Behavior Selection Method for Entertainment Robots Using Intelligence Operating Architecture

Woo-Ri Ko and Jong-Hwan Kim

Abstract To get and hold a user's attention, entertainment robots should be able to think and behave like a human being to show various responses in a certain situation. For this purpose, this paper proposes a behavior selection method for entertainment robots using intelligence operating architecture (iOA). The iOA consists of five parts and 15 modules to implement the robot intelligence, which is motivated by the key functions of human brain. In the internal state part, the strengths of the robot's internal states, i.e. motivation, homeostasis, and emotion, are updated. The sensory information is converted to contexts in the context module. Considering both the internal states and contexts, a behavior which is composed of four expressions, i.e. facial expression, gesture, movement, and voice expression, is selected in the problem solving module. To show the effectiveness of the proposed method, a software entertainment robot is implemented for simulations. The simulation results show that entertainment robots with different characteristics can be created and they can generate various behaviors by the proposed behavior selection method.

Key words: Behavior selection, entertainment robot, robot intelligence, fuzzy integral, fuzzy measure

1 Introduction

Nowadays, a number of entertainment robots have been developed, which aim to interact with users in a variety of situations. To receive attention from the user and hold the attention, entertainment robots should think and behave like a human being to show various responses depending on its internal state even in the same situation.

W.-R. Ko and J.-H. Kim
Department of Electrical Engineering, KAIST, 291 Daehak-ro, Yuseong-gu, Daejeon, 305-701, Republic of Korea,
e-mail: {wrko, johkim}@rit.kaist.ac.kr

For this purpose, there has been much research on the behavior selection method for entertainment robots. An expressive gesture generation algorithm was presented for storytelling application [1]. A system for a spontaneous speech recognition, multi-modal dialogue processing and visual perception of a user was developed for a natural human-robot interaction [2]. A composite facial expression generation method was devised to reflect the robot's emotion [3].

In this paper, the intelligence operating architecture (iOA) is used to construct a behavior selection method for entertainment robots [4]. The iOA consists of five parts and 15 modules to implement the robot intelligence. The robot's internal states, i.e. motivation, homeostasis, and emotion, are updated in the internal state part. In the context module, external situation is defined as a context. Considering both the internal state and the context, the next proper behavior is selected in the problem solving module, which is composed of facial expression, gesture, movement, and voice expression. To diversify the robot's responses, the characteristics of the robot is formed by assigning different preference degrees for each and every internal state and context.

This paper is organized as follows. Section II presents the intelligence operating architecture (iOA), which explains how the intelligence of robots is generated. The behavior selection method using the iOA is proposed in Section III. The effectiveness of the proposed method is illustrated through the simulations in Section IV. The concluding remarks follow in Section V.

2 Intelligence Operating Architecture (iOA)

As shown in Fig. 1, the iOA consists of five parts and 15 modules to implement the robot intelligence [4]. It is motivated by the key functions of the human brain, such as problem solving in the frontal lobe, actions from the motor cortex, etc. In the perception layer, internal and external sensor data are gathered in the internal and external sensing modules, respectively, and the gathered data are converted to context data in the perception module. The three modules in the internal state part, i.e. motivation, homeostasis, and emotion modules, work for controlling internal states. All data including the internal state strengths and the context are shared with other modules through the memory.

There are three kinds of memory: short-term memory (STM), working memory (WM), and long-term memory (LTM). The STM deals with the data required to be remembered for a short time period. Therefore, sensory data are stored in the STM. The LTM, on the other hand, lasts for a few days to a whole lifetime of a robot. It is composed of three sub memories; episodic memory (EM), semantic memory (SM), and procedural memory (PM). The EM stores the events that are related to time and can be described explicitly. However, the SM holds the knowledge which is factual and concept-based. The PM consists of the information about the way of using a certain object or conducting a specific action. Lastly, the WM contains

