Behavior Selection and Learning for Synthetic Character

Yong-Duk Kim and Jong-Hwan Kim
RIT Lab., Dept. of EECS, KAIST
373-1, Guseong-dong, Yuseong-gu, Daejeon
305-701, Republic of Korea
Email: {ydkim, johkim}@rit.kaist.ac.kr

Yong-Jae Kim
Intelligent Robot Lab.,
Institute of Intelligent System Mechatronics Center,
Samsung Electronics Co., Maetan3-dong, Paldal-gu,
Suwon-si, Gyeonggi-do, Republic of Korea. 442-742
Email: yjkim@vivaldi.kaist.ac.kr

***Abstract*—This paper proposes a novel behavior selection mechanism and interactive learning method for a synthetic character. Synthetic character can decide its behavior by itself based on its own internal states (motivation, homeostasis, and emotion), and external sensor information. A behavior is selected by both probabilistic and deterministic methods. The probabilistic method uses the internal states and external sensor information, and the deterministic method which imitates animal's instinct, uses only external sensor information. Both methods are complementary to each other. A user can teach the synthetic character a desired behavior by an interactive training method. To select a desired behavior among many behaviors, behaviors are grouped into analogous behavior sets. The learning algorithm includes the emotional parameters by which the training efficiency is affected. The performance of the synthetic character, Rity, developed with the proposed mechanism at RIT Lab., KAIST is demonstrated in a 3D virtual world.**

I. Introduction

A synthetic character is an autonomous agent which behaves based on its own internal states, and can interact with a person in real-time [1]. Recently, synthetic characters have been heavily studied throughout the world, along with the development of robots and computer technologies. Synthetic characters can be applied to entertainment robot and service robot as well as game, movie, and education as a software agent. That is to say, any area, which needs an autonomous function, an artificial emotional model, and learning skill, requires a synthetic character.

Study of a synthetic character can be divided into three topics: an implementation of artificial emotional model [2][3][4][5][6], a study of behavior selection and a study of learning and adaptation [7]. This paper focuses on a study of behavior selection and learning. There have been many studies on this topic, but most of them, based on various ethological hypotheses, proposed and implemented a behavior based architecture in which a behavior is selected and learned by a motivation [8][9][10].

A method to select a proper behavior can be categorized into two groups according to a character of behavior [11]. First is the behavior which has to be carried out deterministically, in accordance with an external input. Second is the one which needs an input as well as internal states (emotion, motivation, and homeostasis) to select a behavior. The behaviors of the second category are those which have to be performed surely according to an external input, behaviors according to a user's command, behaviors related to a synthetic character's survival and instinctive behaviors. On the other side, when there is no special command, the variety of behaviors can be selected based on the internal states and external inputs. Here, an important thing is how to select various behaviors given the same environment. The capability of generating varied behavior by the synthetic character in consideration of internal states is important because without this, the user may become bored after repeatedly observing the same behavior for a given input.

An "adaptation" and "learning" ability is also important, together with a mechanism of a behavior selection. This paper defines learning as a synthetic character's ability to change to perform a proper behavior according to user's command. An adaptation is also defined as internal states' changing according to an input and an influence to a behavior selection as a result. A synthetic character which imitates an animal needs an ability to perform a proper behavior according to user's command, and to adapt to the environment. However, problems arise when a user instructs the synthetic character on how to respond in a desired manner. If synthetic character has so many available behaviors, it is very difficult or even impossible for the synthetic character to learn what to do. It may require a prohibitive amount of time to teach what to do among many available behaviors.

This paper proposes a novel mechanism of behavior selection. A deterministic and a probabilistic behavior selection method are used in the proposed behavior selection mechanism. Both methods are complementary to each other, so these guarantee a behavior which will be surely carried out, and make a synthetic character be able to perform various behaviors. This paper also proposes a novel mechanism of learning and adaptation for a synthetic character. To select a desired behavior among many behaviors, behaviors are grouped into analogous behavior sets. The learning algorithm includes the emotional parameters by which the training efficiency is affected. The performance of the synthetic character, Rity, developed with the proposed mechanism at RIT Lab., KAIST is demonstrated in a 3D virtual world.

