

# 동기와 계층화된 감정에 기반한 로봇의 행동결정

## Model Behavior selection based on the motivation and hierarchicalized emotions.

안형철, 박명수, 최진영  
서울대학교 전기컴퓨터공학부

Hyoungh Chul Ahn, Myoung Soo Park, Jin Young Choi  
School of Electrical Engineering and Computer Science, Seoul National University  
E-mail : mspark@neuro.snu.ac.kr

### ABSTRACT

본 논문에서는 인간과 엔터테인먼트 로봇의 상호작용을 위해, 동기(motivation)와 계층화된 감정(hierarchical emotion)에 기반한 행동결정 모델을 설계하였다. 감정모델은 계층화되고 학습 가능하도록 하여, 인간의 행동결정과 유사하게 동작하도록 하였다. 감정모델을 통해 로봇의 행동에 대한 인간의 반응이 학습되는데, 그 결과가 행동결정에 영향을 주어 로봇의 행동에 반영되도록 하였다. 감정모델과 함께 동기가 행동결정에 영향을 주는데, 초기에는 외부에서 주어지는 동기가 주로 행동을 결정하고, 감정모델이 학습될수록 점차 감정의 영향이 증가하여 동기와 계층화된 감정을 함께 고려하여 행동을 결정하도록 하였다. 그럼으로써, 인간과의 상호작용을 통해 정보를 축적하고 인간의 반응에 적응해나갈 수 있게 하였다.

**Key words :** 인간과 로봇의 상호작용, 동기, 계층화된 감정, 로봇 행동선택, 감정 학습.

## I. 서 론

사람과 로봇의 상호작용은 로봇의 행동과 이에 대한 사람의 반응을 통해 이루어진다. 외부/내부적인 자극은 내부에 동기(motivation)라는 내부신호를 유발시키고, 동기에 따라 행동(behavior)이 선택된다. 센서를 통해 얻어지는 사람과 환경에 대한 정보는 외부적인 자극이고, 하드웨어의 상태, 예를 들어 충전지의 상태 등은 내부적인 자극이 된다. 로봇은 인간의 반응에 따라 적절한 행동을 선택하게 되고, 사람은 로봇의 행동을 원하는 방향으로 이끌기 위한 자극을 가함으로써 상호작용이 이루어진다.

일반적으로 이러한 상호작용은 반사적인 행동(reactive behavior)에 해당한다. 자극이 주어지면 이에 대해 미리 결정된 어떤 미리 행동이 고정적으로 선택되는 것이다. 그러나 상대방에 따라 이러한 행동의 선택은 적절하지 않을 수 있다. 어떤 사람의 경우는 보다 친근한 동작을 선호할 수 있고, 어떤 사람은 로봇이

가급적 조용하게 있는 것을 선호할 수 있다. 이러한 상대방에 대한 정보를 행동결정에 반영하기 위해 감정(emotion)을 도입할 필요가 있다. 사람이 공포스러운 상대방에게는 보다 조심스럽게 반응하는 것처럼, 로봇의 감정은 동기와 더불어 보다 상황에 적합한 행동결정에 영향을 준다.

이 논문에서는 로봇에 감정을 도입한 몇 가지 모델에 대해 살펴보고, 보다 다양하고 복잡한 반응에 응용할 수 있는 감정모델 및 행동결정모델을 제안하였다. 제안된 모델을 이용한 감정변화는 시뮬레이션을 통해 관찰하였다.

## II. 본 론

### 2.1 감정모델에 대한 기존 연구

Minsky[1]에 의해 인공지능에 있어서 감정의 중요성을 지적하였고, Damasio[2]는 신경학적인 증거를 통해 감정지능의 존재와 역할을 지적하였다. Goleman[3], Strongman[4],

Kleinginna 등은 감정을 정의하고자 하였고, Picard[5], Ortony[6], Rosman[7], Bates [8], Reily[9] 등은 감정 모델링을 시도하였다. 로봇의 행동과 관련하여 감정모델을 이용한 대표적인 연구로 Velasquez[10], Breazeal[11], Fujita[12]에 의한 감정모델들을 들 수 있다.

