TELKOMNIKA, Vol. 11, No. 8, August 2013, pp. 4498~4504
e-ISSN: 2087-278X

Artificial Emotion Engine Benchmark Problem based on Psychological Test Paradigm

Wang Yi*, Wang Zhi-liang
School of automation and Electrical Engineering, University of Science and Technology Beijing
Beijing, China, 100083
*Corresponding author, e-mail: bkcool@163.com

Abstract

Most of testing and evaluations of emotion model in the field of affective computing are self-evaluation, which aims at the application-specific background, while the research on the problem of the Benchmark emotional model is scarce. This paper firstly proposed the feasibility of making psychological test paradigm a part of artificial Benchmark engine, and with taking versatility and effectiveness as the evolutional factor to judge the engine by testing psychological paradigms. In addition, an emotional hidden Markov model is built and tested based on the Benchmark theory. The detailed simulation process of the experiment is given. The testing resultants are coincide with the real world's situation.

Keywords: *affective computing, artificial emotion, benchmark, hidden markov model*

1. Introduction

The first one who proposed the idea of making computers with artificial emotion is American professor Minsky In 1985 in his book "The society of Mind". The question is whether intelligent machines can have some emotions, and which can be intelligent as humankind. the problem has become a classic referenced model by most researchers in this field. In 1995, Rosalind W. Picard from MIT Media Lab published his original monograph becomes a kind of early form of affective computing technology. After several years, affective computing and artificial emotional issue attracted the attention of the world. Since 1990s, Japanese have begun the research work of Kansei Engineering, which as a kind of technological science, combines human emotion with engineering together to design goods manufacturing.

Nowadays, Japan has formed the most advanced system of Kasey engineering. In 2000, Professor Wang Zhi-liang from University of Science and Technology Beijing in China developed the concept of artificial psychology, which is a new exploration in the field of affective computing. Many associations of China have started doing the relevant research and reached higher level, such as the celebration of the first session of affective computing intelligent interaction international academic conferences and the establishment of artificial psychology and artificial emotion Professional Committee of China Association for Artificial Intelligence indicates.

Because the emotional thinking is measurable, there are many researchers try to simulate the generation and changes of mood from different disciplines that psychology, cognitive science and information science included [1]. Meantime, due to the complexity of the emotions as well as the imperfectness of the research of humanity emotional variation, this work is complicated, and there is also a number of emotional theory buildings coexist. Currently, there have been a lot of models of the mood. Of course, it is not easy to make these models achieve an exact quantitative description of human and animal emotions perfectly, but at least some of the models achieve a limited mimicry from the functional angle. For example, the OCC emotion model, this was put forward by Ortony, Clore and Collins at The Cognitive Structure of Emotions in 1988. It is an original emotional cognitive model and is the first one of the practical models towards the studies of human emotions, which is not only cater to the requirement of Computer-implemented development, but also summarizes and concludes the corresponding relation by using a regular model. According to analysis of various events, a series of emotion is triggered by other entity interaction [2]. The OCC model divided the course

Received February 16, 2013; Revised May 15, 2013; Accepted May 25, 2013

of emotion into three categories: the results of the event, the action of the agent and the perception of the object. The Kismet emotional model [3] that is designed by the MIT C.Breazeal to apply to a humanoid robot Kismet is applied to a humanoid robot Kismet. The model is constituted by four parts: emotional stimuli, emotional evaluation, emotional activation and emotional expression. European state space-based emotion model is an emotional space which represents the basic emotions' factors as several elemental vectors, emotional state described as discrete points in this state space [4]. Affective model [5] based on the probability space provides a new method for affective computing machine automatically. Theoretical study generated through Markov Chain and the HMM emotional transfer simulation consists of four steps that the transfers of mood stimulate the mood spontaneous transfer, transfer of emotional stimuli and emotional spontaneous metastasis. The emotional model as an artificial emotion engine plays a key role in the function of making intelligent machines intelligent. But evaluation factor of the engine is still not exact enough. This evaluation process needs a set of generic standards as a foundation for the purpose of getting the judge results of analysis and synthesis. That's to say the evaluation of the performance of the Affective generation engine needs a group of studies on Benchmark problem.

