Using Conversational Agents to Support Science Learning From Animated Television Shows

Ying Xu (ying.xu@uci.edu) and Mark Warschauer, University of California, Irvine

Science-oriented television programs can be an important source of STEM learning for young children. However, the educational benefits of these programs have long been limited by children not being able to interact with the content in a meaningful way[1]. One common approach to providing children with interaction during their television/video watching is to embed conversational prompts in the media sources themselves, as exemplified in *Dora the Explorer*. These conversational features typically consist of a character creating faux eye contact, asking the viewer a question, pausing a set amount of time, and responding in a way that does not actually acknowledge the viewer's specific answer. One major limitation of such conversational features is the lack of contingency and responsiveness, which may discourage children from answering the questions or lead children to misgauge their own learning comprehension or doubt the character as a reliable source of information[2].

With the rapid development of conversational agents (CAs), however, contingent interaction between children and media is now possible. CAs have the affordances to understand unconstrained natural language input, allowing for complex dialogue and potentially mimicking human-to-human spoken conversation. These technologies are now prevalent in many homes, and children readily interact with and accept these digital conversational partners in their daily lives. This points to the feasibility of integrating CAs into educational television programming.

This research uses a CA to enable verbal interaction between children and the main character in a children's science animation series (Figure on the left), who can play the role of an engaging and responsive video co-viewer for children. The verbal interaction is designed to promote active viewing, in which children are "minds on," or actively thinking and reasoning rather than passively watching the video. In addition, socially interacting with onscreen characters provides a fun experience, which may enhance children's enjoyment and motivation during video watching.

Summary of System

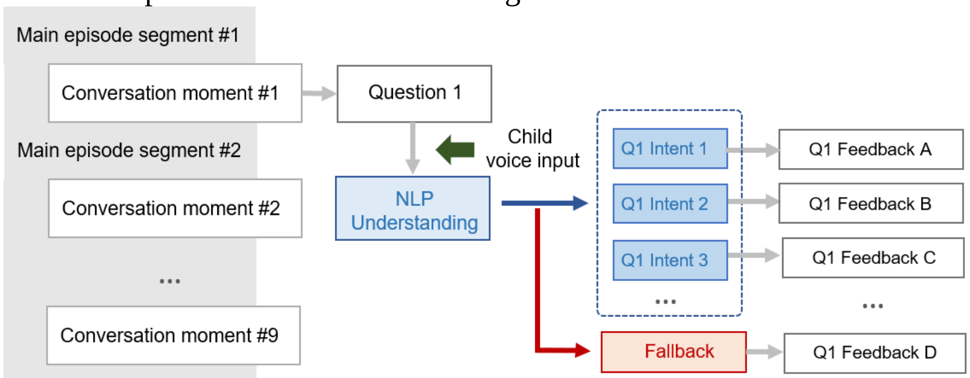
We have developed a mobile application consisting of multiple episodes integrated with conversational agents based on a popular PBS KIDS science animation, "Elinor Wonders Why." Each episode contains about ten conversation moments, in which the show's main character, Elinor, asks children questions and gives feedback to children's responses.

Elinor Initiating Prompts

The conversational prompts, or questions, are aligned with the Next Generation Science Standards and can be categorized into one of three types: small talk, fact-based questioning, and problem solving. The small talk prompts are included at the beginning of the episode and evoke children's curiosity about the scientific concepts while relating the concepts to children's everyday life experiences. In our video, Elinor opens up the conversation by asking children whether they like honey and how it feels when they touch it. Fact-based questions seek to clarify children's understanding of basic scientific facts or explain scientific terms when they are introduced in the episode. For example, Elinor asks children how bees turn nectar into honey. Lastly, the prompts ask children to apply what they have learned in the video to help solve a similar problem. Most questions were asked in an open-ended format (i.e., starting with "what," "how," "why"), as research suggests that open-ended questions encourage children to put in cognitive effort and express their thinking more freely as compared to closed questions that pose options for children to choose from.

Elinor Providing Feedback

The CA underlying Elinor, built on Google's cloud services, performs end-to-end language processing that classifies speech utterances into semantic intents. The agent's natural language understanding module was based on Google's generic pre-trained models, then refined with utterances specific to the conversational moments in our video. With a small set of training utterances, the agent can naturally extract the semantic meaning of these training phrases and thus understand other phrases with similar meanings.



For each conversational prompt, we defined intents that we want Elinor to classify children's utterances into (Figure above). Given that children can respond to a particular question in a variety of distinct ways, each opportunity for a child to respond is usually associated with more than one predefined intent. These intents were created based on the predictions of the research team, as well as on the field-testing mentioned above. During this field-testing, a human experimenter followed the same dialogue script used by Elinor but paused the video and asked the children the question in the script. Even with the field-testing, however, it is unlikely that we could exhaust all the possible responses due to the open-endedness of the questions. To address this, we included a fallback intent that is triggered if a child's utterance does not match any of the predefined intents or if the child does not respond to the prompt at all. This fallback intent is important for preventing breakdowns and ensuring the conversation flow will be maintained regardless of whether the child's response can be classified. Elinor then provides feedback based on the child's intent as categorized by the agent following the child's response.

Potential Impact and Relevance to Equity and Inclusion

We have carried out a randomized control trial to evaluate the efficacy of this kind of interactive videos with children aged four to six years from both high and low socioeconomic households. We found that integrating a conversational agent into science animations enhanced children's understanding of the science concepts introduced in the video through eliciting children's relevant conversation around the STEM topics. This finding is particularly relevant for children from under-resourced households because television programs, especially those aired on public broadcasting services, have been valuable and accessible learning resources. As such, this application has the potential for addressing the gap in early science learning.

[1] Strouse, G. A., & Samson, J. E. (2021). Learning From Video: A Meta-Analysis of the Video Deficit in Children Ages 0 to 6 Years. *Child Development, 92*(1), e20-e38.

[2] Troseth, G. L., Saylor, M. M., & Archer, A. H. (2006). Young children's use of video as a source of socially relevant information. *Child Development, 77*(3), 786-799.