2023 IEEE Symposium Series on Computational Intelligence (SSCI)
Mexico City, Mexico. December 5-8, 2023

A Definition and a Test for Human-Level Artificial Intelligence

Deokgun Park*, Md Ashaduzzaman Rubel Mondol†, Aishwarya Pothula‡, Mazharul Islam§
Department of Computer Science and Engineering
University of Texas at Arlington
Arlington, Texas USA
Email: *deokgun.park@uta.edu, {†mdashaduzzaman.mondol, ‡aishwarya.pothula, §sxi7321}@mavs.uta.edu

***Abstract*—Although AI research aims to build human-level artificial intelligence, it was not clearly defined. Furthermore, many tests for HLAI have been proposed, but those are not practical and thus are not used in evaluating AI research. 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 behavior policy with verbal descriptions as if they had experienced it first-hand. 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. We built a simulated environment to conduct this test practically, and we hope other researchers can use it to facilitate the research on HLAI.**

***Index Terms*—I.2.0 General, I.2.0.b Philosophical foundations, I.2.6.h Language acquisition**

I. Introduction

Despite recent advances in AI, the limitations of the current state-of-the-art are most apparent in the context of robotics. When a layperson or popular culture imagine AI, it is frequently associated with a butler robot that can do many services a human butler could provide. The robot would converse with other humans and robots to do more tasks. If someone asks for a new dish, it might search the Internet for a recipe and learn how to prepare it. AI is thought of as a software part for such a robot. It might be convenient if there is a specific name for this aspect for AI research because the term AI nowadays has a broader meaning. Alternative terms such as true AI, strong AI, or artificial general intelligence (AGI) [1] are often used, but they are not clearly defined.

This paper suggests naming a subfield of AI research for something like a butler robot as human-level artificial intelligence (HLAI). We provide a formal definition and a test as theoretical common ground for the HLAI research. Specifically, we try to answer the following questions.

- What is the difference between human intelligence and other animals?
- What does it mean to learn with the language?
- How can we test whether an agent has HLAI?
- How can we administer such a test practically to aid the model development?

Let us begin by explaining what distinguishes human-level intelligence from the rest.

II. Level of Intelligence

It would be helpful for our discussion to clarify a few terms, such as intelligence, instinct, learning, language, and human-level intelligence. These definitions draw from an examination of biological actors, an earthworm, a rabbit, a monkey, and a human baby, to distinguish different levels of intelligence.

Let us examine the nature of *intelligence* with a concrete question of whether an earthworm is intelligent. The answer will depend on the definition of intelligence. Legg and Hutter proposed the following definition for intelligence after considering more than 70 definitions from psychology and computer science [2], [3]:

> *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. Maximizing gene spreading, or *inclusive fitness*, is accepted as the ultimate goal of biological agents [4]. Earthworms have light receptors and vibration sensors. They move according to these sensors to avoid the sun or predators [5]. These behaviors increase their chance of survival and inclusive fitness [6]. Therefore, we can say that earthworms are intelligent. If we agree that an earthworm is intelligent, we might ask again if it has a *general intelligence*. Considering that it feeds, mates, and avoids predators in a diverse environment, it has general intelligence. However, we would not be as interested in replicating an earthworm-like intelligence. That is why we suggest using HLAI as a term for our community's goal instead of more established terms such as artificial general intelligence (AGI).

However, there are differences in intelligence between earthworms and more advanced agents such as rats and humans. *Behavior policy* is a function that maps a sensory input with the appropriate action. The behavior policy of an earthworm is hard-coded and updated only by evolution. In other words, it is *instinct* [7] that is innate and does not change with experience. For rats and humans, the behavior policy changes

This work was supported by Electronics and Telecommunications Research Institute(ETRI) grant funded by the Korean government. [23ZS1100, Core Technology Research for Self-Improving Integrated Artificial Intelligence System]

with experience, which is *learning*. In this paper, we propose levels of intelligence based on how learning is achieved in agents. Table 1 shows a summary of this idea.

a) Level 1 Intelligence: In this categorization, earthworms have Level 1 intelligence, where there is no learning occurring at the individual level. Their behavior policy have a hard-coded mapping from sensory input to the corresponding action that is instinct updated with evolution [7].

b) Level 2 Intelligence: The problem with Level 1 intelligence is that adaptation with evolution is slow. For example, if there is abrupt climate change due to a meteor crash, agents with Level 1 intelligence will have difficulty adapting to the new environment in a timely manner. Furthermore, the behavior policy is encoded in the genetic code. If a species wants to adjust to various settings, such as diverse climates, the behavior policy has to be encoded in the genetic code, which is costly. If an agent can update the behavior policy during its lifetime by *learning* new rules such as a new type of food or shelter, it would increase inclusive fitness and reduce the amount of genetic code for diverse environments.

