H. Samani

MULTIMODAL COGNITIVE PROCESSING USING ARTIFICIAL ENDOCRINE SYSTEM FOR DEVELOPMENT OF AFFECTIVE VIRTUAL AGENTS

Samani H. **Multimodal Cognitive Processing Using Artificial Endocrine System for Development of Affective Virtual Agents.**

Abstract. In this paper a comprehensive architecture for emotional and affective process in a virtual agent is presented . By fusing video, audio and text emotion of the users as affective sources to the system, the virtual agent can appraise the mood of clients. To emulate the influence of the human hormones in the virtual agent, the proposed system employs Artificial Endocrine System (AES) in the aspects of moods and biological needs, by controlling the concentration level of the influential hormones. The agent affective processor engages AES, personality and mood modules to manage the internal state. Intelligent virtual agent would interact with clients according to its affective state circumstances.

The proposed system presents a complete platform to capture emotional channels through the network to analyze and process them in an affective engine in order to determine the emotional quality of the response.

Keywords: Multimodal, Emotional Agent, Cognitive Robotics, Affective Computing, Artificial Endocrine System.

1. Introduction. Intelligent Virtual Agents are human-like embodied characters [3]. These autonomous artificial characters have applications in many fields, such as computer game [58], customer relationship management [61], human-machine interface [62], virtual pets [63] and chatterbot [4]. In order to achieve an effective emulation, virtual agents must display realistic behavior [51]. With great behavioral responsiveness, the user will have a sensation of interacting with an agent [48,51]. The virtual agent with strong artificial intelligence should possess sapience and reasoning abilities [5]. For instance, in [49] the virtual agent who displays sense of humor and has learning ability to respond has presented. Some models have the ability to display emotions while interacting with users [52]. They can also be equipped with modules like voice recognition and language learning abilities [60]. In other words, virtual agents should respond to human interaction in real-time and with suitable behavior, probably with emotions rather than predetermined, highly contextual and behaviorally subtle [53].

In the proposed system, the online virtual agent operates in a network to capture messages of the clients by utilizing the web camera, microphone and keyboard as sensors. Our aim is that the virtual agent gives human-like responses to the clients. That goal is achievable if the virtual agent has internal

affective state which resembles to the mood of the human being. In daily life, our responses not only depend on the interaction with one specific conversational and emotional channel, but also rely on the overall mood at that time. Hence, we employ the concept of the internal affective state for the virtual agent to emulate the emotional states of humans.

The internal state of the virtual agent changes according to interactions with all of the users and its responses are based on its overall mood instead of direct responses which merely depend on the corresponding client. We consider both emotional and biological related hormones in order to develop the realtime structure that provides the emotional weight of the response for the overall decision making module of the virtual agent by developing the Artificial Endocrine System.

In Section 2 we described the structure of the Sentimental Architecture in details according to several layers in Figure 1. We presented an overview of the experimental development of that platform to demonstrate the practical functionality of the proposed architecture in Section 3. The paper is concluded in Section 4.

2. System Architecture. We propose ae Sentimental Architecture as the multi-layer and multi-module platform for the emotional processing of the virtual agent in this paper. The overall schematic of this platform is presented in Figure 1.

Each of the next sections of this paper corresponds with one layer in the system architecture in Figure 1 to provide the detailed description regarding the functionality of that layer. System layers include: input, perception, data fusion, AI and output.

2.1. Input Layer. While clients chat with the virtual agent over network, the input layer captures data from three different sources of video, audio and text of all the interacting users and transmits that channel information to the perception layer. The technical details of this layer depend on the structure of the network. We have considered the centralized and parallel network in our architecture.

2.2. Perception Layer. The role of the perception layer is to pre-process, filter and classify the emotional data from input layer. This layer computes the emotional value of the clients through $n \times 3 \times 6$ channels, where n is number of the clients over the network; Video (V), Audio (A) and Text(T) are 3 sources of data acquisition; and Happy (H), Sad (S), Surprise(U), Disgust(D), Anger (A) and Fear (F) are 6 basic emotions.

2.2.1. Visual Module. Facial expression is the excellent source for identifying the emotional state of the human being. The visual module of the
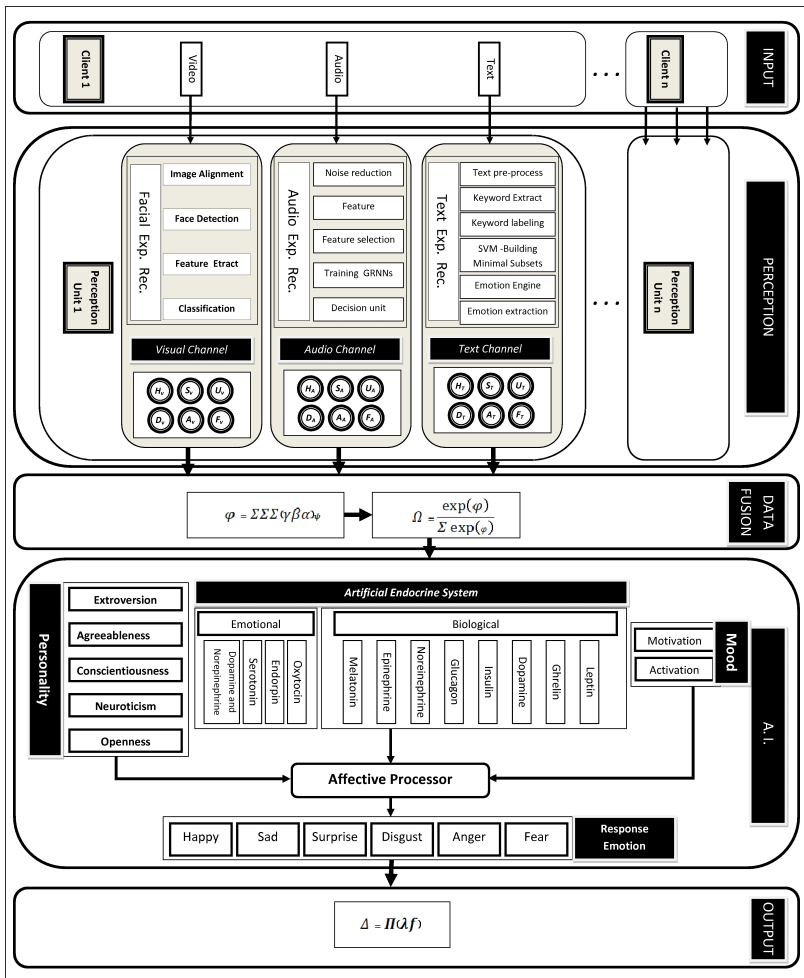
Fig. 1. Sentimental Multi-Layer Architecture

