Turn-taking Supports Humanlikeness and Communication
in Perceptual Crossing Experiments

-Toward developing human-like communicable interface devices-

Hiroyuki Iizuka, Davide Marocco, Hideyuki Ando and Taro Maeda

Graduate School of Information Science & Technology
Osaka University
2-1 Yamadaoka, Suita, Osaka, 565-0871, Japan
{iizuka, hide, t_maeda}@ist.osaka-u.ac.jp

Centre of Robotics and Neural Systems,
University of Plymouth
Drake Circus, Plymouth PL4 8AA, United Kingdom
davide.marocco@plymouth.ac.uk

Abstract—Our aim of this paper is to investigate the human communication in terms of two following questions. One is how human can know the fact that an interacting partner is human. Another is how non-communicative behaviour can become communicative. To answer these questions, we performed two experiments exploiting the idea of perceptual crossing experiments. As a result, we will show that the turn-taking structure supports humanlikeness and human communication in the primitive non-verbal interaction. Our results will be discussed with ambient interface technology.

Keywords-Ambient information society, behavioural turing test; emergence of communication; perceptual crossing experiment;

I. INTRODUCTION

The concept of ubiquitous information society is to establish sensory-spaces where a number of and various kinds of sensors are distributed, to collect all information of the space and to display collected useful information to everybody. All users are always connected information society and “anywhere, anytime, anyone” can access to information in a flexible manner. The concept of ubiquitous has a lot of potential in our future society. However, human capability to understand such information is limited and the amount of information can easily go beyond user’s handling ability. Recognizing these situations, a slightly different concept has been proposed, which is ambient information society [1]. It is a similar concept but differs in a following important point. As well as the ubiquitous information society, the sensor-spaces are required but those collected information is given to users personally and intuitively. It means that the ambient information society provides desired information to a user according to their circumstances. Therefore, the concept can be “information is given to users according to current time, place, and personal preferences”.

The difference of required technologies between ambient and ubiquitous information societies is that the former needs information of the behavioural context to provide user’s desired information at a proper timing for a user. Therefore, the integrated system for the ambient information society needs to actively interact with users somehow to read their intention or states rather than passively collecting information.

978-1-4673-1246-2/12/$31.00 ©2012 IEEE
target of the second experiment is how human can establish an explicit communication from the simple bodily interactions. Finally, our results will be discussed with how to develop an interface device that can interact with human.

II. Behavioural Turing Test

The turing test is known as a test of a computer’s ability to demonstrate human intelligence in a natural conversation. In this section, we performed behavioural turing test to clarify how human can know the fact that an interacting partner is human in the non-verbal interaction.

A. Experimental Setup

A schematic view of the experimental setup is shown in Fig. 1. Our experimental setting was based on the perceptual crossing experiment by Auvray [3]. Subjects could move left or right by his/her finger along a rail on the desk that restricted the motion precisely on the one dimensional space and made it possible for the CCD laser displacement sensor to detect the position of the finger without lost. A voice coil attached on the subject’s nail gave input signals by vibrations when the avatar touched objects in the virtual space, i.e., subjects received only all-or-none information. This experimental environment is different from the original perceptual crossing experiments in terms of reliability of the space and inseparation of sensation and motion. In the original experiment, the correspondence between physical and virtual position is easily broken since a computer mouse is used and also the stimulus given to another hand. On the other hand, there is no spatial gap and both happen in a single finger.

B. Task

The subjects were asked to answer if the partner had been controlled by human or not after 30 seconds interactions with an unknown object. They knew that a single object existed and it did not disappear during the experiment.

In the experiment, the subjects interacted with two kinds of objects. One was the avatar that was controlled by another subject. Another one was an object that replayed the partner’s previous motions. A trial consisted of 20 interactions, 10 of them were interactions with the human avatar and another 10 were with the recorded motions. After each trial, the subject could know how much he/she could correctly answer human or not-human. The recorded motions were made from previous trial results where the subject had answered that the partner was a human.

C. Human-Human Interaction

Fig. 2 shows the results of correct answer rates by two pairs of subjects. At the beginning of the experiment, subjects were not used to this low-level environment. They did not know how to interact with each other. Consequently, both players could not tell if the interacting partner was human player or not-human player. The rate became around chance-level (50%). During 4th or 5h trials, the rate did not go beyond the threshold that indicates significant difference from random choice, but after 5th or 6th trial, both subjects started understanding something with each other in the partner’s behaviour and discriminating human player’s behaviour with non-human’s. At the tenth trial, player 2 can achieve the discrimination perfectly in either pair.

