Learning to Communicate through Imitation in Autonomous robots*

Aude Billard and Gillian Hayes

University of Edinburgh, Dept. of Artificial Intelligence
5 Forrest Hill, EH1 2QL Edinburgh
E-mail:{audeb, gmh}@dai.ed.ac.uk

Abstract. Communication is a desirable skill for robots. We describe a method of how these skills could be learned. A control architecture of connectionist model combining *life-long* learning and predefined behaviours is developed and implemented in a physical system of two autonomous robots. A teaching scenario based on movement imitation is used to teach a basic non grammatical language. Teaching is done from a teacher robot to a student robot. While following the teacher, the student robot learns a basic vocabulary concerning its movements and location.

1 Teaching and learning a language

Communication is a social skill and as such may be desirable for autonomous agents that are expected to interact with other agents or humans. In particular, communication seems a desirable skill for learning agents to possess as it would provide us with an easy means for teaching or directing them. Moreover, a flexible language which is acquired through a learning process would be advantageous over a fixed communication protocol as it would make it more adaptable to the tasks the robots can carry out and the types of robots engaged in those tasks. This works investigates how a common communication protocol could be transmitted from a teacher agent to a learner agent, in terms of the agents' behavioural and cognitive capabilities.

By opposition to previous simulation studies on language acquisition ([1], [6], [8]), we take a physical approach to the problem of the language understanding. We focus on the learning and understanding of a language at the level of the individual rather than at the level of the society and propose a method, in terms of the agent's behavioural capacities for transmitting the language from a teacher agent to a student agent. Movement imitation is the means by which the language is taught. A student robot follows and thus imitates the teacher's movements. The student attaches a meaning to the teacher's words in terms of its sensor measurements and motor states. In this context, the 'language' consists of a simple vocabulary, where the 'words' are (radio) signals that label specific configurations of sensor-motor state. However, it is our belief that the method

* In the proceedings of ICANN97, 7th International Conference on Artificial Neural Networks, Springer-Verlag editors, Lausanne-CH, Oct 97.

proposed could enable learning of more complex forms of communication, for which sharing a common context is a basic requirement [3].

Other works on movement imitation for teaching movement related tasks are ([5], [4],[7]). Movement imitation, as a means to learn a language, is interesting compared to a conventional supervised learning method, as used e.g. in [9], for at least two reason. Firstly, it enables the learning agent to share a perception of the world similar to the teacher's. Secondly, it enables the teacher to direct the student's learning.

2 Controller Architecture

A controller architecture is developed to direct the behaviour and the learning of the student agent. The general framework of the DRAMA (Dynamical-Recurrent-Associative Memory Architecture) architecture, described previously in [2], was used to enable the learning. The principal properties of this learning architecture are: 1) It combines two connectionist models types, associative memory (Willshaw type) and recurrent networks, in one single network structure for achieving learning of temporally coincident patterns. 2) The network is fully connected with self-recurrent connections on each node. 3) Each connection in the network is associated with a weight and a confidence factor. The weights represent the temporal correlation between events while the confidence factors keep a memory of the frequency of a pattern occurrence. 4) Weights and confidence factors are updated following Hebbian rules, providing an associative type of learning. The closer the temporal occurrence of the events, the stronger the correlation. 5) The structure of the network is dynamically updated, at each processing cycle (life-long learning). 6) Encoding of inputs and outputs is binary. 7) The self-recurrent connections on the units provide a short-term memory of one event per unit. 8) In addition, the recurrent connections between different units enable learning of sequential and simultaneous correlations. 9) Data retrieving depends on the value of the weights and confidence factors associated with the connections. The weights play the role of a first threshold on the time of occurrence of the event, only the event 'older' than the memorised temporal correlation are considered. The unit output is activated when the sum of confidence factors of the different active inputs passes a minimal threshold.

Figure 1 presents a schematic description of the learning architecture. The learning process used in the experiments is described in the following.

Learning Process Learning is *life-long.* At each processing cycle, all the robot's sensor-motor information is compared to the measurements collected at the previous cycle. Whenever a variation[2] in the measurement is detected (*event recognition*), the new event is presented to the associative architecture (DRAMA) to be correlated with all simultaneous and previously recorded events in the same

[2] Sensor data are collected as binary bytes (hardware definition). A variation is an inversion of one bit.

