Self-Organizing Incremental Neural Network (SOINN) as a Mechanism for Motor Babbling and Sensory-Motor Learning in Developmental Robotics

Tarek Najjar and Osamu Hasegawa

Tokyo Institute of Technology,
Dept. of Computational Intelligence and Systems Science,
4259 Nagatsuta-cho, Midori-ku, Yokohama, 226-8503, Japan
http://www.titech.ac.jp/english/

Abstract. Learning how to control arm joints for goal-directed reaching tasks is one of the earliest skills that need to be acquired by Developmental Robotics in order to scaffold into tasks of higher Intelligence. Motor Babbling seems as a promising approach toward the generation of internal models and control policies for robotic arms. In this paper we propose a mechanism for learning sensory-motor associations using layered arrangement of Self-Organizing Neural Network (SOINN) and joint-egocentric representations. The robot starts off by random exploratory motion, then it gradually shift into more coordinated, goal-directed actions based on the measure of error-change. The main contribution of this research is in the proposition of a novel architecture for online sensory-motor learning using SOINN networks without the need to provide the system with a kinematic model or a preprogrammed joint control scheme. The viability of the proposed mechanism is demonstrated using a simulated planar robotic arm.

Keywords: Developmental Robotics, SOINN, Self-organizing Neural Network, Motor Babbling, Sensory-Motor Learning, Incremental Learning.

1 Introduction

Inspired by Both Developmental Psychology and Cognitive neuroscience, developmental robotics has gained considerable interest among roboticists recently, [1]. The basic concern in this discipline is to formulate embodied Artificial Agents that are capable of autonomous mental development[2],which is the ability of the agent to adapt and grow mentally in the way it perceive, represent and process its experiences and the way it acts in the world around . This development must take place through interaction with the environment, using the agent's sensors and actuators, in a continuous life-long and open-ended manner[3].

Evidence from developmental psychology literature[4][5] suggests the presence of exploratory learning processes in the behavior of infants during the first

I. Rojas, G. Joya, and J. Cabestany (Eds.): IWANN 2013, Part I, LNCS 7902, pp. 321–330, 2013.
© Springer-Verlag Berlin Heidelberg 2013

months of motor-ability development. During the repetitive random motion of the arm, that is considered as a characteristic pattern of infant motor behavior, babies are believed to keep their hand constantly in visual field, which is supposed to serve the goal of building internal associations between actions and consequences in one's own body [6]. So Motor Babbling is described as the exploratory learning process of generating sensory-motor associations through continues random motions with ballistic trajectories. These motions serve the purpose of sampling representative data points that bootstrap the learning system into incremental generation of internal model and implicit control policy for the system at hand.

Many roboticists have attempted to mimic this developmental process using robotic platforms. An example is found in the work of the group[7], here a gradient descent method is used in order to enable the system to learn some of the unprovided elements of the system's kinematic model where the rest of the elements were already provided and preprogrammed. A more efficient approach than gradient descent was taken by group in[8] where the system starts off by a population of candidate possible models then, and through interaction with environment, the system evolve in approximating a more accurate model that represents the system in hand. Beside the explicit dependency, in this system, on artificial visual tags that are attached to segments of the robotic arm , this approach make use of Bayesian learning and Gaussian regression, the mechanism actually is very expensive on the computational side. A rather different approach was taken by the group[9]. Here a camera calibration based method were adopted together with open loop mechanism for generation of an implicit body schema model, this system made use of look-up table learning mechanism which naturally requires longer time for learning. The research group in[10]used a more biologically inspired approach by incorporating concepts like population code and equilibrium-point hypothesis in order to enable the system to achieve reaching tasks. In a different approach[11]the research group used both Bayesian belief functions and social learning mechanisms to facilitate learning-by-imitation competence. This approach actually made use of hard-wired motor primitives that were encoded manually into the robot. A Reinforcement learning approach together with imitation methods using locally weighted regression was facilitated by[12]where a robot was taught specific motor primitives, that are specific to given task sittings, then the robot generated policies that enable it to learn those primitives in an episodic manner . Although the robot managed to perform the given tasks but it seemed like the system was kind of a task-specific oriented in the way it learned each motor primitive.

2 Methodologies

The mechanism we are proposing is based on the idea of autonomous, incremental generation of implicit system model and control policy using layers of self-organizing maps and joint-egocentric representation of reaching experiences. The robot is not provided with any control models or methods for calculating Inverse and Forward kinematics.