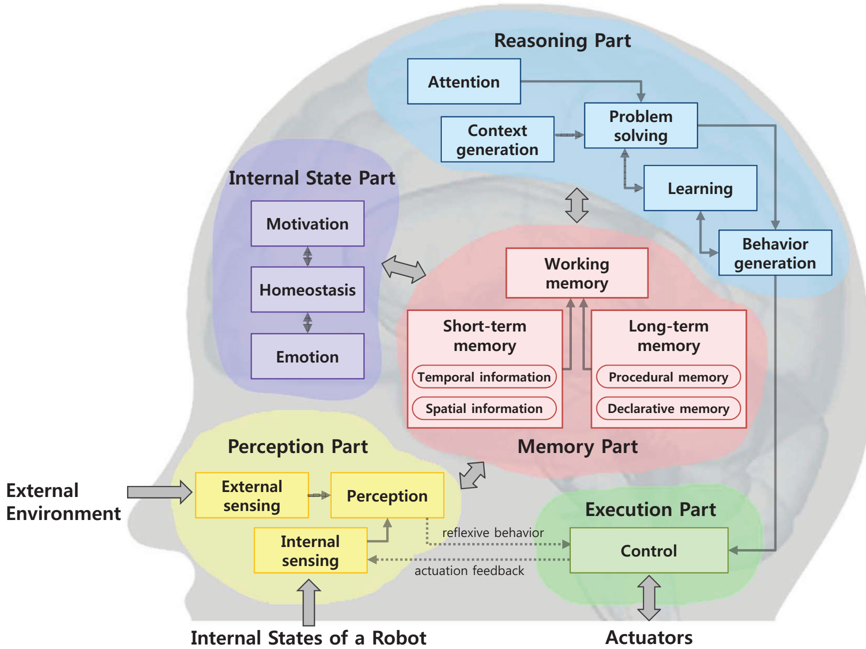


Fig. 1 The intelligence operating architecture (iOA).

the information currently required to carry out tasks including reasoning, learning, understanding, etc.

A behavior decision procedure is carried out on the reasoning part. By help of the WM, task planning and scheduling are processed in this part. In the problem solving module, the most proper behavior for the robot is selected. Then, this decision is passed to the execution layer and the selected behavior is generated to the exact action by actuating the motor system in the control module. In this paper, some modules in the iOA are realized to develop a behavior selection method for entertainment robots.

3 Behavior Selection Method

In this section, a behavior selection method based on the iOA is described. The key modules for the behavior selection, namely motivation, homeostasis, emotion, context, and problem solving modules are described in the following.

3.1 Motivation, Homeostasis, and Emotion Modules

In the motivation, homeostasis, and emotion modules, the internal states of an entertainment robot are defined and the strengths of the internal states are updated. As shown in Table 1, there are eight internal states of the robot. In the motivation module, four motivations, i.e. "curiosity," "power," "social contact," and "tranquility," are defined based on 16 human basic desires theory [5]. In the homeostasis module, "battery" is defined as the robot's homeostasis. In the emotion module, an emotion of the robot is represented as a point in a three dimensional coordinates of "arousal," "valence," and "stance," [6].

Table 1 The eight internal states of the entertainment robot.

Module	Internal state
Motivation	Curiosity (m_1) Power (m_2) Social contact (m_3) Tranquility (m_4)
Homeostasis	Battery (h_1)
Emotion	Arousal (e_1) Valence (e_2) Stance (e_3)

At time t, the strength of the jth motivation $m_j(t), j = 1,2,\ldots,l$, where l is the number of motivations, is updated by [8].

$$m(t+1) = m(t) + \alpha_j(\overline{m_j} - m_j(t)) + S^T \cdot M_j(t) + \delta_{ij}(t), \tag{1}$$

where α_j is the difference gain, $\overline{m_j}$ is the steady-state value of the jth motivation, S is the stimulus vector, M_j is the strength vector between stimulus and the jth motivation, and $\delta_{ij}(t)$ is the amount of change of the jth motivation strength caused by the previous ith behavior. If the previous ith behavior affects positively on the jth motivation, $\delta_{ij}(t)$ is a positive value. If the previous ith behavior affects negatively on the jth motivation, $\delta_{ij}(t)$ is a negative value. In the same manner, the strengths of homeostasis and emotion are updated. Note that the strengths of motivation and homeostasis are scaled to be in $[0,1]$ and those of emotion are scaled to be in $[-1,1]$. The normalized strengths of internal states are used in the problem solving module.

