Zhorai: Designing a Conversational Agent for Children to Explore Machine Learning Concepts

Phoebe Lin[1], Jessica Van Brummelen[2], Galit Lukin[3]
Randi Williams[4], Cynthia Breazeal[5]
[1]Harvard Graduate School of Design, Cambridge, MA 02138
[2]MIT CSAIL, Cambridge, MA 02139
[3]MIT ORC, Cambridge, MA 02139
[4,5]MIT Media Lab, Cambridge, MA 02139
phoebelin@gsd.harvard.edu, jess@csail.mit.edu, glukin@mit.edu, randiw12@mit.edu, cynthiab@mit.edu

Abstract

Understanding how machines learn is critical for children to develop useful mental models for exploring artificial intelligence (AI) and smart devices that they now frequently interact with. Although children are very familiar with having conversations with conversational agents like Siri and Alexa, children often have limited knowledge about AI and machine learning. We leverage their existing familiarity and present Zhorai, a conversational platform and curriculum designed to help young children understand how machines learn. Children ages eight to eleven train an agent through conversation and understand how the knowledge is represented using visualizations. This paper describes how we designed the curriculum and evaluated its effectiveness with 14 children in small groups. We found that the conversational aspect of the platform increased engagement during learning and the novel visualizations helped make machine knowledge understandable. As a result, we make recommendations for future iterations of Zhorai and approaches for teaching AI to children.

Figure 1: Zhorai platform. Children converse with Zhorai through the computer's built-in microphone.

Introduction

Artificial intelligence (AI) curriculum is increasingly being incorporated into what is now considered standard 21st century computing education. AI for K-12 initiatives have driven the development of a diverse range of resources, including formal curricula guidelines and accessible online directories for all educators (Touretzky et al. 2019). At the same time, children are also frequently interacting with intelligent agents, such as Siri and Alexa. Frequent interaction fosters familiarity and natural interaction between children and agents, which enables us to design learning opportunities around these conversations. Currently, commercially available in-home voice assistants are designed to be like black boxes for ease of use and adoption. Thus, children often have limited knowledge for how they work. Our work helps unveil the black box using a simplified, teachable agent, such that children can learn about the agent's underlying processes while teaching it.

The Zhorai platform and curriculum consists of the conversational agent Zhorai, four sequential AI modules, and an assessment for children to explore machine learning and knowledge representation. In this paper, we discuss how we designed the system and curriculum, and the results of a small scale study with 14 eight- to eleven-year-old children. We investigated how well children learned and their engagement levels with Zhorai's conversational interface. Although prior knowledge of conversational agents and other factors may impact a child's ability to understand, we hypothesize that all children can benefit significantly from conversational interaction with Zhorai and the visualizations. From that, we form two specific hypotheses:

1. Interacting with Zhorai through conversation helps increase engagement in the overall activity.
2. Zhorai's visualizations and curriculum leads to children's learning and understanding of how machines learn.

Background

AI in Education

While many of the K-12 AI initiatives are recent calls to action, the idea of introducing children to AI concepts dates back to work by Seymour Papert and Cynthia Solomon

using LOGO programming and Turtle robot (Papert and Solomon 1971). It has served as the foundation for much of current work. Many platforms teach AI by having children program in block-based languages including Cognimates (Druga 2018), Machine Learning for Kids (Lane 2018), and eCraft2Learn (Kahn and Winters 2018). Other platforms introduce AI within the context of robotics, such as Popbots (Williams, Park, and Breazeal 2019), and application development, such as Scratch AI Extensions (Resnick et al. 2009) and MIT App Inventor AI Extensions (Van Brummelen 2019; Zhu 2019). However, Popbots is geared towards preschool age children and the AI Extensions are geared towards middle to high school students.

Conversational Agents for Learning

Conversational agents are also used in education, often as intelligent tutors and learning companions (Kerry, Ellis, and Bull 2009). Design of these conversational systems vary widely, from having text-based inputs and outputs (Aleven, Popescu, and Koedinger 2001; Heffernan and Croteau 2004) to embodied agents that can display emotion (Graesser et al. 2005). Our work uses a simple alien avatar with both speech input and output, which has been used in systems of other contexts ((Kerly, Ellis, and Bull 2008), (Litman and Silliman 2004). However, these interfaces were not designed to teach AI and did not use visualizations.

A key feature of conversational agents is the use of natural language as a medium for learning and expression. Using natural language is something nearly all learners are familiar with, and could allow learners to devote more cognitive resources to the learning task (Beun, de Vos, and Witteman 2003). Although we did not measure the effect of learning through natural language dialogue, we built upon these findings to create an interactive and educational activity through which children can learn about AI. Lastly, conversational agents provide learners with a social environment. Social contexts can support motivation for knowledge. It has been shown that a one-to-many relationship between the conversational agent and learners is effective (Kumar and Rose 2010). Furthermore, multimedia environments can help children be more engaged and motivated to learn (Druin and Solomon 1996).