2.1 SOINN

The core associative learning mechanism that is adopted in this research is based on Self-Organizing Incremental Neural Network (SOINN)[13]. SOINN is a Self-organizing map that does not require any presumption to be made about the topology or the distribution, of data.

Basically SOINN works by propagating network topology in a way that would self-organize as to resemble the "hot zones" of perception. For example, if a new data point is presented to SOINN then the algorithm would find the closest two network nodes to this newly presented data point ,Fig. 1.a. once these, most closest, nodes are found , SOINN determines whether the newly presented data point is within the coverage zone of these nodes. If yes then these nodes would be now connected by an edge to make up a single cluster of nodes and then they would be altered as to reflect the current blobs of persistent activity,Fig. 1.b.

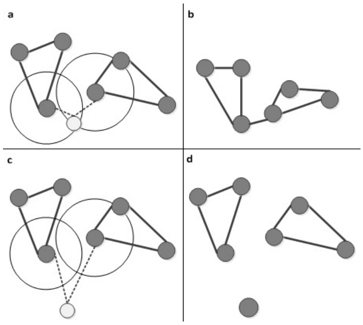

Fig. 1. SOINN dynamics

In the other case where the newly presented data point is out of the coverage zone of the closet nodes, Fig. 1.c, then this data point itself would be stored by SOINN as a node that represents a possible independent zone of activity, Fig.1.d. For a detailed explanation of the algorithm see[13].

SOINN has the feature of eliminating noisy and non-stable representations by checking the level of activity of each stored cluster of nodes and then discarding those stored clusters that doesn't represent regions of input space with high activity. So if a cluster of nodes has not been referenced frequently as being a coverage zone for input data points, then this cluster would eventually fade away and removed from the network.

2.2 The Architecture

First of all it is important to mention that each sensory-motor experience is represented and learned as a pairing between joint angle and the resultant gripper location in space. This pairing is joint-related, or joint-egocentric, i.e. for a given joint this sensory-motor learning experience would be $[\theta_i, L_i]$ where θ_i is the angle of joint i, and L_i is the resultant location of the gripper represented in relevance to the joint i ,hence, in the Peripersonal space of Joint i. Representing the location in the Peripersonal space of a given joint could be achieved by using a receptive field or mathematical transformation method. The purpose of this joint-egocentric representation is to make sure that learning is achieved on the joint level, where each joint would learn, the required associations, in manner that is independent from the other joints.

Each joint has its own associative learner, implemented as self-organizing map (SOINN)Fig. 2. This learner is responsible of learning sensory-motor pairs of the form $[\theta_i, L_i]$ that are related to the joint to which the self-organizing map belong to ,as mentioned above .

When a new target is presented to the system, the location of this target is represented in relevance to the first joint. Then the system would ask the self-organizing map, of the first joint, for the best angle that would achieve as close gripper location as possible to the given target . Depending on the joint's previous experience, the self-organizing map would respond by retrieving the joint angle that is associated with the closest gripper location to the target.

Now this angle, would be used to actuate the first joint of the manipulator even before passing the control to the next joint. This means that after the system has found out the suitable joint angle for the first joint, the target perception would be altered for the rest of joints on the manipulator, so in order for the next joint "$joint_{i+1}$" to ask its associative learner for suitable joint angle, θ_{i+1}, the robot must check the new altered location of target, L_{i+1}, in relevance to the next joint i.e. in the Peripersonal space of the next joint.

Fig. 2 reveals the iterative nature of the solution proposed here, where the problem of finding the best set of joint angles for multi-joint manipulator is solved by breaking down the reaching task into smaller sub problems, each handled by an independent subsystem that consist of single joint with its own perceptual space and its own associative learner.

2.3 From Exploration into Coordinated Reaching

In the approach we are proposing, training and learning take place in a real-time manner. The system itself decides when an exploration action is needed and when actual goal-reaching can be performed while the system is being trained continuously in both cases. So initially when we run the robot for the first time, the robot actions would be random ballistic trajectories similar to the ones performed by infants at early stages of motor development[5]. During this random motor babbling behavior the robot starts to generate an internal model for the control policy of its joints, through action-consequences coupling, which result in an increased ability to control these joints in coordinated manner, hence a less resultant error in reaching a target. To control the balance between motor babbling and target-reaching behaviors the following equation is used:

$$P(rnd) = 0.5 + \xi(m_{cp} - m_{fp}) \tag{1}$$

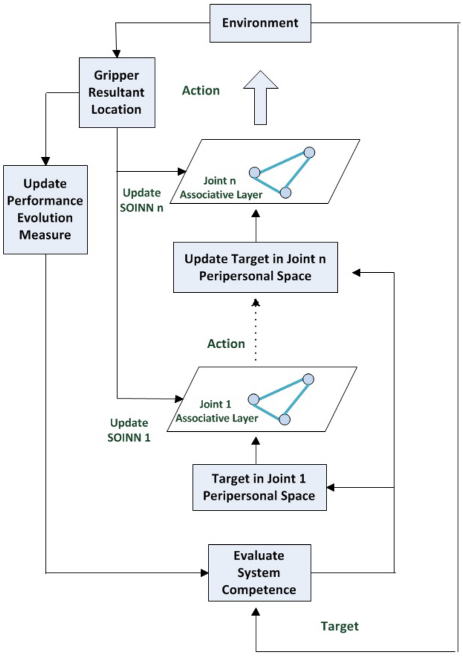

Fig. 2. The system Architecture

Where $P(rnd)$ is the probability of performing a random action, and $\xi(x)$ is the normalized value of x. The quantity m_{cp} is the mean error in the close past and m_{fp} is mean error in the far past.

The concept of close past and far past is generated by making the system maintains, at each time step, a list of measured error, described as the distance between the target location and the resultant gripper location, during the last n steps. This list then is divided in two halves. The most recent half, which consist of set of errors between $j = t$ and $j = t-(n/2)$, is considered as a set of errors in the close past. The other half, that consists of set of errors between $j = t-(n/2)$ and $j = t - n$, is considered as a set of errors in the far past.

Dividing the most recent n time steps into close past and far past serves the goal of altering the frequency and the necessity for random actions. So when error is reducing, and the robot performance is getting better, a negative value of $\xi(x)$ would result, which would decrease the random action probability, $P(rnd)$. on the other hand, when the error is increasing, a positive value of $\xi(x)$ would be generated resulting in higher motor-babbling probability, equation 1.

3 The Experiments

In this experimental setup, a simulated 2DoF planar robotic arm is used to demonstrate the developmental sensory-motor learning process, starting by random motor-babbling actions and then shifting gradually toward performing more coordinated target-reaching trajectories.

It is crucial to mention that the robot was not provided with any knowledge about how to control its joints, besides no action-consequence model was preprogrammed by the designer beforehand of learning.

A red ball is used as the target that the robot is required to reach at any given time. The ball location is generated randomly and then the robot is asked to reach it with its end effector, then, after the robot trail to reach the target, a new location is generated whether the robot has managed actually to reach the target or not, Fig. 3.

As mentioned above, and illustrated in Fig. 2, the trajectories that are performed by the robot, whether target-directed or random, are always used as a training signal for the learning system, which implies a continues adaptation and learning of the generated implicit model of control.

In Fig. 4, a gradual decrease in error is noticed with more practicing of the learned model that was initiated by the babbling actions.

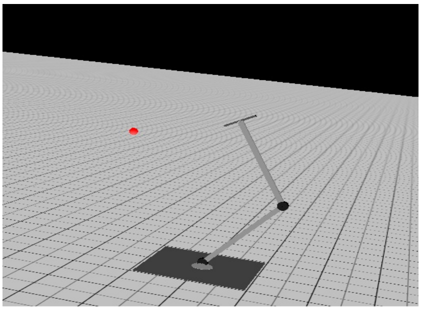

Fig. 3. The Experimental setup

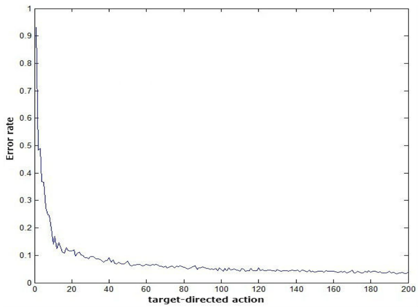

Fig. 4. resultant error during learning

The robot performance starts with high error rate. But with more training experiences the multilayer architecture of self-organizing map, SOINN, starts gradually to capture the contingencies behind joints angles and resultant end effector location. This incremental self-organizing process results in the observed decrease of the anticipated error of generated actions.

3.1 A Sudden Change

In this second experiment we demonstrate the system's reaction to a sudden unexpected change in the physical structure of the robot. This sudden change could account for a breakage in a joint, increased length of a link or a displacement of the end effector location in relevance to the arm links.

