Prenatal to Postnatal Transfer of Motor Skills Through Motor-Compatible Sensory Representations

Timothy A. Mann and Yoonsuck Choe
Department of Computer Science and Engineering
Texas A&M University
College Station, TX 77843
Email: {mann23, choe}@tamu.edu

***Abstract*—How can sensory-motor skills developed as a fetus transfer to postnatal life? We investigate a simulated reaching task by training controllers under prenatal conditions (i.e. confined space) and evaluating them based on postnatal conditions (i.e. targets outside of the confined training space). One possible solution is to identify a sensory representation that is easy to extrapolate over. We compared two kinds of sensory representations: world-centered sensory representation based on Cartesian coordinates and agent-centered sensory representation based on polar coordinates. Despite similar performance under prenatal conditions, controllers using agent-centered sensory representation had significantly better performance than controllers using world-centered sensory representation under postnatal conditions. It turns out that the success of the agent-centered sensory representation is (in part) due to being complementary to the action encodings. Further analysis shows that the action encodings (i.e. changes in joint angles) were highly predictive of the change in state when agent-centered sensory representation was used (but not world-centered). This suggests that a powerful strategy for transferring sensory-motor skills to postnatal life involves selecting a sensory representation that complements the action encodings used by an agent.**

I. Introduction

In addition to learning throughout life, an interesting (and yet often overlooked) possibility is that biological organisms may have evolved to begin development of sensory-motor skills before birth. Prenatal skill development may provide a head start when compared with learning systems that do not develop skills *in utero*. In fact, Robinson and Kleven [1] demonstrate that modifications to prenatal conditions have lasting effects.

Although there is tactile stimulus, sound, and even light available to a developing fetus [2], there are significant physiological and environmental differences between life *in utero* and life after birth. For example, movement of the arms and legs is restricted (see figure 1). Kuniyoshi [3] found evidence that self-organizing neural models are able to learn sensory and motor structure of the body during the prenatal period despite physical restrictions, and furthermore such restrictions may actually be beneficial. This raises an interesting question: How can sensory-motor skills developed under prenatal conditions scale to novel situations that will be experienced in postnatal life?

To investigate how a sensory-motor skill can scale to postnatal conditions, we consider learning to reach targets with a two-link arm in confined space (i.e. prenatal condition) and evaluate learned controllers at reaching for novel targets (outside of the trained range) when the arm is not confined (i.e. postnatal condition) (see figure 2). This is similar to the problem faced by a fetus learning to reach with its arm. While in the uteris, the body is restricted so that some movements cannot be practiced (such as fully extending the arm). Our objective is to find clues that point to how a fetus might learn basic reaching behavior during gestation that scales to novel situations that will be encountered later in life.

This problem seems naturally phrased as an extrapolation problem, because the postnatal condition completely encompasses all prenatal situations. Unfortunately, without prior knowledge, for the general case of nonlinear surfaces, extrapolation is an unsolvable problem.

In artificial intelligence research, selection of an appropriate representation has long been considered a critical part of constructing successful intelligent solutions. One possible solution to the extrapolation problem may be to select a sensory representation that is easy to extrapolate over. Unfortunately, in sensory-motor tasks, the properties of a sensory representation that make it easy to extrapolate over are not well understood, because scalability is affected by the size of the training set, properties of the action encoding, and function approximation techniques. We compare a world-centered sensory representation (WC) based on Cartesian coordinates with an agent-centered sensory representation (AC) based on polar coordinates (see figure 3). The world-centered sensory representations relate the target to hand by explicitly specifying target and hand coordinates with respect to an arbitrary origin in the world. The agent-centered sensory representations, on the other hand, relate the target to hand by explicitly providing polar coordinates (i.e. angle and distance) for target and hand with respect to the agent's shoulder.

We applied a reinforcement learning algorithm, Q-learning [4] with nonlinear function approximation, to learn controllers for both world-centered and agent-centered representations.