Zhorai Curriculum Design

We built the Zhorai platform and curriculum around three "Big Ideas" in AI, as they guide the high impact concepts we want children to understand about AI (Touretzky et al. 2019). At the same time, these ideas also allow us to explore the utility of a conversational interface. The three ideas are:

- **Representation and Reasoning** Children are expected to understand how Zhorai learns and represents new information. Zhorai generates two different visualizations to show its knowledge representation.
- **Learning** Zhorai also demonstrates the concept of how machines classify concepts. Children witness instances when Zhorai might succeed or fail in its learning and make attempts to correct it.
- **Social Impact** The curriculum emphasizes the ethics and societal impacts of AI through structured discussion and thinking about implications for bigger-picture contexts.

The Zhorai platform is an online web interface that can be accessed via a web browser at *https://zhorai.csail.mit.edu* and the source code can be accessed at *http://github.com/jessvb/zhorai*. Children engage with Zhorai through conversation in a small group setting led by an adult facilitator. Children learn that Zhorai is an alien visiting Earth that wants to learn about all of Earth's life (i.e., ecosystems on earth). The curriculum is focused on Earth's ecosystems because children can describe ecosystems, defined as places where animals live, without much prior knowledge. Additionally, many state science standards discuss the concept of ecosystems in our target grade band of 3-5 (NGSS 2016).

The AI curriculum consists of four modules: "What Does Zhorai Know?", "Teaching Zhorai", "Witnessing Machine Learning", and "AI and Ethics". Before entering the first module, children learn to interact with Zhorai through an introduction, in which Zhorai greets and converses with them.

Module 1, "What Does Zhorai Know?" introduces children to knowledge representation and reasoning using concept maps. In this module, children ask Zhorai about the five ecosystems that it knows: deserts, oceans, grasslands, rainforests, and tundras. The representation method Zhorai uses is a concept map or "mind map", which children can analyze to determine ecosystem attributes. Positive attributes (what an ecosystem has) are visualized as blue circles, and negative attributes (what an ecosystem does not have) are visualized as red circles. Children also analyze the corpus that Zhorai is given to form the mind maps, and thus draw connections between natural language sentences and details of the corresponding mind map.

In the next module, **Module 2, "Teaching Zhorai"**, children are tasked to provide Zhorai with data about three animals of Zhorai's choosing. Each child contributes as many sentences for each animal as they want. Children can give Zhorai any detail about the animal they would like, as long as they do not reveal the ecosystem the animal lives in. This allows them to center the AI learning experience on their own knowledge and interests. Zhorai functions as a less-knowledgeable AI that is teachable, and children help Zhorai construct mind maps for each animal. The mind map concept is introduced in Module 1 and built upon in the following modules.

Module 3: "Witnessing Machine Learning", is where children observe Zhorai's learning and reasoning process. They ask Zhorai to guess which ecosystem it thinks the animals they previously taught it about are from. Zhorai computes word similarity scores for each ecosystem by using the words representing the animal and the words representing the ecosystems. Zhorai then displays these scores using a bar graph. For each animal, Zhorai chooses the ecosystem that has the highest similarity score among the five ecosystems. It is important for children to understand when and

why Zhorai may guess incorrectly by drawing connections between similarities within the mind maps of Module 2 and the scores of Module 3, and how to manipulate this connection.

In the last module, **Module 4, "AI and Ethics"**, facilitators lead a discussion about how conversational agents and agents that learn from data are used in society with positive and negative consequences. These discussion questions are scaffolded in the Teacher Resources section on the website. Children reflect on instances when Zhorai made mistakes, and are asked questions like, "Would Zhorai know whether what we teach it is correct or not?" and "How would you feel if Zhorai learned something untrue about you?". The goal of this module is to empower children with the tools to design AI with ethics in mind. We also probe children what societal impacts of mistakes made by AI can be, and how we can mitigate harm (Payne 2019).

System Design and Implementation

The system was designed with elementary school students, teachers, and ease-of-use in mind. It can be run using Google Chrome anywhere with internet access. The main components of the system are the **(1) speech synthesizer**, **(2) speech recognizer**, **(3) semantic parser**, **(4) word map comparator**, and **(5) website visualizer**, as follows:

1. **Speech synthesizer**: Synthesizes Zhorai's voice using the Web Speech API's `webkitSpeechRecognition` interface (Contributors to the API Specification 2019). Note that the particular voice from the `webkitSpeechRecognition` interface was chosen to be high (like a child's voice) and as gender-neutral as possible. (We refer to Zhorai as "it" instead of a gendered "he" or "she".)

