

Critical Consideration of Applications of Affective Robots to Mental Therapy from Psychological and Sociological Perspectives

Tatsuya Nomura* and Noriyuki Tejima†

* Faculty of Management Information, Hannan University
5-4-33, Amamihigashi, Matsubara, Osaka 580-8502, Japan

† Department of Robotics, Ritsumeikan University
Noji-higashi 1-1-1, Kusatsu, Shiga 525-8577, Japan

E-mail: * nomura@hannan-u.ac.jp, † tejima@se.ritsumei.ac.jp

Abstract

This paper deals with the contemporary situation that affective robots are applied to the field of mental therapy, and critically considers problems of these applications by referring to the research works in psychology and sociology on mental reaction of human for computers, social situations on reception of computers, and extreme sensitivity for emotions in the modern society. Based on this consideration, this paper focuses attention on applications of these robots to animal assisted therapy and activity, and narrative therapy, and then considers the problems of these applications.

1 Introduction

There has recently been some researches on applications of robotics to the fields of rehabilitation and therapy. These research works include physical therapy for stroke victims using robot arms [11], therapy aids for children with autistic disorders using a humanoid robotic doll [5], and mental support for people with disabilities and the elderly using interactive robots with functions of emotions [17, 18, 7]. Moreover, in Japan there has recently been a cultural trend on healing, called “Iyashi”. This trend means that many people in Japan require specific goods, human relations, and psychological techniques to maintain their mental stability, implying that the mental capacity of self-recovery of the contemporary people became lower. In addition, some robots implementing emotional behaviors and internal mechanisms are advertised as products for “Iyashi”, as if these two trends synchronize each other. These robots are often referred to with some positive opinions such that they may overcome the problems of more care for more increasing elderly population, lack of therapeutic animals, and sanitation management of them in medical

facilities.¹

Exploring reasons why these trends happened in Japan strictly belongs to a sociological research and we do not deal with it in this paper. This paper aims at considering possibility of mental support by interactive robots with functions of emotions (affective robots) from some viewpoints in psychology (section 2) and sociology of emotions (section 3). In particular, we consider the possibility that making clients of mental therapy interact with affective robots may not lead to aids for them (section 4).

2 Psychological Researches Related to Affective Robots

How robots represents their emotions is roughly formalized as follows:

1. Input data through image, sound, and touch sensors from a user is classified into some categories given in advance or obtained through learning techniques.
2. The current emotion of the user is predicted based on the categories, background knowledge, and rules between the categories and emotions. These rules are given in advance or learned through interaction between the robot and user.
3. The current internal state as an emotion of the robot is determined based on predicted emotions of the user, the previous internal states of the robot, and the background knowledge of the robot

¹For examples, see the following sites:

<http://www.cnn.com/TECH/ptech/9903/25/robocat.idg/>
<http://www.technocopia.com/robots-19990915-astroboy.html>
<http://www.spectrum.ieee.org/INST/nov01/frobots.html>
<http://www.cbsnews.com/stories/2002/03/08/tech/main503340.shtml>
http://news.bbc.co.uk/1/hi/english/sci/tech/newsid_1829000/1829021.stm

in some cases.² The determination rules are given in advance or learned through interaction between the robot and user.

4. The output action for the user is selected based on the emotion and background knowledge, and represented as the motion or face expression.

Even if explicit emotional states do not exist like in case of neural networks, we can regard the above procedures as basic ones. This module structure reflects cognitive approaches for emotions, that is, the idea that emotions are evoked through cognitive evaluation of the current situations, and connection between the evaluated situations and internal states. Moreover, the existence of background knowledge in determination of emotions reflects social constructionism of emotions, that is, the idea that emotions are evoked according to social situations [4]. In these senses, a variety of psychological knowledge can be introduced to implementation of affective robots.

One of problems to be considered is whether affective robots implemented based on psychological knowledge can actually affect mental states of humans. Another problem is whether our societies can receive these robots without mental resistance.