Our experimental results show that, despite similar low error levels during training, learning systems using world-centered sensory representations had higher error under postnatal conditions than systems engaging with agent-centered sensory representations. Interestingly, further analysis shows that the action encodings used (i.e. changes in shoulder and elbow

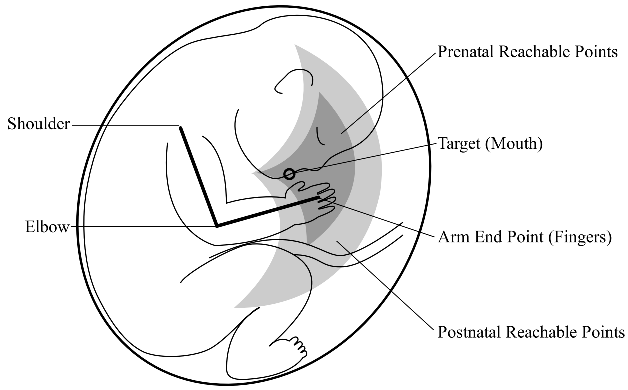


Fig. 1. Depiction of a human fetus under prenatal conditions. Under prenatal conditions movement of the arm is restricted by its own body and the uterine wall, while under postnatal conditions the arm has greater freedom of movement.

angles) are highly predictive of the sensory change when coupled with agent-centered sensory representations, but not with world-centered sensory representations. This suggests that the advantage of agent-centered sensory representations over world-centered is not so much a property of the representation as it is an emergent property achieved when complementary sensory representation and action encodings are used together.

The main contribution of this work is that it demonstrates the importance of complementary sensory representation and action encoding in the development of sensory-motor skills during gestation that scale well to postnatal life without additional training.

This paper is organized as follows. Section 2 provides background and related work about target reaching tasks. Section 3 describes reinforcement learning and the methods used for experimentation. Section 4 introduces experimental results. Section 5 analyzes results using a second experiment. Section 6 discusses the implications of the results and addresses several criticisms of this study. Section 7 provides concluding remarks.

II. Background

Controlling a physical arm is a difficult task. In addition to learning the joint angles that position the hand at the desired target point, correct muscle activation is required to maintain the arm at the desired location and account for gravitational force. However, empirical evidence from studies with human subjects suggests that the brain solves this problem by dividing it into an inverse kinematics problem (i.e. positioning the arm) and a dynamics problem (e.g. dealing with gravitational forces) [5], [6]. For the purposes of this paper, we will focus only on the inverse kinematics problem.

From the context of a developing fetus, the inverse kinematics problem is achieving joint angles for the shoulder and elbow (i.e. a pose) that place the hand at (or as close as possible to) a target point specified in a more flexible coordinate system capable of describing positions that may not be reachable by the hand.

The solution to the inverse kinematics problem depends on properties of the arm, such as length of the links, limits on the joint angles, etc. Differences in arm properties and natural body growth rule out the possibility of a genetically inherited, static solution for biological learning systems. Biological learning systems use an adaptive mechanism for arm control that adjusts to changes ([7]).

Previous work on learning inverse kinematic control has developed many solutions for reaching multiple targets (e.g. [5], [8], [9], [7], [10], [11]). Several researchers have solved the problem using value-based reinforcement learning methods [10], [11]. There are many different approaches, however, we identify some common themes in the literature. Three common themes are that solutions

1) have almost identical training and testing conditions,
2) use joint angles as input describing the pose of the arm, and
3) use nonlinear function approximators with large numbers of parameters.

Our investigation differs from previous solutions to learning inverse kinematics because (1) our emphasis is on how well a solution scales from prenatal to postnatal conditions (i.e. different training and testing conditions), (2) we consider representations for learning that do not explicitly use joint angles to represent the arm's pose, and (3) we use a constrained nonlinear function approximator with few parameters.

Because our experiments consider two link arms, we found, through experimentation, that explicit representation of joint angles were unnecessary to achieve reasonable performance so all compared representations omit explicit joint angle information.

Although there are other studies that use polar coordinates to represent the position of the target (e.g. [12]) we are unaware of any other work that also uses polar coordinates to represent the position of the hand.

Finally, many works have considered generalization to novel target points while controlling an arm (e.g. [7], [10], [13]). However, we believe this is the first work to directly consider extrapolation to target points outside of the original training set (i.e. where interpolation is not possible).

