(s_i|s_k)\Theta(t_k), \tag{2}$$

where k is a neutral stimulus; $P(s_i|s_k)$ is the conditional probability of an action-triggering stimulus i given the neutral stimulus k that is the output of a predictor; t_k is the elapsed time from occurrence of stimulus k; and $\Theta(t_k)$ is a temporal function that outputs a value when it is time for the proactive action. A detailed explanation of $P(s_i|s_k)$ and $\Theta(t_k)$ will be given in the next section.

To select proactive and goal-oriented actions, BMs with predictors are hierarchically composed as a tree structure. An example of a BM

[3] PF is a function to compute the degree of matching for an action-triggering stimulus from sensory information. The PF value gives the strength of the action-triggering stimulus

[4] An action pattern is a sequence of primitive actions.

tree in a coffee-delivery task is shown in Figure 3. A robot has to serve coffee or a meal when a human generates the action-triggering event, order-meal or order-coffee. In the right section of Figure 3, labeled Goal-Oriented Action-selection, situation-adequate actions can be generated. For example, when a human orders coffee, the robot will approach a table when a human sits there, then it will receive a coffee and deliver the coffee to the table. During the task execution, any exceptional case also be handled by reactive planning.

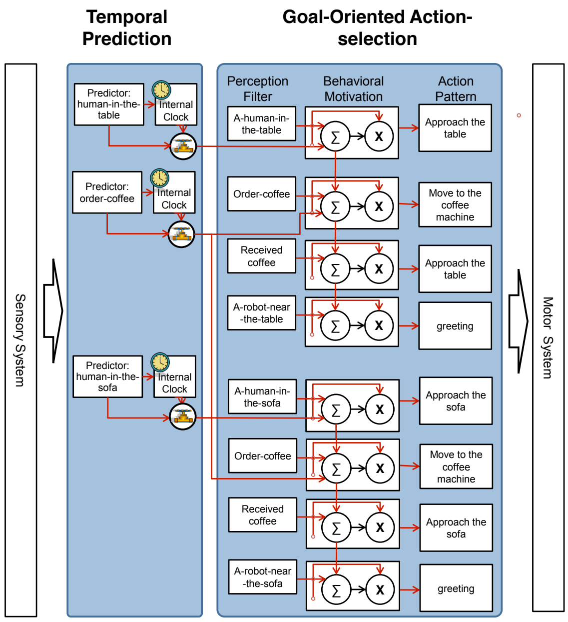


Figure 3. Action selection architecture taking into account both goal-orientedness and proactivity.

As another example, we can assume that the robot will deliver coffee when a human orders coffee immediately after entering the room through the door. Although the human does not sit at a table, the next sequence can be generated by selecting the maximum valued BM. However, these situation-adequate actions keep the human waiting, so the robot seems stupid. By using predictor in the left section of Figure 3, labeled Temporal Prediction, proactive actions can be generated. Before ordering coffee or a meal, the human passes some site such as a kitchen, desk, or table. When the human passes these sites, the robot can receive the neutral stimuli corresponding to the movement of the human. By predicting an action-triggering stimulus from these neutral stimuli, the robot can execute proactive actions to minimize the waiting time.

3 Probabilistic Temporal Prediction of Action-triggering Stimulus for Proactivity

To predict the time and probability of an action-triggering stimulus when observing neutral stimuli, we need to know the a causal probability of $P(s_i|s_k)$ and the best action-time of $\Theta(t_k)$ as shown in (2). Many robotic applications are complex; the robot is in a multi-stimulus situation, and the temporal relations between an action-triggering stimulus and the neutral stimuli are stochastic. Therefore, a probabilistic prediction method is necessary to learn the causal and temporal relations among multiple stimuli. It is more difficult to find these relations when there are multiple stimuli rather than a single deterministic stimulus. Figure 4 shows the causal relations between an order-coffee stimulus and neutral stimuli in the delivery task of Figure 3.