In order to see what produces the clues for the discrimination, the behavioral patterns and input signals during an interaction by the right pair in Fig. 2 is shown in Fig. 3. The results of the both-human-player case are shown. At the beginning of the experiment, their motions were not organized and it was difficult to know how they understood the fact that the partner was human (Fig. 3(top)). Actually, the rates of correct answers were still around chance-level as the subjects reported that they were not sure if the answer was correct or not. Around 10-80 interactions, the player 2’s behaviour strategy became completely passive. The player tried to see the partner’s reaction while moving to one direction constantly (Fig. 3 (the second graph from the top)). The partner attempted to interact with the behaviour but it was not reactive. Therefore the player 1 could not tell if the partner was human player or not since both behaviours were not reactive. Furthermore, this strategy did not produce good results all the time for the passive player itself, either. If the passive player did not have diversity in his behaviour, the partner’s behaviour would not have diversity as well. In turn, it became difficult for the passive player to achieve the discrimination. For a long time, player 2 used the passive strategy but after 6th trial the strategy changed (Fig. 6 (the second from the bottom)). Both players moved more actively than previous examples and also the behaviours were clearly different from irregular motions as the initial stage. After
about 120 interactions, both players’ motions were organized and we could see turn-taking behaviors where roughly they have two fixed roles such as moving and staying, and those roles were alternately exchanged. Player1 touched player2 by oscillating while player2 stayed at a certain place and observed player1’s behavior. After a while, the roles were exchanged. Player1 stopped oscillating and stayed at a certain place and player2 started oscillating. At the beginning of the interaction, they tried to establish this role switching and it continued during the experiment. The turn-taking behaviours were clearly shown in Fig. 6(bottom) at the end of experiment.

D. Emergence of turn-taking

In order to clarify the emergence of turn-taking, we calculated the performance of turn-taking behavior as follows. First, behavioral patterns of each subject were classified into moving and staying behaviors by simply using a threshold. The turn of a player was defined as the state where only the player moves while another stops. If both players move or stop, it means that no one takes turns. The turn-taking performance was calculated by simply multiplying times of each turn. Fig. 4 shows the evolution of the average turn-taking performance at the end of human-human interactions of each trial. It is clear that the successful pairs of subjects established turn-taking behaviors over trials and that the pairs who could not achieve the task did not establish turn-taking behaviours.

III. Emergence of Communication

The objective of the second experiment is to favour the emergence of a minimal communication system in the most open and unconstrained way possible. We investigate: (a) the interaction between communicative and non-communicative behaviour, and (b) to observe how simple sensorimotor interactions become a form of communication.

A. Task

A pair of subjects were assigned a collective categorisation task in which they were asked to judge whether the figure perceived by one subject was identical to the figure perceived by the partner. To complete this task, they interacted with each other in the experimental environment whose setup was same as the first experiment.

Square and Sharp were used as the displayed figure in the experiment. In particular, to accomplish the task, the subjects were required to stay in the same portion of the environment (stay together), i.e. the two avatars must occupy the same position in space, when they think that the two figures are the same, and occupy a different portion of the environment (stay away) when they think the two figures are different.

At the beginning of trials, subjects moved their avatars to an initial position, and a single figure illustrated by white lines on a black background was shown on each monitor for 30 s. During the 30 s, subjects were free to interact with each other to complete the task. After the 30-s interaction, the next trial started.
B. Emergence of turn-taking for communication

Fig. 5 shows example behaviour of the best pair recorded during an interaction game. At the beginning of the experiment, the subjects had no idea how to move or how to communicate with each other. After a few rounds, their behaviour started to become organized. Subject 1 touched subject 2 by oscillating in his position, while subject 2 stayed at a certain place without moving and observed subject 1’s behaviour. Soon, the roles were exchanged. Subject 1 stopped his oscillatory behaviour and remained fixed in a certain place, and subject 2 started to exhibit oscillatory behaviour. It means that the structure of the turn-taking behaviour had emerged, i.e. the subjects moved in an alternating manner and took turns to exchange signals, and we could also observe some encoding of the signals. At this point, they could correctly communicate with each other, while the structure of the turn-taking behaviour was maintained.

The observed turn-taking structure was crucial to establishing the communication protocol. It determined the direction of the information flow out of the non-directional embodied interaction experienced by the subjects. Without turn-taking, in fact, the subjects could not understand why the vibration occurred. They did not know if this had happened because they touched their partner while moving, because their partner touched them, or because they were moving together. However, due to the turn-taking interaction, they knew both who was moving and when they should stay in a certain position by doing nothing. In the turn-taking behaviour, we could recognise active and passive modes: under the active mode, a subject actively touched the partner with the intention of sending a message, whereas in the passive mode, a subject can only receive a message while they are being passively touched by their partner.

Communication can occur by appropriately modulating the two modes.

Fig. 6 shows the correlation between turn-taking and task performance. Regardless of success or failure, all data are plotted in the graph. A correlation was found between them ($r = 0.55$, $p < 0.05$), indicating that the communication tasks were more successful when the turn-taking structure was established.

IV. DISCUSSION

The turn-taking structure was observed in two different experiments. In the first experiment, turn-taking emerged when they discriminated human player with non-human. It is considered that they try to established implicit communication system to detect the difference between human and non-human. In the second experiment, the objective of subjects is that they need to send messages to know what the partner is watching. It means that they are required to establish an explicit communication system.

In order to develop the ambient interface, somehow interface devices need to interact with users to communicate and to detect their intentions or contexts. Our results suggest that the turn-taking structure might be able to induce the feeling of humanlikeness and that it is important to naturally communicate with users even on the non-verbal interactions.

ACKNOWLEDGMENT

This work was supported by “Global COE program: Founding Ambient Information Society Infrastructure” of the Ministry of Education, Culture, Sports, Science and Technology in Japan.

REFERENCES