III. Methods

We describe the reaching task within a reinforcement learning framework. The learning algorithm applied is Q-learning [4] with a neural network function approximator that estimates the long term cost of taking a particular action given the current state. The agent receives information about the environment and then selects an action. Whenever the agent acts it receives updated information about the state of the environment and the critic assigns a cost to the agent's previous action.

The function approximator used was linear in its parameters but accounted for interactions between the input variables. For example, a two variable input vector $\vec{x} = [x_1, x_2]^T$ would first be transformed to the vector $\vec{x'} = \left[x_1, x_2, (x_1)^2, x_1x_2, (x_2)^2\right]^T$, then $\vec{x'}$ is used as input to a

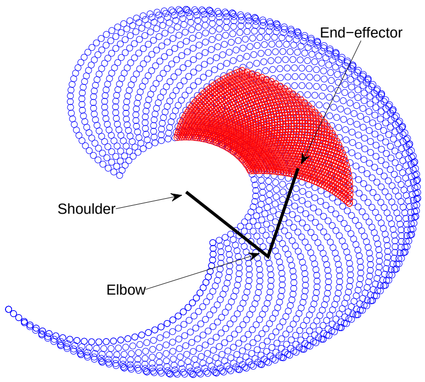


Fig. 2. Valid prenatal condition targets (red) and postnatal condition targets (blue).

standard linear perceptron. This style of approximator has few tunable parameters and was found to have low training complexity (i.e. the number of training samples required to learn a task) when compared with multi-layer perceptrons and radial basis function networks.

The simulated arm is constructed from two links of equal length (figure 2). The total length of the simulated arm (i.e. sum of the two links) was chosen to be 10cm because that is about the average length of a newborn infant arm (10.5cm for females and 10.7cm for males [14]). The first link connects to the origin of the environment by a hinge joint representing the shoulder, while the second link connects to the end of the first link by a hinge joint representing the elbow (see figure 2). In the underlying simulation an arm pose is represented by two joint angles $\theta = (\theta_S, \theta_E)^T$ where θ_S and θ_E correspond to the angle of the shoulder and elbow, respectively. The Cartesian coordinate for the hand was computed by

$$\begin{pmatrix} x_H \\ y_H \end{pmatrix} = \begin{pmatrix} l_1 \cos(\theta_S) + l_2 \cos(\theta_S + \theta_E) \\ l_1 \sin(\theta_S) + l_2 \sin(\theta_S + \theta_E) \end{pmatrix} \quad (1)$$

where $l_1 (= 5\text{cm})$ and $l_2 (= 5\text{cm})$ are the length from the shoulder to the elbow and the elbow to the hand, respectively. The scalar values x_H and y_H are the first and second components of a Cartesian coordinate.

Actions, representing the change in joint angles, are represented by a vector $\dot{\theta} = (\dot{\theta_S}, \dot{\theta_E})^T \in \{-2.86°, 0°, 2.86°\} \times \{-2.86°, 0°, 2.86°\}$ (a total of nine actions). When an action is applied to the arm, for each joint $i \in \{S, E\}$

$$\theta_i = \begin{cases} \lambda_i & \lambda_i > \theta_i + \dot{\theta}_i \\ \theta_i + \dot{\theta}_i & \lambda_i \leq \theta_i + \dot{\theta}_i \leq \Lambda_i \\ \Lambda_i & \theta_i + \dot{\theta}_i > \Lambda_i \end{cases} \quad (2)$$

where λ_i and Λ_i are the lower and upper joint limits (respectively) and $\dot{\theta}_i$ is the desired change in joint angle. The arm's joint angles are adjusted by the angles specified in the action vector unless the sum of the action vector and previous joint angles is outside of the arm's joint limits. If any of the resulting angles fall outside of the arm's limits, the angles are clipped to the closest valid angle.

Under postnatal conditions, joint angle limits were set to reflect the limits of an adult human arm. Shoulder limits were set from $-140°$ to $90°$ and elbow limits were set from $0°$ to $145°$ [15]. This provides the arm with a large amount of freedom (see the blue region in figure 2).

