# A Definition and a Test for Human-Level Artificial Intelligence

**Md Ashaduzzaman Rubel Mondol** , **Aishwarya Pothula** , **Deokgun Park***
University of Texas at Arlington
{mdashaduzzaman.mondol, aishwarya.pothula}@mav.uta.edu, deokgun.park@uta.edu

## Abstract

Despite recent advances in many application-specific domains, we do not know how to build a human-level artificial intelligence (HLAI). We conjecture that learning from others' experience with the language is the essential characteristic that distinguishes human intelligence from the rest. Humans can update the action-value function with the verbal description as if they experience states, actions, and corresponding rewards sequences firsthand. In this paper, we present a classification of intelligence according to how individual agents learn and propose a definition and a test for HLAI. The main idea is that language acquisition without explicit rewards can be a sufficient test for HLAI.

## 1 Introduction

We made a lot of progress in artificial intelligence (AI). Despite this, the limitation of the current state of the art is most apparent in robotics. When laypersons think about an AI robot, they expect to verbally interact with it to get many services like a human butler. However, we do not know how to program such a robot yet.

In this paper, we try to answer following questions.

- What is the fundamental difference between human intelligence and other animals?
- What does it mean to understand the language?
- How can we test whether an agent has the capability for HLAI?

We also introduce our ongoing effort to build a language acquisition environment. We explain why such an environment is required and how it differs from existing language acquisition environments. Let us begin by explaining what distinguishes the human-level intelligence from the rest.

## 2 Level of Intelligence

Let us start our discussion with the following question:

> Is an earthworm intelligent?

The answer depends on the definition of intelligence. Legg and Hutter proposed the following definition for intelligence after considering more than 70 prior definitions [Legg and Hutter, 2007; Legg *et al.*, 2007].

> Intelligence measures an agent's ability to achieve goals in a wide range of environments.

This definition is universal in the sense that it can be applied to a diverse range of agents such as earthworms, rats, humans, and even computer systems. For biological agents, maximizing gene replication, that is called *inclusive fitness*, is generally accepted as the ultimate goal [Dawkins, 2016]. Earthworms have light receptors and vibration sensors. They move according to those sensors to avoid the sun or moles [Darwin, 1892]. It increases their chance of survival and inclusive fitness [Hamilton, 1964]. Therefore, we can say that earthworms are intelligent.

However, there are differences in intelligence between earthworms and more advanced agents such as rats and humans. In this paper, we propose three levels of intelligence to guide the AI research based on how learning is achieved in agents. Table 1 shows a summary of this idea.

**Level 1 Intelligence** In this categorization, earthworms have Level 1 intelligence, where there is no learning occurring at the individual level. Their central nerve system (CNS) or brains have a hard-coded mapping from sensory input to the corresponding action. This hard-coded function is often called as an instinct and updated with evolution [Tinbergen, 1951]. The problem with this approach is that adaptation is very slow because the update to the neural circuit happens through evolution. For example, if there is an abrupt climate change due to the meteor crash, agents with Level 1 intelligence will have difficulty adapting to the new environment.

**Level 2 Intelligence** The next level in intelligence is individual-level learning. Relying on evolution for new rules is too slow. If an individual agent can learn new rules such as a new type of food, it would increase the probability of successful survival and gene spreading. Agents with Level 2 intelligence can learn new rules during its lifetime, showing higher intelligence than Level 1 intelligence.

To enable learning at the individual level, two functional modules are required. The first is a memory to store newly developed rules. The second module is a reward system to

*Contact Author

judge the merit of the state. We stated that the goal of a biological agent is to spread genes. However, the correct assessment is not possible at the individual agent level. For example, an agent may lay eggs in a hostile environment that no descendant will survive. Still, the agent cannot know this because it would perish long before this happens. Therefore, an agent with level 2 intelligence requires a function to estimate whether the current stimulus or state is good or bad during an agent's life. The reward system serves this purpose by providing a proxy for the value of the state.

We point out that the environment does not provide a reward. Instead, it is an agent that produces a reward signal, which is the agent's estimate of the value of the current state. A dollar bill can be rewarding for some cultures but might not generate any reward for a tribal human who has never seen any money before. As for another example, when we eat three burgers for lunch, the reward for the first and third burger will be different, even though it is the same object for the sake of the environment.

