

An empirical test of the IAC model: an extension of the Babybot experiment

Stephen St.Vincent and Benjamin Turner

03 March 2006

Abstract

In this paper, we present an empirical test of the Intelligent Adaptive Curiosity (IAC) model for adaptive robotics. The IAC model is a mechanism for self-motivation in autonomous developmental robots. The purpose of this experiment was to observe the staged development typical of the IAC model as observed in other experiments, as well as to push its limitations with regards to object permanence. By running a simple experiment in simulation using an implementation of an IAC brain in Py-robot, we were unable to find evidence of staged learning, but are unable to claim definitively that learning did not occur. We were also unable to demonstrate any notion of object permanence by the IAC brain.

1 Introduction

Developmental robotics is an emerging subfield of artificial intelligence research in which a robot's behavior is determined by intrinsic motivations and by the robot's interaction with its environment. This approach is intended to allow for more general, flexible robot control than traditional AI methods, which must be fine-tuned to each individual task. Robots attempt to learn the causal relationships between their actions and changes in the world around them. Several models for intrinsic motivation have been proposed (see [3, 5]). In this paper, we present the results of an experiment which uses the Intelligent Adaptive Curiosity model [4].

1.1 The IAC model

Intelligent Adaptive Curiosity attempts to direct the robot into situations in which it can maximize its learning progress. By taking its parameter space as its input, an IAC brain attempts to divide the solution space into a series of regions, each of which has an associated expert, to predict the next state of the world given the sensorimotor inputs at a given timestep. An action is then chosen from a finite set of candidates based on which candidate is expected to maximize learning progress. The overarching goal is to avoid situations which

are either too chaotic or too simplistic, gradually seeing challenging tasks as simple and more complex tasks as within reach. See [4] for a more detailed description of this method.

1.2 The Babybot experiment

Our experiment is based on a developmental experiment by Marshall et al. [3] using simple recurrent networks (SRNs) [2]. The setup there was straightforward: an untrained robot (the Babybot) is placed in the center of a world that is empty save for another robot. This additional robot (the target robot) moves in a circular pattern around the perimeter of the world. The Babybot’s movements were restricted to rotation only; its sensory input was pixel values from an on-board camera. Learning is accomplished through two SRNs: one chooses a motor action, and the other attempts to predict what the robot will see as a result of this action. The Babybot is rewarded for choosing motor actions which focus its view on situations in which there is prediction error in the center of the camera image.

The main goal of the Babybot experiment was to demonstrate emergent behavior which can result from this simple intrinsic motivation; namely, to have the Babybot track the movement of the target robot as it circled the world. Our experiment attempts to replicate this emergent behavior through the use of an IAC brain (rather than the neural network brains of the original) and without the focus on the error in the center of the image.

1.3 Layout of this paper

In Section 2, we describe the setup for our experiment, including the specifications of the simulated world, parameters specific to our implementation of the IAC brain, and how learning performance of our robot was measured. In Section 3, we present the results of our experiments. Finally, in Section 4, we discuss the meanings of these results.

2 Experimental setup

2.1 World description

The environment for our experiment, inspired by the Babybot environment, consists of a large square room (see Figure 1); at the center of the room is the developing robot (DR). Two target robots (T_{inner} and T_{outer}) travel around the room in circular patterns in the same direction at different radii from DR. Four wall segments are placed in the room between the orbits of T_{inner} and T_{outer} , such that DR’s view of T_{inner} will never be obscured, but DR’s view of T_{outer} occasionally will be. T_{inner} is colored blue, T_{outer} is colored green, and the wall segments are colored black (the same color as the bounding wall).

