Innate Sociability: Sympathetic Coupling

Chisato Numaoka
Sony Computer Science Laboratory - Paris
6, rue Amyot
75005 Paris, FRANCE
chisato@csl.sony.fr

Abstract

In this paper, we propose a new approach to emotional interaction of software agents with human users. The basic idea comes from Wallon's theory of human primary development. By assuming an innate structure used in common with both sympathy and social conditioning, we propose a fundamental agent architecture. With this architecture, we make a few experiments, basically for testing conditioning in a social context.

Introduction

Computers in their origin are no more than tools for people. As for tools, we are always conscious of a boundary between computers and people and it leads us to the study on Human-Computer Interface. It was convenient to separate users and computers in those days when we used computers only for limited use.

Vera and Simon (Vera & Simon 1993) say: "All human behavior is social. First and foremost, it is social because almost all the contents of memory, which provide half of the context of behavior, are acquired through social processes - processes of learning through instruction and social interaction." Nevertheless, it is difficult to classify interaction with computers as "social interaction" (although Vera and Simon include computers in a group of social artifacts). This is partially because we face with computer through a given interface.

The recent success of "TAMAGOTCHI [1]" reminds us of the importance of removing a boundary between people and computers. When people use this device in their daily lives, it seems that people forget about the fact that it is artifact. People even cry when a TAMAGOTCHI dies. This appears to prove that people have an emotionally strong empathy with it. In a sense, people have a feeling of identification with their TAMAGOTCHI. This is a kind of feeling: "We play an important role for maintaining its life." or "We are living together." The more we interact with it, the stronger the feeling will become. This is a kind of positive feedback in both directions: people to TAMAGOTCHI and TAMAGOTCHI to people. Why do people do this? I think that this is because people, at least initially, feel sympathy for the TAMAGOTCHI. An ability to feel sympathy may be the most fundamental feeling for humans (and possibly for other mammals). I will discuss this further in the following section.

In this paper, we will seek for a plausible shape of human-computer interaction by looking at human development in infancy and propose an innate architecture of sociability. First we provide an overview of the fundamental ideas of sociability. We are especially interested in Wallon's idea in this paper. Next we present a model of sympathy which explains why we feel sympathy; what could be an origin of sympathy. In succession, we give some examples where we implement software agents that can arouse a person's sympathy. Then, we discuss our viewpoint by comparing it with the preceding ideas in AI or cognitive science. Finally, we conclude this paper.

Prologue: Infants as Immature Social-being

Piaget thinks that humans experience the process of becoming "social" during their growth (Piaget 1948). According to his theory, sociability emerges after we have developed the ability to build representations of the perceived world. On the other hand, Wallon believes that humans are by nature social-beings (Wallon 1945). In other words, we are innately social and this affects all our interaction with the world.

Piaget focuses on the activity of adaptation. Babies, however, cannot survive without adults' support in spite of their attempts to adapt. They cannot even move their arms or hands intentionally. All that they can do is to change the internal states of their body. Wallon, therefore, believes that the primary activity of babies must be to change the state of their body through emotional behavior such as crying. This emotional behavior, then, attracts adult's interest and causes the adults to take care of the baby.

Assimilation and adjustment are important notions

[1] TAMAGOTCHI is a trademark of ©Bandai.

in Piaget's theory whereas attitudes and emotion are central to Wallon's ideas. Wallon does not ignore the observation of infants trying to assimilate the world and to adjust their schemes if necessary, but argues that it is not a primary activity of babies in the early stages of life. He does not think that humans become social at the end of this type of activity as Piaget suggests. He stresses rather activity as a social-being, which is emotional behavior. He thinks that human development is a process of developing a better relationship for mutual understanding through emotional behavior since our own growth levels depend on individuals.

To summarize, Piaget focuses on a highly intelligent skill of social interaction whereas Wallon considers primitive interaction as a sign of social interaction. In software agents, both levels of sociability have to be considered, although we hitherto have dealt mainly with the former level of sociability. It is worth considering sociability as an innate characteristic of humans to compensate for their inability to survive in the world alone.