3.2 Context Module

In the context module, four contexts are defined for entertainment robots, as shown in Table 2. Since the entertainment robots should entertain more people and respond to the user's voice, "the number of attentive people," "distance to the nearest person," "loudness," and "direction of the loudest sound" are defined as the contexts.

Table 2 The four contexts of the entertainment robot.

Context	Scale	Unit
The number of attentive people (c_1)	0, 1, ..., 10	people
Distance to the nearest person (c_2)	[0, 5]	meter
Loudness (c_3)	[0, 100]	decibel (dB)
Direction of the loudest sound (c_4)	[-180, 180)	degree

3.3 Problem Solving Module

In the problem solving module, considering both the robot's internal states and external contexts, the most proper behavior is selected by employing the fuzzy integral. The fuzzy integral is one of the well-known aggregation methods for multi-criteria decision making and it can also reflect the redundancy (negative interaction) or synergy (positive interaction) effects between criteria, i.e. internal states and contexts [7]-[9]. A behavior is composed of four expressions, i.e. facial expression, gesture, movement, and voice expression. A list of expressions is presented in Table 3. The global evaluation value of each expression is calculated by the fuzzy integral with respect to 1) the fuzzy measure values of criteria sets, i.e. internal state and context sets, 2) current internal state strengths and context information, 3) and the partial evaluation values of expressions over criteria. The detailed procedure of behavior selection is described in the following.

Table 3 A list of expressions of the entertainment robot.

Expression	Behavior
Facial expression	Normal, anger, disgust, fear, happiness, sadness, surprise, fatigue
Gesture	Raise left hand, raise right hand, raise both hands, lower both arms, wave left hand, wave right hand, wave both hands, hold out left hand, hold out right hand, hold out both hands, rest chin on left hand, rest chin on right hand
Movement	Stand, bow, turn left, turn right, go forward, step back, turn head left, turn head right, bend knee, lean
Voice expression	Hello, hi, nice to meet you, who are you?, bye, have a good day, I am tired

3.3.1 Fuzzy Measure Identification of Criteria Set

To measure the preference degree of a criteria set, i.e. internal state and context set, ϕ_s transformation method is employed for an efficient fuzzy measure identification [10]. In this method, the fuzzy measure values are calculated using a hierarchy diagram of criteria which represents hierarchical interaction relations among criteria. A fuzzy measure $g(A)$, where A is the subset of criteria, is identified as follows:

$$g(A) = \phi_s(\xi_R, \sum_{P \subset R} u_P^R), \tag{2}$$

where R is the root level in the hierarchy diagram, ξ_R is the interaction degree between the criteria sets in the R, ϕ_s is a scaling function [11], and u_Q^P is defined as follows:

$$\phi_s(\xi, u) = \begin{cases} 1, & \text{if } \xi = 1 \text{ and } u > 0 \\ 0, & \text{if } \xi = 1 \text{ and } u = 0 \\ 1, & \text{if } \xi = 0 \text{ and } u = 1 \\ 0, & \text{if } \xi = 0 \text{ and } u < 1 \\ \frac{s^u - 1}{s - 1}, & \text{other cases} \end{cases} \tag{3}$$

$$u_Q^P = \begin{cases} d_i, \text{where } i \in Q & \text{if } |Q| = 1 \text{ and } i \in A \\ 0 & \text{if } |Q| = 1 \text{ and } i \notin A \\ \phi_s^{-1}(\xi_P, \phi_s(\xi_Q, \sum_{V \subset Q} u_V^Q) \times T_Q^P) & \text{other cases} \end{cases} \tag{4}$$

where $s = (1-\xi)^2/\xi^2$, d_i is the preference degree of the ith criterion, and the value of $\phi_s^{-1}(\xi, r)$ is u, which satisfies $\phi_s(\xi, u) = r$. The conversion ratio T_Q^P from Q to P, is computed as

$$T_Q^P = \frac{\phi_s(\xi_P, \sum_{i \in Q} d_i)}{\phi_s(\xi_Q, \sum_{i \in Q} d_i)}, \tag{5}$$

where P is the upper level set and Q is the lower level set in the hierarchy diagram.