Prenatal conditions were simulated by restricting the joint angle limits of the simulated arm. Shoulder limits were set from $-50°$ to $25°$ and elbow limits were set from $90°$ to $145°$. This restricted the region of valid targets. See figure 2 to see the region of valid targets for both the prenatal and postnatal conditions.

The previous few paragraphs describe the model underlying the simulation. Sensory representations, on the other hand, can take many different forms. In this study, a world-centered sensory representation (WC), relative world-centered sensory representation (RWC), agent-centered sensory representation (AC) and relative agent-centered sensory representation (RAC) were compared (see figure 3).

WC describes the state of the environment by a 4-dimensional vector (x_T, y_T, x_H, y_H) composed of two Cartesian coordinates: target point (x_T, y_T) and hand position (x_H, y_H). Notice that the origin of the coordinate system can be arbitrarily assigned to any point in the world. The reinforcement signal (i.e. immediate cost) used during training was

$$R(s_t, a, s_{t+1}) = (x_T^{s_{t+1}} - x_H^{s_{t+1}})^2 + (y_T^{s_{t+1}} - y_H^{s_{t+1}})^2 \quad (3)$$

where $s_{t+1} = (x_T^{s_{t+1}}, y_T^{s_{t+1}}, x_H^{s_{t+1}}, y_H^{s_{t+1}})$ was the state transitioned to from state s_t after selecting action a.

RWC is similar to WC except that the state description is a 2-dimensional vector $(x_{(T-H)}, y_{(T-H)}) = (x_T - x_H, y_T - y_H)$ describing only the difference between the target and hand represented as Cartesian coordinates.

AC (like WC) describes the state of the environment by a 4-dimensional vector $(\varphi_T, d_T, \varphi_H, d_H)$. However, unlike WC, it is composed of two polar coordinates: target position with angle φ_T and distance d_T and hand position with angle φ_H and distance d_H. Notice that unlike WC or RWC, the origin must be placed at the shoulder. In other words, the world is represented from an agent-centered perspective. The reinforcement signal (i.e. immediate cost) used during training was

$$R(s_t, a, s_{t+1}) = (\varphi_T^{s_{t+1}} - \varphi_H^{s_{t+1}})^2 + (d_T^{s_{t+1}} - d_H^{s_{t+1}})^2 \quad (4)$$

where $s_{t+1} = (\varphi_T^{s_{t+1}}, d_T^{s_{t+1}}, \varphi_H^{s_{t+1}}, d_H^{s_{t+1}})$ was the state transitioned from state s_t after selecting action a.

RAC is similar to AC except that the state description is a 2-dimensional vector $(\varphi_{(T-H)}, d_{(T-H)}) = (\varphi_T - \varphi_H, d_T - d_H)$ describing only the difference between the target and hand represented in polar coordinates.

Learned controllers were evaluated based on the distance of the arm's hand from the target after the last step of an episode averaged over 100 episodes. A low value (e.g. 1cm) indicates that the controller is able to solve the inverse kinematics

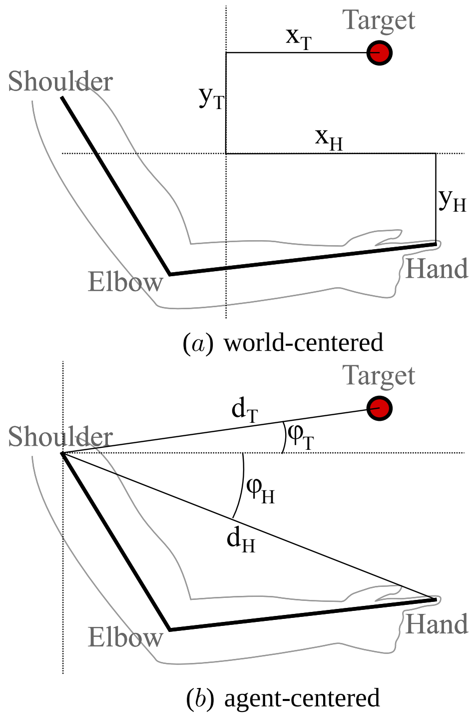


Fig. 3. Two possible sensory representations for the target reaching task. For the world-centered sensory representation (WC) (a) the relationship between the hand and target is expressed by the difference between two Cartesian points: one for the hand and one for the target. The agent-centered sensory representation (AC) (b) relates the hand and target by the difference between two polar coordinates: one for the hand and one for the target.

problem, while high values indicate that the controller is unable to either reach the target or maintain the hand at the target once it has been reached. Due to the granularity of the actions and ϵ-probability of selecting a random action all controllers have at least small error.