Section 2 describes the proposed behavior selection mechanism, and Section 3 presents a new learning method. The

0-7803-8515-2/04/$20.00 ©2004 IEEE

performance of the synthetic character, Rity is shown in Section 4, and the result will be presented in Section 5. Finally, in Section 6 concluding remarks follows.

II. Behavior Selection

The behavior system is composed of the 'behavior selector,' which chooses a behavior probabilistically, and the 'inherent behavior logic,' which chooses a behavior deterministically.

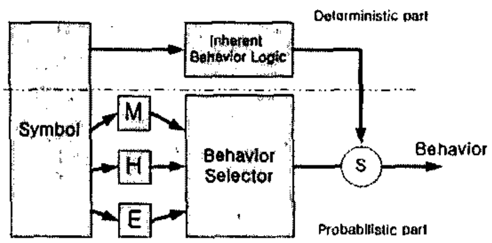


Fig. 1. The behavior selection mechanism

A. Behavior selector

Behavior selector is used to chooses a proper behavior, using synthetic character's internal states: the motivation vector M, and the homeostasis vector H. The internal states are characterized by the symbol vector of which values are based on the external sensor information, which are vision, auditory, IR, touch sensors. When there is no special command, various behaviors can be selected from this behavior selector. The behavior selection is done probabilistically based on voting values. The algorithm is described as follows:

1) Determine 'voting vector' V using M and H.
2) Adjust V considering other factors.
3) Calculate a behavior's probability, $P(b)$, using V.
4) Select a proper behavior b from $P(b)$.

Initially, the synthetic character determines the voting vector. The voting vector is calculated from the motivation and the homeostasis as follows:

$$\begin{aligned}\mathbf{V}_{temp}^T &= (\mathbf{M}^T\mathbf{D}_M + \mathbf{H}^T\mathbf{D}_H) \\ &= [Vtemp_1, Vtemp_2, \cdots, Vtemp_n]\end{aligned}$$

with

$$\mathbf{D}_M = \begin{pmatrix} d_{M11} & d_{M12} & \cdots & d_{M1n} \\ d_{M21} & d_{M22} & & \vdots \\ \vdots & & \ddots & \\ d_{Mi1} & & & d_{Min} \end{pmatrix} \qquad (1)$$

$$\mathbf{D}_H = \begin{pmatrix} d_{H11} & d_{H12} & \cdots & d_{H1n} \\ d_{H21} & d_{H22} & & \vdots \\ \vdots & & \ddots & \\ d_{Hj1} & & & d_{Hjn} \end{pmatrix}$$

where n is a number of behaviors, V_{temp} is the temporal voting vector, D_M and D_H are weights connecting the motivation and the homeostasis to behaviors.

Various masking is carried out to the temporal voting vector. Thus far, only internal status values have been considered when determining the voting vector. As a result, the synthetic character may behave oddly. For the purpose of considering internal states and external information, a process called masking is performed. For example, three kinds of masking are implemented to the temporal voting vector in our realization. These three kinds of masking are 'masking for attention,' 'masking for voice command,' and 'masking for emotion.' For example, if the creature does not see the ball, 'masking for attention' is carried out. As a result, behaviors related to the ball are masked out and they are not carried out. In a similar way, if the creature's dominant emotion is 'sadness,' behaviors connected with 'happiness' are masked out. Thus the masking process prevents the character from carrying out unusual behaviors.

