Proceedings of the Post-Graduate Conference on Robotics and Development of Cognition
10-12 September 2012, Lausanne, Switzerland

Developing Motor Skills for Reaching by Progressively Unlocking Degrees of Freedom on the iCub Humanoid Robot

Salomón Ramírez-Contla[1]*, Angelo Cangelosi*[1]*, Davide Marocco*[1]

[1]University of Plymouth, Adaptive Behaviour and Cognition Lab
Drake Circus, Plymouth, England

salomon.ramirez-contla@plymouth.ac.uk, angelo.cangelosi@plymouth.ac.uk,
davide-marocco@plymouth.ac.uk

Abstract

To explore development of motor skills for reaching in the iCub robot, we test the capabilities for a neural network controller to learn progressively by locking some degrees of freedom (DOF) of the robot's arm before allowing it to explore the space with more DOF's. We consider exploration and bio-inspired mechanisms can aid in the development of control of the iCub robot arm. Results suggest the advantage of progressive development over an initial full training, also, these pointed out the importance of interaction with the world and the necessity of trial and error occurring in a time lapse for developing of reaching skills.

Index Terms: degrees of freedom, motor skills, development, epigenetic robotics

1. Introduction

Proposed by Bernstein, the degrees of freedom problem [1] poses difficulties for autonomous skills learning and has drawn attention recently in the psychology field [2, 3]. Recent research on robotics [5, 4] has addressed this problem as well and tried to implement some of the ideas proposed by Bernstein due to the nature of recent advances in robotics and the need of developing controllers for redundant robot arms, specially of those of humanoids. Current cognitive robotics research has focused on the importance of the embodiment of an agent in order to richly interact within a world plenty of stimuli and cues that can aid in processes and reduce workload for a central controller such as the brain. The body plays an important role for this interaction and roboticists constantly look for new and better ways to control it.

Studies with evolutionary robotics approaches have been carried out with success for reaching and maipulation tasks. Massera et al. [6, 7] successfully evolved networks capable of fine-grained interaction with objects by exploiting the morphological contraints of a robotic arm. In this work, however, we are interested on the epigenetic development of such tasks.

Development of the human body flows from the top and centre of the body to the limbs. The spinal cord is the starting point, arms, legs, hands and toes take longer to develop. It is said that it follows a proximo-distal and cephalo-caudal direction [8] and this can be appreciated in infants: younger infants move their limbs in broad uncontrolled movements because only the most proximal joints of the limbs have been developed, like the shoulders. Later on, it can be seen that the elbow and wrist also come into play. Also, control of the lower part of the body comes after that of the upper part. In experimental psychology and motor development of humans there is evidence indicating that for learning new skills, adults freeze

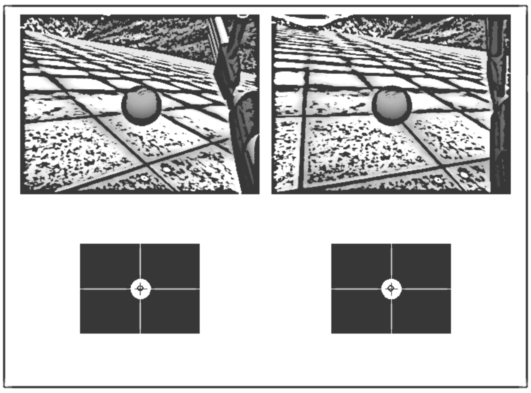

Figure 1: *Images from the robot once the controller has foveated the target. Above, the original images. Below, the low-res colour-segmented images.*

some of the distal joints involved in the new task until some degree of performance has been achieved, then some more degrees of freedom are used for achieving better performance [9, 10]. In the present work we test if that interaction with the world along with experience limited by constraints imposed by the physical characteristics of the arm, can help the learning process if this is segmented. We use a simulated iCub robot with neural controllers for the arm.