Experience is a sequence of sensory inputs (states) and agent actions. A reward is a particular sensory input given by the internal reward system conditioned by the state. We call those agents capable of learning with experience **level 2 intelligence**.

c) Level 3 Intelligence: Contrary to our devotion to learning (machine, supervised, unsupervised, reinforcement, self-supervised learning, etc.), most behaviors of Level 2 intelligent agents are not based on learning but on instincts.

Let us consider a rabbit that has never seen a wolf before. If the rabbit tries to learn the appropriate behavior by randomly experimenting options when it does encounter a wolf, it is too late to update its behavior policy based on the outcome of random exploration. Instead, the rabbit should rely on instinct or Level 1 intelligence. Natural environments are too hostile to use learning as the primary method of updating a behavior policy. Therefore, the range of behavior policy that Level 2 intelligence can learn with direct experience is limited.

Level 3 intelligence overcomes the limitation by learning from others' experiences. Bandura pioneered the social learning theory [8], and learning through observation, so called observation learning, is found in several species, including non-human primates, invertebrates, and reptiles [9]. For example, if you give monkeys locked boxes that contain food, they will try to open them. When one monkey finds how to unlock it, other monkeys observe and imitate it to open their boxes.

d) Level 4 (Human-Level) Intelligence: The limitations of Level 3 are also apparent. In the example of the rabbit and the wolf, the Level 2 rabbit relied on direct experience. But for the Level 3 rabbit to learn the proper behavior, it must observe that its peer rabbit is eaten by wolves, which is also rare. Therefore, even Level 3 cannot learn a lot because they rely on the presence of the example case to be observed.

However, humans are the only known species that use language as a tool for social learning. The verbal and written language uses a sequence of abstract symbols to transfer knowledge, relieving the burdensome requirements of observational learning, such as presence in demonstrations. Thus, we can think of human-level intelligence as **Level 3 intelligence with language**. In humans, language is a tool to learn from others. Human 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 to enable it.

TABLE I
LEVELS OF INTELLIGENCE

Level	Features
1	• No individual learning • Evolution-based refinement • Ex) earthworms
2	• Learning from direct experience • Reward-based refinement • Ex) rats, dogs
3	• Learning from indirect experience • Social, observation-based refinement • Ex) primates, invertebrates, birds
4 (Human-level)	• Learning from symbolic experience • Language-based refinement • Ex) humans

III. Clarifying Language Skill

Language has many aspects. For example, dolphins use a verbal signal to coordinate [10]. Monkeys have been taught sign language [11]. But can we classify the language behavior of monkeys as human-level? Similarly, many previous work in AI has demonstrated various aspects of language skills. Agents have been trained to follow verbal commands to navigate [12], [13]. GPT-3 by open AI can generate articles published as an Op-Ed in Guardians [14], [15]. Some models can do multiple tasks in language in the GLUE benchmark or DecaNLP [16], [17]. Do these models have human-level intelligence?

We claim that learning from others' experiences is the language's essential function, differentiating humans' language use from other animals. We will explain this with a concrete example and then introduce a formal definition. Let us say that you have never tried Cola before. Now, for the first time, you see dark sparkling liquid that looks dangerous. You have a few options for actions, including drinking or running away. You might randomly select to drink. It tastes good. It rewards you.

After this experience, your behavior policy for the situation will change, so it is more likely that you will drink it the next time you see cola. The behavior policy was updated by direct experience. This is how agents with Level 2 intelligence learn.