2. **Speech recognizer**: Converts user speech to text using the Web Speech API's `SpeechSynthesisUtterance` interface (Contributors to the API Specification 2019).

3. **Semantic parser**: Performs natural language processing using the NLTK (Loper and Bird 2002) and Stanford CoreNLP toolkits (Manning et al. 2014). The NLTK toolkit stems and lemmatizes the words, while the Stanford CoreNLP toolkit extracts the parts of speech (Toutanova and Manning 2000) and parses the sentences received from the speech recognizer. The parser executes the following three tasks used throughout the activity.

   (a) *Name identification*: As an introduction to the activity, Zhorai asks children for their name and location. The parser parses the input received after each question, extracting all of the nouns, proper nouns, and foreign words. The identified name or location is the last proper noun recognized.

   (b) *Topic identification*: In the first module, the parser identifies which ecosystem the user is asking about. Since the number of ecosystems Zhorai knows about is limited, the parser searches for these ecosystems. If a known ecosystem is found, the parser returns the identified ecosystem and the pre-compiled mind map is shown to the users. These mind maps are built offline using the parser's third ability.

   (c) *Mind map constructor*: In Modules 1 and 2, Zhorai displays its knowledge of ecosystems and animals through mind maps which are generated based on sentence structure and parts of speech. The parser identifies the topic of a sentence (an animal or ecosystem) and the correlation between descriptive words and the topic (e.g., negative correlation between *water* and *desert*), and sends this information to the website visualizer.

4. **Classifier**: Classifies the animal into an ecosystem by comparing the animal's knowledge representation to that of the ecosystems. The comparison is done using NLTK's Wordnet Interface (Loper and Bird 2002; Fellbaum 1998). For each ecosystem, all of the words describing the ecosystem are compared to the words describing the animal. The WuPalmer similarity score (Wu and Palmer 1994) is calculated for each combination of word pairs (one word corresponding to the animal and the other to the ecosystem) given that both words have the same correlation sign with their topics. The occurrence of a word is taken into account linearly. Once the similarity between all pairs has been calculated, the overall similarity between an ecosystem and the animal is its weighted average similarity score. The classifier returns a normalized score for each ecosystem, denoting the ecosystem with the highest similarity score to be the ecosystem that the animal resides in.

5. **Website visualizer**: Generates mind maps and histograms based on predefined ecosystem sentences as well as user-defined animal sentences using `D3.js` (Bostock 2019).

To illustrate how the components interact, in the introduction, the system prompts the user to speak; sends the user's voice response to the browser-based speech recognizer, which converts the response to text; saves the text response in the user's local session data; sends the text response to the server-based semantic parser, which processes the text for important information; returns the processed information to the browser-based speech synthesizer, which causes Zhorai to speak, prompting the user to speak again; and repeats the process until the conversation completes. This process is illustrated in Figure 2. Module 1, 2, and 3 include the mind map visualizations, which display the predefined sentences about ecosystems and/or users' sentences about animals, as shown in Figure 3. Module 3 also contains a histogram visualization, which displays the similarity between the animal and ecosystem sentences. An example histogram is shown in Figure 4.

Pilot Study

We conducted a small pilot study at a local elementary school to determine whether children understood the visualizations. We recruited five eight-year old children (three boys and two girls) and presented them with three different visualizations for knowledge representation (Module 1 and 2) and three different visualizations for machine learning (Module 3) shown in Figure 5.

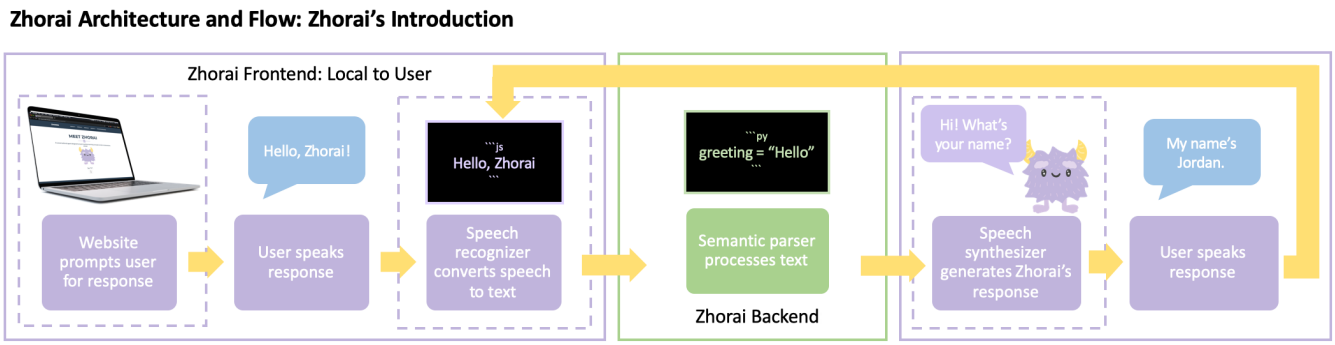

Figure 2: A representation of the introduction to Zhorai in terms of architecture and user flow.

