

SEDRo: A Simulated Environment for Developmental Robotics

Aishwarya Pothula, Md Ashaduzzaman Rubel Mondol, Sanath Narasimhan, Sm Mazharul Islam, Deokgun Park

Computer Science and Engineering

University of Texas at Arlington

Arlington, Texas USA

{aishwarya.pothula, mdashaduzzaman.mondol, sanath.narasimhan, sxi7321}@mavs.uta.edu, deokgun.park@uta.edu

Abstract—Even with impressive advances in application specific models, we still lack knowledge about how to build a model that can learn in a human-like way and do multiple tasks. To learn in a human-like way, we need to provide a diverse experience that is comparable to human's. In this paper, we introduce our ongoing effort to build a simulated environment for developmental robotics (SEDRo). SEDRo provides diverse human experiences ranging from those of a fetus to a 12th month old. A series of simulated tests based on developmental psychology will be used to evaluate the progress of a learning model. We anticipate SEDRo to lower the cost of entry and facilitate research in the developmental robotics community.

Index Terms—Baby robots, Sensorimotor development, Embodiment

I. INTRODUCTION

Imagine a robot that can work as a butler. It can handle many tasks and talk with other butler robots to do even more tasks. Alas, one cannot buy or build one today even with an unlimited budget. The reason a butler robot is not available is because we do not know how to program it. Current approaches require huge data to teach a single skill [1], and the data requirement grows exponentially with the number of tasks. While we have made remarkable progress in solving tasks with well-defined structures such as when explicit rewards or ground truth exist, we do not know how we can generalize this capability for a single task to multiple tasks. Turing suggested [2]:

Instead of trying to produce a programme to simulate the adult mind, why not rather try to produce one which simulates the child's?

Humans are born with a vast blank memory and a mechanism for filling it. Let us call this mechanism *the learning mechanism* in this paper. With diverse experiences as input, the mechanism fills the contents of the memory as shown in Fig. 1. After a few years, we can do many things in multiple domains such as perception, motor, social, language, and physics. We claim that there were the following issues in previous approaches that made the search for the learning mechanism difficult and propose a new approach to mitigate those issues.

- **Targeting a single skill rather than diverse skills (the main issue)** – While a human child can learn to do many things simultaneously, we have focused largely on developing models that can do only a single task. This

approach had resulted in overfitted solutions that cannot be generalized to diverse tasks.

- **Use of refined and focused datasets rather than diverse and noisy datasets (the first common pattern)** – Because the focus is to teach one skill, we tend to build a refined dataset or an environment that contains only task relevant information. This resulted in *spoon-fed human-edited sensory data* [3]. Compare this with how humans learn from unstructured data such as visual and auditory senses and find underlying structures and apply these structures to many domains [4].
- **Relying on explicit rewards rather than on other mechanisms (the second common pattern)** – While operant conditioning is a powerful mechanism [5], we tend to rely on explicit rewards to guide learning. Designing a reward mechanism might be easy for a single task. However, it becomes exponentially difficult as the number of target tasks increase. If we compare the language acquisition abilities of humans and robots, robots can learn to navigate according to the verbal instruction quickly [6]–[8] but do not know how to generalize this to other tasks such as cooking. On the contrary, human infants cannot follow verbal instruction for a very long time. Unfortunately, you cannot give a treat to an 8-month infant for toilet training when he goes to a bathroom himself. But slowly around two years when they acquire language, they can do many tasks with it [4]. One key difference is that while robots are trained using explicit

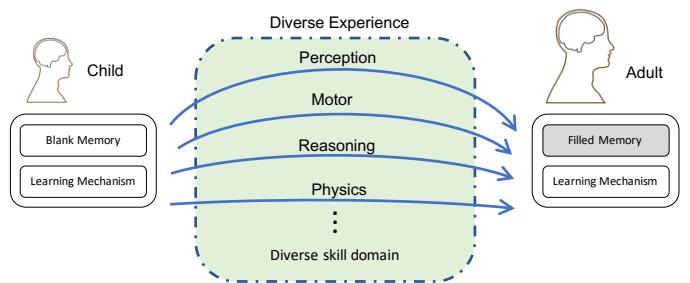


Fig. 1. A child begins with vast blank memory and a learning mechanism. The learning mechanism uses diverse experiences to fill memory. There can be many different skill domains, such as perception, motor, reasoning, and so on.

rewards, it is not the case with infants.

- **Too many necessary components rather than a sufficient set of the learning mechanism (the third common pattern)** – Finally, we tend to find individual necessary mechanisms rather than suggesting a set of the sufficient mechanisms. The learning mechanism is a system of multiple components. Some might classify the components into two different categories: 1) innate or built-in mechanisms versus 2) universal principles that drive learning. Examples of innate mechanisms are reflexes, hippocampus, or limbic systems. Universal principles explain the driving force behind learning and can be usually written as succinct mathematical formulation such as intrinsic motivation [9], [10], Bayesian statistics [11], or the free energy principles [12]. As we can see, there are many candidate components, and we anticipate that the learning mechanism will be a set of multiple components. However, for a single application, a single or small subset of these components might do the job. The problem is that we cannot linearly concatenate the solutions from multiple domains because they are not independent. Therefore, a more critical but neglected question is what is a sufficient set of components for all problems humans can solve.

As a summary, we tend to build models for single tasks resulting in overfitted solutions that cannot be generalized to multiple tasks. In this perspective, we need a regularization. Regularization by sharing is an effective pattern as demonstrated in convolutional neural network (CNN) or recurrent neural network (RNN) [13]. We claim that we need to regularize by enforcing the use of the same learning mechanism to conduct multiple tasks as Allen Newell suggested in his unified theories of cognition [14].

Then why has the focus of past researches been on developing models for individual tasks? Imagine that a researcher has decided to build an agent that can perform many tasks like a human can. The first problem she encounters is that there is no simulated environment that can provide the diverse experiences required to acquire skills across multiple domains.

To solve this problem, we introduce our ongoing effort to build a Simulated Environment for Developmental Robotics (SEDRo). SEDRo provides diverse experiences similar to that of human infants from the stage of a fetus to 12 months of age. SEDRo also simulates developmental psychology experiments to evaluate the progress of intelligence development in multiple domains.