Learning with language means that it should change your behavior policy similarly, when you hear someone say "Cola is a black, sparkling drink. I drank it, and it tasted good." Figure 1 shows this with the notation in Markov decision process (MDP) (MDP) [18]. Humans use language for learning, which distinguishes human-level intelligence from other animals with language. In this sense, we can define human-level artificial intelligence (HLAI) as follows;

*Definition 1 (**Human-level artificial intelligence (HLAI)**):* An agent has human-level artificial intelligence if there exists a

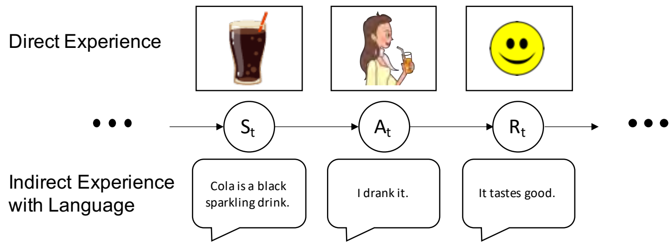

Fig. 1. Learning with language means that the symbolic description brings the same changes to the model comparable to direct experiences.

sequence of symbols (a symbolic description) for every feasible experience such that the agent can update the behavior policy equally, whether it goes through the experience or receives only the corresponding symbolic description.

We can define this more formally with Markov decision process (MDP). Let $\mathcal{S}$ denote a set of all states, and $\mathcal{A}$ denote a set of all actions. The stochastic behavior policy is given as $\pi(a|s) = P[a|s]$ where $a \in \mathcal{A}, s \in \mathcal{S}$. When the behavior policy, $\pi_{old}(a|s)$ is updated with a sequence of states and actions, h, we represent the updated policy as $\pi_{new}(a|s;h)$. Given an original behavior policy, we can derive two policies $\pi(a|s;h_a)$ and $\pi(a|s;h_b)$ that are updated with two different experience sequences h_a and h_b. We can measure the distance $Dist$ between two experience sequences using the KL divergence [19].

$$Dist(h_a, h_b) = \mathbb{E}_s[D_{KL}(\pi_a(a|s;h_a)||\pi_b(a|s;h_b))] \quad (1)$$

Considering s can be large, we might approximate the difference with the restricted set of states $s \in \mathcal{S}' \subseteq \mathcal{S}$, where we choose $\mathcal{S}'$ to be relevant scenarios. Let $\mathcal{D}$ represent the set of all sequences of states and actions that a biological agents can experience firsthand and $\mathcal{T}$ represent the set of all sequences of terms in language. We define a *set of language* to aid the discussion as the following.

*Definition 2 (**A set of language**):* A set of language is a set whose element is a tuple of an experience and a symbolic description, where the agent can update behavior policy equally either by going through the experience or by receiving the symbolic description.

$$\mathcal{L} = \{(h_d, h_l) \in (\mathcal{D}, \mathcal{T}) | Dist(h_l, h_d) \leq \delta\} \quad (2)$$

Using a set of language, we can define HLAI as an agent with an unbounded language set, $\mathcal{L}_{human}$.

$$^{\forall} h_d \in \mathcal{D}, ^{\exists} h_l \in \mathcal{T} \text{ s.t. } (h_d, h_l) \in \mathcal{L}_{human} \quad (3)$$

It might lead to a philosophical debate whether a language set of human is indeed unbound. Authors claim that it is not bounded because it can be extended as needed. For example, an integer is infinite. No human can see every feasible integer in their lifetime. But when required, they can use any of those integers even if they have never seen them before. As an example of language, a typical English person will have only a few words to describe the shades of snow, while an Eskimo might have more words. But if an English person happens to spend 10 years with Eskimo people, he might also acquire more language set for experience related to snow. Or what about a sentence, "He flew through the cheese holes." Even though it is unlikely that someone has seen this sentence before or experienced what the sentence describes, we have no difficulty understanding the sentence or imagining some experience that would justify the sentence.

However, one problem with implementing a test according to this definition will be ensuring a symbolic description exists for every feasible experience.

IV. A Test for HLAI

There are many tests for AI. However, the challenge is to find a sufficient but tractable one. There are many tests that are sufficient but intractable, including the Turing test, robot college student test, kitchen test, and AI preschool test [20]. For example, the robot college student test asks an agent to register, take classes, and get passing grades by doing assignments and exams. Unfortunately, they are rarely conducted in current research, and when they are there is a controversy about the validity of the study [21].