INPUT

Client 1 | Video | Audio | Text | . . . | Client n

PERCEPTION

Perception Unit 1

Facial Exp. Rec.
- Image Alignment
- Face Detection
- Feature Etract
- Classification

Visual Channel
H_a S_t U_t
D_t A_a F_c

Audio Exp. Rec.
- Noise reduction
- Feature
- Feature selection
- Training GRNNs
- Decision unit

Audio Channel
H_a S_a U_a
D_a A_a F_a

Text Exp. Rec.
- Text pre-process
- Keyword Extract
- Keyword labeling
- SVM -Building Minimal Subsets
- Emotion Engine
- Emotion extraction

Text Channel
H_T S_T U_T
D_T A_T F_T

. . . Perception Unit n

DATA FUSION

$$\varphi = \Sigma\Sigma\Sigma \gamma\beta\alpha_\phi$$

$$\Omega = \frac{\exp(\varphi)}{\Sigma \exp(\varphi)}$$

A.I.

Personality
- Extroversion
- Agreeableness
- Conscientiousness
- Neuroticism
- Openness

Artificial Endocrine System

Emotional
- Dopamine and Norepinephrine
- Serotonin
- Endorpin
- Oxytocin

Biological
- Melatonin
- Epinephrine
- Noreinephrine
- Glucagon
- Insulin
- Dopamine
- Ghrelin
- Leptin

Mood
- Motivation
- Activation

Affective Processor

Response Emotion
Happy | Sad | Surprise | Disgust | Anger | Fear

OUTPUT

$$\Delta = \Pi \lambda f\rangle$$

system analyzes the facial expressions of the user to recognize the emotion of the user based on the video information which is provided by the web camera.

Proper utilization of dynamic facial motion information can is invaluable and critical to the process of emotion recognition and interpretation [14]. Several facial expression recognition techniques are available nowadays [8-11].

By the use of neural network it is possible to develop the Multi Layer Perceptron (MLP)system which Primary, hidden and output layers of MLP correspond to the sensory data, facial action units and the classification layers respectively. The output of such MLP can classify facial expressions in to 6 basic emotions: Happiness, Sadness, Disgust, Surprise, Anger, and Fear. Details of such model is presented in others related research [2].

2.2.2. Audio Module. The process of the vocal aspect in the communication has inherent complexity in virtual agent. Firstly, in this field we are interested in daily conversional type and contains short informal utterances. Secondly, we need an emotion recognizer which is able to handle voices of various people. The system should be speaker-independent and capable of handling different voices. Furthermore, speech emotion recognition in such situations is out of reach for current systems.

The existing work used data mining and machine learning methods such as neural networks, Support Vector Machines or decision trees and they use a wide variety of voice features (mean, max, min, max-min, variance, of the pitch and intensity distribution, length of phonemic or syllabic segments or pitch rising segments) to meet these goals [31, 33-35]. They achieved good results in their experiments. However, their systems are speaker-dependent which is not desired for the virtual agents.

For our proposed system, we used an approach to do the recognition of human emotion in speech speaker-dependently. To achieve this aim, the input audio is pre-processed to remove the noise from the audio. Briefly speaking, we remove the leading and the trailing edge to clear the input .The volume is also normalized to make the recognition procedure optimal. Unvoiced sounds are cut and the result of this step will be passed to the next level to extract the features from the audio. we extract the potential features from each utterance. The discriminatory power of these features is then analyzed using GRNN [30] and K-nearest Neighbors classifier [28].

The GRNN is a memory based neural network based on the estimation of a probability density function. The main advantage of such system over the conventional multilayer feed-forward neural network is that unlike the multi layer feed-forward neural network which requires a large number of iterations in training to converge to a desired solution, GRNN needs only a single pass of learning to achieve optimal performance in classification [30]. In mathematical

terms, if we have a vector random variable x, a scalar random variable y, let X be a particular measured value of x, then the conditional mean of y given X can be represented as:

$$\hat{Y}_{(x)} = \frac{\sum_{i=1}^{n} Y_i \exp(-\frac{D_i^2}{2\sigma^2})}{\sum_{i=1}^{n} \exp(-\frac{D_i^2}{2\sigma^2})}, \tag{1}$$

where D_i is defined as:

$$D_i^2 = (X - X_i)^T (X - X_i). \tag{2}$$

In the above equations , n denotes the number of samples. X_i and Y_i are the sample values of the random variable x and y.

After choosing the best features, these selected features are used to train the main neural network which contains six sub-neural networks, one for each of the emotions. The outputs of the six sub-neural networks will be passed to the decision unit to make the final emotion. An overview of this system can be shown in the Audio module of Perception module in the Figure 1.

2.2.3. Text Module. Generally, there are two main approaches for emotion extraction from the text, keyword spotting and statistical classification [17]. The first one is the most popular and naive method for the emotion recognition [18]. In this approach, text is processed using search engine which places tags and intensity of the emotional words. It is considered a straightforward way of classifying words into six emotional categories. This method has two main problems: Firstly, sometimes emotions are hidden in the concept of the sentence and not just the words. Secondly, this technique fails to tag grammatically complex negations, ironic dialogs or slang. The second approach is the statistical classification that can reflect the useful features in addition to previous emotional keywords [19]. There are different ways of feature extraction in this method, based on the machine learning methods, such as support vector machines [22] and conditional learning. For example in [21] the authors used conditional probability as the salient function to automatically learn and extract the keywords. In [19] Salton theory has been employed to automatically recognize the emotion from the text [27]. In this work, we used support vector machine since the training of the system is simple and has no extra complexity based on local minimas. Furthermore, using SVM makes the system clearly dependent on the most informative features of the input [23]. Our text-emotion extraction system has two main modules, the training modules which processes the training text to extract the keywords. These keywords are labeled and used to make the attributes minimal subsets using SVM. Our extraction engine is

actually made of these attribute sets. The other module is the test module which we use in real time to extract the emotions. Basically, when a new input is fed to this module, the keywords are extracted and labeled using the same procedure in training phase. After this step, we can use the extraction engine to get the emotions [24, 25].

2.3. Data Fusion Layer. The total emotional input to the system at any time is the combination of three perceptional modules of video, audio and text from all the clients. The data fusion module integrates the value of the emotional channels over the perception layer in order to supply the union emotional array to the artificial intelligence layer.