A attention masking matrix $\mathbf{Q}^f(s_i)$ is obtained by the attention symbol, s_i. Each attention symbol s_i has its own making value and the matrix is defined as follows:

$$\mathbf{Q}^f(s_i) = \begin{pmatrix} q_1^f(s_i) & 0 & \cdots & 0 \\ 0 & q_2^f(s_i) & & \vdots \\ \vdots & & \ddots & \\ 0 & & & q_n^f(s_i) \end{pmatrix} \qquad (2)$$

where n is a number of behaviors, s_i is the focused symbol. In the same way, other masking matrices (voice command, and emotion) are calculated. Considering these three masking matrices, the behavior selector obtains a final voting vector as follows:

$$\begin{aligned}\mathbf{V}^T &= \mathbf{V}_{temp}^T\mathbf{Q}^f(s)\mathbf{Q}^v(c)\mathbf{Q}^e(e) \\ &= [v_1, v_2, \cdots, v_n]\end{aligned} \qquad (3)$$

where Q^f, Q^v, and Q^e are the masking matrices for the attention, the voice command, and the emotion, respectively. Finally, the probability of carrying out a behavior is calculated from the voting values as follows:

$$p(b_i) = \frac{v_i}{\sum_{k=1}^{n}(v_k)} \qquad (4)$$

where b_i is the ith behavior, v_i is the ith behavior's voting value, $p(b_i)$ is the selection probability of the behavior b_i.

B. Inherent Behavior Logic

Even if a behavior is selected by both internal states and external sensor information, there are still some limits on providing a synthetic character with natural behaviors. 'Inherent behavior logic' makes up for the weak points in the behavior selector. This process imitates an animal's instinct. For instance, as soon as an the obstacle like a wall or a cliff is found, the inherent behavior logic makes the character react to this situation immediately. Since this logic uses only external sensory information directly, its decision making speed is faster than that of the 'behavior selector.' This deterministic inherent behavior logic and the probabilistic behavior selector are complementary to each other for realizing a natural behavior.

If the synthetic character is used for a service purpose, this part will play an important part, because a service robot must follow users commands without considering internal states.

C. Switching Behavior

When there is no special command, a behavior is selected in the behavior selector. As mentioned previously, this can make the synthetic character do various behaviors. If there is a special command for a specific behavior, the inherent behavior logic fixes the behavior to be carried out. These two probabilistic and deterministic methods are constructed as a subsumption architecture. A behavior selected by the deterministic method can suppress a behavior chosen by the probabilistic method, like Figure 1. In other words, the deterministic behavior selection has a higher priority than the probabilistic behavior selection.

III. Learning and Adaptation

A. Learning

The interaction between a user and the synthetic character is an attractive learning method. The character's learning ability makes the interactions between humans and the character a sustainable relationship. For example, the synthetic character that learns through human interactions is very appealing to human beings, if it can reflect upon its past experiences. Through this interactive learning, it is characterized to look like a unique being. This learning can be considered as adjusting weighting parameters between inputs and internal logic.

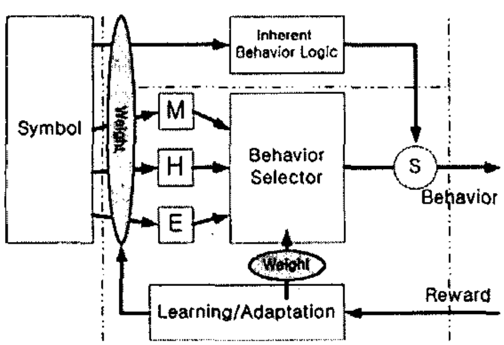


Fig. 2. The Learning Mechanism

In this architecture, however, there are tens of behaviors. Thus, the learning process requires lot of time and it may be difficult to expect desired behaviors from the synthetic character. To solve this problem, analogous behaviors are grouped into a set. For instance, the set 'SIT' is composed of behaviors such as sit, crouch, lie, and so on, similar behaviors to 'sit.' And the set 'DANCE' has dance_with_arms, shake_arms, etc. After grouping behaviors to corresponding sets, the weighs connecting commands and behavior sets are learned. These connections are shown in Figure 3. If a proper behavior is carried out when a certain command is ordered, its corresponding weight is strengthened and vice versa. The update law is as follows:

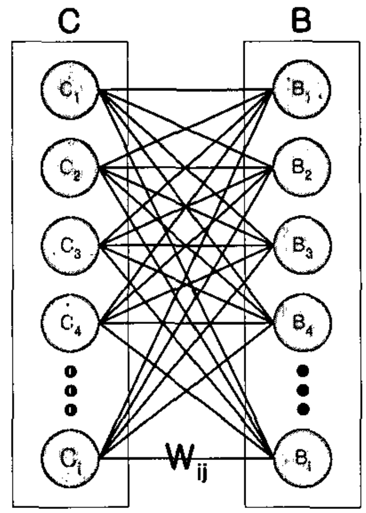


Fig. 3. Weighting connecting commands and behavior

$$
\begin{aligned}
W_{ij}(t+1) &= W_{ij}(t) + \rho R_i \\
R_i &= \begin{cases} +C_r & \text{on rewarding} \\ -C_p & \text{on penalty} \end{cases}
\end{aligned}
\tag{5}
$$

where W_{ij} is a weight between the ith command and the jth behavior set, ρ is an emotion parameter, and R_i is a reward or penalty. The emotion parameter, ρ is a value that influences the learning process. It means that learning efficiency is good when the artificial creature is happy, or bad when unhappy. As a result, the emotion parameters control the learning rate.

Weights between commands and behavior sets were learned, after behaviors had grouped into an analogous behavior set. Now, the relation between a command and each behavior must be adjuested. This can be done by changing the voice-command masking explained in Section II-A.

$$
\begin{aligned}
& q_p^v(c) = \alpha W_{ij} \\
& q_{others}^v(c) = \beta W_{kl} \\
& \text{with } \alpha > \beta > 0\ ,\ k \neq i \text{ and } l \neq j
\end{aligned}
\tag{6}
$$

where $q^v(c)$ is a masking value, p is a behavior number carried out previously. So $q_p^v(c)$ means a masking value of a behavior carried out just now and $q_{others}^v(c)$ indicates others in the same analogous set. And α and β are positive constants. The voice masking matrix in eqn. (3) is updated in proportion to W, but a behavior done just now, and other behaviors get influenced differently by α and β. Since α is bigger than β, the behavior which was performed just now, gets a bigger change of the weight than others in the same group.

Even if the behavior set receives a reward as a result of a behavior, the behavior is influenced more than others, in the process of adjusting a weight between a command and behavior. Therefore, this creates a difference in learning speed.

In summary, the proposed learning method is described as follows:

1) Since there are many behaviors to be learned, behaviors are categorized into the analogous behavior sets which are composed of analogous behaviors.
2) When a user commands the synthetic character and it carries out a proper behavior, it then receives a reward. If not, it receives a penalty.
3) A reward or penalty adjusts the weights between the command and behavior sets. At this time, the speed of adjusting the weights are affected by emotional parameters.
4) The voice-command masking matrix is adjusted based on the weight of the command and the behavior set. The behavior performed previously is affected by a reward or a penalty more than others, even though behaviors are in the same analogous set.

B. Adaptation

The 'preference learner' is the part which teaches the character's likes and dislikes for an object. If we give the character a reward or a penalty, the connected weight from the symbols to internal states is changed accordingly as follows:

$$\mathbf{W}_i(t+1) = \begin{cases} \mathbf{W}_i(t) + \Delta\mathbf{W}_i^r & \text{on rewarding} \\ \mathbf{W}_i(t) + \Delta\mathbf{W}_i^p & \text{on penalty} \end{cases}$$
$$\Delta\mathbf{W}_i^r = \begin{bmatrix} \Delta w_1^r \\ \Delta w_2^r \\ \vdots \\ \Delta w_l^r \end{bmatrix} \tag{7}$$
$$\Delta\mathbf{W}_i^p = \begin{bmatrix} \Delta w_1^p \\ \Delta w_2^p \\ \vdots \\ \Delta w_l^p \end{bmatrix}$$

where $\mathbf{W}_i(t)$ is a weight from the symbols to the motivation or emotion, Δw_l^r is a reward, Δw_l^p is a penalty of the lth weight. As the weight varies, symbols' influences to internal states are modified, and the character's likes or dislike for an object can be changed from interaction with the user.