However, this is different from the standard Markov Decision Process (MDP) framework for reinforcement learning, where a reward is determined from the environment. Legg and Hutter used a standard MDP framework for the formal definition of universal intelligence. However, they admitted that a more accurate framework would consist of an agent, an environment, and a separate goal system that interpreted the state of the environment and rewarded the agent appropriately. Another way to resolve this conflict is how we view an agent. An agent might represent whole rats or humans. But for the sake of AI research, we are mostly interested in the subset of the brain where learning occurs. Therefore, we might call this subset as an agent. In that case, an environment might include other parts of the body where learning is not happening, such as the body, sensory organs, a reward system, and the old brain.

**Level 3 Intelligence** While learning with reward is better than using evolution to improve brains, an agent must experience the stimulus to learn from it. However, there is a limitation in learning with direct experience. For example, a rabbit cannot try random action in front of the lion to learn optimal behavior. It would be too late for the rabbit to adjust the action-value function, and this experience cannot be transferred to others. Level 3 intelligence overcome this limitation by learning from others' experiences. A language is a tool for learning from others. Humans' technological achievements were possible because we can learn from others and contribute new knowledge. Isaac Newton said, "If I have seen further, it is by standing on the shoulders of Giants." Language is an invention that enabled this. Therefore, the main feature of level 3 intelligence is learning from other's experiences using language.

Please note that higher level intelligence includes the skills from the lover level intelligence. Natural environment is usually dangerous. Therefore, most of the behaviors in level 2 intelligence agent will be determined from the hard-coded instinct which is the main mechanism for level 1 intelligence. In the previous example, a rabbit will respond to the mere hint of the lion by sound and shape to run away. There is very small area available for learning for individual level 2 agents. Based on this viewpoint, current reinforcement learning can be thought as the mixture of level 1 and level 2 intelligence. For example, Ha and Schmidhuber used evolutionary algorithm to optimize the policy in the context of the reinforcement learning [Ha and Schmidhuber, 2018].

| Level | Features |
|---|---|
| 1 | • No individual learning<br>• Evolution-based refinement<br>• Ex) earthworms |
| 2 | • Learning from direct experience<br>• Reward-based refinement<br>• Ex) rats, dogs |
| 3 | • Learning from indirect experience<br>• Language-based refinement<br>• Ex) humans |

Table 1: Three levels of intelligence

# 3 Clarifying Language Skill

However, we need to clarify what we mean by learning with language. For example, dolphins are known to use a verbal signal to coordinate [Janik and Sayigh, 2013]. Monkeys have been taught sign language [Arbib *et al.*, 2008]. Are dolphins and monkeys level 3 intelligence? Similarly, there have been many previous works that demonstrated various aspects of language skills. Voice agents can understand the spoken language and can answer simple questions [Kepuska and Bohouta, 2018]. Agents have been trained to follow verbal commands to navigate [Hermann *et al.*, 2017; Chaplot *et al.*, 2018; Chen *et al.*, 2019; Das *et al.*, 2018a; Shridhar *et al.*, 2020]. GPT-3 by open AI can generate articles published as Op-Ed in the Guardians [Brown *et al.*, 2020; GPT-3, 2020]. Some models can do multiple tasks in language as evaluated in the GLUE benchmark or DecaNLP [Wang *et al.*, 2018; McCann *et al.*, 2018]. Models exhibit superior performance in all categories than humans except Winograd Schema Challenge [Levesque *et al.*, 2012], where models perform slightly less than humans [Raffel *et al.*, 2020]. Do these models have level 3 intelligence?

Using language has many aspects. In this paper, we claim that learning from others' experiences is the language's essential function. We will explain this with a simple example and then formalize it in the context of reinforcement learning.