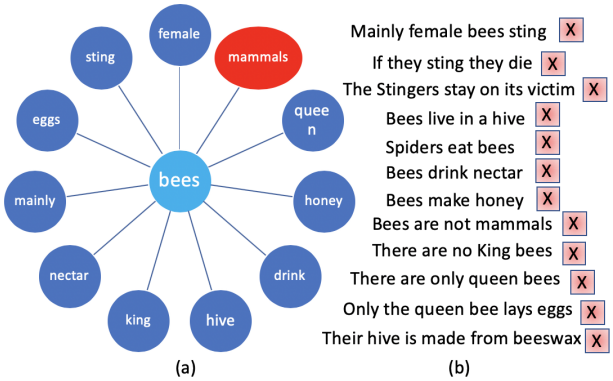

Figure 3: (a) The mind map generated from the sentences in (b). This was shown in Module 2 on the Zhorai website. (b) Sentences users told Zhorai during the study.

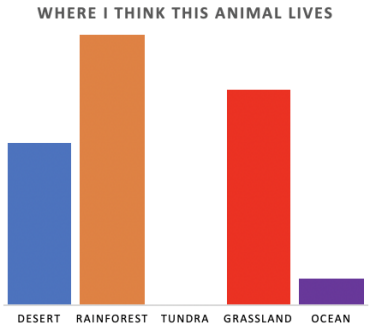

Figure 4: As shown by the orange bar in this histogram visualization, Zhorai would guess that the animal lives in the rainforest.

Children were asked to add to each visualization given the sentences, "Deserts are hot" and "Deserts are not cold", and then to select the visualization that they thought best represented Zhorai's mind. All children successfully added accurately-colored attributes to the mind map. Four out of five children chose the mind map as the best representation, and also drew pictures of a desert. As a result, we selected the mind map to use for both Modules 1 and 2 as the first visualization. Children were then asked to choose the ecosystem Zhorai thinks a camel would live in based on the visualizations on the right in Figure 4. Four out of five children successfully chose the correct ecosystem for camel (desert) using the scatterplot visualization, while three out of five children successfully did so using the histogram visualization. We decided to use the histogram despite lower performance compared to the scatterplot because it better reflects the implementation of underlying system.

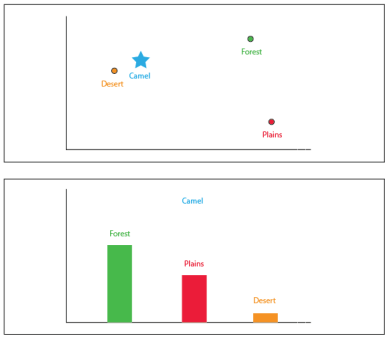

Figure 5: Sample of two visualizations tested for machine learning. We used the histogram for the activity because it more accurately reflects the implementation of the underlying system.

User Study

We tested whether the learning experience with the Zhorai platform and curriculum promotes children's understanding of how machines represent knowledge and learn. Specifically, we evaluated the child's level of engagement, measured by the number of sentences used in conversation with Zhorai, and whether it correlated with their level of understanding.

Method

Participants. Fourteen children participated in the study (6 boys, 8 girls, age range 8-11, M = 9.43, SD = 1.09), recruited from a mailing list. Every participant and their parents gave informed consent to participate in compliance with our institution's ethics review board. All child participants were invited to a small tour of the authors research lab, and the

opportunity to interact with Jibo, an intelligent robot, after each session.

Procedure. We followed a consistent experiment protocol across different sessions. Session sizes ranged from one to four children, each with a facilitator. Parents were invited to contribute ecosystem sentences in sessions with only one child participant. First, we informally assessed participants' prior knowledge on voice assistants like Siri and Alexa with an icebreaker. After this, we did pre-assessments followed by an introduction to Zhorai. Modules 1 through 4 were then completed in order. The entire activity lasted ∼60-80 minutes depending on length of discussion. Children performed all assessments individually on paper and with minimal interference from researchers. In addition to quantitative data from the assessments, we video recorded the sessions, logged sentences participants contributed, and recorded children's responses to the assessment questions.