The fused emotional input value can be defined to mix three emotional inputs of six basic emotions for all the clients as equation (3):

$$\varphi_{(m_i,t)} = \left(\sum_{k=1}^{n}\sum_{j=1}^{3}\sum_{i=1}^{6}\gamma_{(k)}\beta_{(jk)}\alpha_{(ijk)}\right)\psi_{(m_i,t)}, \tag{3}$$

where $i = Happiness$, $Sadness$, $Disgust$, $Surprise$, $Anger$ and $Fear$ presents six basic emotions,

$j = Vision$, $Audio$ and $Text$ refers to three different source of emotional input,

n is number of clients who interact with agent at time t,

k is the counter for number of clients,

$\alpha_{(ijk)}$ is the emotional coefficient for the ith emotion, through jth source in the kth client channel,

$\beta_{(jk)}$ is the source coefficient for jth source through kth client channel,

$\gamma_{(k)}$ is the channel coefficient for kth client channel,

$\psi_{(m_i,t)}$ is the emotional parameter for six basic emotions at time t and

$\varphi_{(m_i,t)}$ is the fused emotional value for six basic emotions at time t which includes 6 values of emotional input from all the clients at the current time by considering 3 different weights of α, β and γ for emotion, source and channel respectively.

The Softmax activation function is used for the input emotional values in the network to be interpretable as posterior values for six different emotional categories. In this way results lie between zero and one and the sum of them would be one:

$$\Omega_i = \frac{\exp(\varphi_{(m_i,t)})}{\sum_{i=1}^{6}\exp(\varphi_{(m_i,t)})}, \tag{4}$$

where Ω_i represents the normalized value of input emotions.

2.4. Artificial Intelligence (AI) Layer. The Artificial Intelligence (AI) layer processes the internal state of the virtual agent by considering the multiple emotional values of the clients and internal affective parameters to change the characteristic parameters of the virtual agent accordingly. This layer consists of several modules: Personality including five main character dimensions, Mood module as long lasting affect of the agent, AES which manages biological and emotional hormones and the Affect Processor module which computes mentioned modules to generate the values for Response Emotion module.

2.4.1. Artificial Endocrine System (AES). Natural endocrine system is viewed as a network of glands that works with nervous system to secrete hormones directly into the blood so as to control the activity of internal organs and coordinate the long range response to external stimuli [37]. Hormones which are chemicals released by components of the endocrine system affect other parts of the body. Hormones play a significant role in the endocrine system so as to preserve homeostasis. Here, we will introduce the relation of hormone with human emotion and behavior and implement the idea into our virtual agent.

Virtual biological systems considered as research field of biological inspired computing. Artificial Neural Network (ANN) is one of the well known tools in computational intelligence techniques. In the same way, The endocrine system could be also very useful tool, but there was not any interest in that apart from some basic systems like [42, 43]. Timmis and Neal [42] first proposed an artificial endocrine system (AES) in the module of a broader conceptual framework which incorporates artificial neural networks (ANN) and artificial immune systems (AIS). Later Vargas et al. [43] proposed an artificial homeostatic system based on the previous work. It focuses on mimicking the some important mechanisms in endocrine system, such as hormone mechanism. The system they proposed includes three modules which are hormone level which is to record level of hormones, hormone production controller serves to control the generation of hormones according to the variation of the internal states and external stimulus. and endocrine glands means to generate the required amount of hormones after receiving the input from the controller. Any change in the internal and external will trigger the activities in ANN and AES [38].

This paper focuses on the hormones which are related to emotions and biological qualities.

For emotion-related hormones, we consider four hormones namely Dopamine, Serotonin, Endorphin and Oxytocin, The level of these hormones is related to the emotional situation of the human being as it is presented in Table 1.

Table 1. Emotional Hormones

Hormone	Affected Emotion	Effects	Virtual function
Dopamine	Excitement, Alertness	High dopamine makes people more talkative and alert.	More Energetic and Talkative
Serotonin	Happiness, Depression, Anxiety, Fear, Apathy, Feeling of worthlessness, Fatigue, tension	High Serotonin helps maintain happiness, relieve depression, anxiety and tension.	Happier and more Confident
Endorpin	Contentment	Pain Killer	Improves the Sense of Well-Being
Oxytocin	Trust, Empathy, Generosity, Love	High oxytocin increases trust and reduces fear. It affects generosity by increasing empathy during perspective taking.	Trust the other party and show empathy over others.

Increasing level of Norepinephrine with Dopamine make people more focused, more talkative and alert. Increasing amount of Serontonin and norepinephrine are used as anti-depressants to relieve depression. Low Serotonin level may lead to depression, anxiety, fear, feeling of worthlessness, fatigue, insomnia. Endorphin which resemble the opiates are known to produce analgesia and a sense of well-being. Oxytocin are the hormone of love. Its role is to maintain healthy interpersonal relationships, high level of Oxytocin increases trust and relieves interpersonal stress. In our artificial endocrine system, we intend to implant the above hormones into our virtual agent so that their affective states and behaviour will be affected by the variation in the hormone levels [40, 41].

Furthermore, the affective state and behavior of the virtual agent should be affected by the physiological parameters, such as blood pressure, blood glucose and heart rate. These parameters are influenced by hormones as well. Hence, we introduce a group of hormones which are closely related to the physiological parameters of humans: Melatonin, Epinephrine, Orexin, Norepinephrine, Glucagon, Insulin, Ghrelin and Leptin.

Melatonin cause drowsiness. Epinephrine increase rate of heart beat and raises blood glucose . Both Dopamine and Noreinephrine increase blood pressure.Glucagon raise blood glucose level and Insulin is known as decreasing blood glucose level. Ghrelin stimulates apeptite, but Leptin decreases apeptite.

Table 2. Biological Hormones

Hormone	Effects/Action	Virtual
Melatonin	Drowsiness ↑	Drowsy, Sleepy
Norepinephrine	Blood pressure ↑	Excited, Active
Epinephrine	Heart Beat Rate ↑ Blood Glucose ↑	Sick
Glucagon	Blood Glucose ↑	Full, Lethargy
Insulin	Blood Glucose ↓	Hungry, Dizzy
Orexin	Heart Beat Rate ↑ Appetite ↑	stimulated, Motivated
Ghrelin	Appetite ↑	Hungry, Dizzy
Leptin	Appetite ↓	Lazy, Listless

The effect of these 8 biological hormones are explained in more details in Table 2.

In our artificial endocrine system, we intend to implant the above hormones into our virtual agent so that their affective states and behavior would be affected by variation of hormone levels [55].

Such a system is able to generate the respective amount of hormones based on the current hormone concentration level which represents the current emotional state and the clients emotional input as external stimulus.

Based on the current internal states of virtual agent and external stimulations, the virtual agent will signal the glands to generate the required amount of hormones. Hence, the virtual agent will experience change in the emotional state and biological need.

In our system all hormones are considered to be secreted by two parameters:

– The activation function which can be presented by employing the logistic function;

– The gland bustle that should be considered through all the stimuli channels.