There are two generalizable lessons in our work. First, we point that the learning environment should provide experiences for the multiple tasks and provide a proof-of-concept example. Fig. 2 shows screenshots of SEDRo. In our environment, the learning agent has to rely on interactions with other characters such as a mother character, who teaches language as a human mother does. Other characters have to intelligently react to the random babbling of the baby in a diverse but reasonable way. Programming a mother character for all situations is intractable and it becomes increasingly challenging to provide

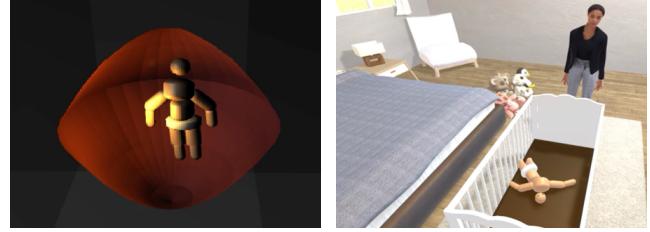


Fig. 2. Screenshots of SEDRo environment. The environment simulates a fetus in the womb(Left). An infant (learning agent) in the crib with a mother character (right).

an experience for open-ended learning when social learning is involved. In our paper, we address this issue by focusing on the earlier stage of development from the stage of a fetus to 12 months of age when a few words are acquired. It is more tractable as the conversations between the mother and the baby tends to be one-directional rather than interactive back-and-forth conversations.

Second generalizable lesson is that we can build upon the prior researches in the developmental psychology to evaluate the developmental progress of non-verbal artificial agent. Because our environment cannot provide sufficient language exposure beyond the first 12 months, the agent cannot acquire advanced language beyond the first few words. Consequently, we cannot evaluate the developmental progress of the agent based on their ability to follow verbal instruction or answer questions correctly. We overcome this challenge by using studies from developmental psychology. There are many experiments revealing developmental milestones for non-verbal infants. We can simulate and make use of those experiments in SEDRo for developmental assessments. As a concrete example, Kellman and Spelke found that babies acquire *perceptual completion* around four months using the habituation-dishabituation paradigm [15]. With SEDRo, models can be computationally evaluated by simulating and running experiments to compare behaviors of the agent to the intellectual progress of human infants. Fig. 3 explains these experiments in more detail and shows screenshots of our simulated environment.

The rest of this paper is arranged in the following manner. In section II, we survey related works which cover different types of simulated environments for developing AI and various evaluation methods for non-verbal agents. Then, in section III, we illustrate our proposed environment, SEDRo. Finally, we draw the conclusion in section IV by pointing out some major limitations of the current version of SEDRo, along with a future plan of actions to resolve these issues.

II. BACKGROUND

We review previous literature for 1) simulated environments for artificial agents and 2) evaluation methods for non-verbal agents.

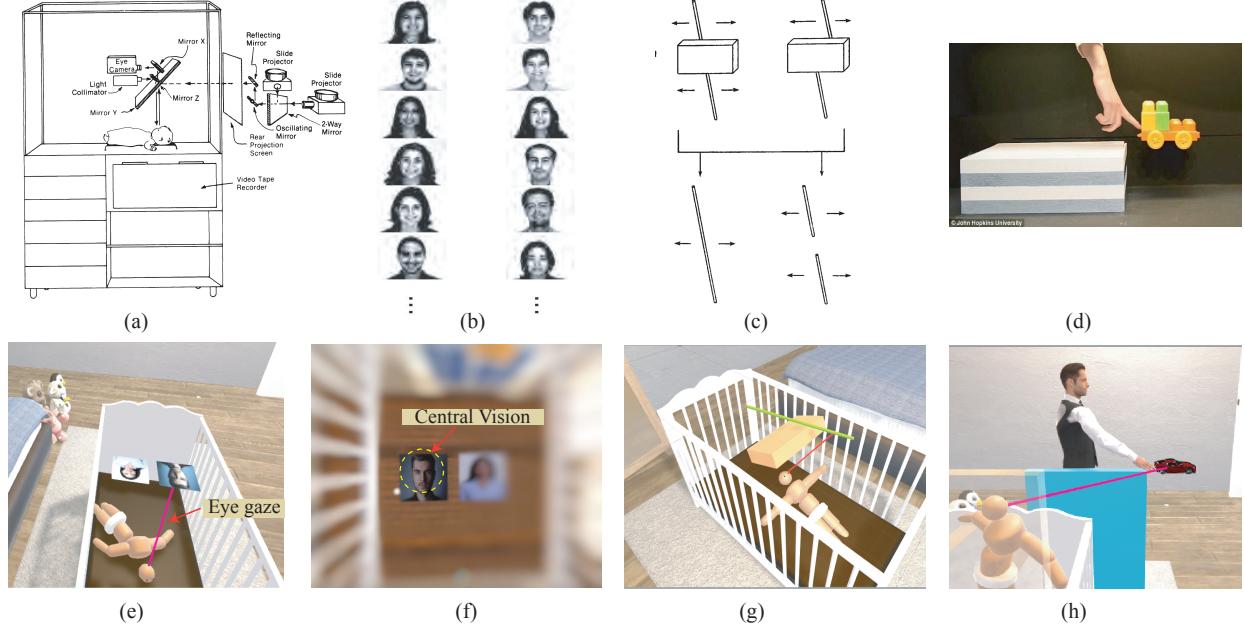


Fig. 3. Example evaluation methods for non-verbal infants and simulated experiments for the artificial agent. (a) An experiment set up to examine visual pattern according to the stimulus [16]. A baby views a series of visual stimuli at two or more locations. Visual patterns, such as looking time and looking preference, of the baby are then analyzed. Depending on the developmental stage, they tend to attend more at novel things [17]. (b) For example, we can examine if the baby can differentiate between male and female faces [18]. (c) When newborn infants, under three months of age, see a rod moving behind a box, they will perceive it as two rods. However, babies, past four months of age, perceive it as a single rod and will be surprised when they are shown two rods [19]. (e) In our simulated environment, we model eye gaze and central vision. (f) The focused area gives a clear view, while peripheral vision gives a blurred image. (g) Simulated object unity-perception task (h) Test for innate physics.