In this experimental setup we still have the same task of reaching a red ball, but now , after the system has learned its own implicit model, we suddenly

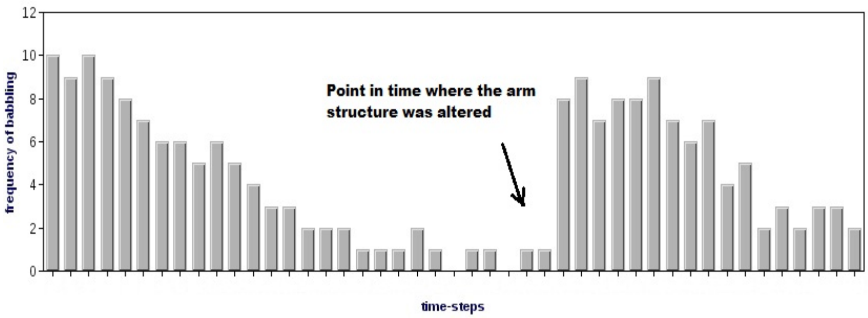

Fig. 5. A real-time reaction to an unexpected change in the physical structure

increased the length of the arm's second link by 10% of its original one. Altering the physical structure of the system means that now the learned implicit model does not accurately reflect the actual system nature. So if the system, before this unexpected change, had already reached a level of stability in term of the frequency of babbling actions, where a lower rate of random motion could be noticed, then now this stability won't last, and the robot would need to re-explore the contingencies of its action-consequence relation.

In Fig. 5, the horizontal axis shows a sequence of groups of time-steps ,each consist of 10 actions, that depicts the transition of the robots performance between motor babbling and target-directed actions. The vertical axis shows the number of babbling motions that was performed in each group of 10 time-steps.

As expected, most of the robot's actions, when it starts learning, are babbling ones and that is because the robot is not aware of its kinematic model. But then gradually this rate of babbling actions would decrease as the system proceed in building an implicit model of its control. Eventually we notice that almost no babbling actions are being performed but rather almost all of the taken actions are goal-directed.

During the robot's performance we altered the second link length, as mentioned above. This change would increase the resultant error in the robot's reaching accuracy because the learned sensory-motor associations does not accurately reflect the actual current status of the system. This increased error would generate a positive difference between m_{cp} and m_{fp} from equation 1, what eventually results in a higher $P(rnd)$, which is the probability of performing a babbling action.

This change in the behavior of the system can be observed in Fig. 5 where a peak in the frequency of babbling actions is clearly noticed around the point in time where the physical structure of the robot was altered. What can be noticed also is that the domination of babbling actions won't last forever, but rather it would be there as long as the system hasn't fully recaptured the Contingent action-consequence relation of its recently altered physical structure.This observation emphasizes the impact of the concept of learning through babbling on the ability of the system to adapt and react to unanticipated changes and conditions.

4 Discussion

A visualization of the first layer of SOINN is depicted in, Fig. 6. A 3-dimensional visualization of this resultant network can be seen in Fig. 6.a, where each node represents a single representative associative sensory-motor pairing of the form $[\theta_1, L_1]$, as described in section2.2. if we look at the topological structure of this network from 2-dimensional perspective, Fig. 6.b, we notice that it captures a very similar structure to the Cartesian work space but spawned across a third dimension of the associated angles of $joint_1$.

Next is a visualization of the learned SOINN network but for $joint_2$, Fig. 7. Again the network to the right, Fig. 7.b, is the 2-dimensional perspective of the 3-dimensional SOINN network, Fig. 7.a, that represents the sensory-motor associative model for $joint_2$. Notice, from Fig. 7.b, the egocentric characteristic

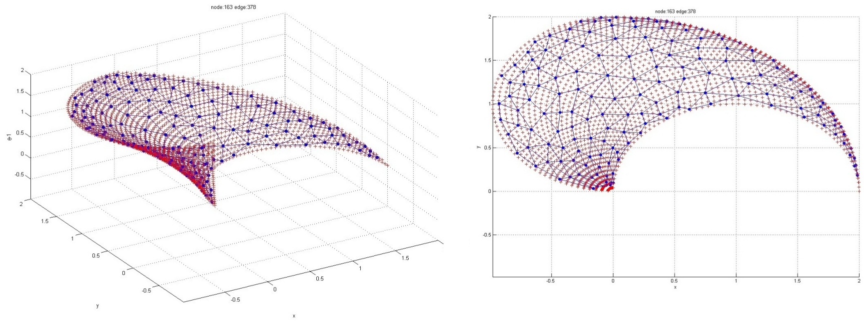

Fig. 6. A 3-dimensional representation of the approximated sensory-motor associations that correspond to $joint_1$ (left). the same network but in 2-dimensional perspective (right)