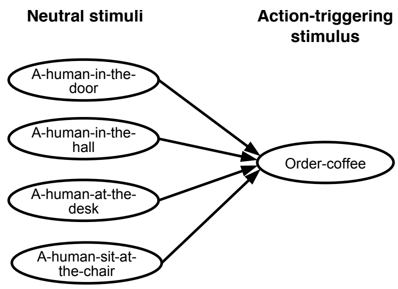


Figure 4. Prediction of an action-triggering stimulus from many neutral stimuli.

Firstly, it is necessary to obtain causal relations between the action-triggering stimulus and neutral stimuli, represented as $P(s_i|s_k)$. These relations are obtained by computing the degree of proximity among those stimuli. In animal timing theories, many researchers have proposed causal relations among many stimuli. In rate estimation theory, Gibbon [17] used a ratio of an overlapping time to the total observation time of the stimulus. If two stimuli are observed at the same time, the ratio will be close to 1. If the overlapping time of two stimuli is small, the ratio will be close to 0. In practical robotic applications, however, there are many relations among stimuli without overlapping time. To compute the degree of proximity among stimuli, we use the concept of a stimulus trace as shown in Figure 5.

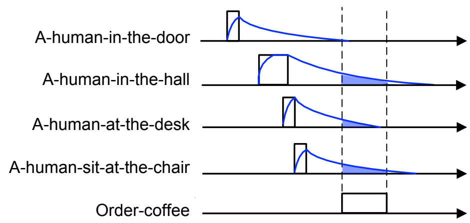


Figure 5. Example of stimulus trace.

The relation between a neutral stimulus and an action-triggering stimulus can be obtained from the degree of temporal contiguity between the two stimuli. If the neutral stimulus disappears before an action-triggering stimulus starts, the two stimuli are not present at the same time. Thus, some form of internal stimulus trace of a preceding stimulus is required as shown in Figure 5. If the difference in temporal contiguity between a preceding stimulus and a target stimulus is small, the overlapping area between the two stimuli is large. Thus, the relation between a human-sitting-on-the-chair and order-coffee is stronger than the relation between a human-detected-at-the-door and order-coffee. In computational models for classical conditioning, many methods to model a stimulus trace have been introduced [18]. We use

$$P(s_i|s_j) = \frac{\int_{t_j=0}^{\infty} Tr_j(t_j)S_i(t_j)dt_j}{\sum_{allk} \int_{t_j=0}^{\infty} Tr_j(t_j)S_i(t_j)dt_j} \tag{3}$$

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

$$Tr_i(t+\varepsilon) = (1-\delta)Tr_i(t) + \delta S_i(t), \tag{4}$$

and

$$S_i(t) = \begin{cases} 1 & \text{stimulus i present} \\ 0 & \text{stimulus i absent} \end{cases}. \tag{5}$$

Here, δ is a real value between 0 and 1 related to the slope of the trace, ε is an interval of sampling time, S_i is an action-triggering stimulus, and S_j is a neutral stimulus.

The next issue is to find the best time for the proactive action. For temporal prediction of the stimulus, we firstly define a temporal random variable T_A for stimulus A. The T_A is a random variable of the elapsed time after stimulus A was observed. Using this temporal random variable, we use the conditional Probability Density Function (PDF) of an action-triggering stimulus A given a neutral stimulus B as $f_{T_A|T_B}(T_A|T_B)$. The conditional PDF can be defined only when stimulus A will eventually present after B is presented. This conditional PDF can be represented as a PDF with another random variable:

$$f_{T_A|T_B}(t_A|t_B) = f_{T_{AB}}(t_A - t_B), \tag{6}$$

where T_{AB} is a relational temporal random variable of A given B. Here, the Normal distribution is used to model the conditional PDF between an action-triggering stimulus and a neutral stimulus. If an action-triggering stimulus is influenced by two or more neutral stimuli, the conditional probability of the action-triggering stimulus given all combinations of its parents is required. However, modeling temporal distributions which involves modeling PDFs in n dimensions is difficult, and the computational complexity becomes too great. Therefore, we need to use parametric modeling of the temporal conditional probabilities to avoid the complexity of modeling and inference. In this paper, the temporal probability of a conditional event in a specific time interval can be modeled as the weighted sum of two or more PDFs.