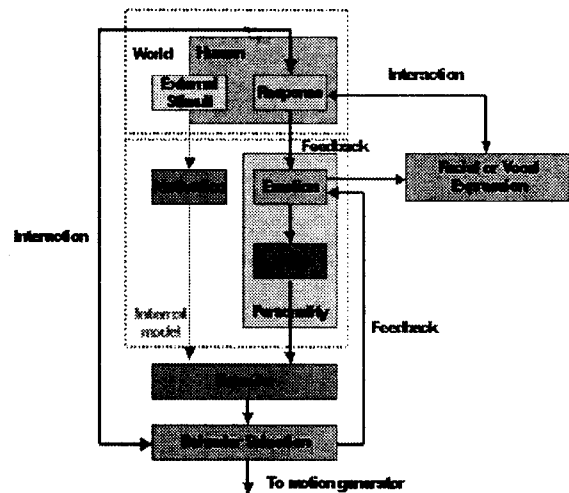
Velasquez는 SIMON이라는 가상환경 상의 캐릭터가, 즐거움/행복함, 고통/슬픔, 두려움, 노여움, 혐오, 놀람 등 여섯 가지 감정을 직접 얼굴의 모습으로 보여주게 하였다. 이러한 행동을 결정하기 위해 Cathexis 모델을 제안하였는데, 여기서 감정은 내부의 다섯 가지 동기(배고픔, 목마름, 온도유지, 피로, 흥미)와 사람에 의해 조작되는 가상환경의 변수에 의해 결정된다. 예를 들어 흥미라는 동기가 클 때, 장난감을 보여주면 SIMON은 행복한 감정을 느끼고 행동으로 표현한다. Breazeal의 경우는 Kismet이라는 로봇의 얼굴표정, 목, 귀의 움직임을 통하여 인간에게 의도를 표현하게 하였다. 카메라로부터 얻어진 영상으로부터 주의, 긍정적임, 개방적임의 3가지 요소를 결정하고 감정을 결정하게 된다. 욕구(사회적 욕구, 자극 욕구, 휴식 욕구)가 감정과 함께 동기를 결정하고 이에 따라 행동이 결정된다. Fujita의 경우는 Sony사의 엔터테인먼트 로봇 AIBO를 이용하였는데, 내부에 로봇의 본능에 관한 여섯 가지 변수(배고픔, 목마름, 배설욕구, 피로, 호기심, 애정)를 구성하고, 이 변수가 적정수준을 만족하면 감정이 유발되게 하였다. 그러면 동기가 결정되고 행동이 따라서 결정된다.

Velasquez와 Breazeal의 모델에서는 행동을 감정 각각에 대응하는 것으로 정의하였다. 따라서 행동에 대한 반응은 곧 감정에 대한 반응이 되어 상호작용이 이루어진다는 장점이 있지만, 감정에 대응하지 않는 행동들에 대해 적용하기는 쉽지 않다는 단점도 가지고 있다. Fujita의 모델에서는, 앞서의 모델에 비해서는 보다 일반적으로 감정에 대응하지 않는 행동들에 대해서도 감정이 영향을 주도록 할 수 있다는 장점이 있다. 그러나, 감정의 계산이 행동에 대한 반응이 아니라, 단순히 사용자의 외부 자극에 의해 이루어지므로, 상호작용이라기보다는 사용자의 일반적인 자극에 따라 감정이 변화하게 된다는 단점이 있다.

본 논문에서는 보다 적절하게 로봇과 인간의 상호작용이 가능하도록 하였다. 우선 감정이 외부의 (행동, 반응)이라는 입력 쌍에 의해 영향을 받도록 하였다. 또한 감정이 동기와 함께 행동결정에 영향을 주도록 하였다. 그럼으로써, 감정에 대응되지 않는 행동들에 대해서도 모델을 적용할 수 있게 하였다.

