USING SOFTWARE AGENTS TO INVESTIGATE THE INTERACTIVE ORIGINS OF COMMUNICATION SYSTEMS

PIETER DE BIE*, THOMAS SCOTT-PHILLIPS†, SIMON KIRBY†, BART VERHEIJ*

* *Artificial Intelligence, University of Groningen*

† *Language Evolution and Computation Research Unit, University of Edinburgh*

pdebie@ai.rug.nl

In contemporary research on the origins of human communication, games are used as a methodological tool (Galantucci, 2005; De Ruiter, Noordzij, Newman-Norlund, Hagoort, & Toni, 2007; Scott-Phillips, Kirby, & Ritchie, 2009). It has turned out that the ample degree of freedom allowed by the proposed games can be an impediment for the fruitful drawing of conclusions from the empirical findings: the creativity and flexibility of experimental participants leads to behavior that is hard to predict up front, allowing only post-hoc analysis (cf. Galantucci (2005)).

In the research reported here, an attempt is made to improve this situation by designing a game that can be played by a human player against a software agent. We hypothesize that this will provide an experimental setting that is sufficiently constrained for the design of experiments on the emergence of a communication system in which the test person's behavior is quantitatively measurable.

To test our hypothesis, a game has been defined that is sufficiently complex to allow for interesting communicative interaction to arise, but that is at the same time sufficiently simple to allow the design of software agents playing different strategies in the game. By the use of software agents, there is control over one of the players in the game, effectively reducing the range of expected behaviour of the second (human) player. The Embodied Communication Game (ECG) by Scott-Phillips et al. (2009) served as starting point for the design of our game. Current results are the design, analysis and implementation of a game and software agents with different strategies. The implemented strategies were derived from theoretical considerations and from the observation of human players playing the game.

In our game, there are a Sender and a Receiver (cf. De Ruiter et al. (2007)). The Sender controls a stick man situated in a box with four squares of one of four colours (red, yellow, blue, green). A colour might occur several times, or not at all. The stick man can travel from square to square. The Sender's goal is to communicate to the Receiver the colour of the square at the end of his turn. The Receiver watches a replay of the moves of the Sender, with all timing infor-

mation removed, and with the squares grayed out. After watching the replay the Receiver has to decide on which of the four colours the Sender ended his turn. If the Receiver chooses the correct colour, both players gain a point. The goal for the players is to score the highest amount of points in succession.

Preliminary experiments have been performed to test whether the game indeed provides an experimental platform allowing the effective measuring of the behaviour of human players. We performed a 24 person experiment and tested two hypotheses. First, a less efficient signal is expected to be easier to recognize as a meaningful signal. Second, a highly repetitive signal is expected to be easier to detect than a signal without repetitions. We predicted higher scores for players paired with an inefficient, highly repetitive agent than for players paired with a highly efficient, not repetitive agent.

Inefficiency was measured as the number of moves used to travel from to the ending square minus the minimal number of moves required. Repetitiveness was measured by creating an algorithm that can detect oscillations, loops, corner oscillations and U-shapes in the moves of the Sender. The hypotheses were tested by creating four agents which all took on the signalling role. The agents used strategies differing in inefficiency and repetitiveness by using different signalling methods. One agent moved as efficient as possible, the others used combinations of oscillations and circles. Contrary to expectations, we found no significant differences in the scores between different agents. It seems that the agents' strategies were too difficult to understand; only one participant was able to score higher than what would be expected by chance.

We expected that replacing one participant with an agent and using a predefined signalling system would make the task easier. In fact, comparison with results from the ECG indicates that this made the task harder. This suggests that interaction is critical in constructing a shared signalling system. One possible reason is that the embodied behaviours of the signaller only make sense once the receiver has actually played that embodied role themselves. If so, researchers looking at these games may need to explore the space of designs more widely to find out what aspects of their games are crucial and what are not. Future work may need smarter agents that play both signalling and receiving roles.

References

De Ruiter, J., Noordzij, M., Newman-Norlund, S., Hagoort, P., & Toni, I. (2007). On the origin of intentions. In P. Haggard, Y. Rossetty, & M. Kawato (Eds.), (p. 593-610). Oxford University Press.

Galantucci, B. (2005). An experimental study of the emergence of human communication systems. *Cognitive Science: A Multidisciplinary Journal*, *29*(5), 737 - 767.

Scott-Phillips, T. C., Kirby, S., & Ritchie, G. R. S. (2009). Signalling signalhood and the emergence of communication. *Cognition*, *113*, 226-233.