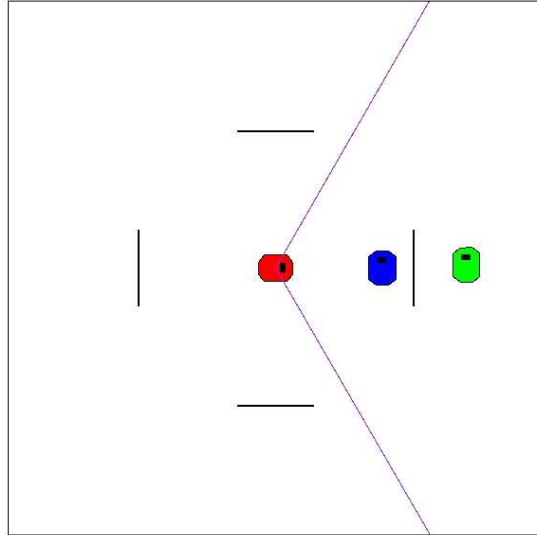


Figure 1: Experimental environment. The developing robot (red, center) sees what is between the two lines extending from its body. The two target robots (blue and green, to the right) trace out circles around the developing robot. The outermost target robot will be obscured from the developing robot’s vision by all four wall segments at one time or another.

2.2 IAC implementation

We used an implementation of IAC written in python as part of the Python Robotics toolkit (Pyro) [1]. Regions were split after 75 exemplars and retrained on every new exemplar. The experts were implemented as back-propagation networks with 31 inputs, 30 outputs, and 15 hidden units (these are all described in more detail in Section 2.3).

2.3 Parameter space

The inputs to the IAC brain were the motor action (a real-valued number between -1 and 1 indicating speed and direction of rotation) and the center row of the image returned from the on-board camera (a 30x1 array of pixels), which was transformed in the following manner: each possible color value (blue, green, red, or black) in the environment was assigned a discrete value between 0 and 1 (see Table 1). As such, the outputs were a 30x1 array of floating-point numbers between 0 and 1, corresponding to the color values in Table 1 and representing the prediction of what the camera will see on the next timestep.

We gave DR the ability to turn maximally as defined by Pyro specifications. The field of view was 120 degrees in width.

Color	R-value	G-value	B-value	Scaled representation
Red	255	0	0	0.3
Green	0	255	0	0.6
Blue	0	0	255	1.0
Black	255	255	255	0.0

Table 1: Pixel parameterization. This table shows the colors that appear in the experimental environment, their RGB values, and their floating-point value when scaled as input to the IAC brain.

2.4 Performance measures

There were four ways in which we measured the performance of DR:

1. Number of timesteps in which T_{inner} was anywhere in DR’s view
2. Number of timesteps in which T_{inner} was in the center of DR’s view
3. Number of timesteps in which T_{outer} was anywhere in DR’s view
4. Number of timesteps in which T_{outer} was in the center of DR’s view

The above representations were strictly binary, being either true or false. In order to look at these binary values meaningfully, we counted the number of times each measure was true over a 50-timestep interval, and plotted these values in histograms (see Figures 2-7).

3 Results

Below we present our results in both qualitative and quantitative forms. There is some uncertainty as to whether the two agree, but they nonetheless benefit from separate discussions.

3.1 Quantitative results

Here we present the results of just one run of this experiment, although it should be noted that all four runs produced similar results. The results presented here are representative of all of the runs.

DR showed no clear signs of staged development. Typically, stages are at least moderately apparent when looking qualitatively at the data. The only place where a stage might be seen is during the first 50(x50) time steps in Figure 5, which measures the DR’s ability to keep the outer robot in the center of its view. There appears to be more interest here. Figures 6 and 7 display results from a different run that more prominently displays this effect. Note the upward trend in DR’s focus on T_{inner} from timesteps zero to 20(x50) in Figure 6, which is accompanied by a lack of focus on T_{outer} during the same time interval

