Emotions in Robots

ABSTRACT. In this paper the emotion model for robots is developed. In order to create the GAI, the emotions in robots should be close to human.

CURRENT STATE. At current there are many papers on human emotions, and less – in robots. In [1] and [2] the emotion model for robots is developed. In accordance with [2] and [6], the emotions in robots can be created in accordance with the formula

$$E(t)_i = W_{0i} + \sum_{j=1}^{t} \gamma_i^{\,t-j} (W_{1i}R^+_{ij} + W_{2i}R^-_{ij}) \qquad (1)$$

where there are six emotions (Fear, Anger, Sadness, Happiness, Disgust, Surprise), each denoted by the subscript i, and there is exponential decay of rewards also.

The subscript j denotes a time instance. R^+ and R^- denote positive and negative rewards, W_{0i} – the steady state of happiness, γ – rate of decay, W_1, W_2, W_3 – magnitude of change

The problems. Models mentioned have little match with reality: there are no current external world EW events, which induce emotions, and there are no individual's values and needs or expectations, created in a lifetime.

SOLUTION. In humans the emotions arouse when they experience EW events. (Here the important notice is appropriate: in accordance with Jeff Hawkins [3] the individual's body for the brain is also the EW. This means that processes in neocortex, limbic brain and in the body create and influence the emotions).

When EW events and processes are experienced, they are confronted with the individual's values, needs and expectations, and the emotions arise.

DEFINITION. Emotions are the brain states and processes, which determine individual's decisions and behavior. First, there are body feelings [4], which often create the quick and unconscious reactions. The sum of body feelings [4] and memories of previous events create emotions.

When events happen, human mind evaluates their utility, joins, associates to every event some value, and the emotions are created. Emotions E emerge some time after the events and decay exponentially, in neural networks for neuron activation the sigmoid-function is used. Rise of emotions is described by the term $1/(1+e^{-\alpha(t-d)})$, where α is the velocity of emotion rise, the decay is described by the term $-1/(1+e^{-\beta(t-\Delta t-d)})$, β is the velocity of decay, t-d is the time after the event, Δt characterizes the emotion duration, but d is time after the previous events, from which some residual level of emotions is remained

$$E(t) = 1/(1+e^{-\alpha(t-d)}) - 1/(1+e^{-\beta(t-\Delta t-d)}) \qquad (2)$$

$$1/(1+exp(5-2x))-1/(1+exp(12-x))$$

http://www.webmath.com
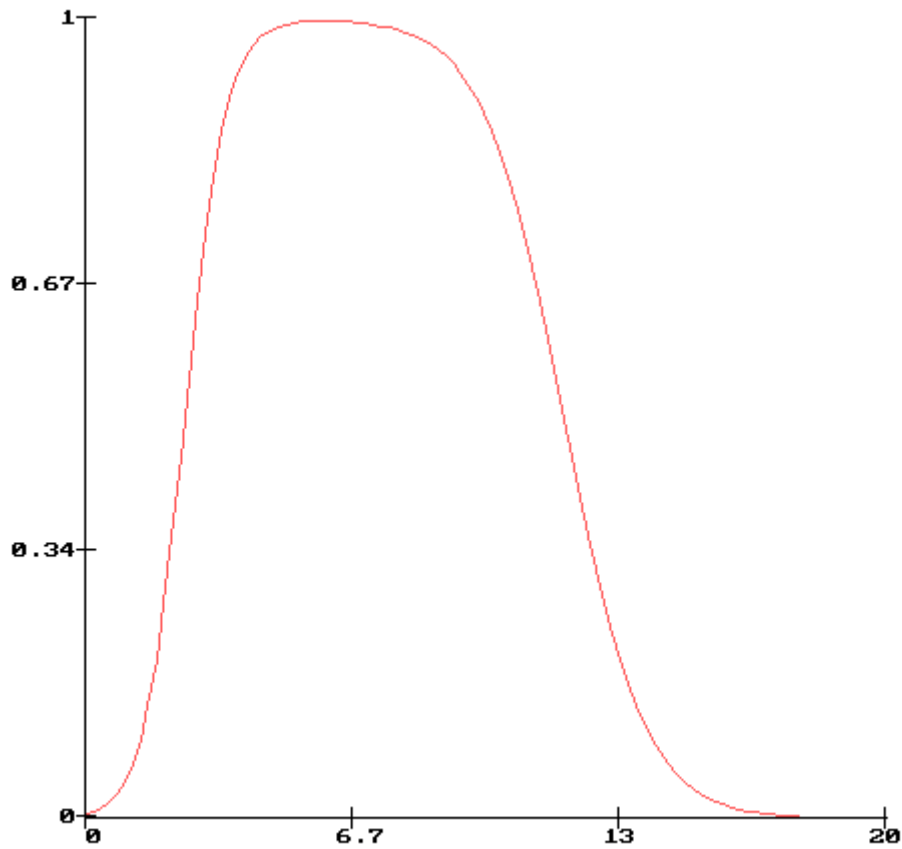
Here it is assumed that t-d = 0 at the start of monitoring, it is the start of robot's 'life', and the time interval [d,…,t] is the time for which the emotions are monitored.