There are a few limitations that make these tests impractical. First, most 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 master StarCraft II player, it took 200 years for machines to masters [22]. 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 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.

At the other end of the spectrum, many tests for AI are tractable but not sufficient for HLAI. While there are models with near-human or super-human level performance in Atari games [23], Go [24], Starcraft II [25], classifying objects from an image [26], or multi-tasks in natural language understanding [27], none would claim that they achieved HLAI. They are effective in proposing a subset of necessary components or mechanisms for HLAI but are not built to study a sufficient set of those mechanisms.

Considering the sufficiency and traceability requirements, we propose a new test for HLAI. If a human infant is raised in an environment such as a jungle where there are no other human, he/she cannot acquire language. It is **environment-limited**. Also, if we have animal cubs and try to raise them as human babies 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 observation, we propose the Language Acquisition Test for HLAI as the following;

*Theorem 1 (**Language Acquisition Test (LAT) for HLAI**):* Given a proper environment, if an agent with an empty set of language can acquire a non-empty set of the language, the agent has the capability for HLAI.

Proof 1: **(Proof by induction)** Suppose an agent can acquire a new element for the set of language that can bring about the same change for a certain experience without relying on the existing set of language. In that case, the agent can keep adding elements to the set of language for a novel experience until it finds the symbolic description for any given experience.

For example, a baby will start learning a single word such as *water* or *mom*. When the baby hears these words, they bring similar effects, such as seeing a cup of water or seeing mom. Although this is a small start, the baby can continue to add vocabulary to become fluent in the language.

A. Practical Administration of the LAT

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 by asking human participants to raise the physical robot agent like a human baby. Turing has suggested this approach [28] and the Developmental Robotics community has actively pursued in many researches [29]–[31]. However, we have already discussed the limitation of human participants: the prohibitive cost and difficulty in reproducible research.

It would be more useful if we could use a simulated environment [32]. There were previous works that used simulated environments for the language acquisition, where agents get rewards by following verbal instructions in navigation [12], [13], [33]–[35] or give correct answers (question answering) [36]. However, previous environments have following limitations 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 reward is not given to the agent but is observed on other agents. Similarly, for human-level 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 [12], [13], [33], [35], [36]. 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*.

1) 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. [37], [38].

2) 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 [39]. 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.

B. An Environment for Language Acquisition

We have been working on a Simulated Environment for Developmental Robotics (SEDRo) for the practical test of HLAI [40]. SEDRo provides diverse experiences similar to those of human infants from the stage of a fetus to 12 months of age. In SEDRo, there is a caregiver character (mother), interactive objects in the homelike environment (e.g., toys, cribs and walls), and the learning agent (baby). The agent will interact with the simulated environment by controlling its body muscles and receiving the sensor signals according to a physics engine. The caregiver character is a virtual agent. It is manually programmed by researchers using a behavior tree that is commonly used in video games to make a game character behave like a human in a limited way. 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 enters more developed stages. The caregiver character teaches language by simulating the conversation patterns of mothers. SEDRo also simulates developmental psychology experiments to evaluate the progress of intellectual development of non-verbal agents in multiple domains such as vision, motor, and social. SEDRo has the following novel features compared to previous work.

1) Open-ended tasks without extrinsic reward: In SEDRo, there is no fixed goal for the agent, and the environment does not provide any reward. Rather than relying on the environment for the rewards, the responsibility of generating rewards belongs to the agent itself. In other words, AI researchers have to manually program a reward system to generate rewards based on the current state. For example, if an agent gets food, the sensory input from stomach will change and the reward system in the agent will generate a corresponding reward.

2) Human-like experience with social interaction: Some studies use environments without explicit rewards, and agents learn with curiosity, or intrinsic reward [41], [42]. However, those environments were arbitrary and non-human such as robot arm manipulation tasks or simple games. Although such simple environments are effective in unveiling the subset of necessary mechanisms, it is difficult to answer what is a sufficient set. In SEDRo, we provide a human infant-like experience, because human infants are the only known example of agents capable of developing human-level intelligence. However, we cannot replicate every aspect of human infants' experience, nor will we try to. There is a subset of experience that is critical for HLAI. Therefore, identifying these essential experiences and finding ways to replicate them in the simulation are two fundamental research questions. Another benefit of a human-like environment is that we can use the experiments from