## 2.2 새로운 감정모델

Wilson[13]의 심리학적인 모델을 확장하여, 이 논문에서는 4가지 단계로 구성된 감정모델을 제안하였다. 감정은 각각 일시적인 감정(momentary emotion), 태도(attitude), 기분(mood), 그리고 성격(personality)로 구분된다. 일시적인 감정은 로봇의 행동과 이에 대한 사람의 반응에 따라 결정되는데, 이 값이 축적되어 태도와 기분을 결정한다. 성격은 기분의 변화 속도를 결정한다.



로봇이 가질 수 있는 일시적인 감정은 심리학적 연구에 기반하여 제안된 Ekman[14], Ushida[15], Miwa[16], Masuyama[17] 등의 모델과 마찬가지로, 슬픔(Sorrow), 노여움(Anger), 혐오(Aversion), 놀람(Surprise), 기쁨(Pleasure), 두려움(Fear)의 6가지로 정하였다. 이러한 일시적인 감정들은 4가지 행동, 즉 손을 내밌(Rai: Raise hand), 다가감(App: Approach), 회피함(Avo: Avoid), 꼬리를 흔들(Wag: Wagging tail)과 3가지 반응, 즉 로봇을 인식함(Recog: Recognize), 로봇을 때림(Attack), 보상을 줌(Have)의 쌍에 의해 결정되는데, 그 관계는 Masuyama의 논문에서 제안된 논리식에 기반하고 있다. 예를 들어, Masuyama의 논문에서 슬픔은 아래와 같이 계산된다.

$$\forall s, x, t_0, sorrow(s, t_0) \equiv \exists x [want(s, have(s, x), t_1) \wedge (t_1 \leq t_0) \wedge recog(s, -have(s, x), t_0)]$$

그러나, 이것은 감정을 제한된 경우에 한해서만 계산할 수 있도록 한다. 대부분의 상호작용에 있어서는 감정이 결정되지 않는다. 보다 많은 상호작용이 일어날 수 있도록, 모든 경우에 대해서 감정이 계산될 수 있도록 식을 아래와 같이 수정하였다.

$$\begin{aligned}
 f_{sorrow} &= 0.25App + 0.25Wag + 0.25(1-Re cog) + 0.25(1-Have) \\
 f_{anger} &= 0.25App + 0.25Wag + 0.25Re cog + 0.25(1-Have) \\
 f_{aversion} &= 0.25Rai + 0.25Wag + 0.25Attack + 0.25(1-Have) \\
 f_{surprise} &= 0.25App + 0.25Avo + 0.25Re cog + 0.25(1-Have) \\
 f_{pleasure} &= 0.25Rai + 0.25App + 0.25Re cog + 0.25Have \\
 f_{fear} &= 0.25App + 0.25Avo + 0.25(1-Re cog) + 0.25Attack
 \end{aligned}$$

각각의 경우에, 입력변수들은 해당하는 행동과 반응이 일어난 경우에 1, 그렇지 않은 경우에 0이 된다.

일시적인 감정들은 태도에 축적된다. 태도는 각각의 감정에 대해 정의되며, 그것은 장시간에 걸친 일시적인 감정들의 평균에 해당한다. 예를 들어 슬픔에 의한 태도는 아래의 식에 의해 결정되며, 다른 감정들에 대해서도 각각 같은 방법으로 계산된다. 아래 식에서 k는 일시적인 감정을 태도에 반영한 반복횟수를 의미한다.

$$Sorrow_{k+1} = \frac{f_{sorrow} + k \times sorrow_k}{1+k}$$

일시적인 기분들은 또한 기분으로도 축적된다. 로봇의 기분은 크게 호기심(curious), 소심함(timid), 친근함(friendly)이라는 세 가지가 있는데, 기분의 전환을 위한 조건은 아래와 같다.