2. Methods

For testing the hypothesis, experiments were planned and carried out on the iCub robot simulator [11]. The iCub robot [12] is a humanoid robot about the size of a four years old child with 53 degrees of freedom designed for cognitive development research. The iCub's head subsystem consists of six degrees of freedom and is capable of vergence (the oculomotor adjustment needed to foveate the same point in space with both eyes). Three degrees of freedom in the head (tilt, pan and eyes' vergence) and four degrees on the arm (two from the shoulder, two from the elbow) were used. The robot head was provided with a visual tracking controller that locates and gazes at a specific target. For the experiments the target was a red ball. The gazing controller performed colour segmenting for the target's colour on the images coming from both eyes. This processing allowed it to track the centre of the target and adjust the position of three joints in the head in order to have the target in the centre of each eyes' field of view (Fig 1). By this mechanism, the robot

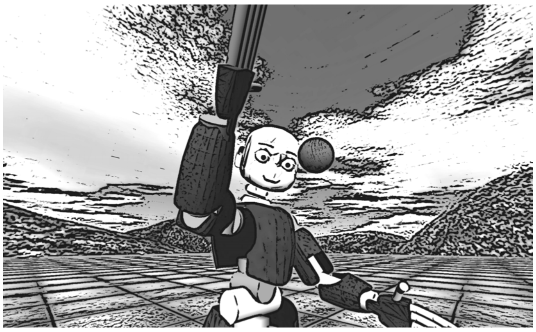

Figure 2: *The iCub performing the reaching task once it has foveated.*

gets information about the depth or distance at which the target is and together with the tilt and pan joints positions, it encodes the positions of the target in space. We use vergence as a depth measure following recent findings [13, 14] that indicate that vergence is in fact one strong signal for depth estimation and programming of prehension movements of humans.

Three different learning conditions (with three networks each) were tested on the robot to test our hypothesis. The first two were: staged learning or development, involving learning control of two DOF's and then the two other (DEV condition), and learning the head-hand associations involving four DOF's from the beginning (NO-DEV condition). For the last condition (NO-TRAIN), a group of three randomly initialized networks were created. These did not go through any learning process and are the control group.

With the help of the gazing mechanism a dataset was captured consisting of joint values of the head and eyes and the joint values of an arm position suitable for locating the end effector (the hand) in the point where the target was. This process can be considered a tutoring stage where the ball was put in the hand every time the robot executed random babbling [15] with the arm, then the gaze controller moved the head for foveating the target. For the cases the head was not able to move to a position were the target could be gazed no data was captured. This train set is equivalent of one acquired by performing random babbling while foveating the target. Reduction of the time required by this process is of course reduced when this kind of tutoring is present, as it happens with infants helped by parents when they start trying to reach objects that are usually out of reach or the baby simply fails to reach.

The controller for the robot arm was a feed-forward network with three inputs (one for each joint of the head controlled by the gazing controller), forty hidden units and four outputs, each of these output units controlled one joint of the arm. Two of these joints are in the shoulder and two in the elbow of the robot. During the initial training set creation, random babbling only occurred for the two most proximal joints of the arm, that is, for the ones in the shoulder. The other two joints were kept in constant values, in positions that we considered natural for an extended, similar to those when performing reaching for objects not very close to the body. Therefore, the positions that can be reached after the initial training are determined by the physical characteristics of the arm and by the generalization capabilities of the network. All learning for the networks was with the backpropagation of error algorithm using a learning rate of 0.01 and a momentum of 0.1.

