As Time Goes by: Representing and Reasoning About Timing in Human-Robot Interaction Studies

Hatice Kose-Bagci, Frank Broz, Qiming Shen, Kerstin Dautenhahn, Chrystopher L.Nehaniv

Adaptive Systems Research Group, University of Hertfordshire,
School of Computer Science,
Hatfield, Herts, AL 10 9AB, U.K.
{h.kose-bagci;f.broz;q.shen;k.dautenhahn;c.l.nehaniv}@herts.ac.uk

Abstract

We summarise the experimental design issues related to timing in several human-robot interaction scenarios investigating turn-taking or synchronization between child-sized humanoid robots and human participants. Our aim[1] *is not to have the humanoid robots just replicate the human's behaviours (e.g. waving, peek-a-boo, or drumming), but to engage in interactions in a socially appropriate manner. From these various studies, we have identified several ways in which time has an impact on interaction. We have also identified practical concerns about data collection for time-dependent interactions and ways to address them. The conclusions drawn from this work is likely to be useful in informing the design of systems which engage in synchronized or turn-taking interactions with people.*

1. Introduction

Timing plays a fundamental role in the regulation of human-robot interaction and communication. We present the experimental design and analysis issues related to timing based on three exploratory studies investigating imitation based interaction games with child-sized humanoid robots and human participants. The primary goal of this work is to achieve (non-verbal) gesture communication and imitation between child-like humanoid robots and human beings, whereby interaction games including drumming and imitation served as a test bed to study key aspects of face-to-face interaction such as turn-taking, synchronisation and non-verbal gestures.

[1] Acknowledgements: This work was conducted within the EU Integrated Project RobotCub ("Robotic Open-architecture Technology for Cognition, Understanding, and Behaviours"), funded by the EC through the E5 Unit (Cognition) of FP6-IST under Contract FP6-004370.

The first presented study is based on *drum-mate*, a drumming game where turn-taking is deterministic and head gestures of the child-sized robot KASPAR[1] accompany its drumming to assess the impact of non-verbal gestures on the interaction [2]. This paper will mostly focus on a modified version of this game, based on emergent turn-taking dynamics; here our aim is to have turn-taking which is not deterministic but emerging from the social interaction between the human and the humanoid [2, 3].

In related research on turn-taking, the interaction history architecture (IHA), a system for continuous behaviour sequence learning, was extended to support the acquisition of turn-taking behaviours (peek-a-boo and drumming) during the course of a simple interaction game with a human participant [4]. While no user studies have yet been conducted with the new version of IHA, the changes to the architecture made to support turn-taking and anecdotal evidence from interactions with the system suggest that a time-based representation of both the robot's and human's action history can play an important role in successful turn-taking behaviours.

Another set of studies focused on imitation. Unlike the turn-taking studies, these are based on synchronisation, which introduces different issues related to timing. Here the robot makes simple body movements such as waving its hand, and the human tries to imitate the robot while the robot evaluates how successful the imitation is [5].

In still another set of studies (currently in the planning stages), our aim is to analyse and model the gaze behaviour of human-human and human-humanoid pairs. Therefore, we need to track the gaze of the participants coming from different sources in real-time and compare them to detect joint and mutual gaze. Additionally, once the data is collected, a suitable representation for the time distribution of the periods of mutual gaze must be chosen.

As stated before, timing is a crucial part of interaction, and delays, gaps and mismatches in responses have a negative effect on the participating human (for a detailed survey on timing please refer to [2]). Human participants have reported negative evaluations if there are delays and gaps in robot's time of response [6]. Likewise, delays between infants' actions and caretakers' responses result in negative reaction from infants [7]. In the rest of the paper, we attempt to categorize several issues in HRI related to the timing, and analyse them within the context of our different experiments as briefly stated above.

2. Issues related to timing

A. Turn-taking issues

We implemented the human-robot drumming game as an example of a *call and response turn-taking interaction.* In the deterministic case, we used predefined fixed time duration heuristics for turn-taking. The human partner started by playing simple rhythms with a toy drum. KASPAR started playing if the human was silent for a few seconds. However, it was not always clear when the robot or human partner should initiate interaction in taking a turn. In the second version of the study (the emergent case), we instead used probability-based computational models to control timing and turn-taking. Three simple models were used to control the starting and stopping of the robot's drumming beats. These models were based on different parameters (i.e., number of beats by the human in a turn, or the duration of the robot's current turn) that were extracted from the previous turn of the human and robot. The output of the models determined the number of beats to be played by the robot for the current turn or the duration the robot would give to the human's turn to play. The temporal dynamics of turn-taking thus emerged from the interaction between the human and the humanoid. We studied how these models impacted the drumming performance of the human-robot pair and the participants' subjective evaluation of the drumming experience.