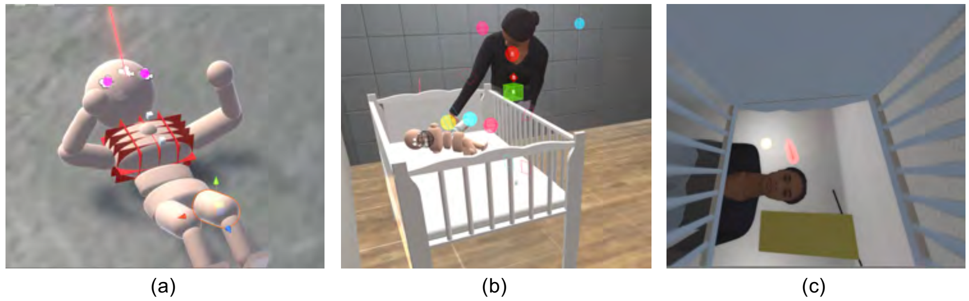

Fig. 2. Screenshot of the SEDRo environment. (a) shows the learning agent that has the physical dimension of a one-year-old human baby. The orange line between the eyes represents the direction of gaze of the eyes. The grid in the torso shows the area for the distributed touch sensors in the skin. (b) shows that a caretaker agent feeds milk to the learning agent. (c) shows the visual input to the agent.

developmental psychology to evaluate the development progress of non-verbal agents.

3) Longitudinal development: SEDRo unfolds agent capabilities according to a curriculum similar to that of human babies. Many studies suggest that humans or agent models learn faster with constrained capabilities [43], [44]. For example, in the first three months, babies are very nearsighted and do not have mobility. This makes many visual signals stationary, and the agent can focus on low-level visual skills with the eyes. In later stages, when sight and mobility increase, babies can learn advanced skills built upon lower level skills.

The final benchmark whether the agent has acquired the language will follow the protocol resembling the previous cola example. We give verbal messages like *"The red ball is delicious(good)"* or *"The blue pyramid is hot (dangerous)"* and check if the behaviour policy toward *the red ball* or *the blue pyramid* has changed accordingly.

V. Discussion

We proposed the definition and test of HLAI. In this section, we discuss the implications of current research on AI. And we discuss the limitations of our approach and alternative options.

A. Agent vs Behavior

The levels of intelligence are used to provide a novel insight into the research of artificial intelligence and not to provide a new taxonomy for the classification of biological agents. There are two limitations to applying this classification to biological agents. First, we do not have complete knowledge of the intelligence of other animals. It is possible that later, we might discover that earthworms learn new skills or that other animals, such as dolphins, have a more sophisticated use of language. In this case, we should adjust which animals belong to which level of intelligence. A more fundamental second limitation is that biological species evolved for a long time, and boundaries tend to be blurry. For example, we might discern mammals from non-mammals with features such as laying eggs or not. But there is a platypus, which is a borderline between mammals and reptiles [45]. Similarly, there can be a gray area between what constitutes as the social learning with language.

Therefore, it is better to have the intelligence level to classify behaviors rather than biological agents. Agents rely on skills from various levels of intelligence. For example, when a baby cries when hungry or shows a stepping reflex, these behaviors are Level 1 intelligence. When they learn to avoid things after experiencing pain, it is Level 2 behavior. Finally, when they observe and mimic caregiver behavior with mobile phones, these behaviors are Level 3 in nature.

B. Limitations and Alternatives of the Test

We proposed using a human-like experience to teach language. The main challenge is that it is difficult to program the caregiver character to enable diverse but reasonable interaction with the random behaviors of the learning agent. It is expected to teach a few first words if we are successful. Some alternatives include 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 have been observed in the reinforcement learning environment with multiple agents [46]–[49]. While we might find clues about the learning mechanism, it may be challenging to apply to human-robot interaction because language is a set of arbitrary symbols shared between members [50].

Another possibility is to transform existing resources into a learning environment. Using Youtube videos to create a diverse experience can be an example. However, Smith and Slone pointed out that these approaches use shallow information about a lot of things, while human infants begin to learn a lot about a few things [39]. 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 the language by having a Chinese tutor in the meeting, they cannot learn by seeing the recorded video of tutoring [51]. Therefore, we assume that we need to acquire the necessary skills before we can learn from recorded video sources.