**Example 1.** *Let's say that you have never tried Cola before. Now for the first time in your life, you see this dark, sparkling liquid that somehow looks dangerous. You have a few available actions, including drinking and running away. Randomly you might select to drink. It tastes good. It rewards you. Now your action value to the same situation has changed such that you will choose to drink it more deliberately. It is the change induced by direct experience.*

*Learning with language means that it should bring a similar change in your mind when you hear someone say, "Cola is a black, sparkling drink. I drank it, and it tasted good." Figure 2 shows this with the notation in Markov decision pro-*

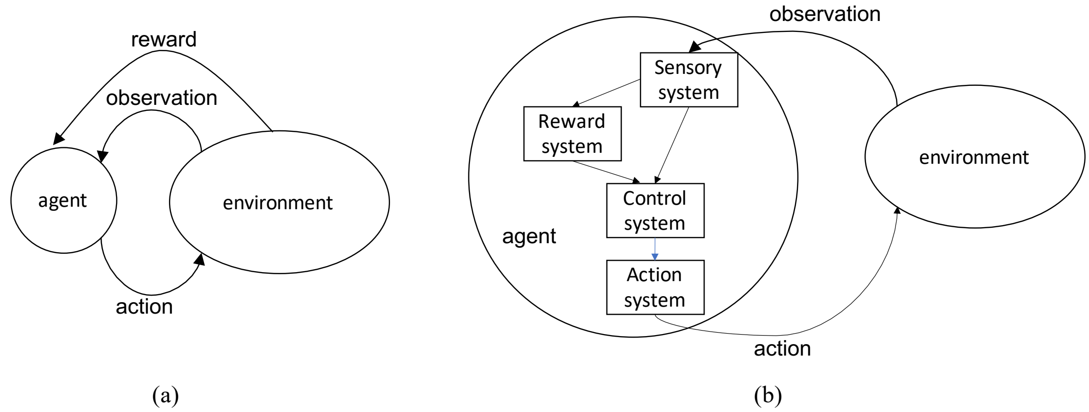


Figure 1: (a) The standard framework for reinforcement learning (b) The revised Relationship of the agent and environment. Environment provides an observation. Some observation is used for the reward system in the agent. The resulting reward signal and the sensory information is fed into the control system.

*cess (MDP) [Sutton* et al.*, 1998].*

Based on this aspect of language use, we can define the HLAI as following;

**Definition 1** (A definition of HLAI)**.** *An agent has human-level artificial intelligence if there exists a symbolic description for every feasible sensory input and agent action sequence, and the agent can update the behavior policy equally whether it receives the symbolic description or it goes through the sensory input and the action sequence itself.*

One challenge with implementing the test of HLAI according to this definition will be how we can make sure that there is a symbolic description for all feasible experience sequences which are not bounded.

## 4 A Test for HLAI

There are many tests for AI, including the Turing test, robot college student test, kitchen test, and AI preschool test [Adams *et al.*, 2012]. For example, Turing test measures if an agent can imitate the human by communicating like one. Robot college student test asks an agent to register, take classes, and to get passing grades by doing assignments and exams. Unfortunately, they are seldom conducted in the current research and when they are conducted, there is a controversy about the validity [Shieber, 1994]. There are a few limitations that make these tests not practical. First, those tests are too difficult. All tests assume that the agent has already acquired the language skill, but we do not know how to program an agent who can learn a language. Second, they require human participants to administer the test. While it takes a few years for humans to be a professional StarCraft II player, it took 200 years of gameplay for machines to masters [Vinyals *et al.*, 2019]. Learning five years of human experience will take a lot of time for training with human intervention. Therefore, using humans is cost-inhibitive and not scalable. Also, interactions with human participants are not reproducible for the validation. Ideally, the test should require the minimum level of intelligence that can pass as human-level intelligence, and it should be cheap to run the test.

To overcome those limitations, we propose a new test for HLAI. If a human infant is raised in an environment such as a jungle where there are no human, he/she cannot acquire language. It is **environment-limited**. Also, if we have animal cubs and try to raise them like a human baby by teaching language, they cannot acquire language. It is **capability-limited**. Therefore, language acquisition is a function of an environment and a capability.

Based on this argument, we propose the Language Acquisition Test for HLAI as the following;

**Theorem 1** (A Language Acquisition Test for HLAI)**.** *Given a proper environment, if an agent without language skills can acquire a subset of the language, the agent has the capability for HLAI.*

*Proof.* Suppose an agent can acquire a new symbolic description that can bring the same change for a certain experience sequence without relying on existing language skills. In that case, the agent can extend the language skills to a novel experience until it finds the symbolic description for any given experience. □

**Example 2.** *A human baby will start learning a single word such as water or mom. When the baby hears these words, they expect similar experiences such as drinking water or seeing mom. Even though this is a small start, the baby will quickly add the vocabulary to be fluent in the language.*

Compared to other tests, it has the least prerequisite. For example, the Turing test or robot college student test assumes that the agent has already acquired language skills, which is a challenging requirement for the current research.

