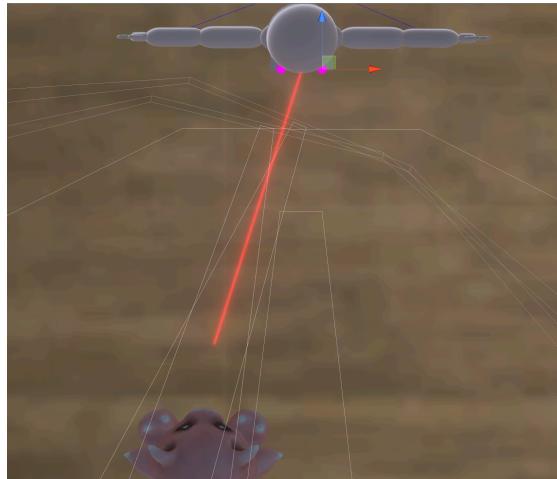


# An Open-World Simulated Environment for Developmental Robotics

SM Mazharul Islam, Md Ashaduzzaman Rubel Mondol, Aishwarya Pothula, Deokgun Park

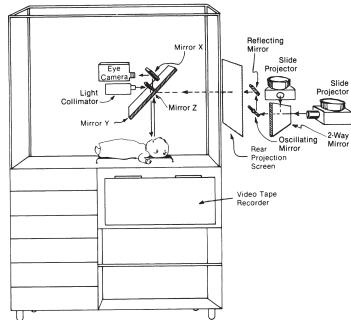
Computer Science and Engineering Department  
University of Texas at Arlington



SEDRo : Simulated Environment for Developmental Robotics

---

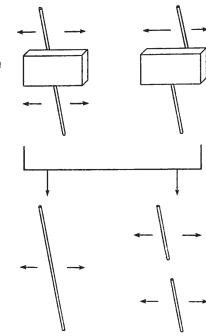
# Evaluation of Non-Verbal Agent



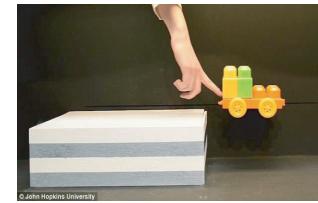
(a)



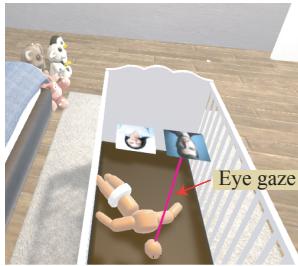
(b)



(c)



(d)



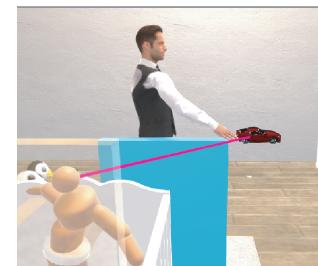
(e)



(f)



(g)



(h)