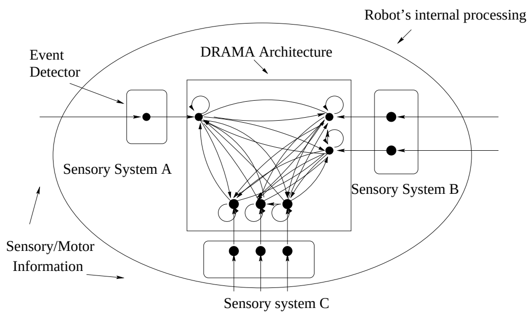

Fig. 1. Learning process: Learning is dynamical. At each processing cycle, all the robot's sensor-motor information is compared to the measurements collected the previous cycle. Whenever a variation in the measurement is detected (*event recognition*), the new event is presented to the associative architecture (DRAMA) to be correlated with all simultaneous and previously recorded events in the same or other sensory system. Event are memorised for fixed period in the self-recurrent connections of the input units. The sensory systems A, B and C in the figure could be interpreted e.g. as the motors, compass and radio systems.

or other sensory system. Learning occurs only when an *event*, i.e. a variation in sensor measurement, has taken place. Events are memorised in the self-recurrent connections of the network input units for a fixed period of time (200 cycles about 5 minutes)[3]. In figure 1, we present a schematic view of the learning architecture. The sensory systems A, B and C in the figure could be interpreted e.g. as the motors, compass and radio systems respectively. The learning process works then as follow: If, at time t, one event is detected by the sensory system A and one by the sensory system B, they will be associated with each other. If then, at time t+n (n $<$ memory length), a new event is detected by the system C, it will be associated with the two previously recorded events in systems A and B. Association consist of 1) increase the confidence factors inversely proportionally to the temporal delay between corresponding to active input bits, following an Hebbian rules 2) update the weights to take in account the temporal delay between the events.

Experimental implementation The control architecture combines *life-long* learning with predefined behaviours. Learning consists of multiple associations

[3] This was preferred to an infinite event memorisation, where e.g. the memory of the event is erased only when a new event is detected, in order to avoid misleading associations. The experimental set-up was done so that the teaching of a word, i.e. sending of a radio signal, occurred at the latest two or three minutes following the event which the word should be associated with.

between all the sensor and motor systems of the robot[4]. Learning of the language more specifically is to associate different radio signals (bit-strings) with distinct sensor measurements and motor states. Association is done between all the sensor (light, infra-red, bumpers, compass and radio) and motor systems of the robot. Learning is achieved once the confidence factors and weights associated with the relevant sensor measurements show a strong correlation, easily distinguishable from the ones with other non relevant sensor measurements. In the experiments we report here, teachings concern the motors state (robot's movements) and the compass measurements (robot's orientation). In [3], we report experiments where other sensory systems were involved, such as light and inclination measurements used for external object recognition.

Three basic behaviours of *wandering, avoiding obstacles* and *following* are predetermined by setting specific sensor-motor connections of the network to appropriate values[5]. The robot's behaviour is directed, at each time step, by retrieving the motor inputs from the network, given the particular sensor measurements.

3 The Experiments

Two autonomous LEGO robots are used for the experiments. The robots are provided with two bumpers and one infra-red sensor for obstacle avoidance and two light detectors for mutual detection, as each robot carries a bright light. In addition, they have a compass that measures bearings (in 45 degrees interval) and a radio transceiver, the means of language transmission. Experiments are carried out inside a 'Dodgem' cage in which, similarly to the system used in the 'Dodgem' game, the robots are continuously recharged through a long stick touching the ground at one end and the roof of the 'Dodgem' cage at the other to complete an electrical circuit. In figure 3 left, we show a picture of the two robots in the Dodgem environment.

Figure 3 right presents a schema of the teaching scenario. The student robot follows the teacher robot that wanders randomly in the arena. Whenever the teacher detects the student behind it, it stops and sends it a radio signal combination of the signal for *stop* and of a signal for the *orientation* given by its compass. Then it starts to move again and sends consecutively a signal for *move*. The total vocabulary consists of 6 words, defined as 6 different radio signals, for [*stop, move, South, North, East, West*].

A set of five learning experiments was carried out. Each experiment lasted for 30min to 1 hour. A run was stopped when the learning had stabilised. In

[4] It is important to note that the associations are not only made between sensor and motor systems but also between the sensory systems themselves, e.g. the radio/compass and radio/light associations. The latter associations provide the agent with a knowledge about its environment functioning, which it may use later to predict new event and to react to it in advance

[5] The control architecture we describe was used only for the student robot. A simple behaviour-based controller was used to direct the teacher robot's behaviour.

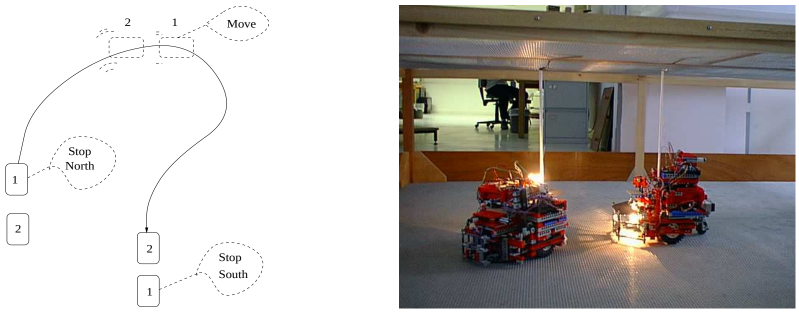

Fig. 2. Right: **The Teaching Scenario.** A teacher robot (1) teaches a student (2) robot a vocabulary to describe its movements and location. The student, while following the teacher, associates the teacher's 'words', radio signals, with its observation. Left: A photo of the two robots, teacher and student, in their Dodgem environment.