2.1 Mental Reaction of Human for Computers

We can answer for the above first problem by referring to a famous work of Reeves and Nass [15]. By application of theories in social sciences on human interaction and experimental methods in psychology, they showed that experiences of human through artificial media including computers are essentially equal to real experiences, that is, human unconsciously reacts to machines in the same way as to human even if it is recognized that those whom he interacts are just machines. Some important results by them are summarized as follows:

- Humans tend to politely react to machines sending polite messages, and prefer machines sending messages of praise to those sending critical messages.
- Humans tend to interpret even image objects consisting of simple lines as ones having characters by using the same dimensions as those for humans (dominant-obedient and cooperative-non cooperative). In addition, dominant persons prefer machines displaying texts in dominant ways and

obedient persons prefer machines displaying texts in obedient ways.

- Humans tend to be affected by social roles assigned with machines such as professionals, teammates, and genders.
- Humans tend to firstly feel good or evil emotions for information even from artificial media, and have a bias to negative information on attention and memory in the same way as the real world.

Important is that these reactions of human are unconsciously evoked. As a source of these phenomena, Reeves and Nass assume that they are a trace of evolution of mental mechanisms in the ancient wild environment. These phenomena imply that affective robots can affect mental states of human by using actions based on their characters and social roles, regardless of positive or negative direction.

2.2 Reception of Computers as “Objects with Mentality” in the Society

For the above second problem, we can answer by referring to research works of Turkle [19]. She investigated minds of people on computers by interviewing with a lot of people in some countries from 1970’s to 1980’s. Some important results by her are summarized as follows:

- As mechanisms of computers became more complex, that is, they lost “transparency”, users gave up trying to understand computers based on the physical functions.
- Furthermore, as interactivity of computers became increasing, people began to understand these interactive and nontransparent objects by analogy of mental states of human, in other words, regard them as objects having mental states, which is not either just physical objects or living ones.
- Researches on artificial intelligence and biologically inspired models such as neural networks have positively been affecting this trend. As one of the results, artificial agent programs for psychiatry have socially been allowed since 1990’s, while synchronized by standardization of diagnosis and treatment in psychiatry.

Her report implies that our society has been prepared for affective robots as agents of therapeutic animals or therapists themselves.

²Some researchers (for example [3]) focus on motivational drives as another internal states in the robots to determine the current emotion.

3 Sociological Researches Related to Affective Robots

In section 2, we suggested a possibility of influence of affective robots to mental states of human and reception of them in the modern society. However, we must consider whether introduction of these robots to mental therapy naively leads to aids for clients.

For this problem, we can answer by referring to research works on the modern culture for emotions in the sociology of emotions. In the traditional sociology, emotions have actively not been dealt with because social phenomena and organizations have been analyzed based on the assumption that important decisions should rationally be made. Moreover, since analogical models from natural sciences such as physics have succeeded in explaining phenomena even in sociology, emotions have been regarded as those parameters which are impossible to objectively observe. The sociology of emotions [10] is an attempt to develop a new perspective in sociology by focusing on emotions.

Although there are several stances in the sociology of emotions, we deal with one of them, called “symbolic interactionism”. One of the important concepts in symbolic interactionism is “feeling rules” proposed by Hochschild [8]. From the perspective of symbolic interactionism, an actor generates his/her emotions after cognitive evaluation and interpretation of his internal state and given situations. However, this interpretation cannot be done based only on subjective arbitrariness. Regularity that exists in emotional experiences and expressions is socially shared and reconstructed beyond individual situations. In other words, feeling rules are a set of socially shared guidelines that direct how we want to try to feel and not to feel emotions according to given situations. They are often referred to in the form of rights and duties (for example, we often speak of “having the right” to feel angry at someone, or we say that we “should feel more grateful” to a benefactor).

Based on feeling rules and given situations, emotion management is executed. Emotion management means to evoke an emotion appropriate for a feeling rule but not being felt, or to suppress an emotion inappropriate for a rule but being felt.

Emotions in the Modern Society

Based on the concepts of feeling rules and emotion management, Hochschild showed an alienation problem of emotional workers such as flight attendants who are required to have a high degree of emotion management [8]. Furthermore, she claimed that in the modern society emotion management has been spread and too strict, and we are always forced to check whether our

internal states are appropriate for the current situations.