References

[1] B. Goertzel and C. Pennachin, *Artificial general intelligence*. Springer, 2007, vol. 2.

[2] S. Legg and M. Hutter, "Universal intelligence: A definition of machine intelligence," *Minds and machines*, vol. 17, no. 4, pp. 391–444, 2007.

[3] S. Legg, M. Hutter *et al.*, "A collection of definitions of intelligence," *Frontiers in Artificial Intelligence and applications*, vol. 157, p. 17, 2007.

[4] R. Dawkins, *The selfish gene*. Oxford university press, 2016.
[5] C. Darwin, *The formation of vegetable mould through the action of worms: with observations on their habits*. Appleton, 1892, vol. 37.
[6] W. D. Hamilton, "The genetical evolution of social behaviour. ii," *Journal of theoretical biology*, vol. 7, no. 1, pp. 17–52, 1964.
[7] N. Tinbergen, *The study of instinct.* Clarendon Press/Oxford University Press, 1951.
[8] A. Bandura and D. C. McClelland, *Social learning theory*. Englewood cliffs Prentice Hall, 1977, vol. 1.
[9] L. Ferrucci, S. Nougaret, and A. Genovesio, "Macaque monkeys learn by observation in the ghost display condition in the object-in-place task with differential reward to the observer," *Scientific reports*, vol. 9, no. 1, pp. 1–9, 2019.
[10] V. M. Janik and L. S. Sayigh, "Communication in bottlenose dolphins: 50 years of signature whistle research," *Journal of Comparative Physiology A*, vol. 199, no. 6, pp. 479–489, 2013.
[11] M. A. Arbib, K. Liebal, S. Pika, M. C. Corballis, C. Knight, D. A. Leavens, D. Maestripieri, J. E. Tanner, M. A. Arbib, K. Liebal *et al.*, "Primate vocalization, gesture, and the evolution of human language," *Current anthropology*, vol. 49, no. 6, pp. 1053–1076, 2008.
[12] 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.
[13] 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.
[14] T. B. Brown, B. Mann, N. Ryder, M. Subbiah, J. Kaplan, P. Dhariwal, A. Neelakantan, P. Shyam, G. Sastry, A. Askell *et al.*, "Language models are few-shot learners," *arXiv preprint arXiv:2005.14165*, 2020.
[15] GPT-3, "A robot wrote this entire article. are you scared yet, human?" *The Guardian*, Sep 2020. [Online]. Available: https://www.theguardian.com/commentisfree/2020/sep/08/robot-wrote-this-article-gpt-3
[16] A. Wang, A. Singh, J. Michael, F. Hill, O. Levy, and S. R. Bowman, "Glue: A multi-task benchmark and analysis platform for natural language understanding," *arXiv preprint arXiv:1804.07461*, 2018.
[17] B. McCann, N. S. Keskar, C. Xiong, and R. Socher, "The natural language decathlon: Multitask learning as question answering," *arXiv preprint arXiv:1806.08730*, 2018.
[18] R. S. Sutton, A. G. Barto *et al.*, *Introduction to reinforcement learning*. MIT press Cambridge, 1998, vol. 135.
[19] J. Schulman, S. Levine, P. Abbeel, M. Jordan, and P. Moritz, "Trust region policy optimization," in *International conference on machine learning*. PMLR, 2015, pp. 1889–1897.
[20] 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.
[21] S. M. Shieber, "Lessons from a restricted turing test," *arXiv preprint cmp-lg/9404002*, 1994.
[22] O. Vinyals, I. Babuschkin, J. Chung, M. Mathieu, M. Jaderberg, W. M. Czarnecki, A. Dudzik, A. Huang, P. Georgiev, R. Powell *et al.*, "Alphastar: Mastering the real-time strategy game starcraft ii," *DeepMind blog*, p. 2, 2019.
[23] J. Schrittwieser, I. Antonoglou, T. Hubert, K. Simonyan, L. Sifre, S. Schmitt, A. Guez, E. Lockhart, D. Hassabis, T. Graepel *et al.*, "Mastering atari, go, chess and shogi by planning with a learned model," *Nature*, vol. 588, no. 7839, pp. 604–609, 2020.