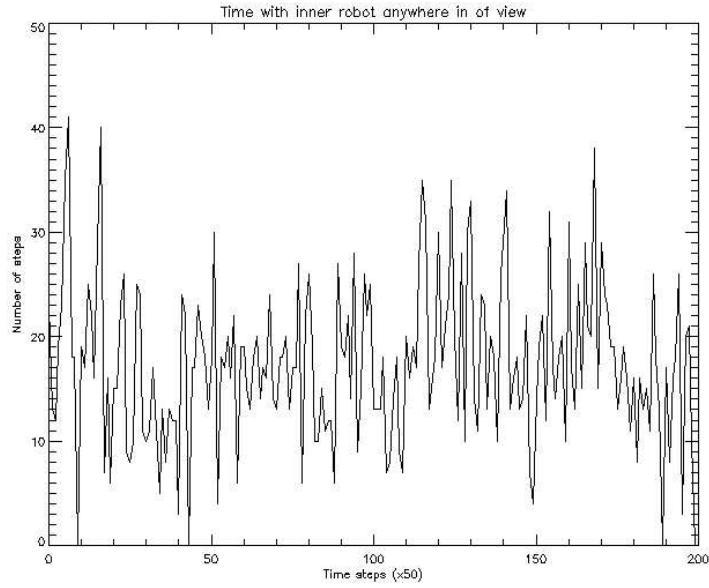


Figure 2: T_{inner} in view. This shows the number of timesteps per 50 in which T_{inner} was anywhere in the view of DR.

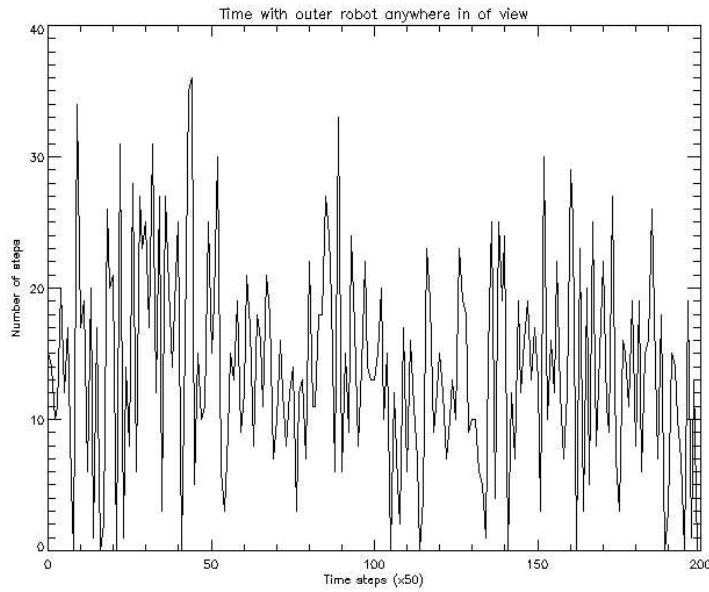


Figure 3: T_{outer} in view. Similar to Figure 2, but for T_{outer} being anywhere in DR's view instead of T_{inner} .

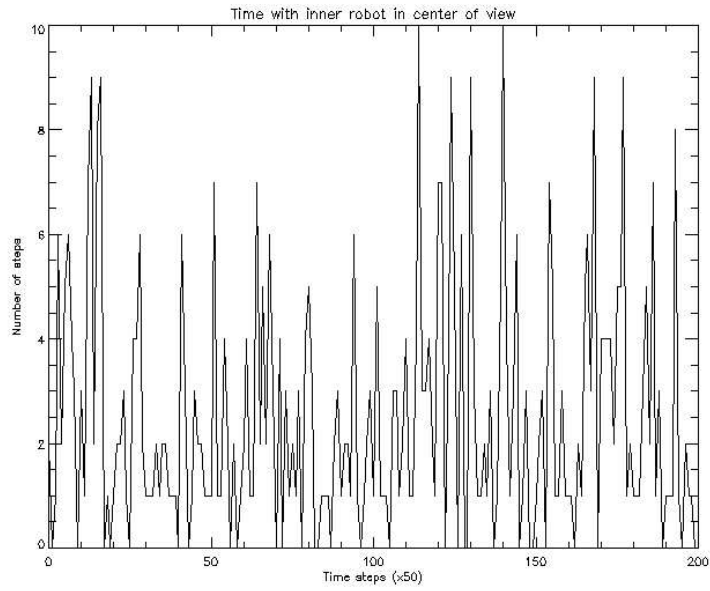


Figure 4: T_{inner} in center of view. Similar to Figure 2, but counting the timesteps in which T_{inner} was directly in the center of the DR's view.