So the glands secretion can be modeled as equation 5:

$$\Lambda_q = \frac{1}{1 + \exp(-aq)} \sum_{q=1}^{18 \times n} \rho_i \Theta_q. \tag{5}$$

Above representing shows that the gland secretion, Λ, is the product of the each gland bustle, Θ_q, by considering ρ_q as the stimuli weight, which can be activated through the nonlinear activation function $\frac{1}{1+\exp(-aq)}$. The gland bustle should be considered over 6 emotional values of 3 different sources, (18), for all the n clients. The coefficient a in the activation function depends on the current volume of the hormone in the system.

Furthermore, The ratio of secreted amount of each hormone is the friction of that hormone to total hormone volume which can be calculated according to the gland activity. Equation 6 shows that, δ_q, which is the hormone ratio for each of 12 hormones is calculated by considering hormone rate as produced over time t, by assuming coefficient ζ for each hormone flow rate.

$$\delta_q = \frac{\zeta_q\left[\frac{1}{1+\exp(-aq)}\left(\sum_{q=1}^{18\times n}\rho_i\Theta_q\right)\right]t}{\sum_{q=1}^{12}\zeta_q\left[\frac{1}{1+\exp(-ab)}\left(\sum_{q=1}^{18\times n}\rho_i\Theta_q\right)\right]t}. \tag{6}$$

In this way, the virtual agent is equipped with basic emotional hormones to control affective situation like being happy, talkative and energetic; also it will have basic biological needs like feeling hungry, sleepy, full and sick. That capability makes it possible to see dynamic and realistic behavior from the agent.

2.4.2. Mood. The mood refers to more long term emotional state. Psychologist have considered two fundamental dimensions for the mood. The circumplex model of the affect is considered with two dimensions of valence and arousal [44]. In the another model the mood is a product of two dimensions, energy and the tension [45].

Following that fundamental concept, in our model, we also considered mood has two main dimensions including activation and motivation. Activation is related to the amount of energy in the mood. For instance, excitement can be considered as high level of activation, The activation level with surprise is higher than happiness. Motivation refers to pleasure and displeasure of a mood. For example, joy represents high motivation in our model, but sad means low motivation.

2.4.3. Personality. Personality is set of characteristics that makes a person distinct from another. The basic five dimensions of human personality are extroversion, agreeableness, conscientiousness, neuroticism and openness [47].

In our model, we adopt the above five big personality dimensions to equip our virtual agents with unique personalities. Extroversion describes the attributes as sociability, talkativeness with high level of emotional expressiveness. Agreeableness includes characteristics like trust, affection and kindness. Conscientiousness describes people with good impulse control and great thoughtfulness. People with high Neuroticism tend to experience anxiety, emotional instability. Lastly, people with high openness are likely to have a broad spectrum of interests and very imaginative, creative. The above mentioned five personality parameters will create the personality module of the virtual agent in the proposed model.

2.5. Output Layer. The affective processor of the system which computes the affective state of the agent decides on the emotional output value of the agent for all the clients. This layer determines the total output of the system based on the value of the AI modules. Equation 7 presents the output generation structure:

$$\Delta_E = \prod_{i=1}^{4} \lambda_i f_i. \tag{7}$$

Δ_E represents the emotional output weight, where f and λ represent the function and corresponding emotional output weight for four decision making modules consist of Emotional input, Mood, Personality and Endocrine system.

3. Experimental Result. According to the described architecture we developed the proof of concept virtual agent in TCP/IP network that interacts with 5 users simultaneously. Clients could interact with the virtual agent in this network whilst administrator control the internal state of the server. Figure 2 shows the sample of the experiment in the Sentimental layout for one client.

To evaluate the performance of the Sentimental Architecture we asked 20 users to participate in interaction experiment with the developed online virtual agent. We put them in four groups of five participants, two with emotional system and two without it, to compare their engagement level according to the duration of conversation in the interaction test. The average interaction time for two groups with the Sentimental Architecture was %35 more than the duration of conversation for two groups using the system without that platform, which shows the significance of emotional processing in the virtual agent.

We also conducted a survey including a questionnaire and interview with each of the clients after using the system. We asked participants to rate the system performance. Participants who used the system with Sentimental Architecture reported %23 more engagement with the system compare to those who used the system without such architecture.

We also asked participants to report if they felt emotions in responses from the virtual agent by rating between 0 for no emotion and 10 for maximum emotion. The Mean Score for the sentimental system was %28 more than a system without that capability. Furthermore, Users expressed that they felt more realistic behaviors during interaction with the agent which is equipped with the Sentimental architecture.

The network included 90 emotional channels for (5 clients) × (3 sources) × (6 emotions). According to this experiment even by using almost equal emotional input coefficients to the system and considering similar affective state, clients receive different emotional responses from the agent. That tallies with the fact that even though the human can interact with few different people
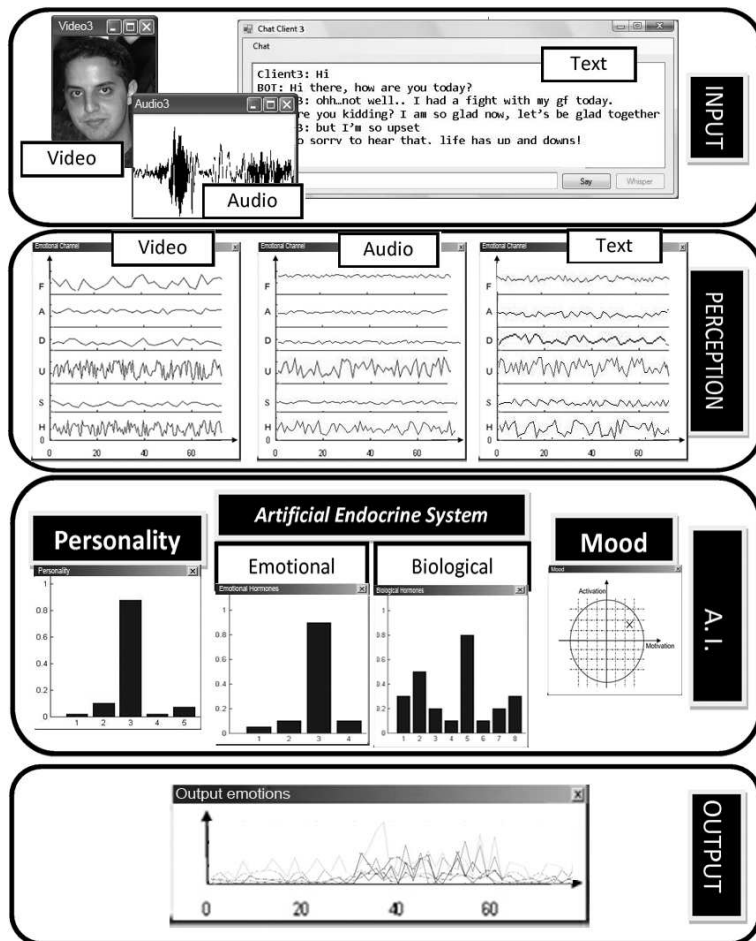
Fig. 2. Experimental result in the form of the proposed Architecture

at the same time, we are able to response with variant emotional values, even having the same mood [1].