Using these modeling methods, the temporal probability of a conditional stimulus can be defined as

$$f_{T_A|T_{B_1},\ldots,T_{B_N}}(t_A|t_{B_1},\ldots,t_{B_N}) = \sum_{i=1}^{N}\left(P(A|B_i)\, f_{T_{AB_i}}(t_A - t_{B_i})\right), \tag{7}$$

where, $P(A|B_i)$ is obtained using (3). By using the PDF of the action-triggering stimulus, we can calculate the best time for proactive action. Here, $f_{T_{Pa}|T_A}$ represents a PDF of proactive action given an action-triggering stimulus to minimize the waiting time, and $f_{T_A|T_B}$ represents a PDF of an action-triggering stimulus given neutral stimuli. If we model these two PDFs, the robot can proactively select an action. In this situation, a human waits for the robot to serve. The waiting time of the human is given by $T_{Hw} = T_{Pa} - T_A$ and the waiting time of the robot is given by $T_{Rw} = T_A - T_{Pa}$. In this case, T_{Rw} is 0. Thus, T_{Hw} is always larger than T_{Rw}. The expectation of the waiting time of the human is given by

$$E[T_{Hw}] = \int_{-\infty}^{\infty} t_{Hw} f_{T_{Hw}}(t_{Hw})\, dt_{Hw}, \tag{8}$$

and the expectation of the waiting time of the robot is given by

$$E[T_{Rw}] = \int_{-\infty}^{\infty} t_{Rw} f_{T_{Rw}}(t_{Rw})\, dt_{Rw}. \tag{9}$$

The best time for proactive action is given by minimizing the sum of these two expectations $E[T_{Hw}] + E[T_{Rw}]$. Using only a Gaussian function and a Dirac delta function for PDFs, the expectation for the waiting time is given by

$$\begin{aligned} E[T_{Rw} + T_{Hw}] &= 2\int_{-\infty}^{\infty} t_{Hw} f_{T_{Hw}}(t_{Hw})\, dt_{Hw} \\ &= 2\int_{-\infty}^{\infty} t_{Hw} f_{T_A} \star f_{T_{Pa}}(t_{Hw})\, dt_{Hw} \\ &= 2\int_{-\infty}^{\infty} t_{Hw} N\left(t_{Hw}; \mu_A - \mu_{Pa}, \sigma_M^2 + \sigma_{Pa}^2\right) dt_{Hw}, \end{aligned} \tag{10}$$

where μ_{Pa} is the mean of the PDF of the ending time of the approach action, and $\star$ is a cross-correlation operator[5]. Therefore, the best action time is $t = \mu_A - \mu_{Pa}$. Finally $\Theta(t)$ is given by

$$\Theta(t) = \begin{cases} 1 & t \geq \mu_A - \mu_{Pa} \\ 0 & t < \mu_A - \mu_{Pa} \end{cases}. \tag{11}$$

4 Experiment

4.1 Experimental Setup

In this section, we will experimentally evaluate the usefulness of the proposed proactive action selection method using a robot-serving scenario. Figure 6 illustrates an experimental environment used for a serving task by a mobile robot, and Figure 7 illustrates the movements of a human and a robot during the experiment. Figure 7(a) presents a path in which the human selected the action of approaching the table and in Figure 7(b) the human selected the action of approaching the desk.

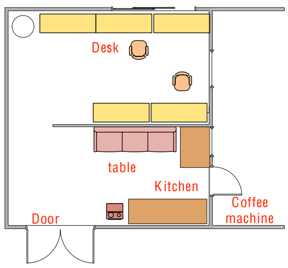
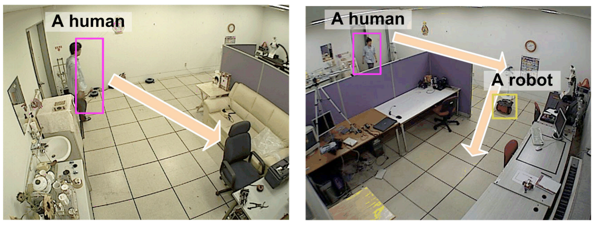