### 4.1 Practical Administration of the Language Acquisition Test

In the Language Acquisition Test, a proper environment means that there are other humans to teach language to the learning agent. A straightforward way to administer the test is

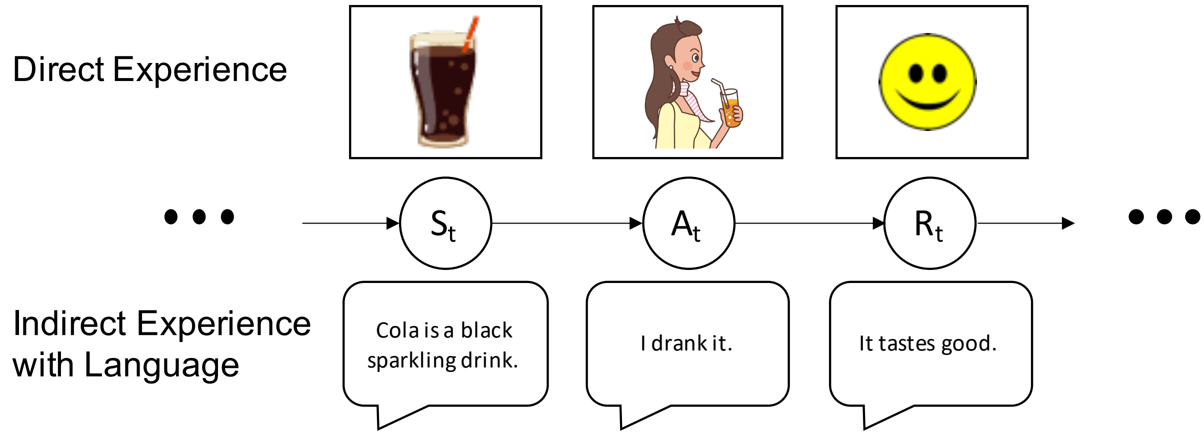


Figure 2: Learning with language means that the symbolic description brings the same changes to the model comparable to direct experiences.

by asking human participants to raise the physical robot agent like a human baby. Turing has suggested this approach [Turing, 1950] and the Developmental Robotics community has actively pursed in many researches [Lungarella *et al.*, 2003; Asada *et al.*, 2009; Cangelosi and Schlesinger, 2015]. However, we already discussed the limitation of the human participants: the prohibitive cost and difficulty in reproducible research.

It would be more useful if we could use artificial environment [Brockman *et al.*, 2016]. There were previous works using simulated environments for the language acquisition, where agents get rewards by following verbal instructions in navigation [Chen *et al.*, 2019; Savva *et al.*, 2019; Chaplot *et al.*, 2018; Hermann *et al.*, 2017; Shridhar *et al.*, 2019] or give correct answers (question answering)[Das *et al.*, 2018a]. However, previous environments have the following limitation for the test of the HLAI.

- **Use of Rewards:** Using reward signals generated by environments will be sufficient for the implementation of level 2 intelligence. However, for level 3 intelligence, the experiencing reward itself should be part of verbal description. In our previous cola example, there is a part related to the explicit reward that is *it tasted good*. In the previous researches, they tend to use explicit reward to teach the concept of the *black sparkling drink* by giving explicit reward when the agent point or navigate to the verbal description. [Chen *et al.*, 2019; Hermann *et al.*, 2017; Chaplot *et al.*, 2018; Das *et al.*, 2018a; Shridhar *et al.*, 2020]. This approach cannot be applied in this case because we need a separate reward mechanism for teaching object concept *black sparkling drink* and the associated reward *it tasted good*.
- **Grounded Language and Embodied Exploration:** The language symbols need to bring changes in the policy. It means that the language symbols need to be grounded with sensory input and the actions in the embodied agents. Some environments that use only the text lack this grounding. [Narasimhan *et al.*, 2015; Côté *et al.*, 2018].
- **Shallow interaction with large number of items and vocabulary:** Previous Environments tend to pour large items and vocabulary into the training. However, as Smith and Slone pointed out, human infants begin to learn a lot about a few things [Smith and Slone, 2017]. We need to build upon basic concepts before we can learn advanced concepts.

Therefore, we claim that we need a new simulated environment for the test of HLAI to overcome these limitations.