For the development condition (DEV), each of the three net-

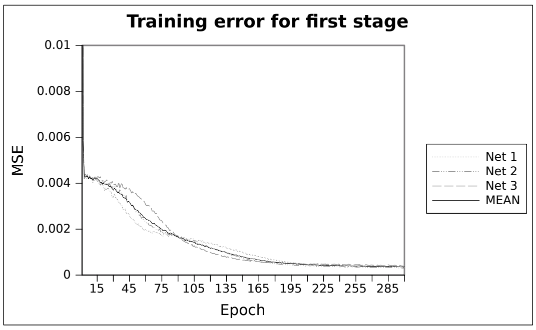

Figure 3: *Error during training of the first stage of development.*

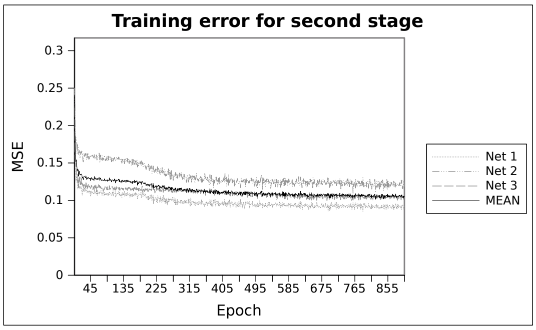

Figure 4: *Error during training of the second stage of development.*

works was trained using the set acquired via tutoring until the mean square error (MSE) became stable. For the three networks this was around the three hundred epochs. Figure 3 shows the training error for this stage of learning. After the initial training for reaching using two degrees of freedom, a test phase was carried out in ecological conditions. For an extended period of time, the robot was presented with the target in different locations, each time, the robot gazing mechanism was used for gazing the target, then the arm neural controller was activated with the inputs coming from the position of the head and eyes. When the robot successfully reached the target, that is, it touched it, the arm went to it's resting position and the next test target position was presented. Otherwise, the two degrees of freedom that were initially locked (remember their values were constant for the first phase of learning) were randomly moved while the two most proximal degrees of freedom were kept constant with the values the neural controller produced. With this movement the robot was sometimes able to reach the target. When that was the case, the position that enabled it to achieve reaching was stored in a new set that was used for later training. This phase will be called from now on "experience phase". Figure 4 shows the training error during the second phase of learning for the three networks and the mean of the three of them.

The training using the new set generated in the experience consisted of 900 epochs. Figure 4 shows the error during training of this second stage of the learning.

The second condition (NO-DEV) consisted of using the training set generated during the experience phase on randomly initialised networks without going through an initial, partial, learning phase nor an experience phase. That is, these controllers were trained with the set that uses four degrees of freedom from the beginning. The training was for nine hundred

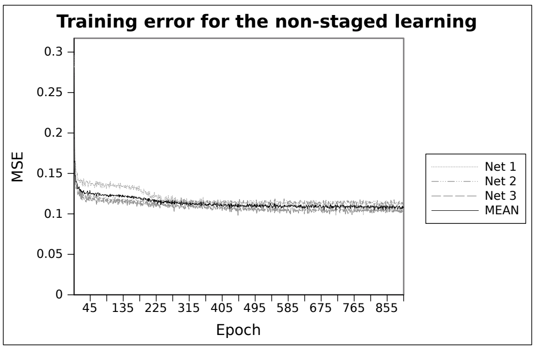

Figure 5: *Error during training of the non-developed controllers.*

epochs, at that point the MSE was stable. Figure 5 shows the error for this no-development learning.

Measurements for comparing the two conditions were performed during the execution of a reaching task similar to the task executed in the experience phase. Figure 2 shows the iCub executing the task once it has foveated the target. Final distance from hand to target was saved for each of the trails of the three controllers. Also, the number of times the controller successfully reached was recorded for having a percentage of success for each of the networks.

3. Results and discussion

Analysis on the output data indicates the controllers belonging to the staged or developmental training performed better in terms of final distance to the target as well as in the percentage of success (Figs 6 and 7). An analysis of variance test was performed to check for statistical difference between conditions (including the non-trained condition). This test reported statistical difference: current effect $F_{2,897}$=850.45, p=0.0000.

This can be due to various factors: following a developmental training, consisting of tutoring, experience during operation in it's environment and learning based on that experience could have shaped the weights of the controller's networks to a stage that was able to find a solution for the second training set. Even when the training error of the final training in both conditions is very similar, in test conditions an advantage of the developed can be appreciated.

Because reaching is an important step in the development of motor and cognitive skills, it is also a skill explored to get an insight of the series of processes emerging in infants [16]. Our work on development of reaching tries also to consider the fact that for acquiring a skill it is necessary to has trial-and-error processes where time constraints cannot be avoided. In our experiments, the generation of the second training set for the staged learning condition, the "experience phase", took considerably longer than any other part of the experiment. But we believe this was a very important step due to the fact that each network will generate different outputs for the same inputs so the set is particular to each of them.