Fig. 1 The drum-mate experiment

During the experiments we tested three basic computational models (linear, threshold, and hyperbolic). The aim was not to find the best model but to see how different models can affect users' behavioral and subjective evaluation.

The detailed analysis of the behavioural data (the amount of human's and robot's drumming and the durations of the drumming, the differences between human's and robot's drumming performance, and turn-taking measures) showed that the different models led to different behaviour for the robot, and thus also impacted the behaviour of the human participants and their evaluation of the games based on timing and the social interaction between the robot and themselves. They did not like the game which gave them the least play time (which led to many turn-taking conflicts) because they thought there was not enough interaction. They also disliked the game which gave them longest time to play, which led to many delays and gaps in turn-taking. Although no significant difference was found between the games in amount of human drumming, the human participants preferred the model which gave them the intermediate play time each turn (the hyperbolic model), which resulted in the least gap between the turns, and found it more natural. It seems that humans preferred the drumming interaction to be reciprocal, with both human and robot drumming for a similar amount of time overall. This suggests that longer-term time-based models of the history of turn durations may be useful for robots that engage in turn-taking with humans, to help produce "fair" turn-taking behavior that people will find enjoyable.

In the deterministic version of the drumming experiments we used simple gestures to help with turn-taking by indicating the beginning and end of robot turns. Surprisingly, the gestures did not seem to reduce turn-taking errors (the game with the highest amount of gestures had the highest error rate). However, they had a large impact on the behavior of the participants, increasing their amount of play, and were preferred to the non-gesture condition. In this particular experiment, we observed that using too many gestures distracted the participants from the main focus of the experiment and caused delays in humans' response, since the humans spent a longer time trying to understand and react to the gestures. Still, a small amount of simple head gestures (blinks, nodes, smiles) could be successfully used in interaction games to help with turn-taking and motivation. People enjoyed the gestures, and may have felt that they were better able to detect the beginning and ends of turns even though they took longer to respond.

The extension of the interaction history architecture to support turn-taking also yielded insight into that types of time representations lead to successful turn-taking interactions. IHA makes action selections based on the reward associated which previously encountered experiences (a collection of sensor and proprioceptive data that is roughly analogous to “state” in state-based decision-making systems). An experience collects the immediate sensor data for a period of time typically roughly on the order of the length of one of the robot’s action executions. While associating the immediate reward with sensor information at this time-scale allows meaningful reward calculations in many cases, it is not sufficient to encourage turn-taking behaviors. In order to introduce rewards relevant to turn-taking, a short term memory module was added to the system. This module collected a sequential history of relevant sensor data and the robot’s actions over a time scale of several robot actions durations. This short-term data and action history was analysed to provide scores for turn-taking performance that could influence the reward calculation for the current experience. The addition of this memory allowed the system to learn to successfully engage in both peek-a-boo and drumming turn-taking interactions.

While the peek-a-boo interaction involves alternating single acts of face-hiding on the part of the robot and the human, drumming has greater variability. Anecdotally, the interactions which seem most successful are those in which the robot learns to play only a small number of drum beats during its turns. This is likely because the robot drums at a much slower rate than a typical human. Humans seem to try to match their drumming turns to the amount of time that the robot spent drumming rather than the number of beats played. This offers support for the theory that models of turn-taking should be time-based rather than just action based. Given the findings about compatible turn lengths over entire episodes of turn-taking interactions in the drum-mate studies, it would be interesting to see the effect of adding longer term information about the relative lengths of past turns to the memory module.

B. *Synchronisation issues*

Synchronisation is another vital issue in timing which we encounter during human-robot interaction experiments. Many interaction games, including imitation of physical motion or dancing (simple rhythmic motion) with music, are based on synchronisation. The second experiment we present was based on the synchronisation of the human and humanoid using a simple arm waving motion. We used magnetic motion trackers to detect humans’ arm motion and compare it with the position of the robot’s arm joints, and tried to detect the synchronisation between them in different scenarios, i.e. waving hands totally in phase or out of phase [5]. We proposed a method based on *information distance* [8] to detect the similarity and synchronisation between the motion of human and the humanoid robot KASPAR2.

During the experiment setup and process, we were faced with several issues related to synchronisation. We had to collect data from the robot and human online and correctly match them in order to compare them and calculate the information distance for the time series.