# Motivation



AlphaStar by DeepMind 2019

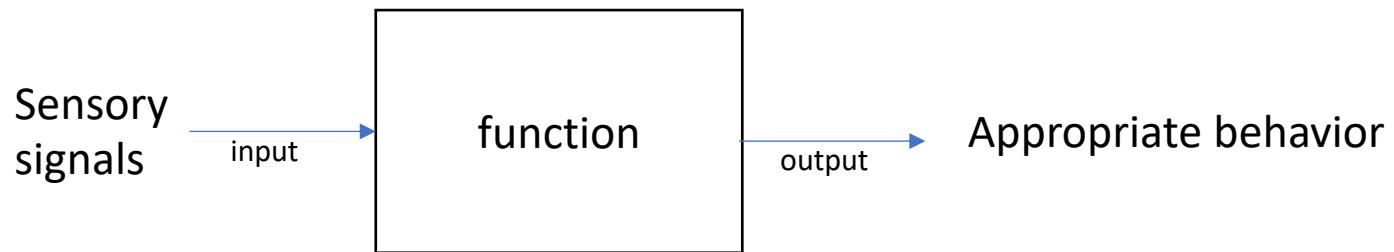
VS



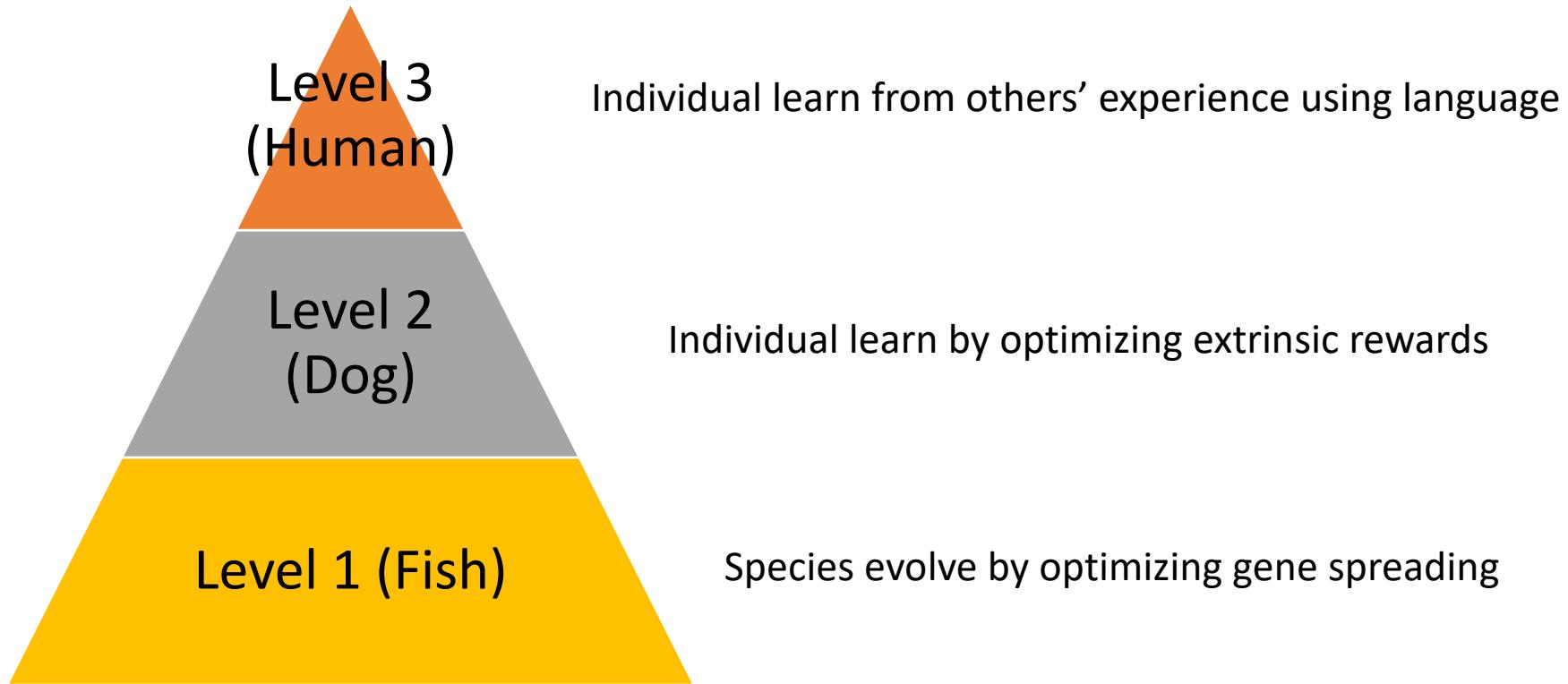
- Targeting a single skill rather than diverse skills
- Use of refined, focused dataset rather than diverse and noisy dataset
- Relying on the explicit reward or ground truth

# Intelligence Simplified

The ability to do appropriate behavior according to the sensory input



# Level of Intelligence



# Park's test for Human level intelligence

- Given the proper **environment**, if the agent can learn **language**, we can say it has a **capability** for human-like artificial intelligence.

Language acquisition = func ( environment, capability)



# Proper environment



Motherese, Infant Directed Speech (IDS)