### 4.2 An Environment for Language Acquisition like a Human Child Does

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 [Turing, 1950]. SEDRo also simulates developmental psychology experiments to evaluate the progress of intellectual development in multiple domains. In SEDRo, there are a caregiver character, surrounding objects in the environment (e.g., toys, cribs, and walls), and the agent. The agent will interact with the simulated environment by controlling its body muscles 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 development stages.

In SEDRo, the agent can learn up to 12th Months' verbal capacity that speaks first words. As a concrete example, let us review how they will learn the word *water*. The agent has a sensor which indicates the thirsty. When the sensor value is larger than the threshold, the agent will choose the crying behavior by the pre-programmed instincts. When the mother hears the crying, she will investigate and bring water, which the agent will drink. At the same time, the mother says sentences such as “Water!”. Therefore, the agent associates the auditory signal, visual signal, action sequence, and the rewards generated by relieving thirst. More specifically, we conjecture that they will learn to predict vectors' sequence

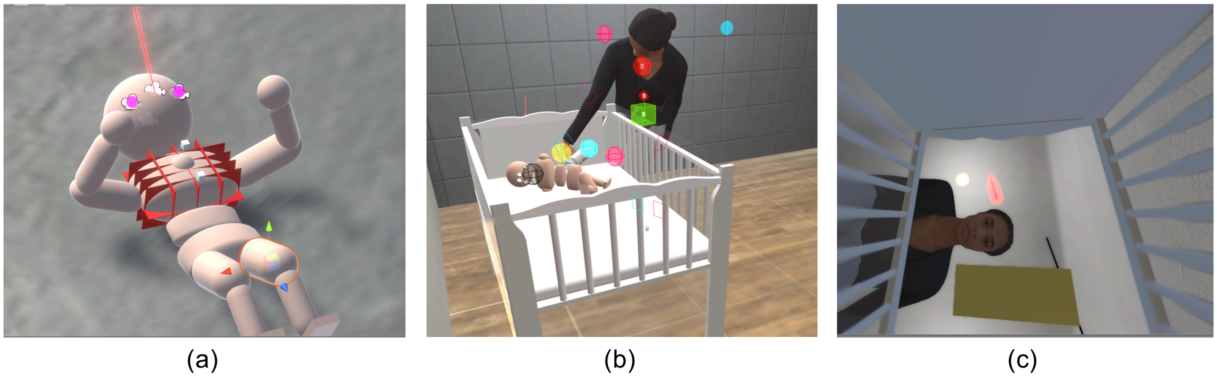

Figure 3: Screenshot of the SEDRo environment. (a) shows the learning agent which has the physical dimension of the one-year-old human baby. The orange line between eyes represents the eye gaze direction. The grid in the torso shows the area for the distributed touch sensors in the skin. (b) shows a caregiving agent feeds milk to the learning agent. (c) shows the visual input to the agent.

where vectors are encoding of the auditory, visual, and somatosensory signals. After enough association has been established, the agent might say "Wada," and the mother brings water. Please note that there are no explicit rewards in this scenario. SEDRo will support this learning to learn the language.

The verbal speech is approximated by the sparse distributed representations (SDR). Speech is encoded to a 512-dimensional vector, where about 10 of them are randomly selected for each alphabet. At each timestep, the corresponding speech signal is represented as the sequence of the vectors. Noise can be added by randomly changing some of the bits.

## 5 Discussion

We compare the definition of HLAI with the related concepts for AI. And we discuss the limitation of our approach and alternative options.

### 5.1 AGI or HLAI

The history of AI is long, and the term AI is used in a broad sense. While AI includes HLAI, it also includes active research area of application-specific AI or machine learning. Interestingly, when the general public thinks AI, they tend to think HLAI, while most academic research is on application-specific AI. Strong or True AI has been used to distinguish the two, but the definition is not clear. Artificial general intelligence (AGI) is also used in a similar context. AGI emphasizes that the agent should be able to do many things as humans do. However, doing many things in a diverse context does not necessarily mean that agents can do what humans do. As a counter-example, a rat can jump around, gather food, mate, and raise a newborn. A virtual rodent by Merel et al. can do multiple tasks depending on the context [Merel *et al.*, 2019]. We might say that this virtual rodent achieved AGI in the simulated environment, but this is not what AGI research targets. In this sense, HLAI might be a better concept for AI research.