2. Relative Works

The so-called Benchmark is a matter of Baseline norm, which is used to establish a set of relatively impeccable models. For high efficient tests and evaluation system under the same issues, a common platform for making comparisons among different kinds of models and systems is built. Nowadays, Benchmark issue has reached many scientific fields. For example, the structural damage identification technology is a principal segment of structural health monitoring (SHM). With deeply improve the capability of these methods the experimental costs decrease. In the International Workshop on Structural Control conference, the research group consisting of Europe, Asia and USA has been proposed dealing with the SHM and the Benchmark structure has been set up by Chen at the same time in order to make direct comparisons among various technologies. The Benchmark structure was raised by Black Ventura after the American SHM research group was set up under the joint-funding of IASC and ASCE. In 1999 for the first time IAC&ASCE research group resolved to focus on the establishment of a well-defined Benchmark issue of the American group thus can do comparative studies about diverse structural damage identification methods by optimizing tools of test data and structural model calculation tools. There is also Benchmark issue in the current testing of steam generator tube research areas. The part-of-speech tagging Benchmark uses CRF++ model and Pocket CRF model. Benchmark, used as an evaluation method in the computer field has a long-term application, and is widely used in the in the hardware such as CPU evaluation, memory, I/O interfaces and peripherals software for operating system evaluation, middleware and application software in machine learning field, which enhanced data processing capability of the database [6]. University of California Irvine (UCI) date sets include 189 groups which contribute a lot to the test and post-evaluation of machine learning algorithms. In the combinatorial optimization issue, the design of broadcasting networks, the switching circuit design, ship transportation route plans, work assignments, goods packing scheme, the shortest path issue, the maximum (minimum) spanning tree issue, the best edge unrelated set, minimum cut sets as well as salesman issue are all belong to Benchmark issues in broad sense. Especially, in the smart calculation field, Levy No. 5 Function, Shaffer's F6 Function, Six-Hump Camel-Back Function and Generalized Schaeffer's Problem are continuous Benchmark issues [7].

It can be seen that in many scientific research fields, Benchmark issue can be interpreted as a model, algorithm, system test and evaluation form. What is more, emotional model studies in current emotional computers fields are mainly for specific application backgrounds such as (Human Robot Interaction, HRI), (Human Robot Cooperation, HRC) and Humanoid Robot and so forth. The reason why we haven't seen pros and cons evaluation outcome of the universe Benchmark issue is that this kind of research is still rare. In view of this situation, in this paper we will use the psychological test paradigm as a part of artificial emotion engine Benchmark issue and use versatility and effectiveness as an evaluation index. Finally, this set of Benchmark issues should be open and have scalability.

The paper is organized as follows: in the third section originally studies the testing of this Benchmark issue on the basis of introducing some Psychological experimental paradigms emotion-generating model. In the section four its effectiveness through experiments is verified.

3. Artificial Emotion Engine Benchmark

Emotional model, as an artificial emotion engine Based on a Psychological Test Paradigm, which plays a key role in the function of putting intelligence into machines but their performance evaluation studies are immature, so in this paper we adopt the new evaluation methods of models by using psychological experimental paradigm as the evaluation standard of models. However, as an artificial emotion engine, the emotional model plays a key role in the function of putting intelligence into machines. Their performance evaluation studies are immature [8].

3.1. Psychological Experimental Paradigms

Recently, the psychology experimental paradigms[9], including object reversal, children's gambling task, delay of gratification, windows task, self-ordered Pointing (SOP or SOTP), tower of Hanoi or tower of London Stroop (Day-Night Stroop) stop-signal task, hand game, Wisconsin card sorting test (WCST), flexible item selection task (FIST), dimensional change card sorting (DCCS), latent variable analysis and so on. In order to make preparations for the follow-up tests, we only focus mainly on the Children's Gambling Task .

The children's gambling task as a method is first created by Bechara to measure the hot executive functions. But, due to its complexity, the method is simplified and improved by Kerr and Zelazot in two aspects. In the task, two decks of cards are used within one set vertical bar patterns and the other deputy polka dots. Happy and sad faces can be seen on the both wrong and right sides of them. But what is different is that the former one deck of cars is always accompanied by happy faces which are occasionally coupled with a sad face. By contrast, the latter one deck of cars has two happy faces on the opposite sides. Sometimes, there will be a few sad faces (4, 5, 6 faces etc.). Happy and sad faces represent getting candies or losing candies separately. Faces number is equal to the number of the candies. Apparently, for each test can only select one card can we get only one candy by choosing vertical-bar pattern cards. On the contrary, by choosing polka-dot cards we get two candies, result in losing more averagely, about four, five, or six candies are lost by a single failure. Consequently, the vertical cards overweight the polka-dot ones in the long run. So the researchers tell the children to get as many candies as they can at the end of the game (under the condition of 50 selections are unknown by children in advance).