Each controller was trained for 5,000 episodes and each episode gave the learning system 100 chances to act with a small probability ($\epsilon = 0.1$) of selecting an action at random. The learning rate used for training the function approximator was set to $\alpha = 0.1$, and the discount factor, a scalar value in the interval $[0, 1)$, required by the Q-learning algorithm, which signifies the amount of emphasis placed on future costs, was set to $\gamma = 0.9$.

IV. Experimental Results

The experimental results presented are for 30 controllers under world-centered sensory representation (WC), relative world-centered sensory representation (RWC), agent-centered sensory representation (AC), and relative agent-centered sensory representation (RAC) (120 controllers total) *trained under the prenatal condition only*.

The error achieved by controllers engaging with the four different representations (i.e. WC, RWC, AC, RAC) appears comparable in the prenatal condition after training (see figure 4a). All controllers, regardless of sensory representation, achieve lower than 1cm average final distance error. This implies that all treatments are able to learn the task. However,

TABLE I
Summary of results from two-sample t-tests on error (smaller is better)

Prenatal Condition		Postnatal Condition	
Alt. Hypothesis	P-Val	Alt. Hypothesis	P-Val
WC > AC	0.0034832	WC > AC	1.1073e-39 ≈ 0
RWC < AC	2.2275e-08 ≈ 0	RWC > AC	6.2491e-34 ≈ 0
RAC < AC	2.7034e-20 ≈ 0	RAC < AC	4.6438e-33 ≈ 0

table I shows that there are statistically significant differences between the treatments under the prenatal condition. For example, treatment AC performs better than treatment WC (i.e. lower error) but not better than RWC, and RAC has the lowest prenatal error.

However, under the postnatal condition, controllers using WC or RWC had significantly higher error than either AC or RAC (see figure 4b). Table I confirms that the results are statistically significant. Again, under postnatal conditions treatment RAC had the lowest error.

The fact that treatment AC has higher error than RWC in the prenatal condition but lower error in the postnatal condition rules out the possibility that there exists an error threshold in the prenatal condition that guarantees low error in the postnatal condition. Thus AC does not have lower error than WC in the postnatal condition just because it had lower error than WC in the prenatal condition.

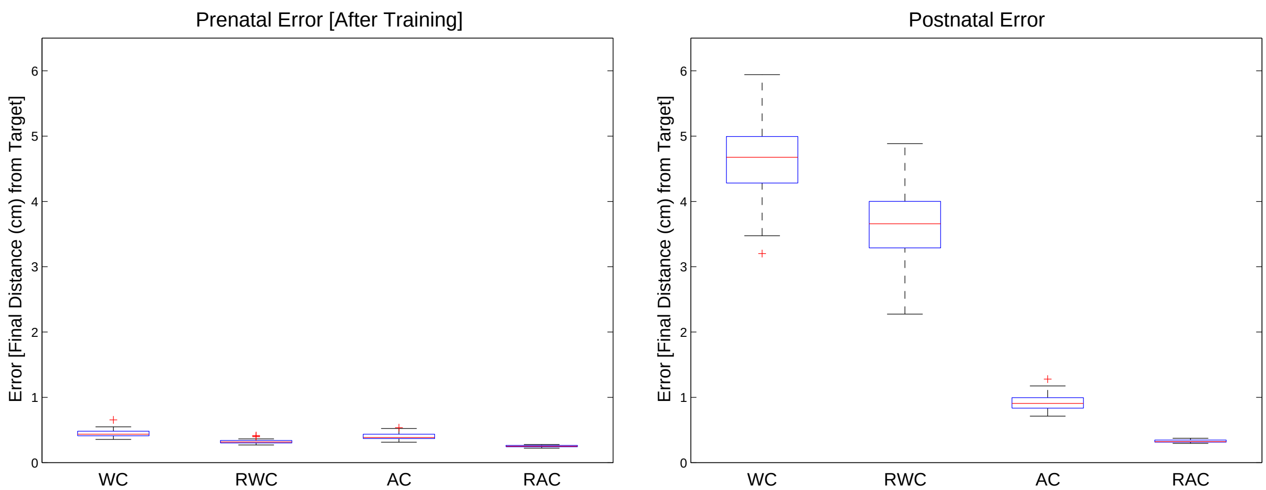
Figure 5 shows two episodes for controllers trained using all four representations. Hand trajectories, for all treatments, are similar when reaching for prenatal valid targets. However, agents using world-centered representations are unable to reach some postnatal valid targets and their trajectories are more complex than the trajectories generated by agents using agent-centered representations.