3.3.2 Global Evaluation of Expressions Using Fuzzy Integral

The global evaluation value $E(expr_i), i = 1, 2, \ldots, n$ of the ith expression $expr_i$, where n is the number of expressions, is calculated by the following Choquet fuzzy integral:

$$\begin{aligned} E(expr_i) &= \int_X h \circ g \\ &= \sum_{j=1}^{n} \{h_{ij} \cdot \Omega_j(t) - h_{i(j-1)} \cdot \Omega_{j-1}(t)\} g(A), \end{aligned} \tag{6}$$

where $X = \{m_1, m_2, m_3, m_4, h_1, e_1, e_2, e_3, c_1, c_2, c_3, c_4\}$ is the universal set of criteria, $A \subset X$ is the subset of criteria, h_{ij} is the partial evaluation value of the ith expression over the jth criterion, $\Omega_j(t)$ is the strength of the ith criterion, and g(A) is the fuzzy measure value of A, identified by (2). One expression in each expression category is selected, which has the highest global evaluation value.

4 Simulations

4.1 Simulation setting

To show the effectiveness of the proposed behavior selection method, the simulations are performed on an software entertainment robot, as shown in Fig. 2. In the perception program, the sensory data is converted into context information and the internal state strengths are updated. Considering both the internal state and context, a behavior which is composed of four kinds of expressions is selected in the expression and behavior selection program. A selected behavior is generated and executed in the behavior generation and execution programs, respectively. The communication of data between programs is performed by the communication server program.

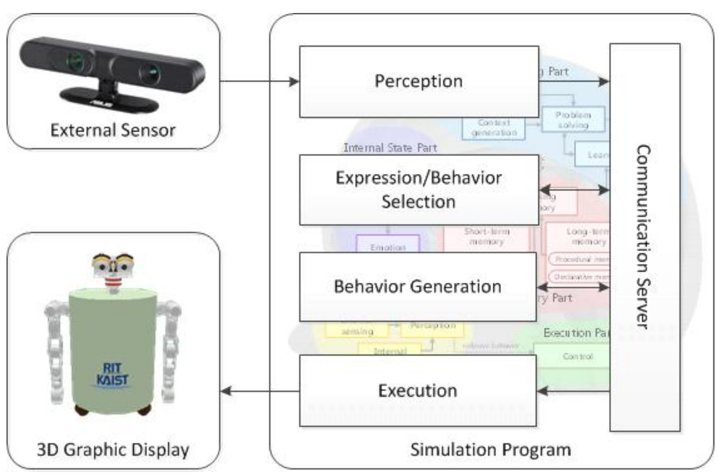
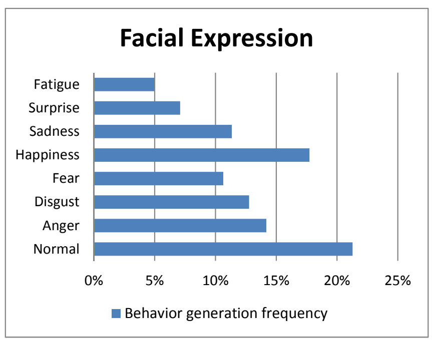


Fig. 2 The structure of the simulation program.

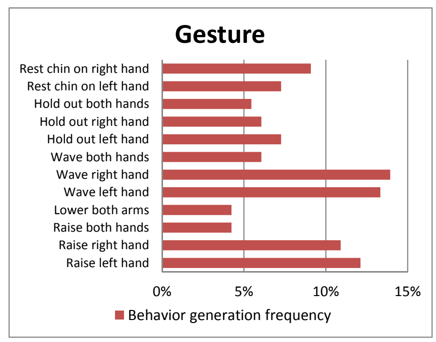
To measure the performance of the proposed method, the mean of behavior generation frequency was calculated from 10 sets of the experiment results gathered for one hour, which is the same as one day in the virtual world. The behavior was selected in every 2.5 sec, and therefore the number of total generated behaviors were 1,440. Note that the partial evaluation values of expressions were pre-given by an expert.