3.2. Emotional Model

In some literatures [10] , on the basis of the probability space, Emotional change process as a random process, the stimuli transfer process of it can be seen as an available description of the hidden Markov model (HMM). Its form is:

$$\lambda = \left(N, M, \hat{\hat{\mathbf{P}}}^{\mathbf{0}}, \hat{\mathbf{A}}, \hat{\mathbf{B}} \right)$$

Where, N is the emotional dimension is the stimulus types; $\hat{\hat{\mathbf{P}}}^{\mathbf{0}}$ is the initial state probability distribution vector; the matrix B is the state transition probability matrix, which is used to describe of the state transition probability. $\hat{\mathbf{A}}$ is the evaluation value, it can be descript as follow:

$$\hat{\mathbf{A}} = \{\hat{a}_{ij}\}_{N\times N} = \begin{bmatrix} \frac{\hat{\theta}\hat{\pi}_1^* - (N-1)}{\hat{\theta}\hat{\pi}_1^*} & \frac{1}{\hat{\theta}\hat{\pi}_1^*} & \cdots & \frac{1}{\hat{\theta}\hat{\pi}_1^*} \\ \frac{1}{\hat{\theta}\hat{\pi}_2^*} & \frac{\hat{\theta}\hat{\pi}_2^* - (N-1)}{\hat{\theta}\hat{\pi}_2^*} & \cdots & \frac{1}{\hat{\theta}\hat{\pi}_2^*} \\ \vdots & \vdots & \cdots & \vdots \\ \frac{1}{\hat{\theta}\hat{\pi}_N^*} & \frac{1}{\hat{\theta}\hat{\pi}_N^*} & \cdots & \frac{\hat{\theta}\hat{\pi}_N^* - (N-1)}{\hat{\theta}\hat{\pi}_N^*} \end{bmatrix} \quad (1)$$

Where, $\hat{\theta}$ is an undetermined parameters, $\hat{\hat{\boldsymbol{\pi}}}^{*} = \left[\hat{\pi}_1^{*} \quad \hat{\pi}_2^{*} \quad \cdots \quad \hat{\pi}_N^{*}\right]$ is a limiting probability, $\hat{\mathbf{B}}_{M\times N}$ is observation probability matrix. Here, when M is equal to N. Considering of the matrix $\hat{\mathbf{B}}_{M\times N}$ is invertible. It can be obtained as:

$$\hat{\mathbf{B}} = \begin{bmatrix} \hat{\hat{\mathbf{B}}}_{1\bullet} \\ \hat{\hat{\mathbf{B}}}_{2\bullet} \\ \vdots \\ \hat{\hat{\mathbf{B}}}_{N\bullet} \end{bmatrix} = \begin{bmatrix} a & b & \cdots & b \\ b & a & \cdots & b \\ \vdots & \vdots & \cdots & \vdots \\ b & b & \cdots & a \end{bmatrix} \tag{2}$$

Where a, b can be expressed as follow:

$$\begin{cases} a = \dfrac{r}{N-1+r}, \\ b = \dfrac{1}{N-1+r} \end{cases} \quad r > 1$$

Where, r is a tunable parameter.

3.3. Benchmark Testing Process

The children's gambling task is subtracted as a Benchmark of the emotion model. The participants are noted by a probability-based description of the finite state machine model. Each state of the model represents a solitaire; between adjacent cards edges extraction of choice should be dispatched [11].

For example: The participants had just extracted a first deck of cards, and plans to extract the next. At this time, He has two choices to continue to be drawn from the first decks of cards as well as to be diverted from another deputy extract.

In this model, if the position of edge is given a certain degree of probability values $P_{ij}, (i,\ j = 1,2)$. These probability values can be calculated. Pumping through comprehensive pre-licensing income and drew income. Extracted from a deck of cards, after a calculation, there are four options. Therefore, each state of the set of edges is $E^{+}\left(Deck_i\right), (i = 1,2)$.

The output is: $\sum_{j=1}^{2} P_{ij} = 1, (i = 1,2)$

The degree is: $\left|E^{+}\left(Deck_{\ i}\right)\right| = 2, (i = 1,2)$

And the probability value of the output fits in Probability $P_{Initial}i$. Select the decks initial probability is only used in the beginning of the task.