Sympathetic Mutual Understanding with Innate Structure

Conditioned reflex

Pavlov discovered that a mental factor affects the secretion of gastric and salivary glands. Based on this discovery he made a famous experiment. He repeated a procedure of giving food just after having applied a conditioned stimulus, in effect, a metronomic sound at a constant frequency. This is the so-called conditioned reflex.

Babies establish an emotionally symbiotic relationship 2 to 3 months after birth. According to Wallon, throughout this time, conditioned reflex is formed based on their two major desires of ingestion and posture change. Here the meaning of conditioned reflex is the same as that of Pavlov except that there is no artificially provided conditioned stimulus.

> When babies feel hunger through interoceptor, there is no other way for them than responding with propriofector or interfector by such an action as crying. ... the representation of posture and emotion is conveyed to his / her mother, and eventually mother satisfies his / her desire. Through the repetition of this activity between a baby and his / her mother, the action of crying becomes a symbol indicating a desire for food. This is clearly conditioning. ... Namely, we can understand Wallon's idea that cooperative relationship with others is formed through classical conditioning regarding posture and emotion (Hamada 1994).

Wallon states: "Infants are tightly connected with their familiar people and can hardly separate themselves from others. " This is actually an important idea

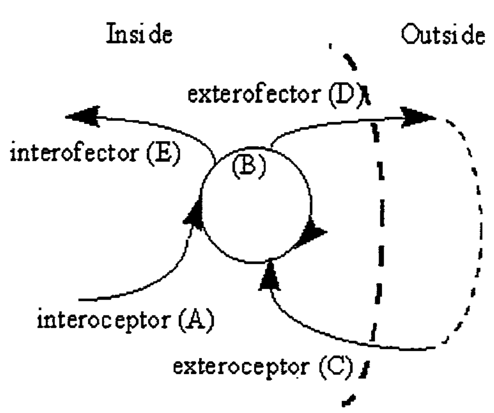

Figure 1: Innate Structure

that we need for next generation human-computer interaction; all participants work in a chain as if they were parts of a system or self. In a sense, this is a basis of cooperation.

Sympathy: a basis of parental affection

Here I attempt to understand why mothers try to satisfy their babies desire (in other words, why people can understand other's desire to be satisfied) in extension of Wallon's idea. First of all, we will present a simplified model of human activity in Figure 1. We can think of a metabolic cycle in the center in association with the hypothalamus [2]. It is reasonable to assume four types of input/output: (A) interoceptor (e.g. a sense of hunger), (C) exteroceptor (e.g. producing expression), (D) exterofector (e.g. looking / hearing / feeling / tasting / smelling), and (E) interfector (e.g. regulation of respiration or heartbeat).

Let us take an example; feeling hunger. When a baby feels hunger, a trigger A stimulates a cycle B. In part, B activates a behavior of crying D as well as applying hormonal/neural regulation E. The voice of crying is heard by him/herself C and this affects B as a positive feedback in relation to hunger. Therefore this initiates conditioning with an unsatisfied experience.

Even after he/she grows up, this conditioning may remain at some level. If this is true, then when an adult hears the voice of an infant, C', crying for nutrition, this may automatically recall an emotional feeling of hunger in B' as if he / she were really hungry (Figure 2). Of course, reality must be much complex. Nevertheless, I believe that conditioning concerning life maintenance may remain consciously or unconsciously in memory.

[2]The hypothalamus is an autonomous nervous system, the most essential organ for the maintenance of life. It is also assumed to be the central cause of emotional expressions.

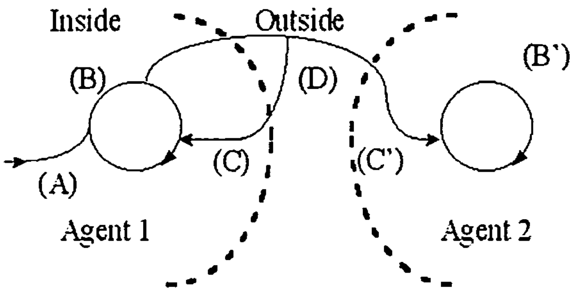

Figure 2: A Mechanism of Sympathy

So, this is a plausible reason why we feel sympathy with each other, particularly for human expressions concerning life maintenance. In analogy, we could create software agents with which human users feel sympathy if these agents have a similar innate structure to condition themselves; with inputs from others, augmented with a signal originally produced by themselves and captured by their own organ.