A. Simulated Environments for AI

Several environments have been developed for AI research and especially for reinforcement learning research [20], [21]. The overarching goal was to provide a common benchmark and to lower the barriers of the entry for researchers. Examples include environments in which agents get rewards by following verbal instructions in navigation [6]–[8], [22], [23] and give correct answers (question answering) [24]. Though we've made substantial progress in reinforcement learning with explicit rewards, it is difficult to transfer these built models to develop artificial general intelligence (AGI). Many previous works were conducted to overcome this limitation. The difficulty in transferring is mainly because humans do not depend on explicit rewards nor labeled data to learn [25], [26]. Principles such as intrinsic motivation [9] and free energy [27] have been proposed to be the underlying mechanism for learning [11], [28]–[31]. A number of simulated environments have been proposed to test these hypotheses in the robotics context. We can classify previous environments into artificial environments and human-inspired environments.

a) Artificial Environments: Oudeyer et al. proposed a mathematical formulation for intrinsic motivation and demonstrated similar observations using both simulation and robots [9]. Similarly, Haber et al. showed that an agent begins by exploring an environment, and then in the later stages begins to interact with objects [32] based on intrinsic rewards.

These works were conducted on 3D simulated environments built using game physics engines. However, previous works usually focussed on developing and testing a single component of the mechanism, such as self-other cognition, imitation, and joint attention [30]. While it is relatively easy to build artificial environments to test a single component, it is difficult to extend the environments for multiple tasks. To overcome this limitation, human-inspired environments were also studied.

b) Human-inspired Environment: For Human-Inspired Environments, there is a benefit in using infant-like environments. We can be rest assured that those experiences are enough for the development of human-level intelligence. Meltzoff et al. elaborated this idea with evidence from developmental psychology, neuroscience, and machine learning [33]. The idea of using human-like experience to nurture AI has been actively pursued in the *Developmental Robotics (DevRob)* or *Epigenetic Robotics* community [34]–[36]. However, it is challenging to simulate the real world. Therefore, researchers used 1) physical robots in the real world, or 2) simulated environment of a simplified real-world, focusing narrow skills. Weng et al. developed SAIL robots that explored the world with humans to themselves acquire skills by in navigation and object perception [3]. Later iCub [37], a humanoid robot that is modelled after human babies was developed and used for developmental robotics research. Using physical robots, studies on perception and physical behaviors with objects

can be conducted [38]–[41]. However, physical robots are expensive and providing same experiences for the reproducible research is an open problem. To lower the cost of entry for research in robotics, many simulators were developed [42], [43]. Environment simulations of human development are modelled after stages as early as the fetus [44], [45] as it is evidenced that fetuses learn auditory [46] and sensorimotor coordination. To tackle the challenge of simulating natural interactions with human users, Murane et al. used virtual reality to allow humans to interact with the robots in the simulation [47], [48]. Using this method, data for human-robot interaction can be accumulated.

B. Evaluation methods for non-verbal agents

There are many tests for human-level intelligence, including the Turing test, robot college student test, kitchen tests, and AI preschool test [49]. However, most tests require a capability for language and cannot be used for evaluating progressive intelligence in diverse domains.

a) *Tests in Developmental psychology:* Researchers in developmental psychology developed various evaluation schemes using behavior patterns related to familiarity and novelty. These includes visual expectation paradigm [16], preferential looking [50], habituation-dishabituation paradigm [51], contingent change of the rate in pacifier-sucking behaviors [46]. For instance, the visual expectations paradigm means that babies look longer and attend more to novel scenes than to familiar scenes. Using these methods, developmental milestones in many skill domains such as visual [52], auditory [53], motor [54], social [55], [56], language [53], physics [15] etc have been studied.

b) *Psychology-inspired Test for AI:* There are previous researches that use human psychological metrics for the evaluation of artificial agents. For example, Leibo et al. used human psychology paradigms such as visual search, change detection, and random dot motion discrimination [57]. However, it tests adult level psychological perception and does not provide developmental milestones. It is also limited to the visual perception domain and does not provide an integrated experience required to learn and perform diverse tasks. Piloto et al. suggested the evaluation of physics concepts that are inspired by developmental psychology [58]. They developed a dataset by examining object persistence, unchangeableness, continuity, solidity, and containment by violation of expectations (VOE) methods. The study of complete and diverse tasks at the human level is challenging. Crosby et al. used various intellectual animal behaviors in the simulated environment [59]. Tests for ten cognitive categories and a playground that can provide an experience to learn those skills are provided in the work. SEDRo builds upon their work to extend those approaches to human-level intelligence.

III. SIMULATING EXPERIENCE OF HUMAN INFANTS

In this section, we discuss about the proposed Simulated Environment for Developmental Robotics or SEDRo. Fig. 4

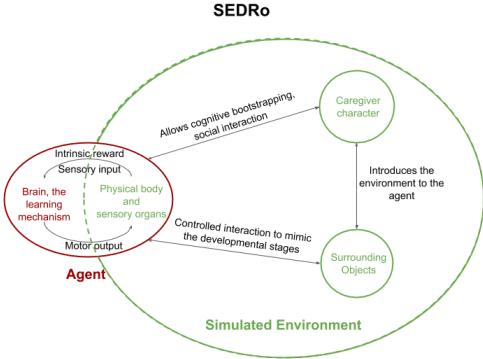


Fig. 4. Ecosystem of SEDRo environment.

illustrates the primary components of SEDRo and their inter-relations. The two main components in SEDRo are the learning agent (with red border), the simulated environment (with green border). Within the simulated environment, there are a caregiver character, surrounding objects in the environment (e.g. toys, cribs, walls etc.) and most importantly the body of the agent. The agent will interact with the simulated environment by controlling the muscles in its body according to the sensor signals. Interaction between the agent and the caregiver allows cognitive bootstrapping and social-learning, while interactions between the agent and the surrounding objects are increased gradually as the agent gets into more developed stages. The caregiver character can also interact with the surrounding objects to introduce them to the agent at the earlier stages of development.

Though there are no rewards that are explicitly awarded by the environment, it does not mean that the reward mechanism does not play a role in the learning. Rather than relying on the environment for the rewards, the responsibility of generating rewards belong to the agent itself. As an example, if an agent can get food from its environment, this input will be given to the agents as a number representing the amount of food in its stomach. It is now the agent's role to generate a negative reward if there is no food in its stomach and positive rewards if new food is given. In this sense, we can say, the body itself is a part of the environment, and what is referred to as the agent is only the brain, which is why the agent's body in Fig. 4 is in green font.