IV. The Synthetic Character: Rity

The proposed behavior selection mechanism and interactive learning method are applied to the synthetic character, 'Rity.' The proposed character is composed of 'perception system', 'internal state system,' 'behavior system, 'learning system' and 'motor system.'The detailed description is presented in [1]

TABLE I
THE INTERNAL STATE SYSTEM

Motivation	curiosity, intimacy, boredom, avoidance, possession, control
Homeostasis	fatigue, sleepy, hungry
Emotion	happiness, sadness, anger, fear, neutral

RIT Lab's synthetic character, 'Rity,'uses the following architecture. Rity has 12 DOFs, 46 symbols and 73 behaviors. The sampling rate of the computational model is set to 0.1 sec, but it can be variable from 0.01 sec to 1 sec.

Rity is implemented in a 3D virtual space as shown in Figure 4. OpenGL is used for the implementation. Figure 5 shows Rity's emotional expressions according to the dominant emotion. Table I shows internal states used for Rity.

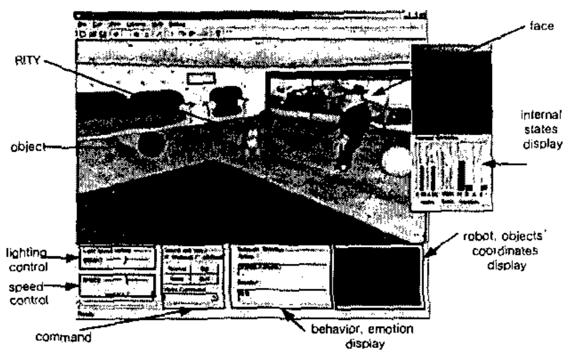
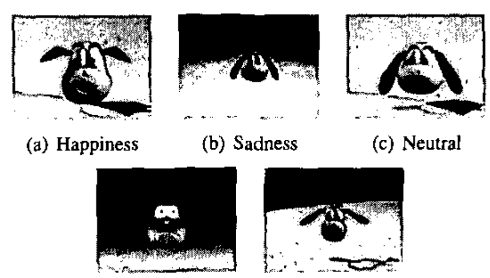


Fig. 4. Implementation of the synthetic character



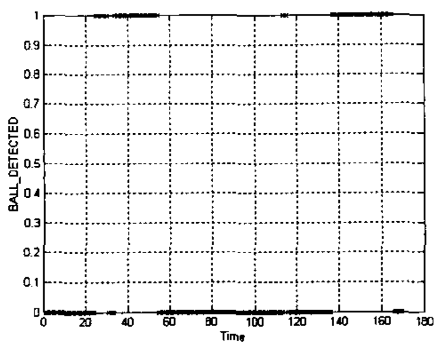
(a) Happiness (b) Sadness (c) Neutral

(d) Anger (e) Fear

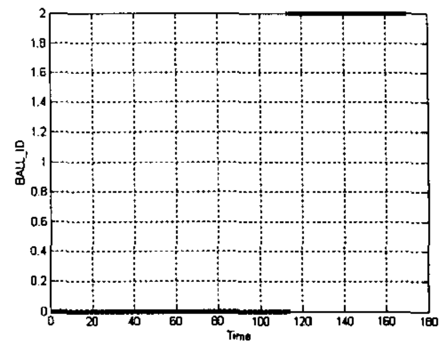
Fig. 5. RITY's facial expression

V. Results

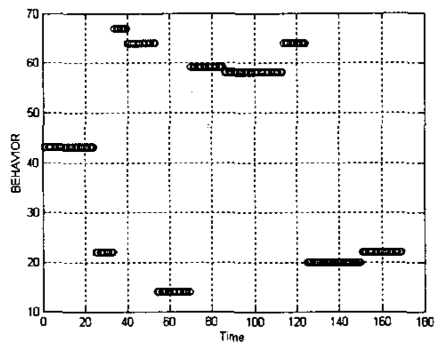
Figure 6 shows Rity's behaviors and emotion states when it interacts with balls. This shows the synthetic character, which uses the proposed behavior selection mechanism, can perform its behavior properly. In Figure 6, Rity was initially sad and stretched itself (the 43th behavior). When it found OBJECT1 (ball_id0) which Rity liked, Rity became happy. When Rity found OBJECT1 at first, it moved back from the object with