Crying is not originally emotional

Following the idea of Wallon, we can postulate that a given emotional behavior, *crying*, will not be bounded to such a particular sense as we describe it as, for example, painful, when a baby is born. From the behaviorist point of view, crying is just a result of making the best effort to maintain its life system. When hungry, a stomach contracts. To avoid hunger for the moment, it might be better to expand the stomach so it does not feel empty. This recovery regulation may cause an action: to take in as much air as possible. This in turn may involve motions of the diaphragm and the trachea and as a result a cry will be produced. Therefore, we could say that the behavior of crying is simply a result of posture regulation.

By repeating the same action and undergoing parental reaction, babies will be conditioned. After that, they may learn how effectively they cry by connecting their observation of the world with their mental condition.

Based on this idea, we could imagine a fundamental innate architecture as in Figure 3. Here, there is a central regulator (which is assumed to be the hypothalamus) and a learning system. The other four modules may be, for example: 1, a stomach; 2, an eating module; 3, a vocal system; and 4, an acoustic system. Arrows a, b, and c indicate respectively a sound output, a sound input, and food. Arrows x, y, and z are signals which activate destination module. Arrows α, β, γ, and χ are the inputs to the learning system for conditioning. This architecture has the following characteristics:

- it has an a and b pair. This pair corresponds to the C and D pair in Figure 1.

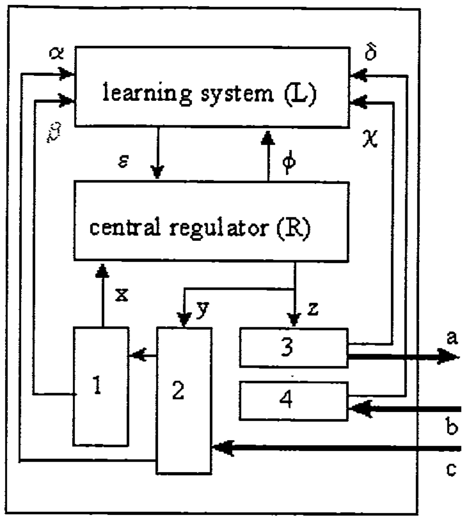

Figure 3: A typical innate architecture

- it produces an output a so as to receive a satisfactory input c.
- it inputs δ as well as χ. As long as it is only this agent that produces an output (which is also received as input b), having these two inputs is for redundancy only. Nevertheless, what we expect is that a user also produces an input that can be received from the the same input port b but as an unrecognizable signal.

Figure 4 shows a typical example of a time series of digital inputs to a learning system as well as that of state of control signals among a central regulator and three modules. Input β indicates a state of satisfaction with the nutrition. Note the relationship between β and the other two inputs α and χ. It is clear that the state of satisfaction β changes at some time while α is 1 and δ is unknown (represented as a symbol x). Of course, it is possible to condition this agent even if we consider only input c. However, it never differentiates reasons of satisfaction; the agent achieves it whether provided by itself or by others. It will always believe that a nutritional need was satisfied because it had a corresponding desire, which is represented by a positive state of regulation signal y. It will never relate the reason to another person's support. It becomes possible to discriminate two cases only by receiving an unknown input from a user. This is the most fundamental contribution of this architecture.