4.2 Simulation 1: The Generation Frequencies of Four Expressions

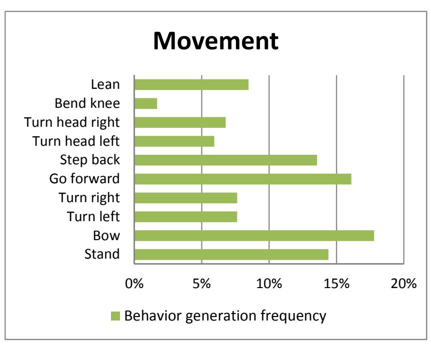
In this simulation, the behavior generation frequencies of four expressions were computed, as shown in Fig. 3. The generation frequency of “fatigue” facial expression was about 5% and that of “normal” facial expression was about 21%. The generation frequency of “raise both hands” gesture was about 4% and that of “wave right hand” was about 14%. The generation frequency of “bend knee” movement was about 1% and that of “bow” movement was about 18%. The generation frequency of “bye” voice expression was about 7% and that of “hi” voice expression was about 22%. In summary, various expressions and behaviors could be generated through the proposed behavior selection method.



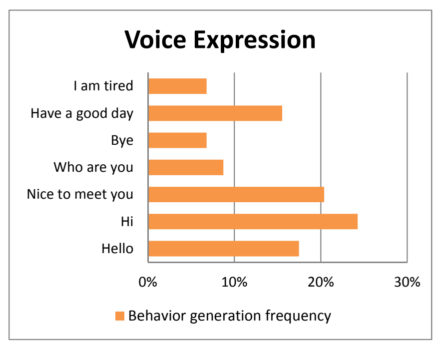
(a) Facial expression



(b) Gesture



(c) Movement



(d) Voice Expression

Fig. 3 The behavior generation frequencies of four expressions.

4.3 Simulation 2: Three Entertainment Robots with Different Characteristics

In this simulation, three entertainment robots with different characteristics were created by assigning different preference degrees for each and every internal state and context, as shown in Table 4. For a cheerful entertainment robot, the preference degrees for “curiosity,” “power,” and “social contact” were given to be higher than the others. For a safety-first one, the preference degrees for “tranquility” and “distance to the nearest person” were given to be higher than the others. For a sentimental one, the preference degrees for emotions, i.e. arousal, valence, and stance, were given to be higher than the others.

Table 4 Three entertainment robots with different characteristics.

Characteristics	Motivation				Homeostasis	Emotion			Context			
	m_1	m_2	m_3	m_4	h_1	e_1	e_2	e_3	c_1	c_2	c_3	c_4
Cheerful	0.174	0.174	0.174	0.043	0.043	0.043	0.043	0.043	0.087	0.087	0.087	0.043
Safety-first	0.050	0.050	0.050	0.200	0.200	0.050	0.050	0.050	0.050	0.200	0.050	0.100
Sentimental	0.050	0.050	0.050	0.050	0.050	0.200	0.200	0.200	0.050	0.050	0.050	0.050

The behavior generation frequencies for the three entertainment robots were computed, as shown in Fig. 4. For the cheerful one, “holding out both hands” was the most frequent gesture expression whose generation frequency was about 16%. For the safety-first one, “stepping back” was the most frequent movement whose generation frequency was about 23%. For the sentimental one, “sadness” was the most frequent facial expression whose generation frequency was about 20%. In summary, different characteristics could be generated through the proposed method.