### 5.2 Limitations and Alternatives

We proposed to use human-like experience to teach language. The main challenge is that it is difficult to program the caregiver character to enable diverse but reasonable back and forth interaction with the random behaviors of the learning agent. It is expected to teach a few first words if we are successful. Some alternative includes using a completely artificial environment that is not relevant to human experience but still requires skills in many domains. For example, emergent communication behaviors that can be thought of as language has observed in the reinforcement learning environment with multiple agents [Eccles *et al.*, 2019; Cao *et al.*, 2018; Das *et al.*, 2018b; Foerster *et al.*, 2016]. While we might find the clues about the learning mechanism, it might be challenging to apply to the human robot interaction because language is a set of arbitrary symbols shared between members [Kottur *et al.*, 2017].

Another possibility is to transform existing resources into an open-ended learning environment. Using Youtube videos to create a diverse experience can be an example. However, Smith and Slone pointed out that those approaches use shallow information about a lot of things, while human infants begin to learn a lot about a few things [Smith and Slone, 2017]. Also, visual information from the first years consists of an egocentric view, and the allocentric view emerges after 12 Months. Another aspect is that humans learn from social interaction. While infants can learn language from having a Chinese tutor in the meeting, but they cannot learn by seeing the recorded video of tutoring [Kuhl, 2007]. Therefore, we assume that we need to acquire necessary skills before we can learn from those sources.

## 6 Conclusion

Even though the goal of the AI research has been building HLAI, it was not clearly defined. Furthermore, many tests for HLAI have been proposed, but those are not practical and thus not used in the evaluation of AI researches. In this paper, we propose a definition of HLAI. This definition emphasizes that humans can learn from others' experiences using language. Based on this definition, we proposed a language acquisition test for HLAI. A version of this test can be approximated by the simulated environment, and we hope that other researchers can use it to facilitate the research on HLAI. The take-away messages in our paper for the researchers who develop or use language acquisition environment are following:

- **RL researchers:** Using reward signal generated by environments will be effective for the implementation of level 2 intelligence. But for level 3 intelligence, the reward system is also part of the agent and using the prediction error or intrinsic errors can be a viable option as the researchers in language models do.
- **Language model researchers:** The language symbols need to bring the changes in the policy. It means that the language symbols need to be grounded with sensory input, the actions, and the corresponding rewards in the embodied agents.

## References


[Adams *et al.*, 2012] Sam Adams, Itmar Arel, Joscha Bach, Robert Coop, Rod Furlan, Ben Goertzel, J Storrs Hall, Alexei Samsonovich, Matthias Scheutz, Matthew Schlesinger, et al. Mapping the landscape of human-level artificial general intelligence. *AI magazine*, 33(1):25–42, 2012.

[Arbib *et al.*, 2008] Michael A Arbib, Katja Liebal, Simone Pika, Michael C Corballis, Chris Knight, David A Leavens, Dario Maestripieri, Joanne E Tanner, Michael A Arbib, Katja Liebal, et al. Primate vocalization, gesture, and the evolution of human language. *Current anthropology*, 49(6):1053–1076, 2008.

[Asada *et al.*, 2009] Minoru Asada, Koh Hosoda, Yasuo Kuniyoshi, Hiroshi Ishiguro, Toshio Inui, Yuichiro Yoshikawa, Masaki Ogino, and Chisato Yoshida. Cognitive developmental robotics: A survey. *IEEE transactions on autonomous mental development*, 1(1):12–34, 2009.

[Brockman *et al.*, 2016] Greg Brockman, Vicki Cheung, Ludwig Pettersson, Jonas Schneider, John Schulman, Jie Tang, and Wojciech Zaremba. Openai gym. *arXiv preprint arXiv:1606.01540*, 2016.

[Brown *et al.*, 2020] Tom B Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, et al. Language models are few-shot learners. *arXiv preprint arXiv:2005.14165*, 2020.

[Cangelosi and Schlesinger, 2015] Angelo Cangelosi and Matthew Schlesinger. *Developmental robotics: From babies to robots*. MIT press, 2015.

[Cao *et al.*, 2018] Kris Cao, Angeliki Lazaridou, Marc Lanctot, Joel Z Leibo, Karl Tuyls, and Stephen Clark. Emergent communication through negotiation. *arXiv preprint arXiv:1804.03980*, 2018.