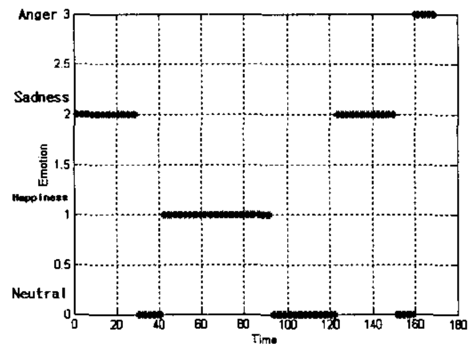
(a) BALL_DETECTED



(b) BALL_ID



(c) Behavior



(d) Emotion

Fig. 6. The behaviors and emotion states

fear (the 22th behavior). Rity looked for the ball (the 67th behavior), Rity approached to the ball (the 64th behavior) when it found the ball and opened its mouth with happiness (the 14th behavior).

Figure 7 shows that the learning method works well, and the synthetic character can learn what to do according to a command through interaction with the user. In this experiment behaviors, which are related to dancing motion, are grouped as follows:

$$B_{DANCE} = \{hurrah, shake\_head, mouth\_open, shake\_arms, dance\_with\_arms\}.$$

There are 10 analogous behavior sets as follows:

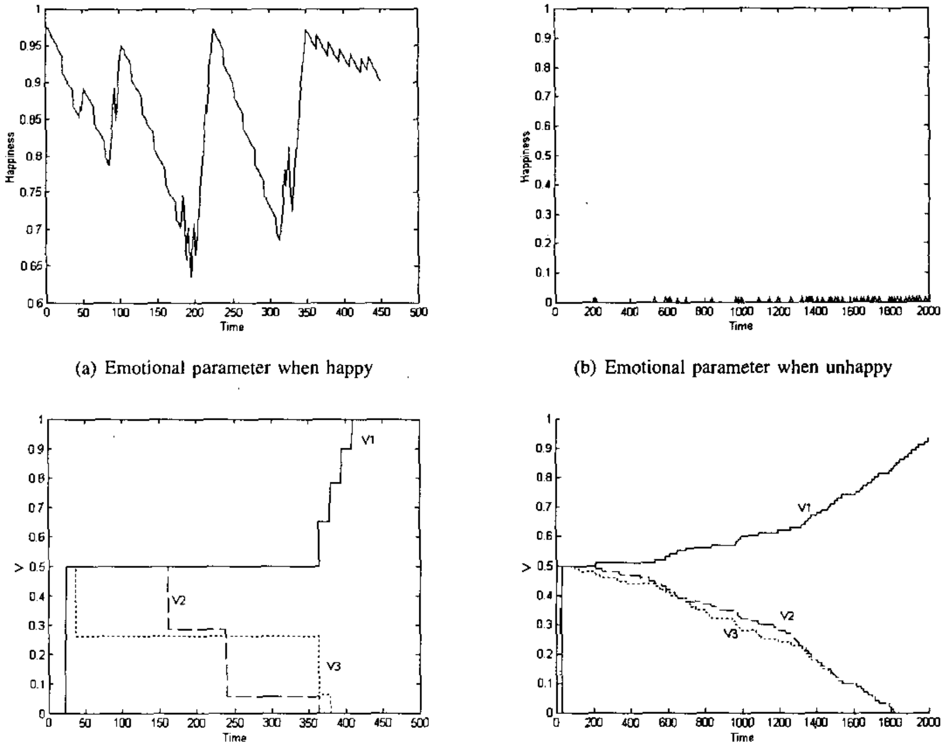
$$B = \{B_{STANDUP}, B_{SIT}, B_{COME}, B_{STOP}, B_{MOVE\_FORWARD}, B_{MOVE\_BACKWARD}, B_{MOVE\_LEFT}, B_{MOVE\_RIGHT}, B_{DANCE}, B_{others}\},$$

where each set contains similar behaviors to be learned together.