The total emotions (Fear, Anger, Sadness, Happiness, Disgust, Surprise), induced by j-th event, is the sum of all exponents, which are initiated by the events at different times d, … t_j, …t.. For simplicity the summing for different types of emotions is omitted. Then the formula for emotions induced by different events j in robots is

$$E(t)=E_0+\sum_j (1-e^{-\alpha_j(t-t_j)})\cdot e^{-\beta_j(t+\Delta t -t_j)}\cdot[N_j(I_a - I)] \quad (3)$$

Here

I_a – the actual information individual has after the event j,

$$I_a=\log_2 1/P_j,$$

I – the mean value of information (considered to be necessary for satisfying the need N_j), the arithmetic mean of probabilities of all events in time interval 0…t,

$$I = \sum_j P_j \log_2 1/Pj.$$

E_0 – some residual level from previous events, the steady state of each emotion,,

$0<N_j<1$, - the need, the expectation for the j-th event from Maslow's pyramid. Emotions are induced by events, they are connected to events, they don't exist without the cause – the events or the processes.

Every type of emotion provoked at a time t_j has its own rate of decay, depends on the need N_j for the event j, the amount of actual information, individual has after the event, and the amount of information, necessary for satisfying the need. Emotions are positive when both $N>0$ and $I_a - I >0$, or when $N<0$ and $I_a - I <0$. The last state means that individual has positive emotions when bad event has not happened.

DISCUSSION. The formula (3) isn't new. In the last century Russian MD V. Levi in his popular-science books wrote that the emotion E equals to the need N, multiplied by the difference of two sets of information I-i, where I – the information the individual has, and the i – the information, necessary for satisfying the need: $E=N(I-i)$. The individual assigns needs or expectations to EW events. Analysis of the formula (3) shows some commonly known truths:

- when the number of events is high, the probabilities of events are small, the information connected to each event is big and so are the emotions $I_a=\log_2 1/P_j$; in other words, people who experience many events, have emotionally rich lives;
- all events induce positive or negative emotions, if their expectations (needs) are small, they create small emotions. If the events are survival-useful or happiness-supporting, the expectations are positive and, if they are important, then close to 1. If the events are harmful and important, their expectations are negative and close to -1.
- from the formula (3) we can notice the commonly known effect of emotion accumulation or overlapping, described in [5]: next events add their emotions while the previous ones have not gone.
- everyone can manage his/her emotions. When the EW events and processes are not changeable, the only one way to manage emotions is to create, to develop or suppress the needs and expectations.

SUMMARY. In order to create the GAI in robots, we need to create the emotions possibly close to the real humans. The created formula mirrors the individual's experience in a real-life conditions, allows to pre-program the libraries of needs and values in robots and attain, create these libraries in robot's real life conditions.

THE PROBLEMS. Rather smart task will be associating the right emotions to current body feelings and the memories from previous events. This task can be solved only by monitoring and measuring the human feelings and corresponding emotions. How the emotions form values and influence decisions and corresponding behavior still are to be discussed.

REFERENCES

[1]. Robb B. Rutledge [a,b,1], Nikolina Skandali [a], Peter Dayan [c], and Raymond J. Dolan [a,b], A computational and neural model of momentary subjective well-being, 2014.
[2]. Lyle N. Long, Troy D. Kelley, Eric S. Avery, **An Emotion and Temperament Model for Cognitive Mobile Robots,** 2015.
[3]. Jeff Hawkins, On Intelligence.
[4]. Mark Solms, The Hidden Spring, A Journey to the Source of Consciousness, Profile Books, London, .
[5]. Thomas Lewis, Fari Amini, Richard Lannon, A General Theory of Love.
[6]. http://www.personal.psu.edu/lnl/papers/Long_Lyle_636.pdf