Development of a Context Dependent Robot Language: Preliminary Results

Holly Yanco *
Artificial Intelligence Laboratory
Massachusetts Institute of Technology
545 Technology Square, Room 741
Cambridge, MA 02139
(617)253-7884
holly@ai.mit.edu

Abstract

The meaning of a word in a context dependent language is determined using the context in which it is spoken. In this paper, we will describe how a group of autonomous agents can develop a context dependent language by using their sensor readings to provide context. Results from a computer simulation of the experiment are presented.

1 Introduction

Words can have many meanings in a context dependent language. For example, in the simple command "Look at this", the word "this" can have a variety of meanings. The object that the speaker is referencing determines the meaning of the word "this." In order for the listener to understand the meaning of the speaker's sentence, the listener must be in the same context as the speaker or must be able to move into that context. If the speaker said "Look at this" while pointing to an object hidden from the listener by a desk, the listener needs to ask for more information or needs to get up and move to where he can see the object being referenced.

In the current work, a collection of simulated robotic agents are being trained to develop a context dependent language. The agents could learn a command such as "Do", where the meaning of "Do" is directly mapped by the sensor readings to a set of appropriate actions to take. For example, agents could do "eat" where there is food, do "gather objects" where objects are present and do "avoid" when predators that are detected. In this case, "Do" would be mapped by the sensor readings {food present, object present, predator present, ...} to the actions {eat food, gather object, avoid predator, ...}. Another command could be "Move away", where "Move away" is mapped by the sensor readings {light, heat, ...} to the actions {go to dark, go to cold, ...}.

The advantage of a context dependent language is the need for fewer signals to represent concepts. (In this paper, *signal* is defined as a word in the context dependent language. *Action* is defined as the meaning of a signal in a given context.) In a robotic system where the number of possible communication signals is limited by the number of bits that can be sent over the radio boards, the number of concepts that can be represented increases as more sensor data is used as a disambiguating context. The number of concepts that can be represented is given by

$$signals \times (\sum_{i=1}^{n} sensor_i \times values_i)$$

where *signals* is the number of signals that the robots can send, $sensor_i$ is one of the sensors being used to determne context, and $values_i$ is the number of values that $sensor_i$ can have.

2 Simulation

The experiments performed in this paper are based upon a language learning experiment described in [Yanco and Stein, 1993]. Two agents learn to communicate with each other in the presence of a trainer. A signal given by the trainer, which we call the *human signal*, is heard by one of the agents who is acting as the *leader*. The leader needs to learn the appropriate action to perform and the appropriate signal to send to the other robots, known as the *followers*. The followers need to learn the appropriate action to perform according to the leader's signal. The instructor provides *task based reinforcement*; all robots must perform the appropriate action in order for the group to receive positive reinforcement. The reinforcement learning algorithm we are using is Kaelbling's Interval Estimation [Kaelbling, 1990].

In the context dependent extension to the experiment, the agents take into account a sensor-reading vector when considering the meaning of a signal (either from the trainer or the leader). Once again, the agents need to map signals (or words) to actions (or meanings for the words). To motivate the agents to create multiple context dependent mappings from signals to actions, the language size is restricted; i.e. there are not enough

*This report describes research done at the Artificial Intelligence Laboratory of the Massachusetts Institute of Technology. Support for the laboratory's artificial intelligence research is provided in part by the Advanced Research Projects Agency of the Department of Defense under Office of Naval Research contract N00014-85-K-0124. The author is supported in part by a scholarship from Digital Equipment Corporation. Support for this research was also provided by the Gordon S. Brown Fund of the Department of Electrical Engineering and Computer Science at MIT.

Number of			Number of Iterations to Convergence		
Signals	Sensor Values	Actions	Average	Minimum	Maximum
2	2	4	219.34	68	555
2	3	6	871.84	297	2676
2	4	8	2287.57	883	5815
2	5	10	5165.33	1887	12600
2	10	20	51860.30	21853	105632

Table 1: Data from experiments with two possible signals from the human instructor to the leader and two possible signals for the robots to use in their language. Each of the languages developed will be optimal; i.e. the fewest signals are used to encode the maximum amount of data. Each experiment was run 100 times on a troupe size of two agents.