The emotion parameter ρ used a value of 'happiness' state and α was 100, β was 50. When Rity was happy, ρ had a value from 637 to 978 and when unhappy, the value was from 0 to 19. Figure 7(c) and 7(d) show the weight between commands and behavior sets. In this case, V_1, V_2, and V_3 were $\mathbf{W}_{MOVE\_LEFT,MOVE\_LEFT}$, $\mathbf{W}_{MOVE\_LEFT,MOVE\_RIGHT}$, and $\mathbf{W}_{MOVE\_LEFT,SIT}$, respectively. A goal of learning was to make V_1, from a 'go left ' command to 'go left ' behavior, its maximum value (1), and to make V_2 and V_3 its minimum value (0). When Rity was happy, the learning executed by 409 samples. There were three penalties and six rewards. When Rity was unhappy, the learning required over 2,000 sampling times and there were 58 penalties and 33 rewards. These results show that Rity can learn a proper behavior among several tens of behaviors by the interactive method and Rity's emotion influences the learning rate.

VI. Concluding Remarks

This paper proposed a novel method of behavior selection and its learning method for a synthetic character implemented as a software agent in the 3D virtual world. The proposed behavior selection method differs from previous approaches for behavior selection that make use of a tree structure. The combination of a deterministic behavior selection and a probabilistic behavior selection was a distinctive characteristic of the proposed behavior selection method. A user could also teach the character to do a desired behavior by an interactive training method. The synthetic character could learn a proper



(a) Emotional parameter when happy

(b) Emotional parameter when unhappy

(c) Learning when happy

(d) Learning when unhappy

Fig. 7. The emotion states and learning

behavior among many behavior pools. The learning process was influenced efficiently by the emotion state.

REFERENCES

[1] Yong-Duk Kim, Yong-Jae Kim, Jong-Hwan Kim and Jong-Rak Lim, "Implementaton of Artificial Creature based on Interactive Learning", *in proceedings of 2002 FIRA Robot World Congress*, pp. 369-373, 2002

[2] Breazeal, C. and Scassellati, B., "How to build robots that make friends and influence people", *IROS' 99*, pp. 858-863, 1998

[3] Breazeal, C., "Robot in society: Friend or appliance?", *Agents99 workshop on emotion-based agent architectures*, pp.18-26, 1999

[4] Velásquez, J. D., "An emotion-based approach to robotics", *Proceedings of IEEE Internaltional Conference on Intellighent Robots and Systems*, pp. 235-240 vol. 1, 1999

[5] Kubota, N., Nojima, Y., Baba, N., Kojima, F., and Fukuda, T., "Evolving pet robot with emotional model", *Proceedings of IEEE Congress on Evolutionary Computation*, pp. 1231-1237, 2000

[6] Miwa, H., Umetsu, T., Takanishi, A., and Takanobu, H., "Robot personality based on the equation of emotion defined in the 3d mental space", *Proceedings of IEEE International Conference on Robotics and Automation*, pp. 2602-2607, 2001

[7] Fujita, M., "Digital creatures for future entertainment robotics", *Proceedings of IEEE Internaltional Conference on Robotics and Automation*, pp. 801-806 vol. 1, 2000

[8] Yoon, S.-Y., Blumberg, B. M., and Schneider, G. E., "Motivation driven learning for interactive synthetic characters", *in Proceedings of Autonomous Agents*, 2000

[9] Isla, D., Burke, R., Downie, M., and Blumberg, B., "A layered brain architecture for synthetic creatures", *Proceedings of the International Joint Conferences on Artificial Intelligence*, pp. 1051-1058, 2001

[10] Ronald C. Arkin, Masahiro Fujita, Tsuyoshi Takagi and Rika Hasehawa, "Ethological Modeling and Architecture for an Entertainment Robot", *Proceedings of IEEE International Conference on Robotics and Automation ICRA*, pp. 453-458, vol.1, 2001

[11] Myung-Jin Jung, Arai F., Hasegawa Y., Fukuda T, "Mood and task coordination of home robotsRobotics and Automation", *Proceedings. ICRA '03. IEEE International Conference on*, p.250-255, Vol.1 , Sept. 14-19, 2003