Experiment

In this section, we introduce a few experiments which embody the innate architecture for sociability. The goal of these experiments is to allow agents to exhibit a kind of conditioned reflex. More precisely, we expect

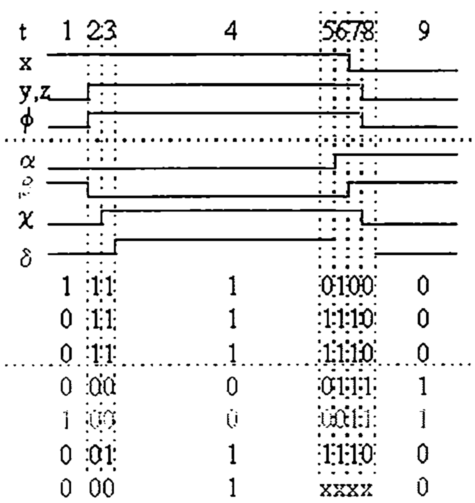

Figure 4: Time series of inputs to a learning system

an agent to exhibit partial satisfaction only with an input b by a user before satisfying an original desire to obtain an input c in Figure 3. In both experiments, we implemented the learning system to produce a conditioned reflex by using a very simple neural network which produced an expected satisfaction (ε in Figure 3 which eventually effect regulatory signals y and z),

Jukebox agent

There is an avatar of an agent in a virtual world which plays MIDI music. The agent wishes to continue to hear a music but the switch to start playing music is not under its control. The only way to start the music is to cry like babies. When the agent finish playing a piece of music, a list appears at a side of the avatar, which shows a selection of music. If the avatar is not in a user's view, it cannot hear a user's voice. What a user can do for the avatar is to change his/her viewpoint and to choose one piece of music in the list.

This agent has basically the same architecture as that described in Figure 3. Its basic components are: 1) a music player, 2) a music menu manager, 3) a sound generator with a speaker system, and 4) a sound receiver with a microphone system, where the reference number corresponds to that in Figure 3. Menu selection is done by clicking a mouse as input c.

Navigation with an agent

Suppose that a user is navigating through a virtual world accompanied by an avatar of an agent. The agent always tries to do its best to keep following the user. When the agent is outside an area where the user is in the center, the agent begins to cry. If the user approaches and the agent again enters the area, it stop crying.

Basic components are now: 1) a wandering system within a limited range of a circle where a user avatar is in the center, 2) a distance tracker and calculator (it generates 1 when the distance between a user and this agent is in a defined range of a circle), 3) a sound generator with a speaker system, and 4) a sound receiver with a microphone system. Input c is always the position of the user.

Discussion

In this section, we compare our idea with some preceding ideas in AI. We would like to emphasize that this is new and how it compensates for other ideas to produce socially intelligent agents.

Behavior-based AI

In a sense, one root of our research can be found in literature on behavior-based AI. Brooks describes behavior-based creatures using a so-called subsumption architecture (Brooks 1986). The important idea is that any kind of behavioral patterns that the creatures show is grounded in sensory inputs. With the least set of primitive behaviors, without symbolic expression, they exhibit apparently very complex behavior. We can see intelligence in their behavior but there is no particular description of intelligent behavior. Brooks calls this: "Intelligence without expression."

In the same sense, we could say: "Emotion without expression." We may feel emotional expression to an agent's attitude. However this is not because agents have an emotional state but because we can have a corresponding emotional feeling. Brooks's creatures have been criticized and our approach faces a similar difficulty; how can we achieve higher levels of emotional attitudes? This is a challenging topic.

Believable agents

Believable agents are animated characters with emotion (Bates 1994). Bates' believable characters have attracted people because they can imitate human-like expressions of emotion and create "illusion of life." Bates is interested in believability of characters. People will have a feeling that they are living together if they feel any kind of sympathy. For this objective, Bates emphasizes the role of the emotional state of the character.

I agree with him in the sense that software agents must demonstrate behavioral patterns, even partially, with which we are familiar in human society. By so doing, people can project their own feeling onto the agents. Some studies emphasize this point (e.g. Neurobaby (Tosa 1993) and ALIVE (Maes 1995)). Nevertheless, I found it difficult to strictly define emotional states. As mentioned before, emotional expression will basically be grounded on some particular needs of bodily regulation.

People may say that the needs of computers are different from those of humans, so it may not be possible to produce similar emotional states. This is true. We do not neglect this point. Our approach is to implement the smallest set of action which metaphorically resembles that of humans (e.g. crying) instead of adopting emotional states directly. Emotional states, in my opinion, arise only when we objectively view someone's attitude.