$$\begin{aligned}
 &curious \rightarrow timid \quad \text{if } (Timid\_condition > W \cdot EMO_{average} \\
 &\quad \wedge Timid\_flag = 1) \\
 &curious \rightarrow friendly \quad \text{if } (Friendly\_condition > W \cdot EMO_{average} \\
 &\quad \wedge Friendly\_flag = 1) \\
 &timid \rightarrow curious \quad \text{if } (Curious\_flag = 1) \\
 &friendly \rightarrow curious \quad \text{if } (Curious\_flag = 1) \\
 &\text{where } EMO_{average} = (Sorrow + Anger + Aversion \\
 &\quad + Surprise + Pleasure + Fear)/6 \\
 &timid\_condition = (f_{sorrow} + f_{aversion} + f_{fear} + f_{anger})/4 \\
 &friendly\_condition = f_{pleasure}
 \end{aligned}$$

Timid\_flag는 슬픔, 혐오, 두려움, 노여움 중의 한 가지 감정이 연속적으로 H번 유발되면 1이 되고, Friendly\_flag는 기쁨의 감정이 H번 연속적으로 유발되면 1이 된다. Curious\_flag는 기분이 소심함일 때 친근함을 유발하는 감정이, 혹은 반대로 기분이 친근함일 때 소심함을 유발하는 감정이 연속적으로 H번 유발되면 1이 된다. 그렇지 않을 경우에는 0이 된다.

W과 H는 로봇의 성격(personality)에 해당하는데, 기분의 변화 속도를 조절한다. W가 크거나 H가 작으면 빠르게 기분이 변하고, W가 작거나 H가 크면 기분이 느리게 변한다. 또한 H는 잡음에 의한 기분의 채터링(chattering)을 방지하는 역할도 한다.

이상에서 정의된 일시적인 감정, 기분, 태도, 성격의 세 가지는 그것이 변하는 시간간격에 따라 구분되며 함께 결합되어 계층화된 감정을

구성한다. 일시적인 감정은 짧은 시간에 변하고, 기분과 태도는 장시간에 걸쳐 변하며, 성격은 처음부터 결정되며 변하지 않는다.

### 2.3. 감정과 동기에 의한 행동결정 모델

계산된 감정과 동기에 따라 값 행동에 해당하는 변수들의 값이 계산된다. 그리고 제일 큰 값을 가지는 변수에 해당하는 행동이 선택되어 실현된다.

행동은 크게 동기와 감정에 기반하여 다음 식에 따라 계산된다. k는 시간(반복횟수)에 해당한다. 첫 번째 항은 동기에 관한 항이며, 두 번째 항은 감정에 관한 항이다. 초기에는 동기에 의한 영향이 크고, 시간이 흘러 다양한 경험에 의해 정보가 감정에 축적된 뒤에는 일정한 정도까지 감정의 영향이 반영되도록 하였다. 우선 감정에 의한 항은 태도에 기분에 해당하는 mode를 변수로 하는 행렬 M을 곱하여, 따라서 기분에 따라 다르게 행동에 다른 영향을 주게 하였다. 소심한 기분일 때는 소심한 행동들이, 친근한 기분일 때는 친근한 행동들이 더 많이 선택되도록 설정되었다.

$$Action_k = \frac{1+\alpha k}{1+k} MOT_k + \frac{(1-\alpha)k}{1+k} M(mode) \cdot EMO_k$$

$$\text{where } Action_k = \begin{bmatrix} Raise_k \\ Approach_k \\ Avoid_k \\ WaggingTail_k \end{bmatrix} \quad MOT_k = \begin{bmatrix} Play_k \\ Affection_k \\ Selfdefense_k \\ Satisfaction_k \end{bmatrix}$$

$$EMO_k = \begin{bmatrix} Sorrow_k \\ Anger_k \\ Aversion_k \\ Surprise_k \\ Pleasure_k \\ Fear_k \end{bmatrix}$$