Assessments

To answer the hypothesis that children successfully learned and can reason about how machines learn, we developed an assessment with five questions:

1. Which sentences could you say to Zhorai to create the following mind map in Figure 3? (This assesses their understanding of how knowledge is represented.)
2. What could you tell Zhorai about monkeys so that it could correctly guess that monkeys live in rainforests? (This assesses their understanding of how Zhorai learns.)
3. The following histogram is what Zhorai thinks about where toucan lives. Based on the histogram in Figure 4, which ecosystem would Zhorai think a toucan lives in? (This assesses their understanding of how Zhorai makes a decision.)
4. Which ecosystems do snakes live in? Why might Zhorai have a difficult time classifying snakes into one ecosystem even if it knew everything there is to know about them? (This is an open-ended question for assessing mistakes Zhorai may make.)
5. Have you tried saying Zhorai to Zhorai? If not, ask the teacher if you can try. Does Zhorai recognize its own name? If not, why do you think it doesn't? Can you think of another name that Zhorai won't recognize? (This is an open-ended question on Zhorai's internal natural language processing.)

We also developed pre- and post-assessments to assess children's self perceptions as an engineer and motivation to learn. Children were asked to rate how much they agree or disagree with various statements on a scale of 1 to 5.

1. I am curious about how a machine learns.
2. I think I can teach a machine.
3. I trust voice assistants like Siri and Alexa.
4. I understand how a machine learns.
5. The activities we did today were useful for learning about how machines learn. (Post-assessment only.)

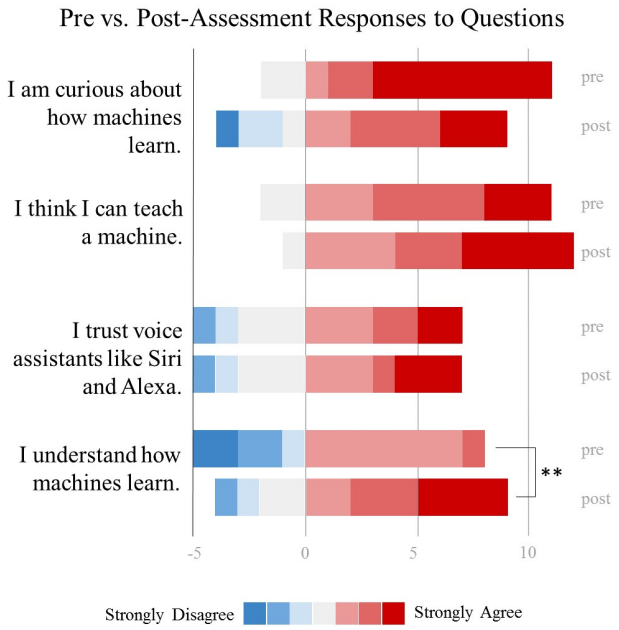

Figure 6: Comparing participant's responses on the pre- and post-assessments. There was a significiant change on the statement 'I understand how machines learn.'

6. The mind maps for each ecosystem and animal helped me think about Zhorai's brain. (Post-assessment only.)
7. The histogram helped me understand how Zhorai made decisions. (Post-assessment only.)

Results

Children's Perceptions of Conversational Agents

In the pre-assessment (Figure 6) we sought to measure participants' perceptions of themselves as engineers and their interest in learning more about conversational agents. We found that participants were not sure that they understood voice agents very well, but they were motivated to learn more. On a 7-Point Likert scale, with 7 being "Strongly Agree," the average score was 6.23 for participants agreeing that they were curious about how machines learned and 3.84 for them understanding how agents learn already. Some participants agreed, but many were not sure that they could teach the machines (Mean: 5.69, Median: 6). One participant said that teaching machines was "something that adults do."

At the end of the Zhorai activity, participants self-reported that they had a much better understanding of how machines learned, see Figure 6. Again, on the pre-assessment the average score for this question was 3.84, however the post-assessment average was 5.46. Using a Wilcoxon signed rank test, we determined that this difference was significant Z=-2.62, $p < 0.01$**). On the post assessment, participants agreed that the mind maps helped them understand how machines learned (Mean: 5.46, Median: 6) and that they would like to teach Zhorai more (Mean: 5.92, Median: 7). Some

participants said, “I want to play games with Zhorai” (P5) and “I want to change the way Zhorai thinks” (P7). However, their belief in their ability to teach a machine did not change significantly, especially when thinking about more complex agents like Siri and Alexa. We did not see any other significant changes in participants’ answers on the pre-and-post assessments.

Trust played a role in participants’ engagement with Zhorai. We observed that children who were intuitively more optimistic about AI enjoyed interacting with Zhorai and wanted to build a deeper relationship with it, whereas children who were more pessimistic about AI enjoyed it less and voiced doubts about whether Zhorai actually learned or only said things because it was programmed to.