Our software agents can produce an association among a bodily need, a related course of actions, and an experience. Once we feel sympathy with an attitude of an agent, then if we observe a similar attitude in the same agent, we may project to it an emotional state that we have individually. An interesting question is how much vocabulary of "emotion"-like attitude agents will produce. This is a topic of further study.

Primate social intelligence

Worden shows us a plausible account of how social intelligence will function based on the assumption that it is expressed as scripts some of which are innate and the rest of which could be learned later. Scripts are simple information structures on which three basic operations may be applied: intersection, inclusion, and unification. By virtue of these operations, a new script can be learned by specializing some innate scripts. According to (Worden 1996), examples of social intelligence typical to primates are: to use kin relations, to habituate to calls, to learn alarm calls, rank and alliances, emotional responses, tactical deception, and so on.

Basically, we do not intend to explain this typical social intelligence in our model of innate sociability. I accept the idea that there must be some innate structures enabling Primates to exhibit social intelligence. Script theory is very sophisticated especially for explanations. Worden's script reminds us of Piaget's scheme.

Nevertheless, I suspect that there must be more fundamental innate structure as explained in this paper. It is an innate structure that enables infants to associate themselves with the surrounding society. we can hypothesize that we (or some higher mammals) have been able to abandon advanced development at birth for immature birth because we have developed a reciprocal nature with which we sympathetically understand the others' suffering or pain and automatically respond with an appropriate reaction. An interesting study would be to deal with individual cases of social intelligence with our model of innate sociability.

Conclusion

In this paper, we have hypothesized that there is an innate structure that contributes to primitive social interaction on which infants strongly rely. Following Wallon's idea of human early developmental process, we have postulated that we must have an innate structure coupled with a bodily regulatory system and by virtue of sharing the same structure, we can feel sympathy with each other. This feeling is especially strong for mothers towards their babies.

Subsequently, we have described a model architecture of software agents that embodies a primary social interaction by which we may have sympathy with them as well as by which the agents can acquire conditioned reflexes to produce social coupling of an agent and a user. This model was implemented on a few experiments by instantiating abstract modules in the architecture. The characteristics of this architecture are that: 1) An agent produces any observable output (e.g. voice) so as to satisfy its desire, 2) it objectively observes the output by itself, 3) In the same channel of observation, the agent receives an input from a user, which is recognized as unknown and uses it together with a desired input for learning or conditioning.

In this paper, basically, I focused on the problem of conditioning as a primary development of social coupling necessary to maintain their lives, and applied the idea to software agents. We mentioned the fact that we need to test sociability of agents by using them in a real context and to extend to the awakening of higher levels of sociability. It might not be so simple. Nevertheless, I believe that the idea of sympathetic coupling between agents and users gives us one of important aspect of a fundamental architecture of social agents.

Acknowledgments

The author wishes to thank Philip Mayen for the use of his source code of Java Script and VRML file. It was a great help in our experiments.

References

Bates, J. 1994. The role of emotion in believable characters. *Communications of the ACM* 37(7):122–125.

Brooks, R. A. 1986. A Robust Layered Control System for a Mobile Robot. *IEEE Journal of Robotics and Automation* RA-2(1):14–23.

Hamada, S. 1994. *Piaget and Wallon (in Japanese).* Minerva Shobo.

Maes, P. 1995. Situated Action: A Symbolic Interpretation. *Communications of the ACM* 38(11):108–114.

Piaget, J. 1948. *La naissance de l'intelligence chez l'enfant (2nd Edition).*

Tosa, N. 1993. Neurobaby. In *Visual Proceedings of SIGGRAPH-93 Tomorrow's Realities.* ACM SIGGRAPH.

Vera, A. H., and Simon, H. A. 1993. Situated Action: A Symbolic Interpretation. *Cognitive Science* 17:7–48.

Wallon, H. 1945. *Les origines de la pensée chez l'enfant.* Presses Universitaires de France.

Worden, R. P. 1996. Primate Social Intelligence. *Cognitive Science* 20:579–616.