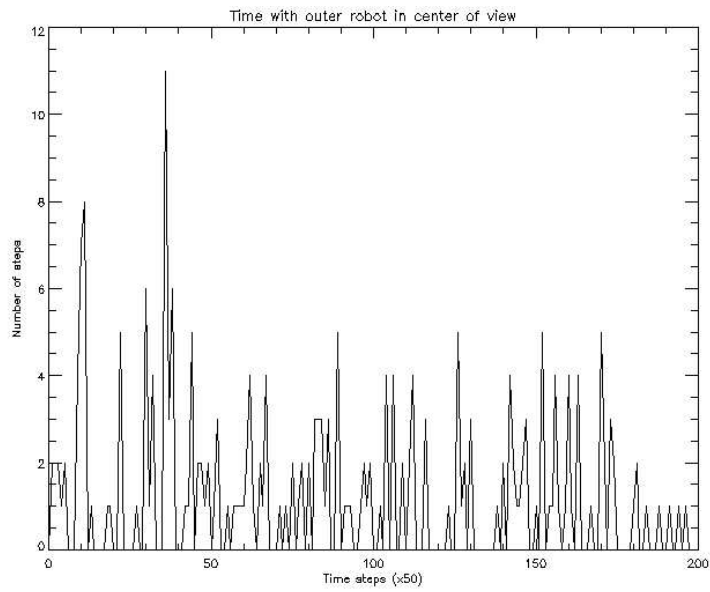


Figure 5: T_{outer} in center of view. Similar to Figure 4, but for tracking T_{outer} .

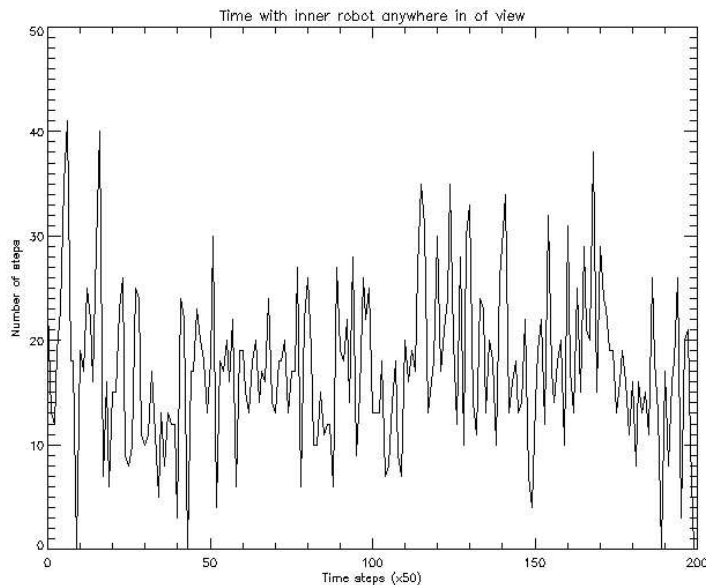


Figure 6: T_{inner} in view (Run 4).

in Figure 7. Similarly, between timesteps 20(x50) and 40(x50) in these same figures, DR appears to show focus on T_{outer} and a lack of focus on T_{inner} .

In order to determine if there was any longterm learning progress in tracking the target robots, linear regression fits were applied to the data in Figures 2 and 3. The results are presented in Table 2. Neither of the two fits had slopes that displayed statistically significant deviation from zero; as such, no learning progress can be inferred by these means.

Target robot	m	Δm	b	Δb
Inner	-0.011	0.026	17.4	1.5
Outer	0.0017	0.0298	15.2	1.7

Table 2: Linear regression fits for figures 2 and 3. Linear regression was of the form $y = mx + b$.