Number of	Num of Iterations to Convergence – Basic			Num of Iterations to Convergence – CD		
Actions	Average	Minimum	Maximum	Average	Minimum	Maximum
4	340.48	53	990	219.34	68	555
10	15011.61	2868	51031	5165.33	1887	12600
20	232267.82	44196	1241767	51860.30	21853	105632

Table 2: Comparison between learning times for basic language learning (where each signal maps to one action as described in [Yanco and Stein, 1993]) and learning times for context dependent language learning. Each experiment was run 100 times using two agents.

signals for the robots to create a one-to-one mapping between signals and actions. We have also performed experiments where the language size is larger than necessary; this is discussed below in the results section.

In these experiments, it is assumed that the agents are communicating in a mutually common context. The problem of learning how to move into context is currently being examined and is discussed in the future work section.

3 Results

We have performed various experiments using the simulation described above. In Table 1, learning times are given for the case where the agents have two signals that are mapped using an increasing number of sensor values. In this experiment, the agents are learning the *optimal* language, which we define to use the fewest number of words to encode the number of actions.

A comparison between learning times for a basic language experiment [Yanco and Stein, 1993] and for the context dependent language experiment is given in Table 2. Note that learning times for the context dependent language are significantly shorter than for the basic language (where each signal maps to one meaning).

When the agents have a number of signals for their language that is greater than the optimal size, it actually takes less time to converge than the optimal language size. An experiment was performed where the agents were given four signals to represent a language that only required two signals (with two sensor values, mapping to four actions). It took an average of 175.27 trials vs. the average of 219.34 for the case where the robots were given two signals and two sensor values for four actions. This occurs since the larger language size allows for the agents to randomly select signals with fewer collisions (defined as a selection made that can not exist with a selection previously made) than in the case where each signal needs to be mapped as many times at it can be as determined by the number of sensor values. When given four signals to create their language, the agents created a two signal language 16 times, a three signal language 70 times and a four signal language 14 times over 100 runs. In order for the robots to create an optimal language in the presence of additional signals, we would need to place a learning restriction that favors smaller languages.

When the sensor data is irrelevant, the robots will still converge upon the solution. However, it takes longer than the case where no sensor data is given; a two robot, two signal, two action experiment described in [Yanco and Stein, 1993] has an average convergence time of 15.24. In an experiment with two agents, two signals, two actions *and two sensor values*, the average convergence time is 91.26. Even though the sensor data is irrelevant, the robots must learn the proper action for each $(signal, sensor)$ pair; i.e. it must fill in the table for the reinforcement learning algorithm.

Another experiment that we have tested is the case where there are four sensor values representing two bits of sensor data. If only one bit is significant, the convergence time is still based upon the fact that there are four sensor values. (For two signals, four sensor values and four actions, the average convergence time is 556.49. Once again, we see that the agents are not able to ignore unnessary input.

4 Future Work

We have seen that the agents lack the ability to detect irrelevant attributes and are thus not able to ignore irrelevant information. In order for the learning to be more efficient, the agents will need to detect irrelevant sensor values and ignore them. We are currently looking into alternate learning algorithms that could help to provide a test for relevance.

Currently, the use of sensor values in the determination of meaning is built in. This has been done by incorporating the sensor values into the reinforcement learning table of $inputs \times outputs$. The sensor values are paired with the signals and used as the input. Later experiments will test the abilities of the robots to learn this capability.

The experiments have only been performed in simulation. The next step is moving them to the robots. The robot hardware is described in [Yanco and Stein, 1993]. Context is provided by actual sensor readings on the robots rather than from a simulated sensor vector. Different areas in the world have various characteristics that the robots are able to detect with their sensors. The robots have sensors that detect light levels, the presence of objects, heat levels and infrared signals. Due to the noise inherent in sensor readings, it is expected that learning times will be slightly larger on the actual robots.

Additionally, when the robots are in the world, the assumption that the communicating robots are in the same context will not be valid. The listener robot will need to figure out what context the speaker robot is in. The robots could solve the context problem by broadcasting context at the start of a message or by moving into the same context.

5 Conclusions

The work described in this paper is still in progress. We have seen that the development of a context dependent language requires a shorter learning time than the development of a basic language where each signal maps to one action. We have also described the need for the ability to ignore irrelevant sensor data to improve the learning times.

References

[Kaelbling, 1990] Leslie Pack Kaelbling. Learning in embedded systems. Technical Report TR–90–04, Teleos Research, Palo Alto, California, June 1990.

[Yanco and Stein, 1993] Holly Yanco and Lynn Andrea Stein. An adaptive communication protocol for cooperating mobile robots. In *From Animals to Animats: Proceedings of the Second International Conference on the Simulation of Adaptive Behavior*, pages 478–485. The MIT Press/Bradford Books, 1993.