V. Analysis

To determine why controllers using agent-centered representations achieve lower error than the world-centered representations we performed a second experiment. The change in joint angles $\dot{\theta}$, change in RAC state variables, $\varphi_{(T-H)}$, $d_{(T-H)}$, and change in RWC state variables, $x_{(T-H)}$, $y_{(T-H)}$ were recorded for 100 steps while actions $\dot{\theta}$ were selected according to a uniform random distribution from the action set described above.

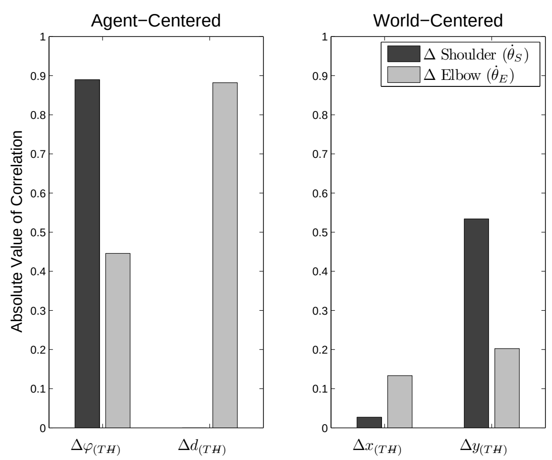
The sample linear correlation between two random variables can be used to determine how well one variable can be predicted by the other. We looked at the absolute value of the linear correlation between the change in shoulder and elbow joint angles and the change in sensory description variables used by relative world-centered and relative agent-centered sensory representations (see figure 6).

The results show that for agent-centered representations, change in shoulder angle is highly predictive of change in the difference between the target and hand polar coordinate angles. Change in elbow is highly predictive of change in the difference between the target and hand polar coordinate distances in agent-centered representations. Changes in shoulder angle and



(a) Prenatal (b) Postnatal

Fig. 4. Error (final distance between hand and target in centimeters) of controllers using world-centered sensory representation (WC), relative world-centered sensory representation (RWC), agent-centered sensory representation (AC), and relative agent-centered sensory representation (RAC) under prenatal (a) and postnatal (b) conditions. The agent-centered representations show better performance than the world-centered representations under postnatal performance, while prenatal performance is similar for all representations.



(a) (b)

Fig. 6. Absolute value of the correlation between the change in shoulder angle ($\dot{\theta}_S$) and change in elbow angle ($\dot{\theta}_E$) with change in sensory variables that measure difference between hand and target. (a) Notice that $\dot{\theta}_S$ is highly predictive of $\Delta\varphi_{(T-H)}$, and $\dot{\theta}_E$ is highly predictive of $\Delta d_{(T-H)}$, variables used in agent-centered sensory representation, (b) while variables used in world-centered sensory representation (last two) are more difficult to predict.

elbow angle, however, are not highly predictive of the changes in sensory variables used by world-centered representations.