Emotional feedback is the decision-making for the emotional impact. This affect revenue extracted Solitaire. Emotional feedback in emotional stimulus intensity calculation model is:

$$\frac{\partial I(t)}{\partial t} = \frac{G(t)}{C(t)} \times r \times \alpha_{\lambda} \quad (\lambda = 1,2) \tag{3}$$

Where, $I(t)$ is the emotional stimulus intensity; $G(t)$ is the capital prior to the draw ; $r \in [0,1]$ is the refused coefficient , which is a constant represented participants sensitive ; α_{λ} is the stimulus intensity gain .The positive and negative emotions can have different values ,if $\lambda = 1$ corresponding positive affect , and if $\lambda = 2$ Should be held to negative.

Concerning of the positive and negative emotional stimulus, the positive transfer hidden Markov model calculated values through the emotional stimulus formula, Recorded as: $p_{Em}^{+}(t)(p_{Em}^{+}(t) \in [0,1])$. Get positive emotions through emotional feedback p_{Em}^{+}. Historical earnings per decks need to be calculated by evaluate the scores of each of the cards.

$$score_j = \frac{(gains_j \times gainMax_j) - (losses_j \times lossMax_j)}{(gains_j \times gainMax_j) + (losses_j \times lossMax_j)} \quad (4)$$

Where, $gains_j$ and $losses_j$ respects the number of obtained positive and negative earnings respectively; $gainsMax_j$ and $lossesMax_j$ respects the single maximum and minimum income respectively.

4. Experiment Process

4.1. Building and Testing of Platform

The test platform constituted of Intel Pentium dual-core T3200. The programming language is VC++6.0 and the relevant compilation configuration. On this basis, the entire process of the child gambling task is simulated. This emotional decision-making process is decrypted as follows [12]:

Step 1: According to the initial probability $P_{Initial}i$ of a deck of cards;
Step 2: Extract of cards;
Step 3: Calculate $G(t)$ and $C(t)$;
Step 4: Update $I(t)$;
Step 5: Calculate the emotional value $p_{Em}^{+}(t)$;
Step 6: Calculate $score_j$;
Step 7: Calculate the probability P_{ij} and normalized;

4.2. The Experimental Results

Based on the above emotion model, firstly test the degree of versatility in order to validate the model. This study assumes that 20 children were selected. The two samples are shown in Table 1.

Table 1. Assumptions Solitaire

Draws	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Card1	1	1	1	1	1	-1	1	1	-1	1	1	1	1	1	1	1	-1	1	1	1
Card2	2	2	2	-4	2	2	-5	2	2	2	2	2	-6	2	2	2	-5	2	2	2