# Environments for Language Acquisition



Chaplot 2018



CHALET, Yan et al 2018

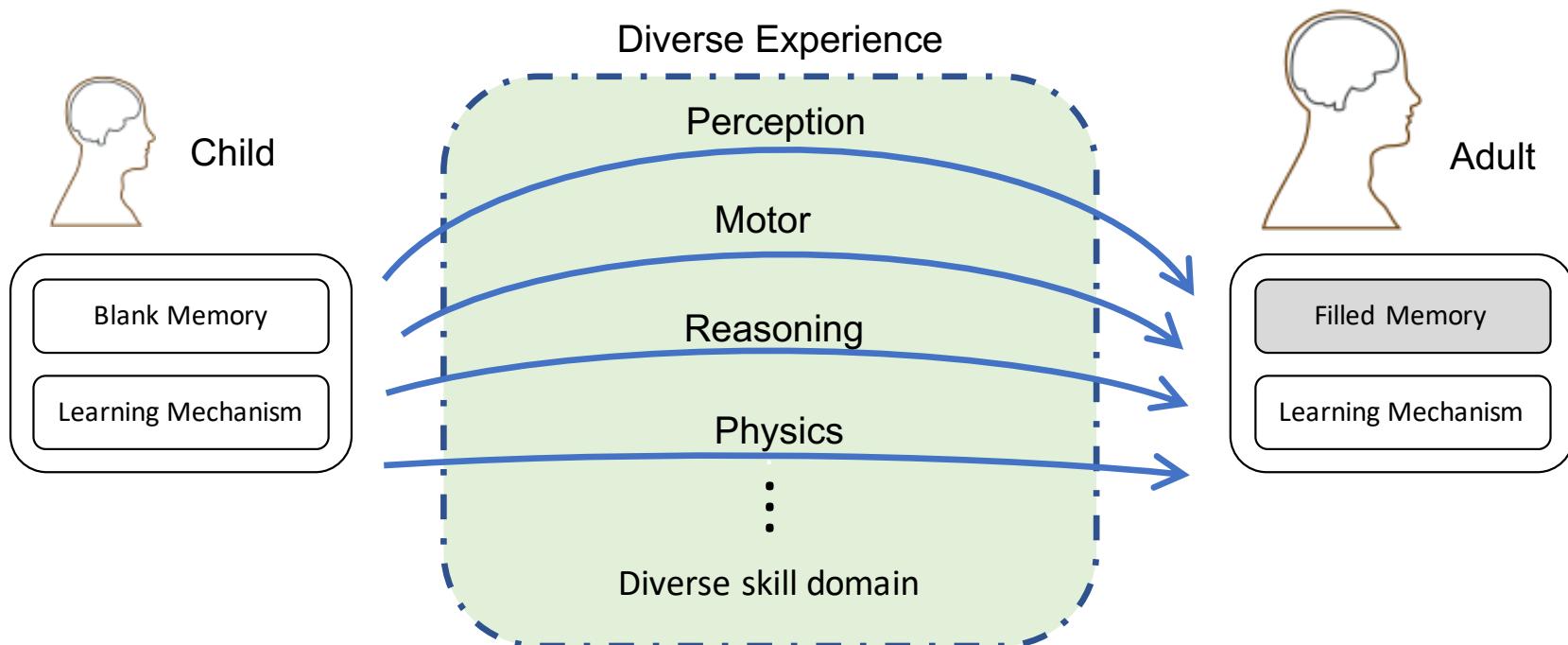
Limitation: Explicit reward and Application Specific Models



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

Alan Turing “Computing machinery and intelligence”, 1950

# Goal



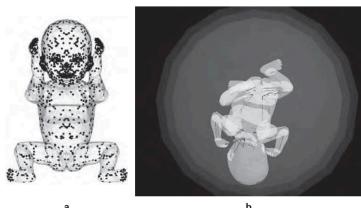
# Two research questions

- How to build an open environment with diverse experience for multi-task learning without explicit rewards?
- How to evaluate the progress of non-verbal agent?

# Environment

- To help progress towards building Artificial General Intelligence, we propose to develop a ‘3D Environment’
- We will develop 4 different environments to represent 4 different learning stages.
- Each environment will test different capabilities of the agent.

Fetus  
stage



1-3 Months  
stage



8-10 Months  
stage



12-18 Months  
stage



# Environment Design

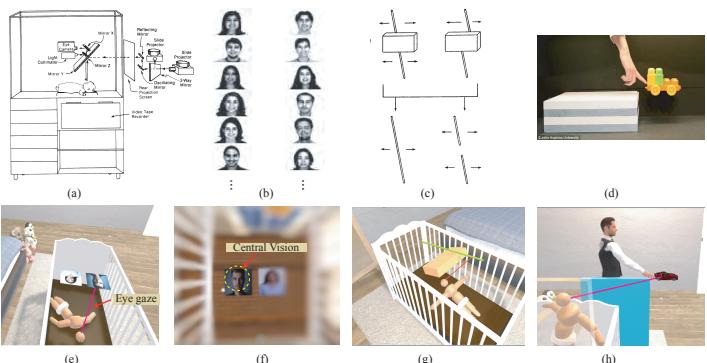
- Womb Environment: For Fetus stage
  - It will be some small spherical area.
  - There will be a minimal (Dim) light.
  - Baby model will be placed inside this area.
- Room Environment : for other 3 stages.
  - It will be a small (e.g. 10m X 10m) room.
  - There will be a crib for the baby.
  - There will be some toy objects.
  - A mother model
  - A child model of different ages.