동기에 관한 첫번째 항은 네 가지 동기 각각에 대한 식에 의해 결정되는데, 각각의 동기는 대응하는 외부자극에 영향을 받는다. Play는 인간이 보였을 때, affection은 친구가 나타났을 때, self defense는 적이 나타났을 때, satisfaction은 먹이가 보였을 때 각각 영향을 받는다. 이 영향을 고려하여, Play라는 동기는 아래 식과 같이 결정된다. 첫 항은 MIN값과 MAX값 사이의 난수를 의미한다. 다른 동기도 같은 식에 의해 계산된다.

$$Play_k = \begin{cases} rand(MIN_k, MAX_k) & \text{when no external stimuli} \\ rand(MIN_k, MAX_k) + MID_k & \text{otherwise} \end{cases}$$

$$MIN_{k+1} = MIN_k + CONVERGENCE\_GAP$$

$$MAX_{k+1} = MAX_k - CONVERGENCE\_GAP$$

이와 같이 동기와 감정에 대한 항을 이용하여 행동을 결정한다. 초기에 감정이 충분히 축적되기 전에는 다양한 경험을 위해 임의로 동

기가 선택될 여지를 부여하다가, 어느 정도 시간이 흐른 후에는 외부의 자극에 의해서만 동기가 결정되도록 하였다.

### 2.4 모의실험

제안된 감정모델과 이를 이용한 행동결정모델을 이용하여 모의실험을 수행하였다. 처음에는 감정을 배제한 동기에 의한 행동선택을 살펴보고, 다음에는 감정이 개입되었을 때 어떻게 행동선택이 영향을 받는지를 확인한다. 또한 성격에 변화를 줌으로써 기분전환에 미치는 효과를 확인한다.

#### 2.4.1. 공통된 설정.

외부에서 50번에 걸쳐서 자극을 제공하고 감정을 학습하는 것으로 한다. 그리고 동기 계산시에 사용되는 변수 MIN, MID, MAX는 각각 0.4, 0.3, 07를 값으로 가지는 것으로 한다. CONVERGENCE\_GAP은 0.01로 설정하였고, 따라서 수렴까지의 학습횟수는 15회가 된다. 행동계산에서의 변수 alpha는 0.4로 설정한다. 성격과 관련된 W\_M과 히스테리시스 구간 H는 각각 중간값이 0.85와 5로 설정하였다.

#### 2.4.2. 동기에 의한 행동결정

그림 1(a)에서와 같이 동기를 주고, 이에 따라 결정되는 행동을 관찰하여 그림 1(b)의 결과를 얻었다. 감정이 배제된 상태에서의 행동은, 동기에 따라 고정적으로 선택된다.

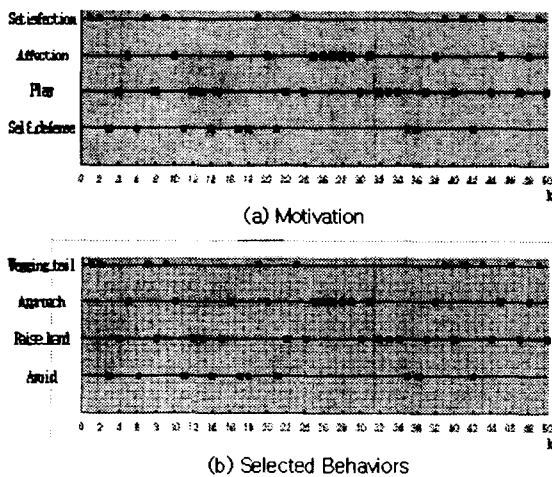


그림 1. 동기에 의한 행동선택

#### 2.4.3. 동기와 감정에 의한 행동결정.

그림 2의 아래 표에 나온 것과 같은 (행동, 반응)쌍을 통해 감정을 조정하여, 그림 2의 위 그래프와 같이 감정이 변화하도록 하였다. 그리고, 이와 같은 감정변화에 따라 기분과 행동선택이 어떻게 변화하는지를 그림 3에 나타내