We have tried to implement what Bernestein [1] suggested for simplifying the degrees-of-freedom problem: in our experiments the robot arm controller goes through a developmental progression in order to find a first but simpler solution to the problem an later on, increasing the complexity of the problem. Other roboticists have implemented similar ideas in recent years. Ivanchenko and Jacobs [17] simulated a three degrees of

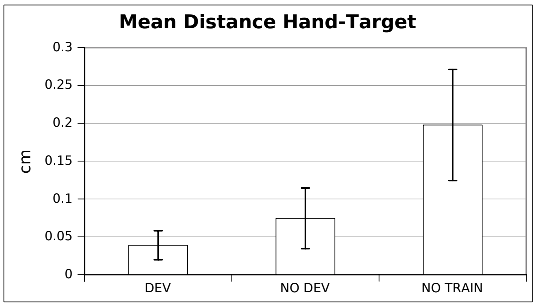

Figure 6: *Mean distance from the center of the palm fo the hand to the center of the target. Each controller was tested with 60 different target locations, none of them belonging to any of the sets used for training. Bard indicate standar deviation for all tests on each condition.*

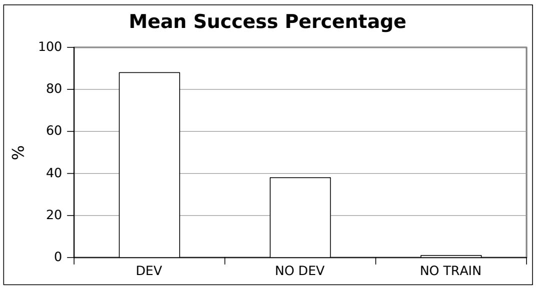

Figure 7: *Results show the percentage of times the end efector touched the target. Each controller was tested with 60 different target locations, none of them belonging to any of the sets used for training.*

freedom robot arm that tries to learn the dynamics of the arm while moving on trajectories on two dimensional space. The difference with our approach is that in our case the architecture of the networks is the same for every condition, it is the presence or absence of experience what shapes the performance at their final stages. Ivanchenko has a special architecture, devised from the idea that this decouples dynamic interactions among the joints and therefore allows to separately train the joints. Unlike Ivanchenko, for our experiments we decided to keep the same architecture. We want to explore uncoupling of dynamics without changing the internal (not directly exposed to the environment) characteristics of the system. In Ivanchecnko's, results indicate that a developmentally trained controllers only outperformed the non-developmentally ones when the developmental path matched the nature of the task executed. In the case of our experiments, training as well as the "experience" phase matched the final task. This could explain the obtained results.

As Massera et al. in [6], we have started this exploration on a robotic arm with just four degrees of freedom. Our approach contrasts with that one in that we are interested on the epigenetic development of the skill instead of an evolutionary one. Moreover, in our case, experiments look to include vision into the development of the task instead of direct pass of coordinates or distances to the system without visual processing, as we consider that working towards the implementation of this type of skill development will require real-life sensory capabilities. The head controller for our experiments employs vision as a simple processing but action-involving task. Schlesinger et al. [18]

have also explored with the freezing of DOF's but again, using a non-realistic vision mechanism and a 2D environment and using evolutionary algorithms. Our work has pushed this type of exploration to a more realistic environment and explores the interaction on fixed architecture systems. We showed that even in this circumstances, a developmental approach can lead to better performance. Using the iCub simulator has proven to be a good test-bed for this type of research, as it allowed to implement and test controllers and visual sensors and explore performance in a controlled environment and free of mechanical strain issues.

3.1. Future work

In this study we have investigated the advantages of a progressive unlocking of joints to achieve better reaching performance. Our system has used two and then four degrees of freedom to explore and then improve a motor skill. However, limbs of natural systems, such a humans, display the property of overcompleteness. Overcompleteness implies that even though only 4 degrees of freedom are required for navigating a limb through three dimensional space [19], limbs on many vertebrates usually exhibit more than 4 degrees of freedom. This property turns the problem of controlling a limb more complex in computational terms (at least for traditional control) but also can represent and advantage in terms of the possibility of finding solutions that allow to reach a target at the same time that an obstacle is avoided. This could keep a relation with the representation of the reachable space. Also, constraints in other sensory or mechanical parts will be explored in further work.

