A Curious Khepera

A. Eppendahl
Department of Artificial Intelligence
Faculty of Computer Science, University of Malaya, Malaysia

K. H. Gan
Department of Engineering Design and Manufacture
Faculty of Engineering, University of Malaya, Malaysia

Abstract

We show that the curiosity algorithm developed by Kaplan and Oudeyer works with the Khepera robot and with a very simple prediction mechanism. We confirm that the behaviour generated by the algorithm accelerates learning and makes the robot appear 'more curious'. This strengthens the claim that the algorithm embodies a motivational principle, analogous to curiosity, that applies generally to any sensor-motor system. The algorithm requires both a predictor and a meta-predictor which predicts the errors of the predictor. We were surprised to get positive results with a predictor and meta-predictor that simply look up the most similar past experience. This supports the idea that it may be simpler to improve learning by clever behaviour generation than by clever learning mechanisms.

1. Introduction

How universal can a brain be? Ideally, a completely universal brain could be plugged into any body and, after a period of adaptation and development, would produce intelligent behaviour. It would do this autonomously and without prior information about the body or its environment.

In strong contrast to biological brains, which have evolved together with their bodies and environment, no signal into the completely universal brain can have a special status or absolute polarity because this would then place constraints on the body and the brain would not be universal. This means a completely universal brain cannot rely on grounded signals like pain or pleasure to guide its actions. The universal brain does, however, have access to its own workings and may base its behaviour on the significance of its own internal states. For example, if the brain has a mechanism that learns to predict the behaviour of its body and environment, it may try to do things that lead to improved predictions.

In a series of papers, Kaplan and Oudeyer have described experiments with such an architecture in simulation and on an Aibo robot. Two subtleties are worth noting. First, to evaluate possible actions, the Kaplan-Oudeyer architecture requires both a predictor and a meta-predictor which predicts the performance of the predictor. Second, even with a meta-predictor it is not obvious how to select actions that lead to continued development. In (Kaplan and Oudeyer, 2003) it is observed that learning tends to stall when actions are selected according to simple-mined criteria such as do-what-is-predicted-to-be-most-unpredictable. In (Oudeyer and Kaplan, 2004) a more sophisticated criterion is proposed that is designed to continually drive learning progress.

Because the aim of such work is to establish methods that apply to any robot in any environment, we took on the (somewhat thankless) task of implementing the Kaplan-Oudeyer algorithm on a Khepera robot to see if we could reproduce their results. Despite the very simple predictor (see below), we were delighted to observe a clear quantitative effect on learning and an obvious qualitative effect on behaviour.

2. The Predictor and Meta-predictor

Each time the robot executes a physical action a (a vector of 2 values from -500 to 500) in sensor situation s (a vector of 8 values from 0 to 1023), the predictor P_H is consulted for a sensor-prediction $\bar{s}$, which is compared with the actual sensor outcome s' to obtain a prediction error $e = d(s', \bar{s})$, where d is the Hamming metric, and the meta-predictor M_H is consulted for a error-prediction $\bar{e}$, which is compared with the actual error e to obtain an meta-prediction error $ee = d(e, \bar{e})$. The whole experience (s, a, s', e, ee) is then added to the history H.

$$H \mapsto H \cup \{(s, a, s', e, ee)\}$$

Given a history H, the predictor P_H predicts the next sensor vector $\bar{s} = P_H(s, a)$ as a function of the current sensor vector s and the current action vector a, and the meta-predictor predicts the error $\bar{e} = M_H(s, a)$ of the prediction $P_H(s, a)$. The predictions $\bar{s}$ and $\bar{e}$ are obtained by identifying the experience $(s_n, a_n, s'_n, e_n, ee_n)$, in the robot's history H, with (s_n, a_n) nearest (s, a) and taking $\bar{s} = s_n$ and $\bar{e} = e_n$.

The meta-predictor errors ee are only stored in the history so as to confirm that the meta-predictor is indeed learning (Gan, 2008). They are not used by the curiousity algorithm.