4. Conclusions. We presented a multimodal sentimental system for virtual agent base on the Artificial Endocrine System in order to improve the affective properties of the agent. We also employed mood and personality in order to reinforce the emotional capability of the agent. The proposed system is capable of interacting with several users simultaneously, however the agent's behaviors depends on the instantaneous affective properties of the agent. Such ability aims to grant realistic behaviors by the agent compare to other systems which behave merely according to the interactive parameters. We tested such system with several participants and user studies show that the proposed sentimental architecture presents an efficient emotional system for interaction with users. The main idea is that current virtual agent can be equipped with emotional units and the artificial inteligent module of virtual agents can add emotional intelligence as well. This system can be applied to various types of virtual agents such as chatbots, virtual avatars and robots.

References
1. Samani H. The evaluation of affection in human-robot interaction. *Kybernetes*. 2016. vol. 45. pp. 1257–1272.
2. Samani H.A., Elham S. A multidisciplinary artificial intelligence model of an affective robot. *International Journal of Advanced Robotic Systems*. 2012. vol. 9. pp. 1–11.
3. Sam T., Silvervarg A., Gulz A., Tom Z. Physical vs. Virtual Agent Embodiment and Effects on Social Interaction. International Conference on Intelligent Virtual Agents. 2016. pp. 412–415.
4. Benton M. et al. Quality in Chatbots and Intelligent Conversational Agents. *Software Quality Professional Magazine*. 2017. vol. 19(3).
5. Samani H. Cognitive robotics. CRC Press, 2015. 215 p.
6. Ivan M. Some Related Article I Wrote. *Some Fine Journal*. 1999. vol. 99. pp. 1–100.
7. Andreas N. A Book He Wrote. Erewhon : His Publisher, 1999.
8. Liu P., Han S., Meng Z. Tong Y. Facial expression recognition via a boosted deep belief network. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2014. pp. 1805–1812.
9. Happy S. L., Aurobinda R. Automatic facial expression recognition using features of salient facial patches. *IEEE transactions on Affective Computing*. IEEE, 2015, vol. 6. pp. 1–12.
10. Evangelos S., Hatice G., Andrea C. Automatic analysis of facial affect: A survey of registration, representation, and recognition. *IEEE transactions on pattern analysis and machine intelligence.* IEEE, 2015 . vol. 37. pp. 1113–1133.
11. Ge S., Samani H., Ong Y., Hang C. Active affective facial analysis for human-robot interaction. The 17th IEEE International Symposium on Robot and Human Interactive Communication(RO-MAN 2008). 2008. pp. 83–88.
12. Abboud B., Davoine F., Dang M. Facial expression recognition and synthesis based on an appearance model. *Signal Processing: Image Communication*. Elsevier, 2004. vol. 19. pp. 723–740.

13. Cohen I., Sebe N., Garg A., Chen L.S., Huang T.S. Facial expression recognition from video sequences: temporal and static modeling. *Computer Vision and Image Understanding*. Elsevier, 2003. vol. 91. pp. 160–187.

14. Krumhuber E., Kappas A., Manstead A. Effects of dynamic aspects of facial expressions: a review. *Emotion Review*. 2013. vol. 5. pp. 41–46.

15. Bartlett M., Littlewort G., Fasel I., Movellan J. Real Time Face Detection and Facial Expression Recognition: Development and Applications to Human Computer Interaction. CVPR Workshop on Computer Vision and Pattern Recognition for Human-Computer Interaction. 2003. vol. 5. pp. 52–53.

16. Fasel B., Luettin J. Automatic facial expression analysis: a survey. Pattern Recognition. Elsevier, 2003. vol. 36. pp. 259–275.

17. Mohammad S. Sentiment analysis: Detecting valence, emotions, and other affectual states from text. *Emotion measurement*. 2015. pp. 201–238.

18. Li W., Xu H. Text-based emotion classification using emotion cause extraction. *Expert Systems with Applications*. Elsevier, 2014 . vol. 41. pp. 1742–1749.

19. Lee C., Lee G. Emotion recognition for affective user interface. The 16th IEEE International Symposium on Robot and Human interactive Communication. 2007. vol. 8. pp. 798–801.

20. Zhe X., Boucouvalas A.C. Text-to-Emotion Engine for Real Time Internet Communication. Proceedings of International Symposium on Communication Systems, Networks and DSPs. 2002. pp. 164–168.

21. Lee C.M., Narayanan S.S. Toward detecting emotions in spoken dialogs. Speech and Audio Processing, IEEE Transactions on. 2005. vol. 13. pp. 293–303.

22. Cristianini N., Shawe-Taylor J. An Introduction to Support Vector Machines, Cambridge University Press, 2000. 204 p.

23. Povoda L. et al. Optimization Methods in Emotion Recognition System. *Radioengineering*. 2016. vol. 25. pp. 565.

24. Saadatian E. et al. Artificial Intelligence Model of an Smartphone-Based Virtual Companion. International Conference on Entertainment Computing. 2014. pp. 173–178.

25. Elham S., Samani H., Arash T., Ryohei N. Technologically mediated intimate communication: An overview and future directions. International Conference on Entertainment Computing. 2013. pp. 93–104.

26. Zhang Y., Ren F., Kuroiwa S. Semi-Automatic Emotion Recognition from Chinese Text. Proceedings of the Ninth IASTED International Conference on Intelligent Systems and Control. 2006.

27. Salton G., Yang C.S. On the Specification of Term Values in Automatic Indexing. *Journal of documentation*. 1973. vol. 29. no. 4. pp. 351–372.

28. Dasarathy B.V. Nearest neighbor (NN) norms: nn pattern classification techniques. Los Alamitos, Calif.: IEEE Computer Society Press, 1991. 550 p.

29. Bhatti M.W., Wang Y., Guan L. A neural network approach for human emotion recognition in speech. Circuits and Systems, 2004. ISCAS'04. Proceedings of the 2004 International Symposium on. 2004. vol. 2. pp. II–181.

30. Specht D.F. A general regression neural network. Neural Networks, IEEE Transactions on. 1991. vol. 2(6). pp. 568–576.