5 experiments out of 6, the 6 words had been learned at the end of the run. In the 6th run, only 5 words were learned. 75 teaching episodes were needed on average before the 6 words were learned. A teaching episode consists of a 3-signal transmission (1 for *stop*, 1 for *move* and 1 for a *compass bearing*). The learning has become stable once the proportion of correct to incorrect teaching examples was greater than 7/3 on average. An incorrect teaching example consists of a situation where the taught word does not correspond to the student's observation. This happens when e.g. teacher and student are not well aligned, thus not sensing the same orientation. The words *stop* and *move* are learned firstly after about 5 to 10 teaching episodes. The words for the compass measurements need more time since they are taught a quarter times as often as the other two and because they depend strongly on the positioning of the two robots at the time of the teaching, which is the result of their random wandering.

In summary, this experiment shows that 1) a number of words are learned and correctly associated. 2) The vocabulary was learned although the experimental conditions create incorrect teachings. 3) The learning is stable once a significant number of relevant correlations have been recorded. One may now question whether these experiments, because of their apparent simplicity (the vocabulary has only six words), are a demonstration of the efficacy of the learning method. Although the sensory system and consequently the vocabulary are quite simple, the learning task is actually rather difficult. Because of their very limited sensor capabilities, the robots mutual recognition and following is very imprecise. Alignment of the two robots is seldom perfect. Consequently, the chances of their sensor measurements being identical (in particular their compass measurements) are quite slim and misleading associations may easily occur. On the contrary, if we had used more sophisticated sensors, giving a finer and thus more informative description of the environment, the number of taught words could have been greatly increased. On the other hand, however, we would have simplified the task of imitation by giving the robot more information on which to rely for adjusting their positions one towards the other. This leads us to think that the levels of difficulty of the learning task in the two cases are comparable. Significant data have to be extracted from a considerable amount of noise that is produced, in the first case, by misleading teachings due to an imprecise imitation and, in the second case, by a more important source of information. Therefore, it seems reasonable to conclude that our experiments, although apparently simple, are relevant as a demonstration of the validity of the system.

4 Conclusion

A method for transmitting a communication protocol from a teacher agent to a student agent was investigated. An imitative strategy, based on movement following, was developed and implemented in a physical system of two robots. A teacher robot taught a student robot a vocabulary about movements and directions. Learning of the language was part of a global associative mechanism, the DRAMA architecture, combining all sensor-motor information. Experiments were successful and demonstrated that a movement imitation-based learning scenario is a possible strategy for teaching a basic non grammatical language.

5 Acknowledgement

An enormous thanks to the technicians for their essential support. Many thanks to John Demiris and Auke Jan Ijspeert for their useful comments on this paper. The research reported here was supported by a grant from the LEMO company and the Swiss Research National Fund.

References

1. M. Oliphant and J. Batali (1996), 'The emergence of coordinated communication', submitted to Cognitive Science.
2. A. Billard (1996), 'Allo Kazam, do you follow me? or Learning to speak through imitation for social robots, MSc thesis, DAI Technical Paper no. 43, Dept. of AI, University of Edinburgh.
3. A. Billard and K. Dautenhahn (1997), 'Grounding communication in social, embodied robots', Proceedings of TIMR conference, Manchester, September 97.
4. J. Demiris and G. Hayes, 'Imitative Learning Mechanisms in Robots and Humans', Proc. of the 5th European Workshop on Learning Robots, Bari, Italy, July 1996.
5. K. Dautenhahn, (1995), 'Getting to know each other – artificial social intelligence for autonomous robots', Robotics and Autonomous Systems, vol. 16, pages 333-356.
6. B. MacLennan, (1993) 'Synthetic ethology and the evolutive cooperative communication', In the journal Adaptive behaviour, 2:161-187.
7. Kuniyoshi Y., Inaba M. and Inoue H. (1994), 'Learning by Watching: Extracting Reusable Task Knowledge from Visual Observation of Human Performance', in IEEE Transactions on Robotics and Automation', Vol. 10, No 6, Dec. 1994.

8. L. Steels,(1996) 'The spontaneous self-organization of an adaptive language', S. Muggleton (ed.): Machine Intelligence 15, Oxford University Press.
9. H. Yanco and L. A. Stein (1993). "An Adaptive Communication Protocol for Co-operating Mobile Robots." In From Animals to Animats 2, MIT Press.