5 Conclusions

This paper proposed the behavior selection method for entertainment robots using the iOA which consists of five parts and 15 modules to implement the robot intelligence. The internal state strengths and context information were obtained in the internal state and context modules, respectively. Considering them, a behavior which is composed of four expressions, i.e. facial expression, gesture, movement, and voice expression, was selected in the problem solving module. Each expression was evaluated by the Choquet fuzzy integral of the partial evaluation values with respect to the preference degrees of the internal states and contexts. The effectiveness of the proposed method was demonstrated through the simulations with a software entertainment robot in the 3D virtual environment. The results showed that the robot

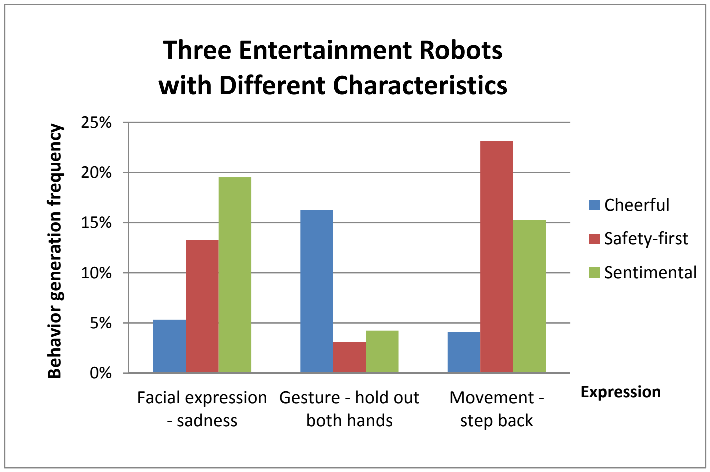


Fig. 4 The behavior generation frequencies for the three entertainment robots.

could generate various behaviors through the proposed behavior selection method. Moreover, by assigning different preference degrees, various entertainment robots with different characteristics could be created.

Acknowledgements This research was supported by the MKE (The Ministry of Knowledge Economy), Korea, under the Human Resources Development Program for Convergence Robot Specialists support program supervised by the NIPA (National IT Industry Promotion Agency) (NIPA-2012-H1502-12-1002)

References

1. Pelachaud C, Gelin R, Martin J-C, and Le QA (2010) Expressive Gestures Displayed by a Humanoid Robot during a Storytelling Application. Paper presented at New Frontiers in Human-Robot Interaction (AISB), Leicester.
2. Stiefelhagen R, Fugen C, Gieselmann P, Hozapfel H, Nickel K, and Waibel A (2004) Natural Human-Robot Interaction using Speech, Head Pose and Gestures. Paper presented at IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Sndai, Japan, pp. 2422-2427.
3. Yoo B-S, Cho S-H, and Kim J-H (2011) Fuzzy Integral-based Composite Facial Expression Generation for a Robotic Head. Paper presented at IEEE International Conference on Fuzzy systems 2011, pp. 917-923.
4. Kim J-H, Choi S-H, Park I-W, and Zaheer SA (2012) Intelligence Technologies for Robots That Think. IEEE Computational Intelligence Magazine. under review.
5. Reiss S (2002) Who am I? The 16 Basic Desires that Motivate Our Actions and Define Our Personalities. Berkley Trade.
6. Breazeal C (2003) Emotion and sociable humanoid robots. International Journal of Human Computer Interaction. in press.
7. Sugeno M (1974) Theory of fuzzy integrals and its applications. PhD dissertation, Tokyo Institute of Technology.

8. Kim J-H, Ko W-R, Han J-H, and Zaheer SA (2012) The degree of consideration-based mechanism of thought and its application to artificial creatures for behavior selection. IEEE Computational Intelligence Magazine, 7(1):49-63.
9. Ko W-R, Hyun H-S, Kim H-J, Choi S-H, and Kim J-H (2011) Behavior Selection Method for Intelligent Artificial Creatures Using Degree of Consideration-based Mechanism of Thought. Paper presented at IEEE International Conference on Systems, Man and Cybernetics..
10. Takagagi E (2008) A fuzzy measure identification method by diamond pairwise comparisons and ϕ_s transformation. Fuzzy Optimization and Decision Making, 7(3):219-232.
11. Narukawa Y and Torra V (2005) Fuzzy measure and probability distributions: Distorted probabilities. IEEE Transactions on Fuzzy System, 13(5):617-629.