31. Dellaert F., Polzin T., Waibel A. Recognizing emotion in speech. Proceedings of the Fourth International Conference on Spoken Language (ICSLP 96). 1996. vol. 3. pp. 1970–1973.

32. Teng Z., Ren F., Kuroiwa S. Emotion Recognition from Text based on the Rough Set Theory and the Support Vector Machines. International Conference on Natural Language Processing and Knowledge Engineering (NLP-KE 2007). 2007. pp. 36–41.

33. Breazeal C.L. Designing Sociable Robots. Bradford Book, 2002. 282 p.

34. Oudeyer P.Y. The production and recognition of emotions in speech: features and algorithms. *International Journal of Human-Computer Studies*. 2003. vol. 59. pp. 157–183.

35. McGilloway S. et al. Approaching Automatic Recognition of Emotion from Voice: A Rough Benchmark. ISCA Tutorial and Research Workshop (ITRW) on Speech and Emotion. 2000. 6 p

36. Purves W.K., Orians G.H., Heller H.C. Life: The Science of Biology: 7th ed. 2003. 1121 p.

37. Straub R.H. Interaction of the endocrine system with inflammation: a function of energy and volume regulation. *Arthritis research and therapy*. 2014. vol. 16. no 1. 15 p.

38. Samani H., Saadatian E., Jalaeian B. Biologically Inspired Artificial Endocrine System for Human Computer Interaction. International Conference on Human-Computer Interaction. 2015. pp. 71–81.

39. Norman A.W., Litwack G. Hormones. Academic Press San Diego, 1997. 806 p.

40. Morrison M.F. Hormones, Gender and the Aging Brain: The Endocrine Basis of Geriatric Psychiatry. Cambridge University Press, 2000. 259 p.

41. Pfaff D.W., Phillips M.I., Rubin R.T. Principles of Hormone/Behavior Relations. Academic Press, 2004. 360 p.

42. Timmis J., Neal M. Once more Unto the Breach: Towards Artificial Homeostatsis. Recent deveopments in Biologically inspired computing. 2005. pp. 340–365.

43. Vargas P. et al. Von Artificial Homeostatsis system : A Novel approach. European Conference on Artificial Life (ECAL 2005). 2005. pp. 754–764.

44. Russell J.A. A circumplex model of affect. *Journal of Personality and Social Psychology.* 1980. vol. 39. pp. 1161–1178.

45. Thayer R.E. The Biopsychology of Mood and Arousal. Oxford University Press, USA, 1989. 234 p.

46. Barrick M.R., Mount M.K. The Big Five Personality Dimensions and Job Performance: A Meta-Analysis. *Personnel Psychology*. 1991. vol. 44. no. 1. p. 1–26.

47. Kim H.J., Shin K.H., Swanger N. Burnout and engagement: A comparative analysis using the Big Five personality dimensions. *International Journal of Hospitality Management*. 2008. vol. 28. no. 1. pp. 96–104.

48. Del B.A., Vicario E., Zingoni D. An interactive environment for the visual programming of virtual agents. Proceedings of 1994 IEEE Symposium on Visual Languages. 1994. pp. 145–152.

49. Alfonsi B. "Sassy" Chatbot Wins with Wit. *IEEE Intelligent Systems.* 2006. pp. 6–7.

50. Herrero P., de Antonio A. Modelling Intelligent Virtual Agent Skills with Human-Like Senses. Conference of Computer Science. Springer. 2004. vol. 3038. pp. 575–582.

51. Del B.A., Vicario E. Specification by-Example of Virtual Agents Behavior. *IEEE Translations on Visualization and Computer Graphics.* 1995. vol. 1. no. 4. pp. 350–360.

52. Heudin J.C. Evolutionary virtual agent. Proceedings of the IEEE/WIC/ACM International Conference on Intelligent Agent Technology. 2004. pp. 93–98.

53. Zhao R., Papangelis A., Cassell J. A dyadic computational model of rapport management for human-virtual agent interaction. International Conference on Intelligent Virtual Agents. 2014. pp. 514–527.

54. Badler N., Allbeck J., Zhao L., Byun M. Representing and Parameterizing Agent Behaviors. *Computer Animation.* 2002. pp. 133–143.

55. Samani H. Lovotics: Loving robots. LAP LAMBERT Academic Publishing, 2012. 168 p.

56. Bloch L.R., Lemish D. Disposable Love: The Rise and Fall of a Virtual Pet. *New Media & Society.* 1999. vol. 1. no. 3. pp. 283–303.

57. Adobbati R. Gamebots: A 3D Virtual World Test-Bed For Multi-Agent Research. Proceedings of the Second International Workshop on Infrastructure for Agents, MAS, and Scalable MAS. 2001. vol. 5. 6 p.

58. Fernandez-Ares A. et. al. Its time to stop: A comparison of termination conditions in the evolution of game bots. European Conference on the Applications of Evolutionary Computation. 2015. pp. 355–368.

59. Jutla D., Craig J., Bodorik P. Enabling and measuring electronic customer relationship management readiness. System Sciences, 2001. Proceedings of the 34th Annual Hawaii International Conference on. 2001. p. 10.

60. Johnson A., Roush, T., Fulton M., Reese A. Implementing Physical Capabilities for an Existing Chatbot by Using a Repurposed Animatronic to Synchronize Motor Positioning with Speech. *International Journal of Advanced Studies in Computers, Science and Engineering.* 2017 . vol. 6. pp. 20.

61. Vieira A., Sehgal A. How Banks Can Better Serve Their Customers Through Artificial Techniques. *Digital Marketplaces Unleashed.* Springer, 2018. pp. 311–326.

62. Folstad A., Brandtzaeg Pe.B. Chatbots and the new world of HCI. *Interactions.* 2017. vol. 24. no. 4. pp. 38–42.

63. Liu Y. et. al. Chatting system, method and apparatus for virtual pet. Google Patents. 2014. US Patent 8645479.

Hooman Samani — PhD in Robotics, associate professor; director of AIART Lab (Artificial Intelligence and Robotics Technology Laboratory), associate professor, Department of Electrical Engineering, College of Electrical Engineering and Computer Science, National Taipei University, Taiwan. Research interests: Robotics, Affective Computing, Artificial Intelligence, System Engineering. The number of publications — 50. hooman@mail.ntpu.edu.tw, www.hoomansamani.com; Number 151, Daxue Road, Sanxia District, New Taipei City, 23741 Taiwan; office phone (+886) 2 - 8674 1111 ext:67736, lab phone (+886) 2 - 8674 1111 ext:66981.