Engaging and Learning with a Conversational Agent

Our first hypothesis is that the use of a conversational agent would be an engaging tool for children to learn about machine learning. In past experiences, we found that children quickly tire providing training examples to text-based machine learning systems while video or sound-based systems, that learn by recording a stream, are much easier. Using a conversational agent, Zhorai, as the training interface we hoped that participants would remain engaged for longer. To measure engagement, we used the total number of utterances that participants used to provide information about an animal. Utterance counts ranged from 5 to 34 (Mean: 11.85, Median: 10) for each participant.

Before knowing the objective of having Zhorai guess which ecosystems animals belonged to, children taught Zhorai about anything that came to mind: (P11 and P12 on teaching Zhorai about birds)

> Birds twitter to talk to each other.
> Birds can fly due to their hollow bones.
> Birds usually have colorful feathers.
> Birds make nests in trees and high places.
> Birds will go down underneath bushes in a certain biomes because there is no lease where they are.

Given that participants worked in groups when interacting with Zhorai, we saw large numbers of training utterances. The number of training examples that groups gave to Zhorai per animal ranged from 4 to 29 (Mean: 10.8, Median: 10).

Understanding How Machines Think

We hypothesized that engaged participants would be able to gain a deeper understanding of how voice agents worked. For this, we used assessments that we designed specifically for Zhorai. Question 1 on the assessment sought to measure participants’ understanding of mind maps by having participants derive the sentences that were used to create a mind map. There were four sentences in total and we scored the question by the number of sentences that participants got correct, so the scores are out of four. As shown in Figure 7, participants seemed to understand the the relationship between the mind maps and input data. The average score on

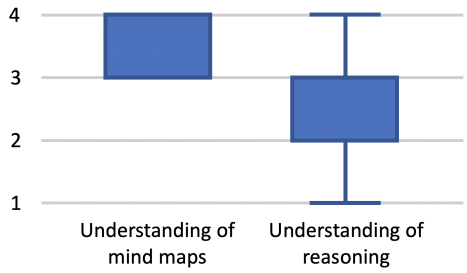

Figure 7: Zhorai activity assessment results for Questions 1 and 2. Participants generally understood the relationship between the mind maps and input data, the average being 3.54 out of 4 for Question 2.

Question 1 was 3.54 out of 4. Twelve out of fourteen participants got three or more of the four sentences, one student seemed to misunderstand the question, and the other refused to answer.

Understanding How Machines Learn

We also hypothesized that the overall experience of training Zhorai and using the mind maps would help participants understand how machines learned. Question 2 on the assessment evaluated participants’ understanding of how Zhorai drew conclusions about which ecosystem animals belong to based on the information it knows about them. We rated participants out of 4 based on the number of sentences they generated that clearly placed an animal in a particular ecosystem. There was a broad distribution of scores on this question (Mean: 2.62, Median: 3). There was not a direct correlation between this question and any of the things we measured. However, engagement with training Zhorai may have been a factor since we saw that participants who did not contribute a lot to Zhorai did not do as well on this question.

In module 3, when Zhorai incorrectly guessed which ecosystem an animal belonged to, participants were eager to go back and teach Zhorai more so that it could guess correctly. Notably, participants experimented with different techniques such as using more negative sentences: “Birds do not live in the water” and “Birds do not live their whole life in the water” (P11 and P12). And in response to teaching Zhorai “Birds do not live in the ocean” three times consecutively, P11 explained the concept of sample size: “Zhorai needs to hear things a lot of times to learn, like humans sometimes do”.

How Children Feel about AI Ethics

Overall, children explained that they would feel bad if Zhorai learned something wrong, especially when talking about people and culture. “[If Zhorai learned something wrong] it would be interesting because if you start a conversation between you and Zhorai and you talk about the thing the person said wrong, like if someone said ’baseball is a bad sport’ she would think baseball is a bad sport, and you would feel bad” (P11). P2 reflected on the topic of over-reliance: “...basically later in the future, humans will start depending and relying on AI a lot more, then they could easily make a mistake and

not realize that the person they're trying to help is not getting helped" but was optimistic about how to reduce harm: "I think if we gave Zhorai more knowledge, like things we would have in a conversation, so like you would teach it about Earth, about people, about cultures, about math."

Discussion

Through interacting with Zhorai and completing four AI modules, children learn how an agent can represent knowledge from data, and how data can be manipulated so that the agent learns different concepts. We found that children's engagement with Zhorai played a significant role in understanding Zhorai's reasoning and the mind map visualizations. Training via conversation and visualizations together were effective in teaching children about machine learning concepts.

In addition to demonstrating that a conversational interface can be a useful addition to AI curriculum for children, this work also includes key design considerations for future teaching approaches.