3. The Curiosity Algorithm

Given a predictor P_H and a meta-predictor M_H, the curiosity algorithm generates a sequence of 5 actions by evaluating 50 randomly chosen sequences and choosing the one with the best score. To evaluate a sequence of actions a_i, the predictor is used recursively to imagine the sequence of sensor states $s_{i+1} = P_H(s_i, a_i)$ that would ensue were the actions executed and the meta-predictor is used to predict a corresponding sequence of errors $e_{i+1} = M_H(s_i, a_i)$. The sum-of-falls, $\sum \text{pos}(e_{i+1} - e_i)$ where $\text{pos}(x) = x/2 + |x/2|$, then gives the score of the sequence a_i.

This is the measure of 'learning progress' studied in (Kaplan and Oudeyer, 2004). The idea is to continually oscillate between actions with known and unknown outcomes so as to surf the edge of predictability as learning progresses. Of course the sum-of-falls is calculated from imagined outcomes and there is no guarantee that the actions maximizing this imagined sum will actually cause the real errors to rise and then fall as desired. In practice, however, it seems that choosing actions in this way does accelerate learning.

4. Experiments

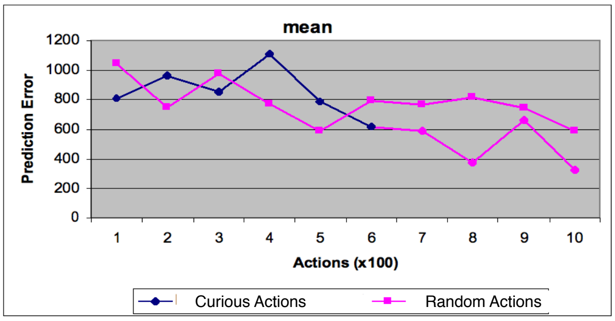
For the experiments, the Khepera robot is placed in a square paper corral that it is able to push around. The corral is about four times the diameter of the robot: if the robot is pushing the corral in one direction, it can move about three body lengths in the opposite direction before it begins to push it the other way. When the robot pushes the corral to the edge of the table, the program is paused and the corral and robot are shifted back the the centre of the table.

The Figure shows the results for a typical run. A run consists of 500 actions generated by the curiosity algorithm, followed by 500 random actions. Predictor error is plotted (averaged over intervals of 100 actions) along with predictor error for a typical run of 1000 purely random actions.

During the first 500 actions the curious Khepera appears to learn at about the same rate as the random Khepera, but this is misleading because learning is being judged with respect to different behaviours. After the first 500 atctions, when both robots are moving randomly, it becomes clear that the curious Khepera has actually learned more.

Subjectively, the curious Khepera spends more time 'investigating the walls' of the corral than the random Khepera. Looking carefully at videos of the robot, there are also signs of the oscillations near the wall addressed in (Oudeyer and Kaplan, 2004).

The full report (Gan, 2008), videos and Khepera code can be obtained from the first author's home page: http://hompage.mac.com/a.eppendahl/work/



References

Gan, K. H. (2008). Implementing the Kaplan-Oudeyer curiosity algorithm on the Khepera. Final year project report, University of Malaya, April 2008.

Kaplan, F. and Oudeyer, P.-Y. (2003). Motivational principles for visual know-how development. In *Proceedings of the Third Epigenetic Robotics Workshop*.

Kaplan, F. and Oudeyer, P.-Y. (2004). Maximizing learning progress: an internal reward system for development. *Embodied Artificial Intelligence*, LNAI 3139.

Kaplan, F. and Oudeyer, P.-Y. (2007a). In search of the neural circuits of intrinsic motivation. *Frontiers in Neuroscience*, 1(1).

Kaplan, F. and Oudeyer, P.-Y. (2007b). Un robot motivé pour apprendre : le role des motivations intrinseques dans le developpement sensorimoteur. *Enfance*, 59(1).

Oudeyer, P.-Y. and Kaplan, F. (2004). Intelligent adaptive curiosity: a source of self-development. In *Proceedings of the Fourth Epigenetic Robotics Workshop*.

Oudeyer, P.-Y., Kaplan, F., and Hafner, V. V. (2007). Intrinsic motivation systems for autonomous mental development. *IEEE Transactions on Evolutionary Computation*, 11(1).