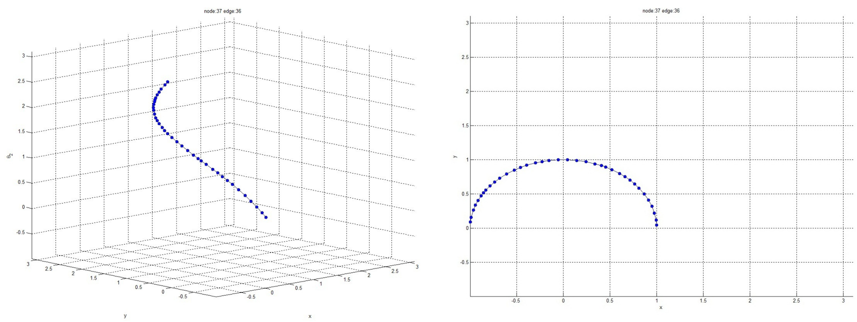

Fig. 7. A 3-dimensional representation of the approximated sensory-motor associations that correspond to $joint_2$ (left). the same network but in 2-dimensional perspective (right)

of the learned model since the Cartesian part of the associative data points does not reflect the whole work space but rather it captures only locations that are taken from the perspective of $joint_2$.

In both learned networks, Fig. 6 and Fig. 7, we notice that SOINN has the ability to cover the whole input training space with consistent distribution of nodes that enables the system to generalize even for unseen data points that was not provided during the process of network generation. This was demonstrated in Fig. 4 where the ball location was generated in continues input space rather than discrete one, but yet the system managed to generate trajectories of decreasing error even without the need for separated training and testing phases.

5 Conclusion

In this paper we have presented an architecture for learning sensory-motor associations for coordinated reaching tasks, using Self Organizing Neural Network (SOINN). The approach that was taking is inspired by developmental psychology where motor learning starts by babbling-like ballistic trajectories, similar to the ones observed during early stages of motor development in human infants, then the robot shifts toward coordinated actions with continuously decreasing error. This Developmental approach toward robot learning was demonstrated by the fact that no preprogrammed control policy was provided beforehand of learning. But rather the robot explored, on its own, the action-consequences contingencies of its joints and then, autonomously, generated an implicit control model through Motor babbling actions.

Acknowledgments. This work was sponsored by the Japan Science and Technology Agency's CREST project.

References

1. Lungarella, M., Metta, G., Pfeifer, R., Sandini, G.: Developmental robotics: a survey. Connection Science 15(4), 151–190 (2003)
2. Weng, J., Hwang, W.: From neural networks to the brain: Autonomous mental development. IEEE Computational Intelligence Magazine 1(3), 15–31 (2006)
3. Stoytchev, A.: Five basic principles of developmental robotics. In: NIPS 2006 Workshop on Grounding Perception, Knowledge and Cognition in Sensori-Motor Experience (2006)
4. Piaget, J.: The origins of intelligence in
5. Van der Meer, A., Van der Weel, F., Lee, D., et al.: The functional significance of arm movements in neonates. Science-New York Then Washington, 693–693 (1995)
6. von Hofsten, C.: Eye–hand coordination in the newborn. Developmental Psychology 18(3), 450 (1982)
7. Hersch, M., Sauser, E., Billard, A.: Online learning of the body schema. International Journal of Humanoid Robotics 5(02), 161–181 (2008)

8. Sturm, J., Plagemann, C., Burgard, W.: Body schema learning for robotic manipulators from visual self-perception. Journal of Physiology-Paris 103(3), 220–231 (2009)
9. Metta, G., Sandini, G., Konczak, J.: A developmental approach to visually-guided reaching in artificial systems. Neural Networks 12(10), 1413–1427 (1999)
10. Caligiore, D., Parisi, D., Baldassarre, G.: Toward an integrated biomimetic model of reaching. In: IEEE 6th International Conference on Development and Learning, ICDL 2007, pp. 241–246. IEEE (2007)
11. Demiris, Y., Dearden, A.: From motor babbling to hierarchical learning by imitation: a robot developmental pathway (2005)
12. Kober, J., Peters, J.: Learning motor primitives for robotics. In: IEEE International Conference on Robotics and Automation, ICRA 2009, pp. 2112–2118. IEEE (2009)
13. Furao, S., Ogura, T., Hasegawa, O.: An enhanced self-organizing incremental neural network for online unsupervised learning. Neural Networks 20(8), 893–903 (2007)