In Japan, Mori [13] also claimed by using the concept of emotion management as follows:

- In the modern society we are always forced to pay attention to our and others’ emotions in order not to hurt our emotions each other (cult of personality).
- Moreover, there is a cultural trend that causes of psychiatric symptoms in individuals are reduced to inner problems in the individuals although the causes may be from social structures and cultural customs (psychologism). Furthermore, therapeutic knowledge has been complementing this trend.
- Cult of personality and psychologism has been complementing each other. Furthermore, psychologism and rationalization in the modern society has also been complementing each other, and as a result we are required to have a high degree of self-control for our emotions.
- Persons executing a high degree of emotion management cannot permit others’ deviation from feeling rules they observe even if it is only a little. This strict observance of feeling rules and difference of the rules between individuals cause disagreement in the modern society.

4 Problems of Affective Robots in Mental Therapy

According to the results mentioned in the previous sections, we can infer as follows:

- Persons facing to affective robots regard them as objects having emotions, and pay attention to the emotions and actions of the robots and themselves through interaction between them, while consciously and unconsciously checking suitability of their emotions and actions for feeling rules that the persons observe.
- However, interaction between humans and robots are different from that between humans themselves, and that between humans and living ones such as animals because robots are regarded as non-living objects having emotions.

Generally speaking, this difference may cause a confusion on emotion management in interaction between humans and robots. As a result, this confusion may lead to a new type of mental cost and unpredictable relations in interaction between humans and affective

robots. It is not guaranteed that this fact has no bad effect for clients in the context of mental therapy (for example, clients mentally tired of emotion management may not prefer to interaction with robots with complex emotion systems).

In the following sections, we critically consider problems that may happen in application of affective robots to animal assisted therapy and activity, and narrative therapy.

4.1 Affective Robots in Animal Assisted Therapy and Activity

Animal assisted therapy (AAT) and animal assisted activity (AAA) mean actions to expect improvement of clients on physiological, mental, and social effects by bringing the clients into contact with animals such as dogs, horses, dolphins, and so on³. As mentioned in section 1, in Japan, affective robots are often advertised as engineering products that can be substituted with therapeutic animals in order to overcome the problems of lack of these animals and sanitation management of them in medical facilities. However, it has sufficiently not been solved yet why AAT and AAA can have some effects for improvement of states of clients, and there are several theoretical assumptions on this problem.

As mentioned in the previous sections, affective robots are regarded as non-living objects having emotions, and are different from living animals that are historically familiar with humans and need to be bred. Thus, researchers aiming to substituting these robots with therapeutic animals should consider from theoretical viewpoints whether this difference influences the effects of AAT and AAA, and which type of clients these robots can contribute to. For example, there are a theory claiming that the states of clients are improved by changing their social roles and actions through breeding living animals, and a theory claiming that mental stability of clients is achieved by satisfaction of an inner desire to touch animals, which is a characteristic culturally hereditary in human minds. From viewpoints of these theories, important is that these animals are living, and naive substitution of affective robots with these animals may not cause either the above change of social roles of clients or satisfaction of the inner desire to touch animals.

On the other hand, there are theories claiming the effects of change of clients' impressions for others and deviation of clients' attention from their problems by the existence of animals in therapeutic contexts, and simply touching. From viewpoints of these theories, there is a possibility of substitution of affective robots

with therapeutic animals since just the existence of them may contribute to therapy regardless of the fact that they are living or non-living. When our extreme sensitivity for emotions is considered, however, affective robots with too complex emotional mechanisms may confuse clients as mentioned in the beginning of this section.⁴ Thus, it should be verified whether and which type of clients emotions of robots can contribute to based on controlled experiments (it is ideal to prepare a controlled group of clients given nothing, an experimental group given therapeutic animals, that given affective robots, and that given robots without emotional mechanisms, and then execute statistical testing for several mental indexes obtained from the experiments).

4.2 Affective Robots in Narrative Therapy

Narrative therapy [12] is one of therapeutic positions in family therapy [6]. Family therapy is originally based on family system theory that caused from cybernetics [2, 20]. In this theoretical framework, a family is a system that consists of its family members including clients and communication between them. It has a kind of homeostasis and the existence of the clients means a result of warped homeostasis in communication. Then, family therapists aim at perturbing the family system to improve states of the system by using autonomous capacity of it, and several therapeutic techniques for it have been developed [6]. However, some family therapists have recently been criticizing meta-positions of therapists for families based on this mechanism and empiricism existing behind the mechanism since power of therapists for clients caused by them oppresses clients themselves. Narrative therapy appeared as a position proposing that therapists stay on equal terms with clients.