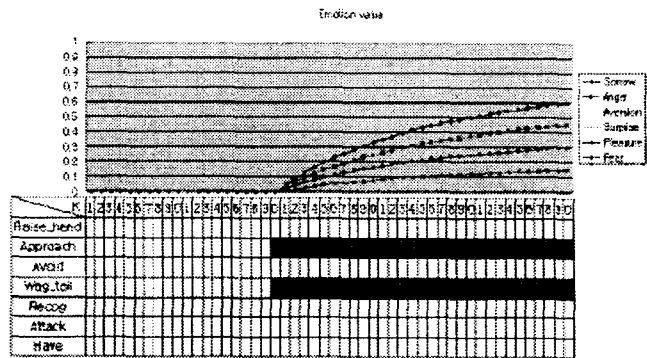


그림 2. 감정설정을 위한 (행동, 반응) 쌍

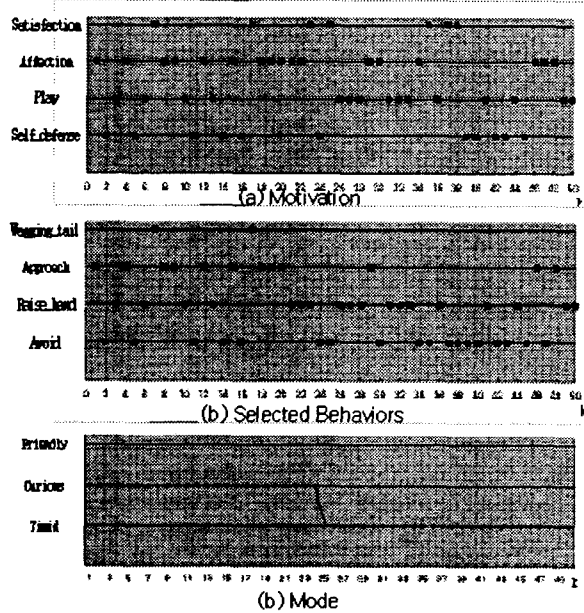


그림 3. 동기와 감정에 의한 행동선택

었다. 슬픔의 감정이 커짐에 따라 기분은 호기심에서 소심함으로 차츰 변화하는데, 이에 따라 행동선택이 그림 1의 경우와 달라지는 것을 알 수 있다. 20회까지는 동기에 따라 행동이 고정적으로 선택되다가, 20회 이후부터는 점차 소심한 행동들, 즉 회피가 더 자주 발생하게 되는 것을 알 수 있다.

#### 2.4.4. 다른 성격을 가지는 경우의 기분전환.

그림 6(a)의 각각의 표는 동일한 (행동, 반응)쌍 하에서 다른 W\_M과 H의 값에 따라 기분이 다른 속도로 변하는 것을 나타낸 것이다. W와 H가 각각 0.85와 5인 (a)의 경우에 비해 0.75와 2인 (b)의 경우는 기분변화가 빠르게 일어나는 반면, 1과 8인 (c)의 경우는 지나치게 기분변화속도가 느려서, 중간 단계의 기분변화가 생략되는 것을 확인할 수 있다.

