Goal-directed imitation through repeated trial-and-error interactions between agents

Bart Jansen[1] Tony Belpaeme[2]

[1]Vrije Universiteit Brussel – Artificial Intelligence Laboratory
Pleinlaan 2, 1050 Brussels, Belgium

[2]University of Plymouth
A318 Portland Square, PL4 8AA Plymouth, United Kingdom

bartj@arti.vub.ac.be, tony.belpaeme@plymouth.ac.uk

1 Introduction

Imitation or learning by demonstration is one of the holy grails in robotics. Wouldn't it be great to show a service robot how to do the laundry or to show an edutainment robot how to clear up the kids' toys at the end of the day? However, the community is still at a stage where we can only dream from such applications. Most if not all researchers have been concentrating on sub problems of imitation, such as imitation using logic programming (Kuniyoshi et al. 1994), imitation between agents with different embodiments (Alissandrakis et al. 2004) or the metrics for deciding what to imitate (Billard, this volume); the work presented in this paper is not any different. In the last years insights from cognitive science and developmental psychology have provided new impetus in artificial imitation research (see among others Dautenhahn and Nehaniv (2002)). It is this work that inspired us to build a new model for agents to imitate the intention of observed behaviour.

In previous work we have demonstrated how a group of agents can agree on imitating gestures (a class of imitation which Byrne and Russon (1998) have dubbed *action-level imitation*) and how we avoided to need for an explicit measure of imitation quality. This was done by letting the initiator judge if the imitator's behaviour corresponded to the demonstrated action (Jansen 2003; Jansen et al. 2004; Belpaeme et al. 2005). In this paper we wish to present some of our ongoing research in *goal-level imitation*, where agents learn to imitate the goal of an action. This is a natural continuation of action-level imitation, and has been recognised as a notoriously difficult problem as the imitator needs to solve two difficult problems. First, it needs to understand the goal of the observed behaviour, something which is immediately obvious to us, but is hard to grasp if you not do have the necessary world knowledge to deduce the goal. If you see someone sweeping the patio with a broom, is the goal then the sweeping motion itself? Or the swishing noise made by the broom? Or the fact that the

patio is cleaner after the sweeping? Of course, already possessing the necessary world knowledge, we know what sweeping is intended to achieve, but how is one to know without this knowledge, and how then can we acquire that knowledge? A second problem with goal-level imitation is that there are many ways to attain the same goal. A clean patio can be achieved by holding the broom stick and sweep, or one could hold the broom itself and start sweeping (being careful not to poke the eye of passers by), or one could even sweep the patio with his bare hands. All would have the same end state, but it would not be reached with the same efficiency. So clearly, only imitating the goal is not correct, the steps towards reaching the goal are as important.

2 The model

We present a simple model to study how the intention, or goal, of an action can be deduced from repeated interactions between agents. This model could in principle be extended to human-robot imitation, but for the time being results only exist for simulated agents.

Figure 1 shows an example of a board with a number of geometric figures. The figures can be shifted in four directions, these are the *actions* that an agent can perform. Each agent has its own private board, instead of one common board for all agents, on which to execute actions. The rationale behind this is that an imitating agent will not be able to simply copy the behaviour of the initiating agent, as the blocks in each agents' board will be in different initial positions. Agents are forced to come up with their own solution for reaching a particular goal.

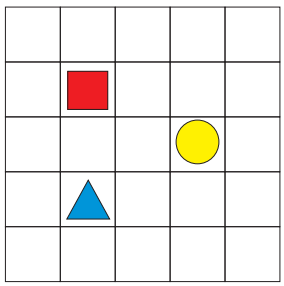

Figure 1: The board on which the agents perform action by shifting blocks.

The goals are defined as a conjunction of two primitives `left-of?` and `above?`. As an example, a possible goal for the board in figure 1 could be `(AND (left-of? block circle) (above? circle triangle))`.

2.1 The imitation game

The interaction between agents is set up around an *imitation game*. During such a game, one agent acts as the initiator and the other as the imitator. The initiator randomly selects a goal from the set of goals it already possesses, and builds a plan for reaching that goal. It then perform the actions in the plan. The

imitator observes this sequence of actions, and selects from its own repertoire of goals the goal that best matches the perceived actions. The imitator now builds a plan to reach this goal on its own board, and executes that plan. The initiator then observes the end state of the imitator's goal, and checks whether the intended goal was attained. The initiator lets the imitator know if the goals match, after which the game ends[1].

2.2 The repertoire of goals

The *repertoire of goals* is crucial to our approach. Each agent has such a repertoire of goals, and initially it is empty. During the playing of imitation games, the agents add goals to their repertoires. Each goal has associated with it a *use* and *success* counter, the use counter reporting how often the goal has been shown of imitation, and the success counter reporting how often imitation succeeded.