[24] D. Silver, J. Schrittwieser, K. Simonyan, I. Antonoglou, A. Huang, A. Guez, T. Hubert, L. Baker, M. Lai, A. Bolton *et al.*, "Mastering the game of go without human knowledge," *nature*, vol. 550, no. 7676, pp. 354–359, 2017.
[25] O. Vinyals, I. Babuschkin, W. M. Czarnecki, M. Mathieu, A. Dudzik, J. Chung, D. H. Choi, R. Powell, T. Ewalds, P. Georgiev *et al.*, "Grandmaster level in starcraft ii using multi-agent reinforcement learning," *Nature*, vol. 575, no. 7782, pp. 350–354, 2019.
[26] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.
[27] P. He, X. Liu, J. Gao, and W. Chen, "Deberta: Decoding-enhanced bert with disentangled attention," *arXiv preprint arXiv:2006.03654*, 2020.
[28] A. Turing, "Computing machinery and intelligence," *Mind*, vol. 59, pp. 433–460, 1950.
[29] M. Lungarella, G. Metta, R. Pfeifer, and G. Sandini, "Developmental robotics: a survey," *Connection science*, vol. 15, no. 4, pp. 151–190, 2003.
[30] 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.
[31] A. Cangelosi and M. Schlesinger, *Developmental robotics: From babies to robots*. MIT press, 2015.
[32] G. Brockman, V. Cheung, L. Pettersson, J. Schneider, J. Schulman, J. Tang, and W. Zaremba, "Openai gym," *arXiv preprint arXiv:1606.01540*, 2016.
[33] 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. 12 538–12 547.
[34] 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.
[35] 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," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2020, pp. 10 740–10 749.
[36] 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.
[37] K. Narasimhan, T. Kulkarni, and R. Barzilay, "Language understanding for text-based games using deep reinforcement learning," *arXiv preprint arXiv:1506.08941*, 2015.
[38] M.-A. Côté, Á. Kádár, X. Yuan, B. Kybartas, T. Barnes, E. Fine, J. Moore, M. Hausknecht, L. El Asri, M. Adada *et al.*, "Textworld: A learning environment for text-based games," in *Workshop on Computer Games*. Springer, 2018, pp. 41–75.
[39] L. B. Smith and L. K. Slone, "A developmental approach to machine learning?" *Frontiers in psychology*, vol. 8, p. 2124, 2017.
[40] A. Pothula, M. A. R. Mondol, S. Narasimhan, S. M. Islam, and D. Park, "Sedro: A simulated environment for developmental robotics," 2020.
[41] S. Singh, A. G. Barto, and N. Chentanez, "Intrinsically motivated reinforcement learning," MASSACHUSETTS UNIV AMHERST DEPT OF COMPUTER SCIENCE, Tech. Rep., 2005.
[42] M. G. Bellemare, S. Srinivasan, G. Ostrovski, T. Schaul, D. Saxton, and R. Munos, "Unifying count-based exploration and intrinsic motivation," in *NIPS*, 2016, pp. 1471–1479.
[43] J. Law, P. Shaw, K. Earland, M. Sheldon, and M. Lee, "A psychology based approach for longitudinal development in cognitive robotics," *Frontiers in Neurorobotics*, vol. 8, p. 1, 2014.
[44] F. C. Keil, "Constraints on knowledge and cognitive development." *The psychological review*, vol. 88, no. 3, 1981.
[45] W. C. Warren, L. W. Hillier, J. A. M. Graves, E. Birney, C. P. Ponting, F. Grützner, K. Belov, W. Miller, L. Clarke, A. T. Chinwalla *et al.*, "Genome analysis of the platypus reveals unique signatures of evolution," *Nature*, vol. 453, no. 7192, p. 175, 2008.
[46] 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.
[47] K. Cao, A. Lazaridou, M. Lanctot, J. Z. Leibo, K. Tuyls, and S. Clark, "Emergent communication through negotiation," *arXiv preprint arXiv:1804.03980*, 2018.
[48] 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.
[49] 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.
[50] 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.
[51] P. K. Kuhl, "Is speech learning 'gated'by the social brain?" *Developmental science*, vol. 10, no. 1, pp. 110–120, 2007.