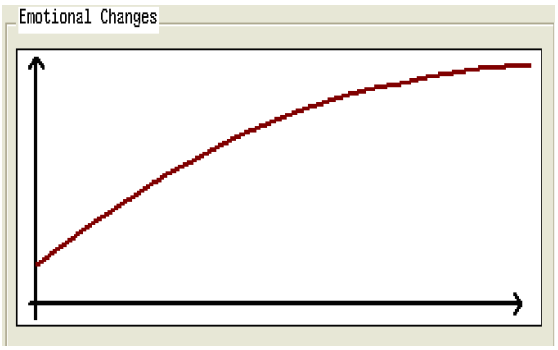

Figure 1. Emotional Changes

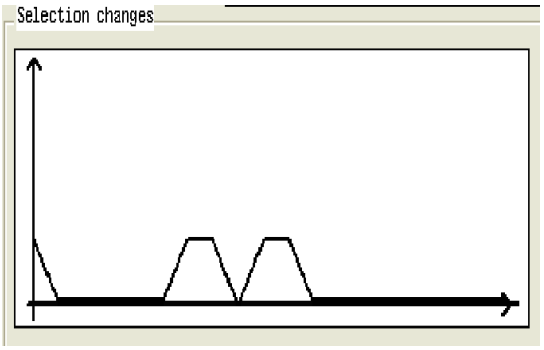

Figure 2. Selection Changes

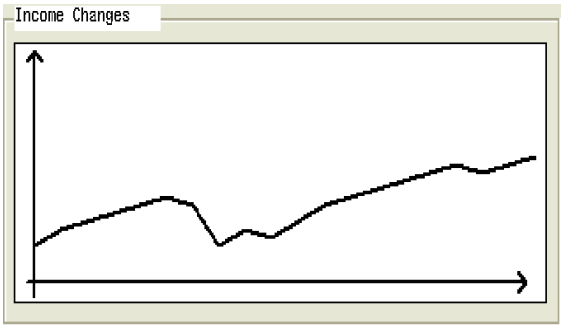

Figure 3. Income Changes

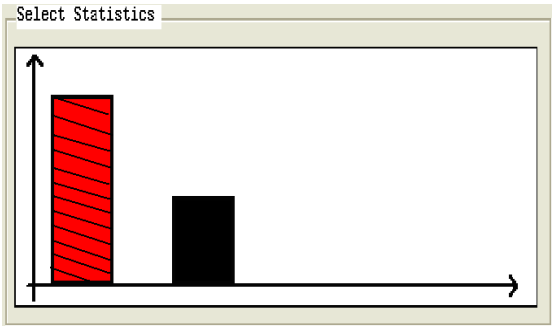

Figure 4. Select Statistics

In Table 1, positive number means that there is a happy face on the back of the card; negative number means a sad one. The big or small of its magnitude represents the number of faces. This assumption is consistent with the testing requirements of the children's gambling task. The experimental result is shown in Figure 1 to 4 from the computer simulation of the process of children's gambling task.

Figure 1 shows a positive emotional value with the discrete-time variation curve, Positive emotional value is calculated by the HMM emotion model $p_{Em}^{+} \in [0,1]$, as is shown in Figure 2. By analog participants in total revenue versus time, it can be seen from these two charts, with the game tasks being carried out, children were sinking in emotional feedback, accumulating and enhancing its revenue and positive emotions. Figure 3 shows the drawing process in dynamic timeshare, reflecting details the choice decks.

Figure 4 is chart of a process on the entire draw STATS, It can be seen that the participants more likely to choose a third decks, because of the overall income of the decks. This simulation process is coincides with the real situation.

5. Conclusion

This article describes some of the psychological test paradigm and an emotional model, and based on this model showing its test Benchmark issue. The experimental results show that the realistic simulate in the entire game in the experiment of Children's Gambling Task is effective. The new evaluation method with experimental paradigm as the evaluation standard of models is adopted suitably.

Acknowledgements

Thanks for the financial supports from fund of the projects as follows:
1) The "National Natural Science foundation of China" (No. 61170117).
2) The "National 973 Program of China" (No.2011CB505402).
3) The "Cooperation Project of Provincial Ministry" (No.2011A090200008).
4) The "Major projects of national science and technology" (No. 2010ZX07102-006).

References

[1] Bing W, Xiaoli W. Perception Neural Networks for Active Noise Control Systems. *TELKOMNIKA Indonesian Journal of Electrical Engineering*. 2012; 10(7): 1815-1822.
[2] Ortony A, Clore G L, Collins A. The cognitive structure of emotions. UK: Cambridge University Press, 1988.
[3] *Breazeal C. Emotion and sociable humanoid robots. International Journal of Human-Computer Studies*. 2003; 59(1-2): 119-155.
[4] Wei Zhehua. Artificial psychology theory research of the emotional robot. Beijing University of Science and Technology. 2002
[5] Teng less winter. Artificial emotion model used in personal robots. Beijing University of Science and Technology. 200.

[6] Chen Dezhi, Shao Yin, Wang Tao. Eddy current testing of steam generator tubes Benchmark model numerical calculation. Nondestructive Testing. 2000; 22(10): 435-438.
[7] Wang Ling. Intelligent Optimization Algorithm and Its Application. Beijing: Tsinghua University Press. 2003.
[8] UC Irvine Machine Learning Repository [OL]. Http://archive.ics.uci.edu/ml/
[9] Psychological Experiments [OL]. Http://www.cnpsy.net.
[10] Sutikno T, Facta M. Brain Emotional Learning Based Intelligent Controller. *TELKOMNIKA Indonesian Journal of Electrical Engineering*. 9(2): 201-202.
[11] Wang Liang. Benchmark Summary of Performance Test. *Computer Engineering and Applications*. 2006; (15): 45-48.
[12] Xian G, Zeng B, Yun Q, et al. Non-Specific Person Continuous Speech Identification in Second Language using BPR. *TELKOMNIKA Indonesian Journal of Electrical Engineering*. 2012; 10(7): 1604-1609.