Figure 6. Experimental environment

We used a camera with wide viewpoint to capture the movement of the human, and a Pioneer 3-DX robot to serve coffee or a meal. At the start of the experiment, the robot is located near the desk. Next, the human enters the room through a door. The preferred paths of the human are as follows:

1. door→kitchen→table→table→order a meal
2. door→kitchen→table→table→order coffee
3. door→desk→order coffee

[5] The cross-correlation of f and g is defined as $(f \star g)(t) = \int_{-\infty}^{\infty} f(\tau) \cdot g(t+\tau)\, d\tau$



(a)Approach to the table (b) Approach to the desk

Figure 7. Experimental setups.

4.2 Experimental Results

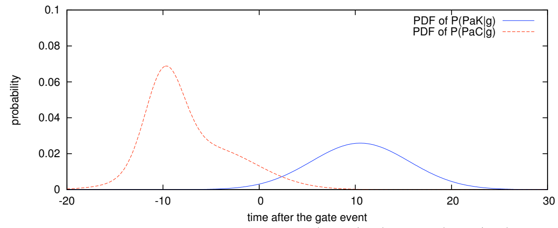
Cause	Effect	μ	σ	P_{AB}
door	kitchen	4.57	1.61	0.5
kitchen	table	7.44	1.73	1.0
table	order a meal	10.8	5.86	0.5
table	order coffee	7.23	4.53	0.5
door	desk	9.26	1.12	0.5
desk	order coffee	5.58	1.65	1.0
order a meal	approaching the kitchen	12.3	1.31	1.0
order coffee	approaching the coffee machine	24.7	2.19	1.0

Table 1. Temporal and static evidence from the experimental data.

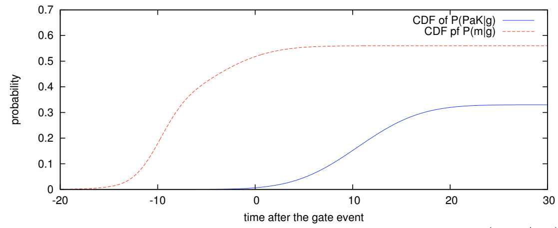
For probabilistic model for the predictor, we must provide CPTs and PDFs for the inference of causal probability as well as temporal probability. In the experiment, we collected the time interval data between events for every trial. Each trial is terminated when the human is successfully served a meal or a cup of coffee by the robot. We made CPTs by counting the frequency of conditional events and learned PDFs for the temporal interval of conditional events for every trials; means and variances of Normal distribution were computed by using the expectation maximization method.

The temporal and causal data is shown in Table 1. From these data, the causal probability and best time for predictors can be obtained. predictor value of an order-coffee given a-human-at-the-door stimulus is 0.57, and for an order-meal given a-human-at-the-door stimulus, it is 0.33. Moreover, their conditional PDFs are given as follows: the PDF of move-to-the-coffee-machine is $0.3N(t;-5.49,5.56)+0.26N(t;-9.86,2.96)$, and the PDF of move-to-the-kitchen is $0.33N(x;10.51,6.45)$. The cumulative distribution functions and PDFs of these two proactive actions are shown in Figure 8.

A time-line expression of all stimuli and actions when a predictor does not apply to the action selection mechanism are shown in Figure 9(a). Here, a-human-at-the-door does not have any effect on the action selection. Figure 9(b) is a time-line expression of all stimuli and actions when predictors do apply to the proactive action selection mechanism. After the onset of the a-human-at-the-door stimulus, the predictor of order-coffee outputs a value for proactivity. Given the value from the predictor, a BM having an action of move-to-the-coffee-machine as a proactive action will be selected. Without proactive action selection, the average overall task-execution time is 71s with proactive action takes 60s. This result shows that the predictor

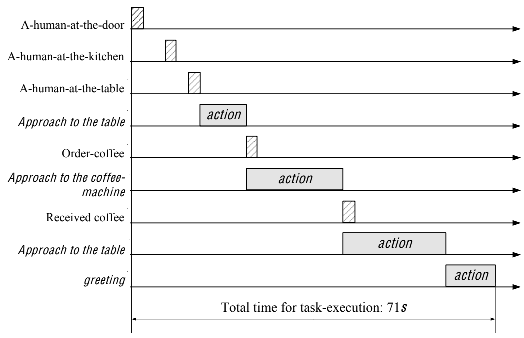


(a) PDFs of two proactive actions;$P(T_{PaK}|T_G)$ and $P(T_{PaC}|T_G)$.