Allow AI to make mistakes. In our activity, children were most intrigued when Zhorai guessed incorrectly. This motivated them to figure out why Zhorai made mistakes and improve its training: "Zhorai, where are you? I want to teach it more" (P10). When this happened, children often said, "Wait what, why did it do that?" and then drew connections in the animal and ecosystem mind maps: "Oh [the ecosystem mind map and animal mind map] both have 'different' in them, and 'wing' and 'bird' are similar so that's what Zhorai thought" (P13). Children observed that manipulating the sentences spoken to Zhorai can change its knowledge representation (mind maps) and learning (histogram). We observed that the process of Zhorai guessing incorrectly in the first iteration, children retraining Zhorai, and then Zhorai guessing correctly in the second iteration was incredibly rewarding for them. Children were delighted that their efforts made a difference and that Zhorai was successful in the end. Through experimentation, children strengthened their mental models on the capabilities and limitations of AI. Future work should emphasize designing curriculum to communicate how agents are trainable and not perfect.

Visualize the system model. We created visualizations to communicate the underlying system model to children. Because these visualizations are true to the system model, we presented an accurate explanation of Zhorai's "black box". This opened the model to interpretation, allowing children to draw their own conclusions, facilitating learning. The visualizations were essential in helping children reason about Zhorai's mental model. The children constantly referred to the mind maps when they were analyzing Zhorai's predictions and teaching it more about particular animals. When they added sentences to Zhorai's animal corpus, they observed the mind maps changing and Zhorai's predictions about the most suitable ecosystem changing, as observed by changes in the histogram visualization. This helped cement the idea that Zhorai's knowledge was manipulable.

Learn by teaching. There is extensive research demonstrating the effectiveness of the learning-by-teaching paradigm (Biswas et al. 2005) and using conversational agents in those environments (Segedy, Kinnebrew, and Biswas 2012). Teaching Zhorai helped children understand what Zhorai could and could not understand. This strengthened their relationship with Zhorai: "I guess I was thinking like Zhorai" (P10). Children wanted Zhorai to learn about themselves: "I want to teach Zhorai about things like 'my favorite color is blue"' (P5). Therefore, we highlight the potential to create enjoyable learning experiences for children with this teachable agent paradigm.

Limitations

The present study has some limitations. First, we recognize that this we had a small number of participants. Further testing could strengthen our claim that Zhorai is effective. Also, we would like to do a future iteration of this study where we compare the effectiveness of the conversational agent interface to another interface, i.e. a text-based interface. Finally, in order to run this study we somewhat controlled how participants interacted with Zhorai. We limited its knowledge of ecosystems to five options and had Zhorai specify which animals it would like to learn about. These were necessary so that participants could train Zhorai in real time, which we felt was necessary to make the overall experience more engaging.

Conclusions and Future Work

In this work, we presented Zhorai—a conversational agent that teaches children about knowledge representation and machine learning. Overall, we observed that Zhorai was engaging and effective. Children taught Zhorai different facts about animals and then observed the agent using its knowledge to infer which ecosystems those animals belong to. By training an agent, witnessing its mistakes, and retraining the agent, children were able to make sense of the agent's intelligence. In the future, we hope to expand Zhorai's curriculum to address more topics in AI. We hope that this work will inspire more AI curricula that use voice interfaces and visualizations to aid children's understanding of AI algorithms.

Acknowledgements

We would like to thank Hal Abelson for his mentorship and guidance in designing this project. We are also grateful to the children and parents who participated in the study for their eagerness and hardword.

References

[Aleven, Popescu, and Koedinger 2001] Aleven, V.; Popescu, O.; and Koedinger, K. 2001. Towards tutorial dialog to support self-explanation: Adding natural language understanding to a cognitive tutor.

[Beun, de Vos, and Witteman 2003] Beun, R. J.; de Vos, E.; and Witteman, C. 2003. Embodied conversational agents:

Effects on memory performance and anthropomorphisation. 315–319.

[Biswas et al. 2005] Biswas, G.; Leelawong, K.; Schwartz, D.; and Vye, N. 2005. Learning by teaching: A new agent paradigm for educational software. *Applied Artificial Intelligence* 19:363–392.

[Bostock 2019] Bostock, M. 2019. D3.js. https://d3js.org/, Last accessed on 2019-09-04.

[Contributors to the API Specification 2019] Contributors to the API Specification. 2019. The possibility of a universal social welfare function. Report, W3C.

[Druga 2018] Druga, S. 2018. Growing up with AI: Cognimates: from coding to teaching machines. Master's thesis, Massachusetts Institute of Technology, Cambridge, MA.

[Druin and Solomon 1996] Druin, A., and Solomon, C. 1996. *Designing Multimedia Environments for Children: Computers, Creativity, and Kids.* ERIC.

[Fellbaum 1998] Fellbaum, C. 1998. *WordNet: An Electronic Lexical Database*. Bradford Books.

[Graesser et al. 2005] Graesser, A. C.; Chipman, P.; Haynes, B. C.; and Olney, A. 2005. Autotutor: an intelligent tutoring system with mixed-initiative dialogue. *IEEE Transactions on Education* 48(4):612–618.