3.2 Qualitative results

Based on direct visual observation of the experimental runs, it did appear at times that when one of the target robots entered DR's view that DR tracked the target robot for a brief time. However, DR was never witnessed to track a target robot for a full orbital period.

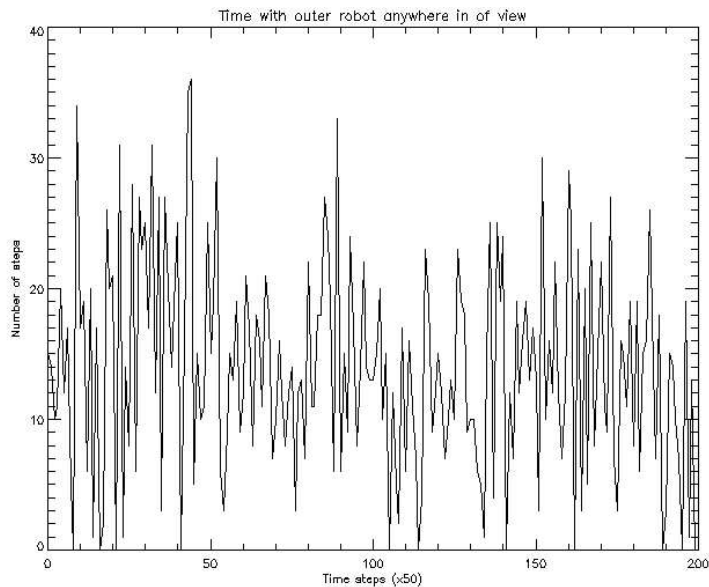


Figure 7: T_{outer} in view (Run 4).

4 Conclusions

Based on the quantitative results presented above, the IAC brain did not show any learning progress that deviated with statistical significance from random, except maybe in the earliest timesteps. Even when looking qualitatively, observations in favor of learning progress were uncertain. Multi-staged learning behavior was not observed, which was the principle goal of this experiment. Certainly, nothing can be inferred with regards to object permanence given the inconclusive nature of the other results.

One possible reason for the lack of learning is that tracking a robot was not significantly challenging to the robot, and as such was quickly learned and then ignored. In the Babybot experiment, the developing robot was rewarded for keeping the highest amount of error in the center of its view. Our experiment lacked this feature, which may have led to its overall failure in observing the desired outcomes. Another possibility is that the environment was too complex for DR to learn, but colleagues running similar experiments on the same IAC brain using the original Babybot world have seen results similar to ours.

Analysis of the errors associated with the experts in each region may reveal some learning progress. If this were to be the case, then varying the environment's complexity may lead to clearer results. In addition, the lack of predictive capability beyond the next time step may be necessary to observe knowledge of object permanence, which was the purpose of having T_{outer} 's orbit be behind wall segments.

References

- [1] D. Blank, L. Meeden, and D. Kumar. Python robotics: an environment for exploring robotics beyond legos. In *Proceedings of the 34th SIGCSE technical symposium on Computer science education*, pages 317–321. ACM Press, 2003.
- [2] Jeffrey L. Elman. Finding structure in time. *Cognitive Science*, 14(2):179–211, 1990.
- [3] J. Marshall, D. Blank, and L. Meeden. An emergent framework for self-motivation in developmental robotics. In *Proceedings of the 3rd International Conference on Development and Learning*. Salk Insititute, San Diego, 2004.
- [4] P-Y. Oudeyer and F. Kaplan. Intelligent adaptive curiosity: a source of self-development. In Luc Berthouze, Hideki Kozima, Christopher G. Prince, Giulio Sandini, Georgi Stojanov, G. Metta, and C. Balkenius, editors, *Proceedings of the 4th International Workshop on Epigenetic Robotics*, volume 117, pages 127–130. Lund University Cognitive Studies, 2004.
- [5] J. Schmidhuber. Developmental robotics, optimal artificial curiosity, creativity, music, and the fine arts. *Connection Science*, 18(2), 2006.