# Environments' design and targets

	Fetus	1-3 Months	8-10 Months	12-18 Months
Description	Represents when the baby is in the womb, but it is developed enough to move arms, legs.	Represents when the baby is 1-3 months old. In this stage the babies moves arms and legs. They have a limited distance vision and can hear the sound.	Represents when the baby is 8-10 months old. In this stage they have a fully developed vision. They try to crawl. Can interact with objects. They try to interact with mother or other persons by making some sounds.	Represents when the baby is 12-18 months old. At this stage babies start speaking small words. They start speaking first meaningful words.
Goal	<ul style="list-style-type: none"> <li>Can move their arms and legs.</li> <li>Can sense their movement.</li> <li>Learns the sense of touching.</li> </ul>	<ul style="list-style-type: none"> <li>All abilities of previous stage</li> <li>Can move Eye</li> <li>Learn to recognize faces.</li> </ul>	<ul style="list-style-type: none"> <li>All abilities of previous stage</li> <li>Can crawl</li> <li>Moves to a toy and interact with it (Grab/Release)</li> <li>Make random sounds</li> </ul>	<ul style="list-style-type: none"> <li>All abilities of previous stage</li> <li>Can walk</li> <li>Can speak first nouns</li> </ul>
Design	<ul style="list-style-type: none"> <li>Baby is put in dark Womb environment mentioned earlier.</li> </ul>	<ul style="list-style-type: none"> <li>Baby is placed in a crib in the fully illuminated 'Room Environment'.</li> <li>A mother face shows up randomly from a short distance and make some sound/speak</li> </ul>	<ul style="list-style-type: none"> <li>Baby is placed in a crib or in the floor in the fully illuminated 'Room Environment'.</li> <li>A mother model present in the room but not at fixed location</li> <li>Toys spread randomly</li> </ul>	<ul style="list-style-type: none"> <li>Baby is placed in a crib or in the floor in the fully illuminated 'Room Environment'.</li> <li>A mother model present in the room but not at fixed location</li> <li>Toys spread randomly</li> </ul>

# Evaluation



Stage	Fetus Stage (Less than 3 Months)	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		Imitate facial motion (new born) [75], Visual expectation of regular pattern(2 vs 3.5 Months) [51, 70, 71, 72], face preference(new born, 1 vs 2 Months) [76, 77, 57, 78], mother face preference(2 days) [54, 79] ,face gender detection(newborn vs 3 Months) [80, 81], depth perception by visual cliff(newborn vs 2 months) [82]	Visual scan pattern (2 vs 11 weeks) [83], Attending face (1 vs 2 Months) [76], tracking occluded objects(4 vs 6 Months) [84], egocentric vision(9 months)	Novelty preference inversion (6-1 months) [85, 86], 1 ability to distinguish faces of different genders(3 months vs months) [81]
Joint attention	Mutual gaze through eye contact [87]	left/right attention manipulation	Gaze angle detection, fixation of first salient object	Fixation of any salient object, declarative pointing, drawing attention
Motor	Hand/face contacts (11 gestation weeks) [46]	Open hand grasping [88, 89]	Recognizing own motion vs others(3 vs 5 Months) [90, 91]	Partial integration visual and motor skills (9 Months) [92]
Memory	Hand/face contacts (11 gestation weeks) [46]	Mobile paradigm(3 Months) [93]		
Language		Differentiate mother tongue and foreign language [47], marginal babbling	canonical babbling	intentional gestures, single word, word-gesture combination [94]
Reasoning		Self-perception at mirror(3 Months) [58]	Fear of heights (after crawling)[95, 96], Allocentric spatial frame of reference (9 Months) [97, 98, 99]	Mark test(10 Months) [10], adapted use hook(12 Months) [10]

# Conclusion