Х. САМАНИ

МНОГОМОДАЛЬНАЯ КОГНИТИВНАЯ ОБРАБОТКА С ИСПОЛЬЗОВАНИЕМ ИСКУССТВЕННОЙ ЭНДОКРИННОЙ СИСТЕМЫ ДЛЯ РАЗВИТИЯ АФФЕКТИВНЫХ ВИРТУАЛЬНЫХ АГЕНТОВ

Самани Х. **Многомодальная когнитивная обработка с использованием искусственной эндокринной системы для развития аффективных виртуальных агентов.**

Аннотация. Представлена всеобъемлющая архитектура эмоционального и аффективного процесса, происходящего в виртуальном агенте. Соединяя визуальные, аудио- и текстовые эмоции пользователей как аффективные источники в системе, виртуальный агент может оценивать настроение клиентов. С целью имитации воздействия гормонов человека в виртуальном агенте в предлагаемой системе используется искусственная эндокринная система (ИЭС) для выявления настроения и биологических потребностей посредством контроля уровня концентрации воздействующих гормонов. Аффективный процессор агента задействует модули ИЭС, параметров личности и настроения для управления внутренним состоянием. Интеллектуальный виртуальный агент взаимодействует с клиентами в соответствии со своими аффективными состояниями. Предлагаемая система представляет собой полную платформу для захвата каналов эмоций в сети с целью анализа и обработки их в аффективном движке для определения эмоциональной окраски ответа.

Ключевые слова: многомодальность, эмоциональный агент, когнитивная робототехника, эмоцио-нальные вычисления, искусственная эндокринная система.

Самани Хуман — к-т техн. наук, доцент, руководитель лаборатории искусственного интеллекта и робототехники (ИИР), доцент кафедры электротехники института электротехники и информатики, Национальный университет Тайбэя. Область научных интересов: робототехника, эмоциональные вычисления, искусственный интеллект, системная инженерия. Число научных публикаций — 50. hooman@mail.ntpu.edu.tw, www.hoomansamani.com; Даксвей роад, 151, район Санься, Синьбэй, 23741, Тайвань; р.т.: (+886)2-8674-1111.

Литература

1. *Samani H.* The evaluation of affection in human-robot interaction // Kybernetes. 2016. vol. 45. pp. 1257–1272.

2. *Samani H.A., Elham S.* A multidisciplinary artificial intelligence model of an affective robot // International Journal of Advanced Robotic Systems. SAGE Publications Sage. 2012. vol. 9. 6 p.

3. *Sam T., Silvervarg A., Gulz A., Tom Z.* Physical vs. Virtual Agent Embodiment and Effects on Social Interaction // International Conference on Intelligent Virtual Agents. 2016. pp. 412–415.

4. *Benton M. et al.* Quality in Chatbots and Intelligent Conversational Agents // Software Quality Professional Magazine. 2017. vol. 19(3).

5. *Hooman S.* Cognitive robotics // CRC Press. 2015. 215 p.

6. *Ivan M.* Some Related Article I Wrote // Some Fine Journal. 1999. vol. 99. pp. 1–100.

7. *Andreas N.* A Book He Wrote // Erewhon : His Publisher, 1999.

8. *Liu P., Han S., Meng Z. Tong Y.* Facial expression recognition via a boosted deep belief network // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2014. pp. 1805–1812.

9. *Happy S.L., Aurobinda R.* Automatic facial expression recognition using features of salient facial patches // IEEE transactions on Affective Computing. 2015. vol. 6. pp. 1–12.

10. *Evangelos S., Hatice G., Andrea C.* Automatic analysis of facial affect: A survey of registration, representation, and recognition // IEEE transactions on pattern analysis and machine intelligence. IEEE, 2015. vol. 37. pp. 1113–1133.

11. *Ge S., Samani H., Ong Y., Hang C.* Active affective facial analysis for human-robot interaction // The 17th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN 2008). 2008. pp. 83–88.

12. *Abboud B., Davoine F., Dang M.* Facial expression recognition and synthesis based on an appearance model // Signal Processing: Image Communication. Elsevier, 2004. vol. 19. pp. 723–740.

13. *Cohen I et al.* Facial expression recognition from video sequences: temporal and static modeling // Computer Vision and Image Understanding. Elsevier. 2003. vol. 91. pp. 160–187.

14. *Krumhuber E., Kappas A., Manstead A.* Effects of dynamic aspects of facial expressions: a review // Emotion Review. 2013. vol. 5. pp. 41–46.

15. *Bartlett M., Littlewort G., Fasel I., Movellan J.* Real Time Face Detection and Facial Expression Recognition: Development and Applications to Human Computer Interaction // CVPR Workshop on Computer Vision and Pattern Recognition for Human-Computer Interaction. 2003. vol. 5. pp. 52–53.

16. *Fasel B., Luettin J.* Automatic facial expression analysis: a survey // Pattern Recognition. Elsevier, 2003. vol. 36. pp. 259–275.

17. *Mohammad S.* Sentiment analysis: Detecting valence, emotions, and other affectual states from text // Emotion measurement. 2015. pp. 201–238.

18. *Li W., Xu H.* Text-based emotion classification using emotion cause extraction // Expert Systems with Applications. Elsevier, 2014 . vol. 41. pp. 1742–1749.

19. *Lee C., Lee G.* Emotion recognition for affective user interface // The 16th IEEE International Symposium on Robot and Human interactive Communication. 2007. vol. 8. pp. 798–801.

20. *Zhe X., Boucouvalas A.C.* Text-to-Emotion Engine for Real Time Internet Communication // Proceedings of International Symposium on Communication Systems, Networks and DSPs. 2002. pp. 164–168.

21. *Lee C.M., Narayanan S.S.* Toward detecting emotions in spoken dialogs // IEEE transactions on speech and audio processing. 2005. vol. 13. pp. 293–303.

22. *Cristianini N., Shawe-Taylor J.* An Introduction to Support Vector Machines, Cambridge University Press, 2000. 204 p.

23. *Povoda L. et al.* Optimization Methods in Emotion Recognition System // Radioengineering. 2016. vol. 25. pp. 565.

24. *Saadatian E.et al.* Artificial Intelligence Model of an Smartphone-Based Virtual Companion // International Conference on Entertainment Computing. 2014. pp. 173–178.

25. *Elham S., Samani H., Arash T., Ryohei N.* Technologically mediated intimate communication: An overview and future directions // International Conference on Entertainment Computing. 2013. pp. 93–104.

26. *Zhang Y., Ren F., Kuroiwa S.* Semi-Automatic Emotion Recognition from Chinese Text // Proceedings of the Ninth IASTED International Conference on Intelligent Systems and Control. 2006.