[Chaplot *et al.*, 2018] Devendra Singh Chaplot, Kanthashree Mysore Sathyendra, Rama Kumar Pasumarthi, Dheeraj Rajagopal, and Ruslan Salakhutdinov. Gated-attention architectures for task-oriented language grounding. In *Thirty-Second AAAI Conference on Artificial Intelligence*, 2018.

[Chen *et al.*, 2019] Howard Chen, Alane Suhr, Dipendra Misra, Noah Snavely, and Yoav Artzi. Touchdown: Natural language navigation and spatial reasoning in visual street environments. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 12538–12547, 2019.

[Côté *et al.*, 2018] Marc-Alexandre Côté, Ákos Kádár, Xingdi Yuan, Ben Kybartas, Tavian Barnes, Emery Fine, James Moore, Matthew Hausknecht, Layla El Asri, Mahmoud Adada, et al. Textworld: A learning environment for text-based games. In *Workshop on Computer Games*, pages 41–75. Springer, 2018.

[Darwin, 1892] Charles Darwin. *The formation of vegetable mould through the action of worms: with observations on their habits*, volume 37. Appleton, 1892.

[Das *et al.*, 2018a] Abhishek Das, Samyak Datta, Georgia Gkioxari, Stefan Lee, Devi Parikh, and Dhruv Batra. Embodied question answering. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pages 2054–2063, 2018.

[Das *et al.*, 2018b] Abhishek Das, Théophile Gervet, Joshua Romoff, Dhruv Batra, Devi Parikh, Michael Rabbat, and Joelle Pineau. Tarmac: Targeted multi-agent communication. *arXiv preprint arXiv:1810.11187*, 2018.

[Dawkins, 2016] Richard Dawkins. *The selfish gene*. Oxford university press, 2016.

[Eccles *et al.*, 2019] Tom Eccles, Yoram Bachrach, Guy Lever, Angeliki Lazaridou, and Thore Graepel. Biases for emergent communication in multi-agent reinforcement learning. In *Advances in Neural Information Processing Systems*, pages 13111–13121, 2019.

[Foerster *et al.*, 2016] Jakob Foerster, Ioannis Alexandros Assael, Nando De Freitas, and Shimon Whiteson. Learning to communicate with deep multi-agent reinforcement learning. In *Advances in neural information processing systems*, pages 2137–2145, 2016.

[GPT-3, 2020] GPT-3. A robot wrote this entire article. are you scared yet, human? *The Guardian*, Sep 2020.

[Ha and Schmidhuber, 2018] David Ha and Jürgen Schmidhuber. Recurrent world models facilitate policy evolution. In *Advances in neural information processing systems*, pages 2450–2462, 2018.

[Hamilton, 1964] William D Hamilton. The genetical evolution of social behaviour. ii. *Journal of theoretical biology*, 7(1):17–52, 1964.

[Hermann *et al.*, 2017] Karl Moritz Hermann, Felix Hill, Simon Green, Fumin Wang, Ryan Faulkner, Hubert Soyer, David Szepesvari, Wojciech Marian Czarnecki, Max Jaderberg, Denis Teplyashin, et al. Grounded language learning in a simulated 3d world. *arXiv preprint arXiv:1706.06551*, 2017.

[Janik and Sayigh, 2013] Vincent M Janik and Laela S Sayigh. Communication in bottlenose dolphins: 50 years of signature whistle research. *Journal of Comparative Physiology A*, 199(6):479–489, 2013.

[Kepuska and Bohouta, 2018] Veton Kepuska and Gamal Bohouta. Next-generation of virtual personal assistants (microsoft cortana, apple siri, amazon alexa and google home). In *2018 IEEE 8th Annual Computing and Communication Workshop and Conference (CCWC)*, pages 99–103. IEEE, 2018.

[Kottur *et al.*, 2017] Satwik Kottur, José MF Moura, Stefan Lee, and Dhruv Batra. Natural language does not emerge'naturally'in multi-agent dialog. *arXiv preprint arXiv:1706.08502*, 2017.

[Kuhl, 2007] Patricia K Kuhl. Is speech learning 'gated'by the social brain? *Developmental science*, 10(1):110–120, 2007.

[Legg and Hutter, 2007] Shane Legg and Marcus Hutter. Universal intelligence: A definition of machine intelligence. *Minds and machines*, 17(4):391–444, 2007.