To recognize the goal in the behaviour of an agent, one needs to interpret the actions of the others as intentional behaviour. Just as humans use external, or rather *extra* information to do this, so do the agents. They assume that (1) an agent only uses relevant actions to reach a goal, so no extra or confusing actions are used and (2) all agents' cognitive abilities are equal.

Additionally, and this is quite crucial to the idea we are presenting here, the agents all interpret the observed behaviour as trying to reach a goal they have already encountered (and which is therefore already present in the repertoire of goals). This has two important implications. One is that it reduces the complexity of interpreting observed behaviour. Observed behaviour can be interpreted in hundreds or even thousands of ways. Limiting the interpretation of behaviour to the goals the agent is familiar with, radically constrains the interpretation process and thus reduces the complexity of goal-level imitation. The second implication is that through the mechanisms of the imitation game the agents have a continuously changing repertoire of goals: as use and success counters increase or decrease, new goals are added and unsuccessful goals are sporadically deleted. The repertoires of goals are dynamically updated through the imitative interaction between the agents.

3 Results

Figure 2 shows the results of ten runs of two agents taking turns in imitating each other. The graph reports the average imitative success and the average number of goals. The imitative success rises and levels at about 70% (base line success is 50%). This shows that the agents acquire a repertoire of goals with which they can successfully imitate each other's intention, instead of each other's actions. The rather low imitative success is mainly due to the difficulty of the task, which still lacks some constraints to have a higher success rate. However, if one takes into account that the agents have equal status —there is no teacher and no learner, both need to agree on goals just through imitation— the 70% success rate is remarkably high. Experiments are underway to see how the model behaves when the agents do not have equal status, but when instead one agent acts as a teacher, starting with a full blown repertoire of goals, and the

[1] For pseudocode and more details on the algorithms the reader is referred to (Jansen 2005).

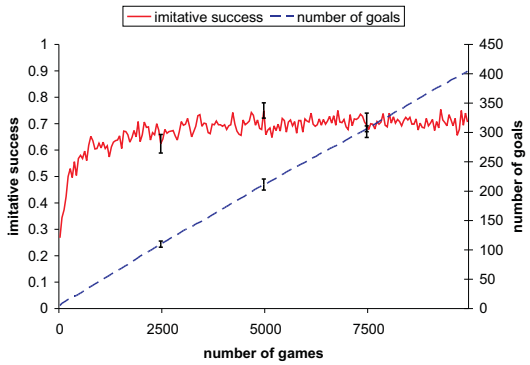

Figure 2: The evolution of the success and number of goals during repeated interactions between two agents.

other as learner, which has to acquire intentional behaviour through imitative learning.

Acknowledgments

Bart Jansen is sponsored by a grant from the Institute for the Promotion of Innovation by Science and Technology in Flanders (IWT). This research was in part sponsored by the EC FET-project Cogniron FP6-002020.

References

Alissandrakis, A., Nehaniv, C. L., and Dautenhahn, K. (2004). Towards robot cultures? Learning to imitate in a robotic arm test-bed with dissimilar embodied agents. *Interaction studies: Social behaviour and communication in biological and artificial systems*, 5:3–44.

Belpaeme, T., de Boer, B., and Jansen, B. (2005). The role of population dynamics in imitation. In Dautenhahn, K. and Nehaniv, C. L., editors, *Imitation and Social Learning in Robots, Humans and Animals: Behavioural, Social and Communicative Dimensions*. Cambridge University Press, Cambridge, UK. In press.

Dautenhahn, K. and Nehaniv, C. L., editors (2002). *Imitation in Animals and Artifacts*. The MIT Press, Cambridge, MA.

Jansen, B. (2003). An imitation game for emerging action categories. In Banzhaf, W., Christaller, T., Dittrich, P., Kim, J. T., and Ziegler, J., editors, *Advances in Artificial Life, 7th European Conference on Artificial Life (ECAL 2003), Dortmund, Germany*, volume 2801 of *Lecture Notes in Computer Science*, pages 800–809. Springer.

Jansen, B. (2005). Goal level imitation with emergent repertoires. In *Proceedings of the AISB Symposium 2005*.

Jansen, B., ten Thij, T., Belpaeme, T., De Vylder, B., and de Boer, B. (2004). Imitation in embodied agents results in self-organization of behavior. In Schaal, S., Ijspeert, A. J., Billard, A., and Vijayakumar, S., editors, *From Animals to Animats: Proceedings of the 8th Conference on the Simulation of Adaptive Behavior*, Cambridge, MA. The MIT Press.

Kuniyoshi, Y., Inaba, M., and Inoue, H. (1994). Learning by watching: Extracting reusable task knowledge from visual observation of human performance. *IEEE Transactions on Robotics and Automation*, 10(6):799–822.