SEDRo provides the diverse experiences of human infants from the stage of a the fetus to first 12 months of life. A new-born brain must learn to control its body. We can compare this with trying to learn to operate a machine using a control panel of 1,000 by 1,000 LEDs and 1,000 by 1,000 buttons. To make this even more challenging, LEDs and buttons are not labeled as shown in Fig. 5. Each LED blinks, maybe sparsely. If you push some buttons, the blinking pattern of the LEDs seems to change and sometimes not; is not easy to track. You need to make sense out of this huge matrix of buttons and LEDs that Piaget called sensorimotor stage [60]. The role of a brain model is to compose an output behavior vector given a sensor

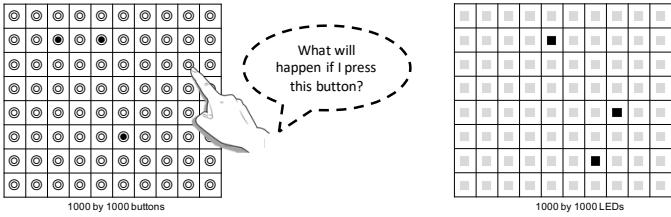


Fig. 5. This diagram illustrates our assumption about the learning mechanism. Pressing some buttons in 1,000 by 1,000 button control panel will affect the flickering patterns in 1,000 by 1,000 LEDs panel. The button panel represents the motor output vector, and the LED panel represents the sensory input vector. Please note that there are no labels on those two vectors and the learning mechanism needs to learn how to operate the body.

vector.

A. Curriculum for Development

To make the learning easier, human infants develop in a curriculum which scaffolds the involved sensory and motor capabilities. [61]–[64]. For example, in the fetus stage there are no visual inputs. A small subset of LEDs and buttons that are available at that stage can be isolated to master new skills such as sucking a thumb or rotating body. In the first three months, babies are very near-sighted and do not have any mobility, which makes many visual signals stationary. At later stages, when babies learn to sit and grasp, they develop alternative strategies of learning using a rotating viewpoints and the contingent verbal speech of caregivers.

In SEDRo, the input output signal changes according to the development of the agent. For example, the agent in the womb stage will not have any visual input signals which will be available after birth. But for the first 3 months, visual signals will represent nearsightedness. Muscles will develop over time. Full force at the early stage will not be enough for an agent to crawl or stand, but it will steadily increase to afford walking in the later stages.

B. Specification for I/O Vectors

The sensory input consists of touch, vision, acceleration, gravity, and proprioceptors. Visual attention plays an important role in the evaluation of non-verbal human infants as explained in Sec. III-D. In SEDRo, the artificial infant can control eye movement with three parameters - vertical angle, horizontal angle, and focal length. To simulate central and peripheral vision with two eyes, four images are generated as central and peripheral vision images for left and right eyes each. Central vision has an 8° field of view (FOV) and a higher resolution, while the peripheral vision has a 100° FOV and a lower resolution. One additional image is provided for the purpose of debugging that will represent reconstructed visual imagery in the brain. Five cameras with different settings have been used in the game engine for implementation.

The touch sensors are spread all over the body but the distribution pattern varies. The face, lips, and hands have a higher density of touch sensors than the torso. In the current

version, there are 2,110 touch sensors. Each body part is segmented into meshes and a touch sensor provides a binary feature which represents if a contact has been made. We have implemented touch feature using collision information provided by the game engine.

The motor output vectors constitute muscle torques, which will determine the 53 motors, including the 9 degree of freedom (DOF) in each hand inspired by iCub [37]. Main loop of the environment runs 100 steps per second motivated by the human biological brain [65]. At each step, the agent will read a sensory input vector and write a motor output vector.

C. Social Event Scenario

Social interaction plays an important role for human development. In SEDRo, we are building scenarios for social interaction according to the following process:

- 1) We start by choosing a meaningful interaction pattern by reviewing developmental psychology literature.
- 2) We write a scenario for the chosen interaction.
- 3) Actors perform that scenario and we capture their behaviors using a motion capture facility.
- 4) We add the recorded scenario into SEDRo along with a schedule for the event.

Building a library of the social event is time consuming and we anticipate the SEDRo environment will expand over the coming years. We will maintain the versioning of SEDRo such that the research using SEDRo can be reproducible.

D. Evaluation Framework for Non-verbal Agents

We have developed an evaluation framework for the development of skills in multiple domains by simulating established experiments from developmental psychology. There are multiple developmental milestones in multiple skill domains. At each stage, there are key milestones that the agent needs to

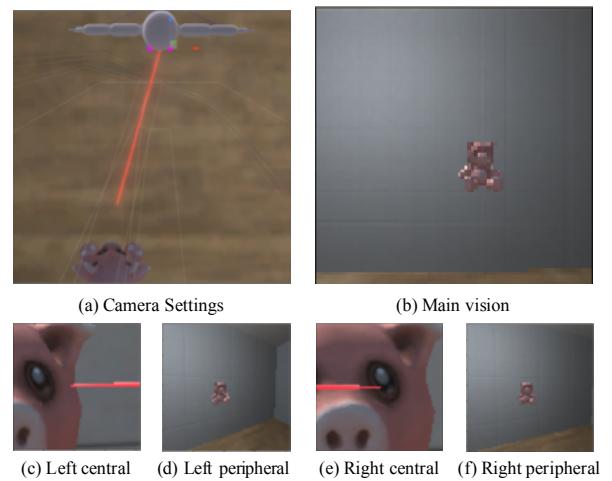


Fig. 6. The visual system of the agents in SEDRo. The orange laser beam in (a) shows visual attention. Each eye has a central (c and e) and a peripheral vision (d and f). For debugging purpose, a main view is provided as shown in (b).