Fig. 2 Waving imitation game

In [9], we studied the effect of music on synchrony for this interaction. In this work, the participants were told to follow a humanoid robot’s simple waving motion (horizontal/vertical) (Fig.2). The different conditions in this experimental setup were contingent (horizontal-horizontal/vertical-vertical) or non-contingent (horizontal-vertical) movement, with/without music. Although we could not find a significant effect of musical conditions on synchronisation, the results showed that the participants were effected by the robot’s behaviour rhythm in the experiments and adapted to it, which confirms previous results on timing adaptation in human-robot interaction experiments [10].

The music chosen for the experiment was simple and well known (we used “baa baa black sheep”), which encouraged the participant to move and was compatible with the tempo of the full cycle of the robot’s motion, in this case horizontal and vertical waving of one hand. If the robot’s motion and the music do not fit rhythmically, it could be distracting for the human, and they might choose to adapt to the music instead of the robot. In this particular experiment the robot’s waving speed was fixed.

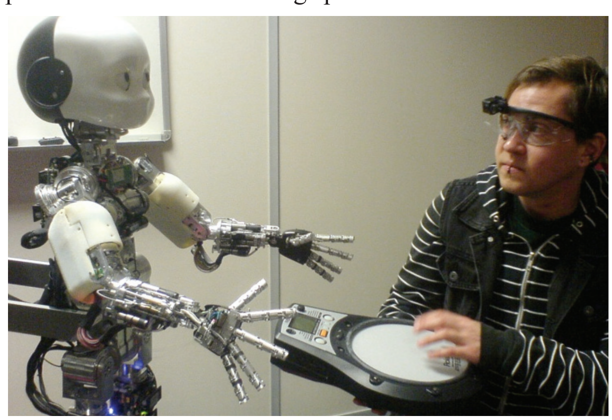

Fig.3 interaction with robot using IHA

C. Issues related to measurement devices

In a study still under preparation, we analyse the gaze habits of two human subjects. Human-human pairs sit at each side of a table looking towards each other and engage in simple dialogues while their gaze direction is recorded using gaze-tracking devices.

Timing is a big issue in using cameras, gaze-trackers, and other sensors to measure body motion for human-robot interaction studies. The data coming from these devices should be time-stamped in order to be able to accurately associate it with data from other sources. If data is collected on different computers, synchronizing time-stamps is not trivial. A Network Time Protocol (NTP) server/client setup should be able to maintain clock accuracy among machines within tens of milliseconds, a resolution which should be adequate for most sensors [11]. Better synchronization is commonly reported, especially across local networks, but may not be reliably achievable without specialized hardware or software [12].

Especially when studying relatively rare events such as mutual gaze behaviour, it is extremely important to properly record the timing of the raw data in order to analyse the timing of events. Human participants do not look at one another very long, with most mutual gaze fixations lasting only 1-3 seconds [13]. Given the short duration of the instances of the behaviour being studied, correct association of data from multiple tracking sources is crucial before analysis can begin.

D. Issues related to post-hoc analysis

Analysis of the drum-mate data showed that human perception of time and rhythm may impact how they perceive turn-taking (as opposed to how it is modelled by the system). The robot's algorithm defined a play session as a turn if after a fixed time it did not detect any drum beats by the human. But sometimes the human participants played very slow rhythms, which were detected by the robot as several short turns but in fact were perceived as a single very long turn by the human playing. Turns where the human didn't play or the robot could not detect the human's drumming were called zero-turns. When the turn-taking interactions were analysed, therefore, it was vital to eliminate the zero-turns caused by this difference in perception and match the rest of the turns using heuristics to take into account both the robot's and human's point of view (For a detailed discussion of possible heuristics please refer to [2]).

E. Issues related to adaptation

Adaptive behaviour is a very important part of our interactive studies, and in the case of producing or detecting such behaviour, timing is crucial. It is important to compare the real-time waving-motion data/ drumming performance/gaze direction of the robot and the human (or human-human pair) to get feedback which will be used in the adaptation. If the data from both participants can not be synchronized correctly, this feedback may be of limited use.

Additionally, creating adaptive behaviour often relies upon designing or learning a computational model of the desired behaviour. This model may be, as in the case of the drumming studies, quite simple, but the realism and interpretability of the behaviour produced is likely to be highly dependent upon its internal representation. The rudimentary representations of timing used in these models were sufficient to establish turn-taking (in part because of humans' natural adaptability). But analysis of the resulting interactions reveals that the humans were sensitive to temporal aspects of the interaction that were not modelled by the systems and affected their enjoyment.