[Heffernan and Croteau 2004] Heffernan, N. T., and Croteau, E. A. 2004. Web-based evaluations showing differential learning for tutorial strategies employed by the ms. lindquist tutor. In Lester, J. C.; Vicari, R. M.; and Paraguaçu, F., eds., *Intelligent Tutoring Systems*, 491–500. Berlin, Heidelberg: Springer Berlin Heidelberg.

[Kahn and Winters 2018] Kahn, K., and Winters, N. 2018. Ai programming by children.

[Kerly, Ellis, and Bull 2008] Kerly, A.; Ellis, R.; and Bull, S. 2008. Calmsystem: A conversational agent for learner modelling. *Knowledge-Based Systems* 21:238–246.

[Kerry, Ellis, and Bull 2009] Kerry, A.; Ellis, R.; and Bull, S. 2009. Conversational agents in e-learning. In Allen, T.; Ellis, R.; and Petridis, M., eds., *Applications and Innovations in Intelligent Systems XVI*, 169–182. London: Springer London.

[Kumar and Rose 2010] Kumar, R., and Rose, C. P. 2010. Architecture for building conversational agents that support collaborative learning. *IEEE Transactions on Learning Technologies* 4(1):21–34.

[Lane 2018] Lane, D. 2018. Machine learning for kids.

[Litman and Silliman 2004] Litman, D. J., and Silliman, S. 2004. Itspoke: An intelligent tutoring spoken dialogue system. In *Demonstration Papers at HLT-NAACL 2004*.

[Loper and Bird 2002] Loper, E., and Bird, S. 2002. Nltk: the natural language toolkit. *arXiv preprint cs/0205028*.

[Manning et al. 2014] Manning, C. D.; Surdeanu, M.; Bauer, J.; Finkel, J.; Bethard, S. J.; and McClosky, D. 2014. The Stanford CoreNLP natural language processing toolkit. In *Association for Computational Linguistics (ACL) System Demonstrations*, 55–60.

[NGSS 2016] NGSS. 2016. Next generation science standards: 5-LS2 ecosystems. https://www.nextgenscience.org/, Last accessed on 2019-11-15.

[Papert and Solomon 1971] Papert, S., and Solomon, C. 1971. Twenty things to do with a computer. artificial intelligence memo number 248.

[Payne 2019] Payne, B. H. 2019. An ethics of artificial intelligence curriculum for middle school students. https://docs.google.com/document/d/1e9wx9oBg7CR0s5O7YnYHVmX7H7pnITfoDxNdrSGkp60/edit?usp=sharing, Last accessed on 2019-09-04.

[Resnick et al. 2009] Resnick, M.; Maloney, J.; Monroy-Hernández, A.; Rusk, N.; Eastmond, E.; Brennan, K.; Millner, A.; Rosenbaum, E.; Silver, J. S.; Silverman, B.; et al. 2009. Scratch: Programming for all. *Commun. Acm* 52(11):60–67.

[Segedy, Kinnebrew, and Biswas 2012] Segedy, J.; Kinnebrew, J.; and Biswas, G. 2012. Supporting student learning using conversational agents in a teachable agent environment.

[Touretzky et al. 2019] Touretzky, D.; Gardner-McCune, C.; Martin, F.; and Seehorn, D. 2019. Envisioning ai for k-12: What should every child know about ai? *Proceedings of the AAAI Conference on Artificial Intelligence* 33:9795–9799.

[Toutanova and Manning 2000] Toutanova, K., and Manning, C. D. 2000. Enriching the knowledge sources used in a maximum entropy part-of-speech tagger. In *Proceedings of the 2000 Joint SIGDAT conference on Empirical methods in natural language processing and very large corpora: held in conjunction with the 38th Annual Meeting of the Association for Computational Linguistics-Volume 13*, 63–70. Association for Computational Linguistics.

[Van Brummelen 2019] Van Brummelen, J. 2019. Tools to create and democratize conversational artificial intelligence. Master's thesis, Massachusetts Institute of Technology, Cambridge, MA.

[Williams, Park, and Breazeal 2019] Williams, R.; Park, H. W.; and Breazeal, C. 2019. A is for artificial intelligence: The impact of artificial intelligence activities on young children's perceptions of robots. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 447. ACM.

[Wu and Palmer 1994] Wu, Z., and Palmer, M. 1994. Verbs semantics and lexical selection. In *Proceedings of the 32nd annual meeting on Association for Computational Linguistics*, 133–138. Association for Computational Linguistics.

[Zhu 2019] Zhu, K. 2019. An educational approach to machine learning with mobile applications. Master's thesis, Massachusetts Institute of Technology, Cambridge, MA.