TABLE I
SUMMARY OF DEVELOPMENTAL MILESTONES (M REPRESENTS MONTHS AFTER BIRTH)

Stage	Fetus Stage	Immobile Stage (Less than 3 Months)	Crawling Stage (4-10 Months)	Walking Stage (11-18 Months)
Description	No vision	Near sighted vision.	Fully developed vision. Sit and interact with objects. Interact with other persons by babbling.	Fully developed muscles. First words
Vision		Visual expectation(0 vs 3M) [16], [66]–[68], face preference(1 vs 2M) [55], [69]–[71], face preference(2 days) [52], [72], gender detection(0 vs 3M) [73], [74], depth perception (0 vs 2M) [75]	Visual scan pattern (2 vs 11 weeks) [76], tracking occluded objects(4 vs 6 Months) [77], lost ability to distinguish faces of different gender(3 vs 9 M) [74]	Novelty preference inversion (6-12 months) [78], [79]
Joint attention		left/right attention manipulation	Gaze angle detection, fixation of first salient object	Mutual gaze through eye contact [80], Fixation of any salient object, declarative pointing, drawing attention
Motor	Hand/face contacts (11 gestation weeks) [45]	Open hand grasping [81], [82]	Recognizing own motion vs others(3 vs 5 Months) [83], [84]	Partial integration of visual and motor skill (9 Months) [85]
Language		Differentiate mother tongue and foreign language [46], marginal babbling	canonical babbling	intentional gestures, single words, word-gesture combination [86]
Reasoning		Self-perception at mirror(3 Months) [56]	Fear of heights (after crawling) [87], [88], Allocentric spatial frame of reference (9 Months) [89]–[91]	Mark test(15 Months) [92], adapted use of hook(12 Months) [93]

satisfy. Researchers may choose to replay relevant experiences if the agent does not achieve those milestones. Consequently, the agent will experience an adaptive experience based on its current capability rather than experiences based on a fixed time schedule. Fig. 3 shows example tasks in the developmental psychology and screenshots of preliminary prototypes simulating those experiments.

In current version, we developed a visual expectation paradigm experiment with a moving rod. The visual attention pattern over the moving rod can be acquired as a separate channel in the gym interface. Table 1 summarizes our plan for evaluation experiments in domains such as vision, motor, attention, and reasoning. Each evaluation has a different expected behavior pattern between two stages of human development. For example, two month old infants cannot predict regular pattern, but at 3.5 months, infants exhibit anticipatory eye movement 200 ms before the actual pattern visual expectation [16], [66]–[68]. We leverage such known developmental milestones to develop suites of simulated experiments for evaluating the development of the artificial agent. The evaluation will conduct multiple experiments and compare the results with those of the human participants.

E. Implementation Detail

We use Unity3D 2018.4 for the development of the environment. Unity ML agent [94] is used to implement Open AI gym interface [20]. To record behaviors of the actors, we use Motive Body software with Opti-track motion capture system with ten Prime 17W cameras.

IV. DISCUSSION

As SEDRo is a work in progress, here we discuss its limitations, a few alternative approaches and future works.

A. Limitations

A major limitation of our work is the lack of back and forth interactive conversation between the caregivers and the infant agent. Currently, only two types of conversations are supported;

- Caregiver initiated conversation that will be played according to a pre-determined schedule, and
- Contingent response that will be played conditioned on infant agent behaviors such as cooing or touching toys.

Despite building diverse scenarios for these two conversation types is a challenge by itself due to the sheer number

of the required diverse experience, they would not provide enough experience for the development of language acquisition beyond the first year level. One potential approach to overcome this limitation would be to ask humans to interact with the artificial agent using a virtual reality technique [47], [48]. Another option would be to use a physical embodied robot and ask humans to take care of it. We claim that SEDRo can be used to test cognitive architectures before the need to perform physical robot experiments, thereby helping in reducing the number of candidate architectures for the expensive physical robot experiments.

In SEDRo, we simulate the human infant experiences, but an alternative is to use a completely artificial environment that is not relevant to human experience but still requires skills in many domains. For example, emergent communication behaviors were observed in the reinforcement learning environment with multiple agents [95]–[98]. Through similar researches, though we might find clues about the underlying human learning mechanism, it might be challenging to apply them to human robot interaction because language is a set of arbitrary symbols shared between members [99].

Another possibility is to transform existing resources into an open-ended learning environment. Using YouTube videos to create a diverse experience is an example. However, Smith and Slone pointed out that these kinds of approaches use shallow information about a lot of things, whereas, on the contrary, human infants begin by learning a lot about a few things [100]. In addition to that, visual information from the first years of human life constitutes an egocentric view of the world. The allocentric view emerges only later, after 12 months of age. Furthermore, humans rely heavily on social interactions to learn. While infants can learn a language by being tutored by an instructor, they cannot learn by seeing a recorded video of an same tutoring [53]. Therefore we think that certain necessary skills have to be acquired before learning from those resources becomes feasible.

B. Conclusion and Future Work

We are building SEDRo, an environment that simulates the early experiences of a human from the stage of a fetus to 12 months of age. The open-ended and unsupervised nature of the environment requires agents to avoid fitting to specific tasks. To evaluate the development of intelligent behaviors of non-verbal artificial agents, a set of experiments in developmental psychology will be simulated in SEDRo. We expect researchers in the AI and robotics community to discover the learning mechanism for artificial general intelligence by testing different cognitive architectures using the open-ended learning environment developed in our project.