Narrative therapists assume that reality surrounding persons do not objectively exist independent from them, but is produced and maintained by "narratives" that are socially constructed through linguistic interaction between them. These narratives give consistency and structures for situations and events in lives of persons, and selves of them. Then, narrative therapists aim at re-organizing narratives on clients' selves that are talked by the clients and produce pain of them, and producing a novel narrative through conversation with the clients on equal terms with them, while removing professional positions of the therapists (for

³In this paper, we referred to [9]

⁴In fact, Dautenhahn mentioned in her research on baby robots for therapy for children with autism that a robot doll that is too humanoid and social could appear confusing and unpredictable for these children because of their specific characteristic [5].

example, there is a term presenting this stance, “not-knowing positions”). Thus, narrative therapy does not mean a concrete therapeutic technique but just an attitude that therapists should have for clients [12].

Criticism for Narrative Therapy and Affective Robots

In Japan, Asano [1] critically considered a position of narrative therapy in the modern society. He claimed as follows:

- The action to talk narratives on selves is one of cultural practices popular in the modern society, that is, there are a lot of increasing people to want to talk narratives on themselves in USA, Europe, and Japan.
- The modern society has a characteristic to produce these people, and industries aiming at satisfying demand of these people like manuals for making narratives on selves, publishers, and so on, called “narrative industries”, have appeared. We should note that narrative therapy is just one of these industries.
- Narrative therapy functions by explicitly drawing things concealed in narratives which clients talk on themselves. However, the desire of people to talk narratives on themselves is also a desire to leave these concealed things concealed. If narrative theorists are not conscious for this fact, narrative therapy has a danger that it only repeats this desire of people

According to Asano’s claim, we can consider a possibility of introducing affective robots to narrative therapy: reception of robots with emotions in the modern society (section 2.2) and narrative industries satisfying desires of people to talk on themselves are combined with each other, and as a result affective robots appear as a product to help people to make narratives on themselves through emotional interaction with people. In addition, in Japan, the cultural trend of “Iyashi” has produced several products for healing and may support appearance of affective robots as tools for narrative therapeutic communication. However, as mentioned above, interaction with affective robots may cause a new type of mental cost for human and repeat desire of people to talk on themselves while leaving concealed things concealed in their narratives, which should be drawn in narrative therapeutic conversation between clients and therapists.

Although technologies of artificial agents are advanced, there seems to be no concrete method to draw concealed things in narratives talked by users of agents

and produce a novel narrative while cooperating with users. For example, although Sengers [16] argued introduction of narrative theories to architectures of artificial agents, no concrete method for realizing it is shown. Moreover, narrative therapy has a stance not to bring objective reality in narratives talked by clients. This stance reflects rejection of natural scientific empiricism under social constructionism. In this sense, introduction of affective robots as scientific and engineering products to narrative therapy is a contradiction in principle.

5 Conclusion and Discussion

In this paper, we critically considered applications of affective robots to mental therapy by referring to the research works in psychology and sociology on mental reaction of human for computers, reception of computers as “objects with mentality” in the modern society, and extreme sensitivity for emotions in the modern society. Moreover, we showed necessity of selection of theories and verification of effects based on controlled experiments in applications to animal assisted therapy and activity, and suggested possibilities of realizing introduction of affective robots to narrative therapy and a bad effect in the therapeutic context.

As future problems, we should empirically investigate whether interaction with affective robots or artificial agents really causes mental cost for human. For this investigation, we are going to plan psychological experiments by interaction between human and artificial agents implementing an architecture consisting of feeling rules and double bind communication [14], to provide an example that negative emotions are evoked through man-machine interaction. Moreover, it seems that most researches on applications of affective robots to mental therapy appear from Japan. We should statistically clarify whether it is real, and analyze causes of it from perspectives of sociological analysis of discourses on both medicine and robotics in Japan. In addition, although we should consider a possibility of selection for these robots in mental therapeutic markets, we should consider this problem discriminating a case that affective robots are supplied through medical organizations from a case that they are supplied as popular products without agency of medical organizations. In particular, a possibility of synchronization between affective robots and the cultural trend of “Iyashi” in Japan should be analyzed in the latter case.