VI. Discussion

Figure 6 suggests that the action encodings simplify predicting the change in state caused by an action when the agent is using agent-centered representations, but the action encodings are not helpful for predicting change in state when the agent is using world-centered representations. This shows that both sensory representation and action encoding are important choices for learning systems that need to learn sensory-motor skills that scale to novel situations.

Learning sensory-motor skills under one set of conditions and applying them to another is an attractive idea (in addition to providing a head start on learning sensory-motor skills) as it can potentially be leveraged to reduce risk in dangerous situations. An agent can learn to behave in dangerous situations without practicing in a dangerous environment.

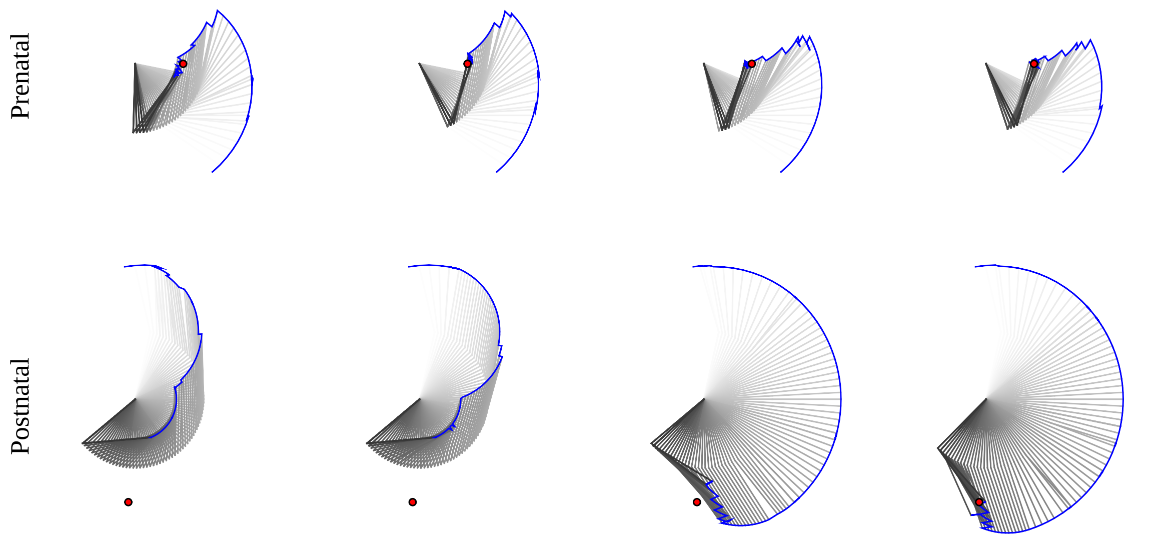
A. Potential Objections

One possible objection to this work is that infants learn to reach after they are born. However, fetuses in the womb have been observed making reaching-like movements such as moving fingers to the lips [2]. Once a baby is born, the neonate must contend with greater force of gravity for which its muscles are too weak and it must learn a dynamics model of the arm. Learning to reach while in the womb may alleviate some of the difficulty of learning an accurate dynamics model.

Another potential issue is that it is unknown how constrained the movements of a developing fetus are. The values chosen for the experiment were selected to place a great deal of restriction on prenatal movement. If there is less restriction during development the extrapolation problem is easier because the training set (i.e. prenatal condition) is more representative of the postnatal condition. Thus the results of this study should still be valid.

B. Future Work

Figure 4b shows that RAC scales significantly better than AC (see table I). Gradient-descent does not find the global optimum. Dimension reduction might be useful for successful transfer of prenatal sensory-motor skills to postnatal life. Future work will investigate what kinds of dimension reduction techniques improve performance under postnatal conditions after training only under prenatal conditions. Preliminary results (unpublished data) show that Principal Component



(a) world-centered (b) reduced world-centered (c) agent-centered (d) reduced agent-centered

Fig. 5. Examples of reaching for a target (red disc) under identical initial conditions with controllers trained using (a) WC, (b) RWC, (c) AC, and (d) RAC. (*Top*) All controllers are able to accurately reach for targets that would be reachable under prenatal conditions and the generated paths are similar. (*Bottom*) Paths generated by WC and RWC are complex, while agent-centered controllers produce smooth paths. RAC is able to reach accurately even when the other controllers fail.