REFERENCES

- [1] B. M. Lake, T. D. Ullman, J. B. Tenenbaum, and S. J. Gershman, “Building machines that learn and think like people,” *Behavioral and brain sciences*, vol. 40, 2017.
- [2] A. Turing, “Computing machinery and intelligence,” *Mind*, vol. 59, pp. 433–460, 1950.
- [3] J. Weng, J. McClelland, A. Pentland, O. Sporns, I. Stockman, M. Sur, and E. Thelen, “Autonomous mental development by robots and animals,” *Science*, vol. 291, no. 5504, pp. 599–600, 2001.
- [4] A. Gopnik, A. N. Meltzoff, and P. K. Kuhl, *The scientist in the crib: Minds, brains, and how children learn*. William Morrow & Co, 1999.
- [5] B. Skinner, “The behavior of organisms: an experimental analysis.” 1938.
- [6] K. M. Hermann, F. Hill, S. Green, F. Wang, R. Faulkner, H. Soyer, D. Szepesvari, W. M. Czarnecki, M. Jaderberg, D. Teplyashin *et al.*, “Grounded language learning in a simulated 3d world,” *arXiv preprint arXiv:1706.06551*, 2017.
- [7] D. S. Chaplot, K. M. Sathyendra, R. K. Pasumarthi, D. Rajagopal, and R. Salakhutdinov, “Gated-attention architectures for task-oriented language grounding,” in *Thirty-Second AAAI Conference on Artificial Intelligence*, 2018.
- [8] H. Chen, A. Suhr, D. Misra, N. Snavely, and Y. Artzi, “Touchdown: Natural language navigation and spatial reasoning in visual street environments,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2019, pp. 12538–12547.
- [9] P.-Y. Oudeyer, F. Kaplan, and V. V. Hafner, “Intrinsic motivation systems for autonomous mental development,” *IEEE transactions on evolutionary computation*, vol. 11, no. 2, pp. 265–286, 2007.
- [10] J. Schmidhuber, “Formal theory of creativity, fun, and intrinsic motivation (1990–2010),” *IEEE Transactions on Autonomous Mental Development*, vol. 2, no. 3, pp. 230–247, 2010.
- [11] A. Gopnik, C. Glymour, D. M. Sobel, L. E. Schulz, T. Kushnir, and D. Danks, “A theory of causal learning in children: causal maps and bayes nets.” *Psychological review*, vol. 111, no. 1, p. 3, 2004.
- [12] K. Friston, “The free-energy principle: a unified brain theory?” *Nature reviews neuroscience*, vol. 11, no. 2, pp. 127–138, 2010.
- [13] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, “Dropout: a simple way to prevent neural networks from overfitting,” *The journal of machine learning research*, vol. 15, no. 1, pp. 1929–1958, 2014.
- [14] A. Newell, *Unified theories of cognition*. Harvard University Press, 1994.
- [15] P. J. Kellman and E. S. Spelke, “Perception of partly occluded objects in infancy,” *Cognitive psychology*, vol. 15, no. 4, pp. 483–524, 1983.
- [16] M. M. Haith, C. Hazan, and G. S. Goodman, “Expectation and anticipation of dynamic visual events by 3.5-month-old babies,” *Child development*, pp. 467–479, 1988.
- [17] R. O. Gilmore and H. Thomas, “Examining individual differences in infants’ habituation patterns using objective quantitative techniques,” *Infant Behavior and Development*, vol. 25, no. 4, pp. 399–412, 2002.
- [18] W. R. Charlesworth, “The role of surprise in cognitive development,” *Studies in cognitive development*, pp. 257–314, 1969.
- [19] A. E. Stahl and L. Feigenson, “Observing the unexpected enhances infants’ learning and exploration,” *Science*, vol. 348, no. 6230, pp. 91–94, 2015.
- [20] G. Brockman, V. Cheung, L. Pettersson, J. Schneider, J. Schulman, J. Tang, and W. Zaremba, “Openai gym,” *arXiv preprint arXiv:1606.01540*, 2016.
- [21] C. Beattie, J. Z. Leibo, D. Teplyashin, T. Ward, M. Wainwright, H. Küttler, A. Lefrancq, S. Green, V. Valdés, A. Sadik *et al.*, “Deepmind lab,” *arXiv preprint arXiv:1612.03801*, 2016.
- [22] M. Savva, A. Kadian, O. Maksymets, Y. Zhao, E. Wijmans, B. Jain, J. Straub, J. Liu, V. Koltun, J. Malik *et al.*, “Habitat: A platform for embodied ai research,” in *Proceedings of the IEEE International Conference on Computer Vision*, 2019, pp. 9339–9347.
- [23] M. Shridhar, J. Thomason, D. Gordon, Y. Bisk, W. Han, R. Mottaghi, L. Zettlemoyer, and D. Fox, “Alfred: A benchmark for interpreting grounded instructions for everyday tasks,” *arXiv preprint arXiv:1912.01734*, 2019.
- [24] A. Das, S. Datta, G. Gkioxari, S. Lee, D. Parikh, and D. Batra, “Embodied question answering,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2018, pp. 2054–2063.
- [25] C. L. Hull, *Principles of behavior*. Appleton-century-crofts New York, 1943, vol. 422.
- [26] R. W. White, “Motivation reconsidered: The concept of competence.” *Psychological review*, vol. 66, no. 5, p. 297, 1959.
- [27] K. Friston and S. Kiebel, “Predictive coding under the free-energy principle,” *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 364, no. 1521, pp. 1211–1221, 2009.