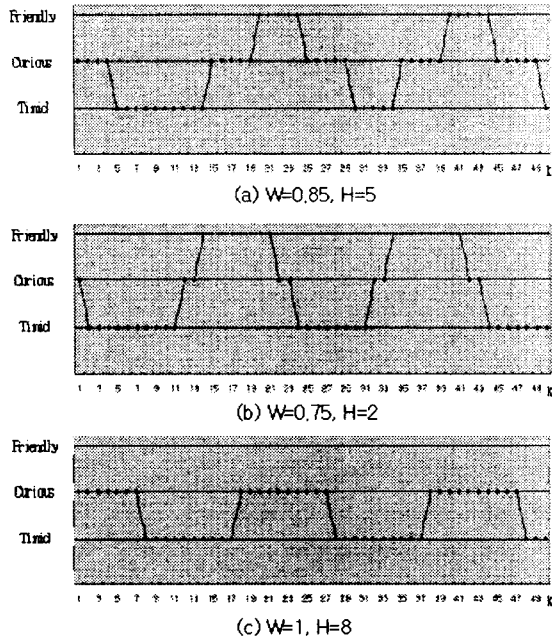


그림 4. 서로 다른 성격에 의한 기본변화

### III. 결 론

본 논문에서는 사람과 로봇의 보다 복잡한 상호작용을 구현하기 위하여 동기와 감정에 기반한 행동결정모델을 제안하였다. 기존의 모델들에 비해 다양한 행동에 감정이 반영될 수 있게 하였고, (행동, 반응)의 쌍에 기반하여 감정이 변화하도록 함으로써 상호작용이 일어나도록 하였다. 감정을 크게 일시적인 감정, 기분, 태도, 성격으로 나누어 장기적/단기적인 감정들의 영향을 고려하고, 성격에 따라 기분변화의 속도를 조절함으로써 로봇의 특징을 설정할 수 있도록 하였다. 모의 실험을 통해 목표로서 하는 반응을 얻을 수 있음을 확인하였다.

### IV. 참고문헌

[1] Stork, D.G., Scientist on the Set : An Interview with Marvin Minsky, In Stork, D.G, ed, HAL's Legacy : 2001's Computer as Dream and Reality, 15-31, MIT Press, 1997.  
 [2] Damasio, A.R., Descartes' Error : Emotion, Reason, and the Human Brain, NY : G.P. Putnam, 1994.  
 [3] Goleman, D., Emotional Intelligence, NY : Bantam Books, 1995.  
 [4] Strongman, K.T., The Psychology of Emotion, John Wiley & Sons, Third Edition, 1987.  
 [5] Picard, W., Affective Computing, MA. : MIT Press, 1997.  
 [6] Ortony, A., Clore, G., and A. Collins. The

Cognitive Structure of Emotions, Cambridge : Cambridge University Press, 1988.  
 [7] Rosman, J., Jose, P.E., and Spindel, M.S., Appraisals of Emotion-Eliciting Events : Testing a Theory of Discrete Emotions, Journal of Personality and Social Psychology 59 : 899-915, 1990.  
 [8] Bates, J., The Role of Emotion in Believable Agents, Communications of the ACM 37 : 122-125, 1992.  
 [9-2] Reilly, S., Believable Social and Emotional Agents, Ph.D, diss., Dept. of Computer Science, Carnegie Mellon Univ., 1996.  
 [10] Velasquez, J., Modeling Emotions and Other Motivations in Synthetic Agents, In Proceedings of the AAI Conference, 10-15, RI, USA, 1997.  
 [11] Breazeal, C., Scassellati, B., How to build robots that make friends and influence people, Proceedings of the International Conference on Intelligent Robots and Systems, 1999.  
 [12] Arkin, R.C., Fujita, M., Takagi, T., Hasegawa, R., An Ethological and Emotional Basis for Human-Robot Interaction, Robotics and Autonomous Systems 42 : 191-201, 2003.  
 [13] Wilson, I., The Artificial Emotion Engine, Driving Emotional Behavior, In AAAI Spring Symposium on Artificial Intelligence and Interactive Entertainment, 2000.  
 [14] Ekman, P. and Davidson, R., The Nature of Emotion : Fundamental Questions, Oxford University Press, New York, 1994.  
 [15] Ushida, H., Hirayama, Y., Nakajima, H., Emotion Model for Life-like Agent and Its Evaluation, In Proceedings of the Fifteenth National Conference on Artificial Intelligence, 62-69, Menlo Park, Calif : AAAI Press, 1998.  
 [16] Miwa, H., Takanishi, A., Takanobu, H., Development of a Human-like Head Robot WE-3RV with Various Robot Personalities, Proceedings of IEEE-RAS International Conference on Humanoids Robots : 117-124, 2001.  
 [17] Masuyama, A., A Number of Fundamental Emotions and Their Definitions, In Proceedings of IEEE Communication, 156-161, Tokyo, Japan, 1994.