27. *Salton G., Yang C.S.* On the Specification of Term Values in Automatic Indexing // Journal of documentation. 1973. vol. 29. no. 4. pp. 351–372.

28. *Dasarathy B.V.* Nearest neighbor (NN) norms: nn pattern classification techniques // IEEE Computer Society Press, 1991. 550 p.

29. *Bhatti M.W., Wang Y., Guan L.* A neural network approach for human emotion recognition in speech // Proceedings of the 2004 International Symposium on Circuits and Systems (ISCAS'04). 2004. vol. 2. pp. II–181.

30. *Specht D.F.* A general regression neural network // IEEE transactions on neural networks. 1991. vol. 2(6). pp. 568–576.

31. *Dellaert F., Polzin T., Waibel A.* Recognizing emotion in speech // Proceedings of the Fourth International Conference on Spoken Language (ICSLP 96). 1996. vol. 3. pp. 1970–1973.

32. *Teng Z., Ren F., Kuroiwa S.* Emotion Recognition from Text based on the Rough Set Theory and the Support Vector Machines // International Conference on Natural Language Processing and Knowledge Engineering (NLP-KE 2007). 2007. pp. 36–41.

33. *Breazeal C.L.* Designing Sociable Robots. Bradford Book, 2002. 282 p.

34. *Oudeyer P.Y.* The production and recognition of emotions in speech: features and algorithms // International Journal of Human-Computer Studies. 2003. vol. 59. pp. 157–183.

35. *McGilloway, S. et al.* Approaching Automatic Recognition of Emotion from Voice: A Rough Benchmark // ISCA Tutorial and Research Workshop (ITRW) on Speech and Emotion. 2000. 6 p.

36. *Purves W.K., Orians G.H., Heller H.C.* Life: The Science of Biology: 7th ed. // 2003. 1121 p.

37. *Straub R.H.* Interaction of the endocrine system with inflammation: a function of energy and volume regulation // Arthritis research therapy. Arthritis research and therapy. 2014. vol. 16. no. 1. 15 p.

38. *Samani H., Saadatian E., Jalaeian B.* Biologically Inspired Artificial Endocrine System for Human Computer Interaction // International Conference on Human-Computer Interaction. 2015. pp. 71–81.

39. *Norman A.W., Litwack G.* Hormones // Academic Press. 1997. 806 p.

40. *Morrison M.F.* Hormones, Gender and the Aging Brain: The Endocrine Basis of Geriatric Psychiatry // Cambridge University Press. 2000. 259 p.

41. *Pfaff D.W., Phillips M.I., Rubin R.T.* Principles of Hormone/Behavior Relations // Academic Press. 2004. 360 p.

42. *Timmis J., Neal M.* Once more Unto the Breach: Towards Artificial Homeostatsis // Recent deveopments in Biologically inspired computing. 2005. pp. 340–365.

43. *Vargas P. et al.* Artificial Homeostatsis system : A Novel approach // European Conference on Artificial Life (ECAL 2005). 2005. pp. 754–764.

44. *Russell J.A.* A circumplex model of affect // Journal of Personality and Social Psychology. 1980. vol. 39. pp. 1161–1178.

45. *Thayer R.E.* The Biopsychology of Mood and Arousal. Oxford University Press, USA, 1989. 234 p.

46. *Barrick M.R., Mount M.K.* The Big Five Personality Dimensions and Job Performance: A Meta-Analysis // Personnel Psychology. 1991. vol. 44. no. 1. pp. 1–26.

47. *Kim H.J., Shin K.H., Swanger N.* Burnout and engagement: A comparative analysis using the Big Five personality dimensions // International Journal of Hospitality Management. 2008. vol. 28. no. 1. pp. 96–104.

48.   *Del B.A., Vicario E., Zingoni D.* An interactive environment for the visual programming of virtual agents. 1994. pp. 145–152.

49.   *Alfonsi B.* "Sassy"Chatbot Wins with Wit // IEEE Intelligent Systems. 2006. pp. 6–7.

50.   *Herrero P. de Antonio A.* Modelling Intelligent Virtual Agent Skills with Human-Like Senses // Conference of Computer Science. Springer, 2004. vol. 3038, pp. 575–582.

51.   *Del B.A., Vicario E.* Specification by-Example of Virtual Agents Behavior // IEEE Translations on Visualizations and Computer Graphics. 1995. vol. 1. no. 4. pp. 350–360.

52.   *Heudin J.C.* Evolutionary virtual agent // Proceedings of the IEEE/WIC/ACM International Conference on Intelligent Agent Technology. 2004. pp. 93–98.

53.   *Zhao R., Papangelis A., Cassell J.* A dyadic computational model of rapport management for human-virtual agent interaction // International Conference on Intelligent Virtual Agents. 2014. pp. 514–527.

54.   *Badler, N., Allbeck, J., Zhao, L., Byun, M.* Representing and Parameterizing Agent Behaviors // Computer Animation. 2002. pp. 133–143.

55.   *Samani H.* Lovotics: Loving robots // LAP LAMBERT Academic Publishing, 2012. 168 p.

56.   *Bloch L.R., Lemish D.* Disposable Love: The Rise and Fall of a Virtual Pet // New Media & Society. 1999. vol. 1. no. 3. pp. 283–303.

57.   *Adobbati R.* Gamebots: A 3D Virtual World Test-Bed For Multi-Agent Research // Proceedings of the Second International Workshop on Infrastructure for Agents, MAS, and Scalable MAS. 2001. vol. 5. 6 p.

58.   *Fernandez-Ares A. et al.* Its time to stop: A comparison of termination conditions in the evolution of game bots // European Conference on the Applications of Evolutionary Computation. 2015. pp. 355–368.

59.   *Jutla D., Craig J., Bodorik P.* Enabling and measuring electronic customer relationship management readiness // Proceedings of the 34th Annual Hawaii International Conference on. 2001. p. 10.

60.   *Johnson A., Roush T., Fulton M., Reese A.* Implementing Physical Capabilities for an Existing Chatbot by Using a Repurposed Animatronic to Synchronize Motor Positioning with Speech // International Journal of Advanced Studies in Computers, Science and Engineering. International Journal of Advanced Studies in Computers, Science and Engineering, 2017. vol. 6. pp. 20.

61.   *Vieira A., Sehgal A.* How Banks Can Better Serve Their Customers Through Artificial Techniques // Digital Marketplaces Unleashed. Springer, 2018. pp. 311–326.

62.   *Folstad A., Brandtzaeg P.B.* Chatbots and the new world of HCI // Interactions. 2017. vol. 24. no.4. pp. 38–42.

63.   *Liu, Y. et al.* Chatting system, method and apparatus for virtual pet // Google Patents, 2014. US Patent 8645479.