- [28] E. Bonawitz, S. Denison, T. L. Griffiths, and A. Gopnik, "Probabilistic models, learning algorithms, and response variability: sampling in cognitive development," *Trends in cognitive sciences*, vol. 18, no. 10, pp. 497–500, 2014.
- [29] J. Hawkins and S. Ahmad, "Why neurons have thousands of synapses, a theory of sequence memory in neocortex," *Frontiers in neural circuits*, vol. 10, p. 23, 2016.
- [30] Y. Nagai, "Predictive learning: its key role in early cognitive development," *Philosophical Transactions of the Royal Society B*, vol. 374, no. 1771, p. 20180030, 2019.
- [31] T. Schmidt and D. Fox, "Self-directed lifelong learning for robot vision," in *Robotics Research*. Springer, 2020, pp. 109–114.
- [32] N. Haber, D. Mrowca, S. Wang, L. F. Fei-Fei, and D. L. Yamins, "Learning to play with intrinsically-motivated, self-aware agents," in *Advances in Neural Information Processing Systems*, 2018, pp. 8388–8399.
- [33] A. N. Meltzoff, P. K. Kuhl, J. Movellan, and T. J. Sejnowski, "Foundations for a new science of learning," *science*, vol. 325, no. 5938, pp. 284–288, 2009.
- [34] M. Lungarella, G. Metta, R. Pfeifer, and G. Sandini, "Developmental robotics: a survey," *Connection science*, vol. 15, no. 4, pp. 151–190, 2003.
- [35] M. Asada, K. Hosoda, Y. Kuniyoshi, H. Ishiguro, T. Inui, Y. Yoshikawa, M. Ogino, and C. Yoshida, "Cognitive developmental robotics: A survey," *IEEE transactions on autonomous mental development*, vol. 1, no. 1, pp. 12–34, 2009.
- [36] A. Cangelosi and M. Schlesinger, *Developmental robotics: From babies to robots*. MIT press, 2015.
- [37] G. Metta, G. Sandini, D. Vernon, L. Natale, and F. Nori, "The icub humanoid robot: an open platform for research in embodied cognition," in *Proceedings of the 8th workshop on performance metrics for intelligent systems*, 2008, pp. 50–56.
- [38] J. Ruesch, M. Lopes, A. Bernardino, J. Hornstein, J. Santos-Victor, and R. Pfeifer, "Multimodal saliency-based bottom-up attention a framework for the humanoid robot icub," in *2008 IEEE International Conference on Robotics and Automation*. IEEE, 2008, pp. 962–967.
- [39] I. Gaudiello, E. Zibetti, S. Lefort, M. Chetouani, and S. Ivaldi, "Trust as indicator of robot functional and social acceptance. an experimental study on user conformation to icub answers," *Computers in Human Behavior*, vol. 61, pp. 633–655, 2016.
- [40] D. Marocco, A. Cangelosi, K. Fischer, and T. Belpaeme, "Grounding action words in the sensorimotor interaction with the world: experiments with a simulated icub humanoid robot," *Frontiers in neurorobotics*, vol. 4, p. 7, 2010.
- [41] B. Serhan and A. Cangelosi, "Replication of infant behaviours with a babybot: Early pointing gesture comprehension," in *2019 Joint IEEE 9th International Conference on Development and Learning and Epigenetic Robotics (ICDL-EpiRob)*. IEEE, 2019, pp. 157–162.
- [42] N. Koenig and A. Howard, "Design and use paradigms for gazebo, an open-source multi-robot simulator," in *2004 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)/IEEE Cat. No. 04CH37566*, vol. 3. IEEE, 2004, pp. 2149–2154.
- [43] O. Michel, "Webots: Professional mobile robot simulation," *Journal of Advanced Robotics Systems*, vol. 1, no. 1, pp. 39–42, 2004. [Online]. Available: <http://www.ars-journal.com/International-Journal-of-Advanced-Robotic-Systems/Volume-1/39-42.pdf>
- [44] Y. Kuniyoshi and S. Sangawa, "Early motor development from partially ordered neural-body dynamics: experiments with a cortico-spinal-musculo-skeletal model," *Biological cybernetics*, vol. 95, no. 6, p. 589, 2006.
- [45] H. Mori and Y. Kuniyoshi, "A human fetus development simulation: Self-organization of behaviors through tactile sensation," in *2010 IEEE 9th International Conference on Development and Learning*. IEEE, 2010, pp. 82–87.
- [46] C. Moon, H. Lagercrantz, and P. K. Kuhl, "Language experienced in utero affects vowel perception after birth: A two-country study," *Acta Paediatrica*, vol. 102, no. 2, pp. 156–160, 2013.
- [47] M. Murnane, M. Breitmeyer, C. Matuszek, and D. EnQel, "Virtual reality and photogrammetry for improved reproducibility of human-robot interaction studies," in *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, 2019, pp. 1092–1093.
- [48] M. Murnane, M. Breitmeyer, F. Ferraro, C. Matuszek, and D. Engel, "Learning from human-robot interactions in modeled scenes," in *ACM SIGGRAPH 2019 Posters*, 2019, pp. 1–2.
- [49] S. Adams, I. Arel, J. Bach, R. Coop, R. Furlan, B. Goertzel, J. S. Hall, A. Samsonovich, M. Scheutz, M. Schlesinger *et al.*, "Mapping the landscape of human-level artificial general intelligence," *AI magazine*, vol. 33, no. 1, pp. 25–42, 2012.
- [50] R. L. Fantz, "A method for studying early visual development," *Perceptual and motor skills*, vol. 6, no. 1, pp. 13–15, 1956.
- [51] P. S. Kaplan and J. S. Werner, "Habituation, response to novelty, and dishabituation in human infants: Tests of a dual-process theory of visual attention," *Journal of experimental child psychology*, vol. 42, no. 2, pp. 199–217, 1986.
- [52] I. Bushnell, "Mother's face recognition in newborn infants: Learning and memory," *Infant and Child Development: An International Journal of Research and Practice*, vol. 10, no. 1-2, pp. 67–74, 2001.
- [53] P. K. Kuhl, "Is speech learning 'gated' by the social brain?" *Developmental science*, vol. 10, no. 1, pp. 110–120, 2007.
- [54] R. K. Clifton, D. W. Muir, D. H. Ashmead, and M. G. Clarkson, "Is visually guided reaching in early infancy a myth?" *Child development*, vol. 64, no. 4, pp. 1099–1110, 1993.
- [55] D. Maurer and M. Barrera, "Infants' perception of natural and distorted arrangements of a schematic face," *Child development*, pp. 196–202, 1981.
- [56] M. L. Courage and M. L. Howe, "From infant to child: The dynamics of cognitive change in the second year of life," *Psychological bulletin*, vol. 128, no. 2, p. 250, 2002.
- [57] J. Z. Leibo, C. d. M. d'Autume, D. Zoran, D. Amos, C. Beattie, K. Anderson, A. G. Castañeda, M. Sanchez, S. Green, A. Gruslys *et al.*, "Psychlab: a psychology laboratory for deep reinforcement learning agents," *arXiv preprint arXiv:1801.08116*, 2018.
- [58] L. Piloto, A. Weinstein, D. TB, A. Ahuja, M. Mirza, G. Wayne, D. Amos, C.-c. Hung, and M. Botvinick, "Probing physics knowledge using tools from developmental psychology," *arXiv preprint arXiv:1804.01128*, 2018.
- [59] M. Crosby, B. Beyret, and M. Halina, "The animal-ai olympics," *Nature Machine Intelligence*, vol. 1, no. 5, pp. 257–257, 2019.
- [60] J. Piaget and M. Cook, *The origins of intelligence in children*. International Universities Press New York, 1952, vol. 8, no. 5.
- [61] L. B. Smith, S. Jayaraman, E. Clerkin, and C. Yu, "The developing infant creates a curriculum for statistical learning," *Trends in cognitive sciences*, vol. 22, no. 4, pp. 325–336, 2018.
- [62] G. Turkewitz and P. A. Kenny, "Limitations on input as a basis for neural organization and perceptual development: A preliminary theoretical statement," *Developmental Psychobiology: The Journal of the International Society for Developmental Psychobiology*, vol. 15, no. 4, pp. 357–368, 1982.
- [63] D. E. Berlyne, "Conflict, arousal, and curiosity." 1960.
- [64] P. H. Mirvis, "Flow: the psychology of optimal experience," 1991.
- [65] J. Hawkins and S. Blakeslee, *On Intelligence*. USA: Times Books, 2004.
- [66] R. L. Canfield and M. M. Haith, "Young infants' visual expectations for symmetric and asymmetric stimulus sequences," *Developmental Psychology*, vol. 27, no. 2, p. 198, 1991.
- [67] S. A. Adler, M. M. Haith, D. M. Arehart, and E. C. Lanthier, "Infants' visual expectations and the processing of time," *Journal of Cognition and Development*, vol. 9, no. 1, pp. 1–25, 2008.
- [68] N. Wentworth, M. M. Haith, and R. Hood, "Spatiotemporal regularity and interevent contingencies as information for infants' visual expectations," *Infancy*, vol. 3, no. 3, pp. 303–321, 2002.
- [69] D. Maurer and P. Salapatek, "Developmental changes in the scanning of faces by young infants," *Child development*, pp. 523–527, 1976.
- [70] M. d. Haan, O. Pascalis, and M. H. Johnson, "Specialization of neural mechanisms underlying face recognition in human infants," *Journal of cognitive neuroscience*, vol. 14, no. 2, pp. 199–209, 2002.
- [71] J. Morton and M. H. Johnson, "Conspec and conlern: a two-process theory of infant face recognition," *Psychological review*, vol. 98, no. 2, p. 164, 1991.
- [72] I. Bushnell, F. Sai, and J. Mullin, "Neonatal recognition of the mother's face," *British Journal of Developmental Psychology*, vol. 7, no. 1, pp. 3–15, 1989.
- [73] M. D. Leinbach and B. I. Fagot, "Categorical habituation to male and female faces: Gender schematic processing in infancy," *Infant behavior and development*, vol. 16, no. 3, pp. 317–332, 1993.
- [74] P. C. Quinn, J. Yahr, A. Kuhn, A. M. Slater, and O. Pascalis, "Representation of the gender of human faces by infants: A preference for female," *Perception*, vol. 31, no. 9, pp. 1109–1121, 2002.