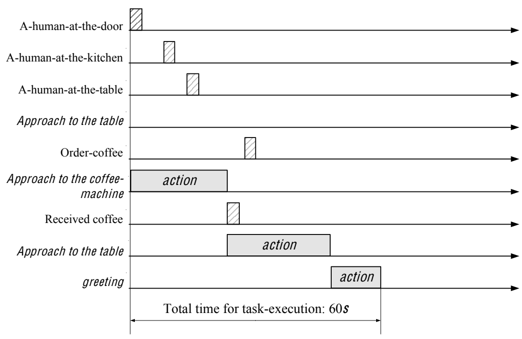


(b) Cumulative distribution functions of two proactive actions;$P(T_{PaK}|T_G)$ and $P(T_{PaC}|T_G)$.

Figure 8. Comparison of two proactive actions given a human-at-the-door event.



(a) Without proactive actions.



(b) With proactive actions.

Figure 9. Temporal expression of stimulus and action-pattern consisting of delivery task.

has a role in reducing the time taken to accomplish a task.

5 Conclusion

For intelligent interactions between a human and a robot, both situation-adequate action and proactive action are necessary. In this paper, we have proposed smart action selection architecture taking into account both goal-orientedness and proactivity that can be used in real robotic applications. By using the proposed method, a robot can predict a future action-triggering stimulus from the currently observed stimuli; by using these predictions, the robot can select a proactive action to minimize both robot and human waiting time. Our experimental results showed that the proposed action selection method reduced the waiting time in a serving task by anticipating a future situation.

ACKNOWLEDGEMENTS

This work was supported by the Technology Innovation Program of MKE/KEIT [KI001849, Development of Context Adaptive Cognition Technology].

REFERENCES

[1] 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.

[2] 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.

[3] D. Tsankova, "Emotional Intervention on an Action Selection Mechanism Based on Artificial Immune Networks for Navigation of Autonomous Agents," *Adaptive Behavior*, vol. 17, no. 2, p. 135, 2009.

[4] 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.

[5] J. Jaafar and E. McKenzie, "A fuzzy action selection method for virtual agent navigation in unknown virtual environments," *Journal of Uncertain Systems*, vol. 2, no. 2, pp. 144–154, 2008.

[6] E. Hogewoning, J. Broekens, J. Eggermont, and E. Bovenkamp, "Strategies for affect-controlled action-selection in Soar-RL," *Lecture Notes in Computer Science*, vol. 4528, pp. 501–510, 2007.

[7] J. Bryson, "Mechanisms of action selection: Introduction to the special issue," *Adaptive Behavior*, vol. 15, no. 1, p. 5, 2007.

[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] 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.

[10] C. Balkenius, "Dynamics of a classical conditioning model," *Autonomous robots*, vol. 7, no. 1, pp. 41–56, 1999.

[11] R. Sutton and A. Barto, "Time-derivative models of Pavlovian reinforcement," *Learning and computational neuroscience: Foundations of adaptive networks*, pp. 497–537, 1990.

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

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

[14] W. Y. Kwon and I. H. Suh, "A Motivation-Based Action-Selection-Mechanism Involving Temporal Prediction," in *International Conference on Service and Interactive Robotics, SIRCon 2009*, 6–7 August 2009.

[15] R. Sutton and A. Barto, "Time-derivative models of Pavlovian reinforcement," *Learning and computational neuroscience: Foundations of adaptive networks*, pp. 497–537, 1990.

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

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

[18] 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.