Because of the importance of timing in the types of interaction we've explored, we believe models of these behaviours should explicitly represent temporal relationships in order to fully capture their fundamental characteristics. What temporal characteristics (e.g, duration, periodicity, tempo) should be represented and how best to represent them (as values, distributions, or functions) is highly dependent on the nature of the interaction, and may have to be determined via trial-and-error or by examining data from humans performing the behaviour in question.

3. Conclusion

We presented the experimental design and the timing related issues resulting from four interaction scenarios with child-sized humanoid robots and/or human participants. Timing plays an important role in human-humanoid interaction, appearing in several different ways in the study of turn-taking, synchronization, real-time interaction, and adaptation. Over the course of this research, the critical importance of accurately capturing the time relationships in the raw data was recognized, and methods to do so were proposed. Another common conclusion from these studies is that humans are highly adaptable and often capable of adapting the timing of their actions to that of the systems they interact with. This may allow for satisfying repetitive or rhythmic interactions with robots that use relatively simple models to generate the timing of their own behavior. However, these studies also suggest that humans may base their satisfaction with such interactions on time-based criteria that encompass the entirety of the interaction (such as a notion of "fairness" in turn-taking). This suggests that robots may improve their performance at these tasks by incorporating temporal information over longer time scales into their internal models.

Selected References

[1] Blow, M.P, K. Dautenhahn, A. Appleby, C. Nehaniv, D. Lee, "Perception of robot smiles and dimensions for human-robot interaction design*", In Proc. IEEE RO-MAN06,* pp.469-474, 2006.

[2] Kose-Bagci, H., K. Dautenhahn, D. S. Syrdal, and C. L. Nehaniv, "Drum-mate: Interaction dynamics and gestures in human-humanoid drumming experiments", Connection Science, 2009, accepted.

[3] Kose-Bagci, H., K. Dautenhahn, C. L. Nehaniv, Emergent dynamics of turn-taking interaction in drumming games with a humanoid robot, in: *Proc. IEEE Int. Symp. on Robot and Human Interactive Communication*, Munich, pp. 346–353, 2008.

[4] Broz, F., H. Kose-Bagci, C. L. Nehaniv, K. Dautenhahn, "Learning behavior for a social interaction game with a childlike humanoid robot", Social Learning in Interactive Scenarios Workshop, Humanoids 2009, Paris, France, 7 December, 2009.

[5] Shen, Q., J. Saunders, H. Kose-Bagci, K. Dautenhahn, "Acting and Interacting like me? A method for identifying similarity and synchronous behaviour between a human and robot", Proceedings IEEE IROS workshop on "From motor to interaction learning in robots", Nice, France. 2008.

[6] Koizumi, S., T. Kanda, M. Shiomi, H. Ishiguro and N. Hagita, "Preliminary Field Trial for Teleoperated Communication Robots"**,** *IEEE International Symposium on Robots and Human Interactive Communications (RO-MAN2006)*, Sep. 6-8, Hatfield, United Kingdom, pp. 145-150, 2006.

[7] Striano, T., Henning, A., & Stahl, D. (2006). Sensitivity to interpersonal timing at 3 and 6 months of age. *Interaction Studies*, 7(2), 251-271.

[8]Crutchfield , J.P., "Information and its Metric," in *Nonlinear Structures in Physical Systems – Pattern Formation, Chaos and Waves*, Springer Verlag, pp 119-130, 1990.

[9] Shen, Q., J. Saunders, H. Kose-Bagci, K. Dautenhahn, "An Experimental Investigation of Interference Effects in Human-Humanoid Interaction Games", IEEE RO-MAN2009, 2009, *accepted*

[10] Robins, B., K. Dautenhahn, R. te Boekhorst, C. L. Nehaniv, "Behaviour Delay and Robot Expressiveness in Child-Robot Interactions: A User Study on Interaction Kinesics." In *Proc. ACM/IEEE 3rd International Conference on Human-Robot Interaction (HRI 2008),* 2008.

[11]Mills, D. L. On the Accuracy and Stability of Clocks Synchronized by the Network Time Protocol in the Internet System. *SIGCOMM Comput. Commun. Rev.* 20, 1 (Dec. 1989), 65-75.

[12]Pásztor, A. and D. Veitch, PC based precision timing without GPS. In *Proceedings of the 2002 ACM SIGMETRICS* June 15 - 19, Marina Del Rey, California, 2002.

[13] Kleinke, Chris L., "Gaze and eye contact: A research review." *Psychological Bulletin*, Vol. 100, No. 1. (1986), pp. 78-100.