- [75] J. J. Campos, A. Langer, and A. Krowitz, “Cardiac responses on the visual cliff in prelocomotor human infants,” *Science*, vol. 170, no. 3954, pp. 196–197, 1970.
- [76] G. W. Bronson, “Infant differences in rate of visual encoding,” *Child Development*, vol. 62, no. 1, pp. 44–54, 1991.
- [77] S. P. Johnson, D. Amso, and J. A. Slemmer, “Development of object concepts in infancy: Evidence for early learning in an eye-tracking paradigm,” *Proceedings of the National Academy of Sciences*, vol. 100, no. 18, pp. 10 568–10 573, 2003.
- [78] B. J. Roder, E. W. Bushnell, and A. M. Sasveille, “Infants’ preferences for familiarity and novelty during the course of visual processing,” *Infancy*, vol. 1, no. 4, pp. 491–507, 2000.
- [79] J. Colombo and C. L. Cheatham, “The emergence and basis of endogenous attention in infancy and early childhood.” 2006.
- [80] F. Kaplan and V. V. Hafner, “The challenges of joint attention,” *Interaction Studies*, vol. 7, no. 2, pp. 135–169, 2006.
- [81] C. von Hofsten, “Developmental changes in the organization of pre-reaching movements.” *Developmental psychology*, vol. 20, no. 3, p. 378, 1984.
- [82] B. L. White, P. Castle, and R. Held, “Observations on the development of visually-directed reaching.” *Child development*, pp. 349–364, 1964.
- [83] L. E. Bahrick and J. S. Watson, “Detection of intermodal proprioceptive-visual contingency as a potential basis of self-perception in infancy.” *Developmental Psychology*, vol. 21, no. 6, p. 963, 1985.
- [84] G. Gergely and J. S. Watson, “Early socio-emotional development: Contingency perception and the social-biofeedback model,” *Early social cognition: Understanding others in the first months of life*, vol. 60, pp. 101–136, 1999.
- [85] N. E. Berthier, B. I. Bertenthal, J. D. Seaks, M. R. Sylvia, R. L. Johnson, and R. K. Clifton, “Using object knowledge in visual tracking and reaching,” *Infancy*, vol. 2, no. 2, pp. 257–284, 2001.
- [86] E. Hoff, *Language development*. Cengage Learning, 2013.
- [87] J. J. Campos, B. I. Bertenthal, and R. Kermoian, “Early experience and emotional development: The emergence of wariness of heights,” 1992.
- [88] R. Kermoian and J. J. Campos, “Locomotor experience: A facilitator of spatial cognitive development,” *Child development*, pp. 908–917, 1988.
- [89] L. P. Acredolo, “Development of spatial orientation in infancy.” *Developmental Psychology*, vol. 14, no. 3, p. 224, 1978.
- [90] L. P. Acredolo and D. Evans, “Developmental changes in the effects of landmarks on infant spatial behavior.” *Developmental Psychology*, vol. 16, no. 4, p. 312, 1980.
- [91] L. P. Acredolo, A. Adams, and S. W. Goodwyn, “The role of self-produced movement and visual tracking in infant spatial orientation,” *Journal of Experimental Child Psychology*, vol. 38, no. 2, pp. 312–327, 1984.
- [92] M. Lewis, *Social cognition and the acquisition of self*. Springer Science & Business Media, 2012.
- [93] L. Van Leeuwen, A. Smitsman, and C. van Leeuwen, “Affordances, perceptual complexity, and the development of tool use.” *Journal of Experimental Psychology: Human Perception and Performance*, vol. 20, no. 1, p. 174, 1994.
- [94] A. Juliani, V.-P. Berges, E. Vckay, Y. Gao, H. Henry, M. Mattar, and D. Lange, “Unity: A general platform for intelligent agents,” *arXiv preprint arXiv:1809.02627*, 2018.
- [95] T. Eccles, Y. Bachrach, G. Lever, A. Lazaridou, and T. Graepel, “Biases for emergent communication in multi-agent reinforcement learning,” in *Advances in Neural Information Processing Systems*, 2019, pp. 13 111–13 121.
- [96] K. Cao, A. Lazaridou, M. Lanctot, J. Z. Leibo, K. Tuyls, and S. Clark, “Emergent communication through negotiation,” *arXiv preprint arXiv:1804.03980*, 2018.
- [97] A. Das, T. Gervet, J. Romoff, D. Batra, D. Parikh, M. Rabbat, and J. Pineau, “Tarmac: Targeted multi-agent communication,” *arXiv preprint arXiv:1810.11187*, 2018.
- [98] J. Foerster, I. A. Assael, N. De Freitas, and S. Whiteson, “Learning to communicate with deep multi-agent reinforcement learning.” in *Advances in neural information processing systems*, 2016, pp. 2137–2145.
- [99] S. Kottur, J. M. Moura, S. Lee, and D. Batra, “Natural language does not emerge ‘naturally’ in multi-agent dialog.” *arXiv preprint arXiv:1706.08502*, 2017.
- [100] L. B. Smith and L. K. Slone, “A developmental approach to machine learning?” *Frontiers in psychology*, vol. 8, p. 2124, 2017.