Analysis (PCA) and other common unsupervised dimension reduction techniques are ineffective at transforming AC into RAC because they preserve irrelevant structure (with respect to the task). We will investigate supervised dimension reduction techniques such as Partial Least Squares (suggested by [9]).

VII. Conclusion

The main contribution of this work is demonstrating how sensory-motor skills learned before birth can transfer to postnatal life, despite the fact that the prenatal training conditions are quite different from the postnatal testing conditions. The difficulty of extrapolation can be avoided by selecting a representation that is complementary to the action encodings simplifying prediction of the sensory consequences of actions.

The results presented in this paper suggest that sensory representation that complements the agent's action encodings may play a key role in transferring sensory-motor skills learned during gestation to postnatal life.

References

[1] S. R. Robinson and G. A. Kleven, *Prenatal Development of Postnatal Functions*. Praeger Publishers, 2005, ch. Learning to Move Before Birth, pp. 131–175.

[2] H. B. Valman and J. F. Pearson, "What the fetus feels," *British Medical Journal*, vol. 280, pp. 233–234, 1980.

[3] Y. Kuniyoshi, "Emergence of embodied behavior: Coupled chaos system and fetal motor development," in *ALIFE X Workshop Proceedings*, 2006, pp. 195–196.

[4] C. Watkins, "Learning from delayed rewards," Ph.D. dissertation, University of Cambridge, 1989.

[5] D. Bullock, S. Grossberg, and F. H. Guenther, "A self-organizing neural model of motor equivalent reaching and tool use by a multijoint arm," *Journal of Cognitive Neuroscience*, vol. 54, pp. 408–435, 1993.

[6] J. W. Krakauer, M.-F. Ghilardi, and C. Ghez, "Independent learning of internal models for kinematic and dynamic control of reaching," *Nature Neuroscience*, vol. 2, no. 11, pp. 1026 – 1031, November 1999.

[7] A. Fagg, N. Sitkoff, A. Barto, and J. Houk, "Cerebellar learning for control of a two-link arm in muscle space," in *Robotics and Automation, 1997. Proceedings., 1997 IEEE International Conference on*, vol. 3, Apr 1997, pp. 2638–2644 vol.3.

[8] T. D'Silva and R. Mikkulainen, "Learning dynamic obstacle avoidance for a robot arm using neuroevolution," *Neural Processing Letters*, vol. 30, pp. 59–69, 2009.

[9] A. D'Souza, S. Vijayakumar, and S. Schaal, "Learning inverse kinematics," in *Proc. IEEE/RSJ International Conference on Intelligent Robots and Systems*, vol. 1, Oct. 29–Nov. 3, 2001, pp. 298–303.

[10] H. Kambara, K. Kim, D. Shin, M. Sato, and Y. Koike, "Learning and generation of goal-directed arm reaching from scratch," *Neural Networks*, vol. 22, pp. 348–361, 2009.

[11] M. Tamosiunaite, T. Asfour, and F. Wörgötter, "Learning to reach by reinforcement learning using a receptive field based function approximation approach with continuous actions," *Biological Cybernetics*, vol. 100, pp. 249–260, 2009.

[12] M. Hülse, S. McBride, and M. Lee, "Robotic hand-eye coordination without global reference: A biologically inspired learning scheme," in *International Conference on Development and Learning*, 2009.

[13] M. Rolf, J. J. Steil, and M. Gienger, "Efficient exploration and learning of whole body kinematics," in *International Conference on Development and Learning*, 2009.

[14] R. L. Copper, R. L. Goldenberg, S. P. Cliver, M. B. DuBard, H. J. Hoffmann, and R. O. Davis, "Anthropometric assessment of body size differences of full-term male and female infants," *Obstetrics and Gynecology*, vol. 81, pp. 161–164, 1993.

[15] V. Grecu, N. Dumitru, and L. Grecu, "Analysis of human arm joints and extension of the study to robot manipulator," in *Proceedings of the International MultiConference of Engineers and Computer Scientists*, March 2009.