[Legg *et al.*, 2007] Shane Legg, Marcus Hutter, et al. A collection of definitions of intelligence. *Frontiers in Artificial Intelligence and applications*, 157:17, 2007.

[Levesque *et al.*, 2012] Hector Levesque, Ernest Davis, and Leora Morgenstern. The winograd schema challenge. In *Thirteenth International Conference on the Principles of Knowledge Representation and Reasoning*. Citeseer, 2012.

[Lungarella *et al.*, 2003] Max Lungarella, Giorgio Metta, Rolf Pfeifer, and Giulio Sandini. Developmental robotics: a survey. *Connection science*, 15(4):151–190, 2003.

[McCann *et al.*, 2018] Bryan McCann, Nitish Shirish Keskar, Caiming Xiong, and Richard Socher. The natural language decathlon: Multitask learning as question answering. *arXiv preprint arXiv:1806.08730*, 2018.

[Merel *et al.*, 2019] Josh Merel, Diego Aldarondo, Jesse Marshall, Yuval Tassa, Greg Wayne, and Bence Olveczky. Deep neuroethology of a virtual rodent. In *International Conference on Learning Representations*, 2019.

[Narasimhan *et al.*, 2015] Karthik Narasimhan, Tejas Kulkarni, and Regina Barzilay. Language understanding for text-based games using deep reinforcement learning. *arXiv preprint arXiv:1506.08941*, 2015.

[Raffel *et al.*, 2020] Colin Raffel, Noam Shazeer, Adam Roberts, Katherine Lee, Sharan Narang, Michael Matena, Yanqi Zhou, Wei Li, and Peter J Liu. Exploring the limits of transfer learning with a unified text-to-text transformer. *Journal of Machine Learning Research*, 21(140):1–67, 2020.

[Savva *et al.*, 2019] Manolis Savva, Abhishek Kadian, Oleksandr Maksymets, Yili Zhao, Erik Wijmans, Bhavana Jain, Julian Straub, Jia Liu, Vladlen Koltun, Jitendra Malik, et al. Habitat: A platform for embodied ai research. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 9339–9347, 2019.

[Shieber, 1994] Stuart M Shieber. Lessons from a restricted turing test. *arXiv preprint cmp-lg/9404002*, 1994.

[Shridhar *et al.*, 2019] Mohit Shridhar, Jesse Thomason, Daniel Gordon, Yonatan Bisk, Winson Han, Roozbeh Mottaghi, Luke Zettlemoyer, and Dieter Fox. Alfred: A benchmark for interpreting grounded instructions for everyday tasks. *arXiv preprint arXiv:1912.01734*, 2019.

[Shridhar *et al.*, 2020] Mohit Shridhar, Jesse Thomason, Daniel Gordon, Yonatan Bisk, Winson Han, Roozbeh Mottaghi, Luke Zettlemoyer, and Dieter Fox. Alfred: A benchmark for interpreting grounded instructions for everyday tasks. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 10740–10749, 2020.

[Smith and Slone, 2017] Linda B Smith and Lauren K Slone. A developmental approach to machine learning? *Frontiers in psychology*, 8:2124, 2017.

[Sutton *et al.*, 1998] Richard S Sutton, Andrew G Barto, et al. *Introduction to reinforcement learning*, volume 135. MIT press Cambridge, 1998.

[Tinbergen, 1951] Niko Tinbergen. *The study of instinct.* Clarendon Press/Oxford University Press, 1951.

[Turing, 1950] AM Turing. Computing machinery and intelligence. *Mind*, 59:433–460, 1950.

[Vinyals *et al.*, 2019] Oriol Vinyals, Igor Babuschkin, Junyoung Chung, Michael Mathieu, Max Jaderberg, Wojciech M Czarnecki, Andrew Dudzik, Aja Huang, Petko Georgiev, Richard Powell, et al. Alphastar: Mastering the real-time strategy game starcraft ii. *DeepMind blog*, page 2, 2019.

[Wang *et al.*, 2018] Alex Wang, Amanpreet Singh, Julian Michael, Felix Hill, Omer Levy, and Samuel R Bowman. Glue: A multi-task benchmark and analysis platform for natural language understanding. *arXiv preprint arXiv:1804.07461*, 2018.