Finally, we mention that we do not totally deny possibilities of applications of robots to mental therapy. There may be some disciplines familiar to robots

in mental therapy (for example, disciplines using syntactical operations in conversation between therapists and clients). However, we should be careful of a naive idea that making robots or software agents behave emotionally like human produces familiarity of human with machines and realization of these robots or agents leads to healing pains of clients in therapeutic contexts. As far as we consider applications of technologies of man-machine interaction to therapeutic fields, we should pay our attention to influences of them in mental, social, and cultural levels. We think that our argument in this paper is not radical but natural.

Acknowledgments

This research was supported by The Grant “Study on Functions of Emotions for Information Exchange in Fields of Social Interaction” from the Institute of Industrial and Economic Research in Hannan University.

References

- [1] T. Asano. *Narrative-Theoretic Approach to Selves*. Keiso-Shobo, 2001. (in Japanese).
- [2] G. Bateson. *Steps to an Ecology of Mind*. Harper & Row, 1972. (Japanese translation: Y. Sato. Shisaku-Sha, 1990.).
- [3] C. Breazeal and B. Scassellati. Infant-like social interactions between a robot and a human caregiver. *Adaptive Behavior*, 8(1):49–74, 2000.
- [4] R. R. Cornelius. *The Science of Emotion*. Simon & Schuster, 1996. (Japanese translation: Y. Saito. Seishin-Shobo, 1999.).
- [5] K. Dautenhahn and A. Billard. Games children with autism can play with robots, a humanoid robotic doll. In S. Keates, P. J. Clarkson, P. M. Langdon, and P. Robinson, editors, *Universal Access and Assistive Technology*. Springer-Verlag, 2002.
- [6] V. D. Foley. *An Introduction to Family Therapy*. Allyn & Bacon, 1986. (Japanese translation: A. Fujinawa, et al. Sogensha, 1993.).
- [7] T. Hashimoto. Emotion model in robot assisted activity. In *Proc. 2001 IEEE International Symposium on Computational Intelligence in Robotics and Automation (IEEE CIRA 2001)*, pages 184–188, 2001.
- [8] A. R. Hochschild. *The Managed Heart*. University of California Press, 1983. (Japanese translation: J. Ishikawa and A. Murufushi. Sekaishishosha, 2000.).
- [9] T. Iwamoto and I. Fukui, editors. *Theories and Practices of Animal Therapy*. Baifukan, 2001. (in Japanese).
- [10] T. D. Kemper, editor. *Research Agendas in the Sociology of Emotions*. State University of New York Press, 1990.
- [11] H. I. Krebs, N. Hogan, M. L. Aisen, and B. T. Volpe. Robot-aided neuro-rehabilitation. *IEEE Transactions on Rehabilitation Engineering*, 6(1):75–87, 1998.
- [12] S. McNamee and K. J. Gergen. *Therapy as Social Construction*. Sage, 1992. (Japanese translation: Y. Noguchi and N. Nomura. Kongo-Shuppan, 1997.).
- [13] S. Mori. *A Cage of Self-Control*. Kodansha, 2000. (in Japanese).
- [14] T. Nomura. Formal representation of double bind situations using feeling rules and triad relations for emotional communication. In *Cybernetics and Systems 2002 (Proc. the 16th European Meeting on Cybernetics and Systems Research)*, pages 733–738, 2002.
- [15] B. Reeves and C. Nass. *Media Equation*. Cambridge Press, 1996. (Japanese translation: H. Hosoma. Shoeisha, 2001.).
- [16] P. Sengers. Narrative intelligence. In K. Dautenhahn, editor, *Human Cognition and Social Agent Technology*, chapter 1, pages 1–26. John Benjamins, 1999.
- [17] T. Shibata. Mental commit robot for healing human mind. *Journal of the Robotics Society of Japan*, 17(7):943–946, 1999. (in Japanese).
- [18] T. Tashima, S. Saito, T. Kudo, M. Osumi, and T. Shibata. Interactive pet robot with an emotional model. *Advanced Robotics*, 13(3):225–226, 1999.
- [19] S. Turkle. *Life on the Screen*. Simon & Schuster, 1995. (Japanese translation: M. Higure. Hayakawa-Shobo, 1998.).
- [20] P. Watzlawick, J. B. Bavelas, and D. D. Jackson. *Pragmatics of Human Communication*. W. W. Norton & Company, 1967. (Japanese translation: K. Yamamoto and J. Ogawa. Niheisha, 1998.).