4. Acknowledgements

This research has been funded by the National Council on Science and Technology of Mexico (CONACyT) and EU FP7 Project ITALK.

5. References

[1] Bernstein, N., "The coordination and regulation of movements", Pergamon, 1967.

[2] Karl M. Newell, David E. Vaillancourt, "Dimensional change in motor learning", Human Movement Science, 20(4–5), pp. 695-715, 2001.

[3] Rohde, M., and Di Paolo, E., "t for Two: Linear Synergy Advances the Evolution of Directional Pointing Behaviour", Proc. Advances in Artificial Life: 8th European Conference, ECAL 2005, pp.262-271, Springer, 2005.

[4] Berthouze, L., and Lungarella, M., "Motor skill acquisition under environmental perturbations: On the necessity of alternate freezing and freeing of degrees of freedom" Adaptive Behavior, 12(1), pp. 47-64, 2004.

[5] Gómez, G., Lungarella, M., Eggenberger Hotz, P., Matsushita, K., and Pfeifer, R., "Simulating development in a real robot: on the concurrent increase of sensory, motor, and neural complexity", Proc. of 4th Int. Workshop on Epigenetic Robotics, pp. 119–122, 2004.

[6] G. Massera, A. Cangelosi, S. Nolfi, "Developing a Reaching Behaviour in a Simulated Anthromorphic Robotic Arm Through an Evolutionary Technique", in Artificial Life X: Proceeding of the Tenth International Conference on the simulation and synthesis of living systems MIT Press, 2006.

[7] Massera G., Cangelosi A., Nolfi S., "Evolution of prehension ability in an anthropomorphic neurorobotic arm", Frontiers in Neurorobotics, 1(4), 2007.

[8] Sharma, S.K., "Principles Of Growth And Development", Isha Books, 2005.

[9] Vereijken, B., Van Emmerik, R., Whiting, H.T.A. and Newell, K., "Free(z)ing degrees of freedom in skill acquisition", Journal of Motor Behavior 24, pp. 133-142, 1992.

[10] McDonald, P. V., Van Emmerik, R. E., Newell, K. M., "The effects of practice on limb kinematics in a throwing task", Journal of Motor Behavior, 21(3), pp. 245-264, 1989.

[11] Tikhanoff, V., Cangelosi A., Fitzpatrick, P., Metta, G., Natale, L. and Nori, F., "An open-source simulator for cognitive robotics research: The prototype of the iCub humanoid robot simulator", in Proc. IEEE Workshop Perform. Metrics Intell. Syst., Washington, DC, 2008.

[12] Metta G., Natale L., Nori F., Sandini G., Vernon D., Fadiga L., von Hofsten C., Rosander K., Santos-Victor J., Bernardino A., Montesano L., "The iCub Humanoid Robot: An Open-Systems Platform for Research in Cognitive Development", Neural Networks, special issue on Social Cognition: From Babies to Robots. 23(8-9), 2010.

[13] Mon-Williams, M. and Dijkerman, H., "The use of vergence information in the programming of prehension", Experimental Brain Research, 128(4), pp. 578–582, 1999.

[14] Wismeijer, D., van Ee, R., and Erkelens, C., "Depth cues, rather than perceived depth, govern vergence" Experimental Brain Research, 184(1), pp. 61–70, 2008.

[15] Von Hofsten, C., "Eye-hand coordination in the newborn", Developmental Psychology, 18(3), pp. 450-461, 1982.

[16] Corbetta, D., Thelen, E. and Johnson, K., "Motor constraints on the development of perception-action matching in infant reaching", Infant Behavior and Development, 23, pp. 351-374, 2000.

[17] Ivanchenko, V. and Jacobs, R. A., "A developmental approach aids motor learning", Neural Computation, 15(9), Pages 2051-2065, 2003.

[18] Schlesinger, M., Parisi, D. and Langer, J., "Learning to Reach by Constraining the Movement Search Space", Developmental Science, 3, pp. 67-80, 2000.

[19] Shadmehr, R. and Wise, S. P., "The Computational Neurobiology Of Reaching And Pointing: A Foundation for Motor Learning", Computational Neuroscience,MIT Press,2005.