Yang, J. C. et al. (Eds.) (2018). Proceedings of the 26th International Conference on Computers in Education. Philippines: Asia-Pacific Society for Computers in Education

A Dialogue Model for Collaborative Storytelling with Children

Dionne Tiffany ONG[a], Christine Rachel DE JESUS, Luisa Katherine GILIG, Junlyn Bryan ALBURO & Ethel ONG[b]
De La Salle University - Manila, Philippines
[a]dionne_ong@dlsu.edu.ph, [b]ethel.ong@dlsu.edu.ph

Abstract: Storytelling is both an entertaining and an educational activity for children. It allows them to express their thoughts about the world around them, while learning about language, values and interacting with others. Thus, storytelling is utilized not only at home but also in the school environment to engage children in sharing information about things that interest them and their everyday life experiences. Conversational agents or chatbots are becoming popular as a form of human-computer interaction. They can engage with their human users in dialogue using natural language to answer queries. In this paper, we describe the design of our dialogue model to enable a conversational agent to act as a collaborative storytelling peer, encouraging children to share their stories through pumps and hints.

Keywords: Conversational Agent, Dialogue, Collaborative Storytelling

1. Introduction

Storytelling helps in enhancing children's communication skills, and develops their vocabulary and language (Isbell et al., 2004) as well as comprehension (Mokhtar et al., 2011). Children are encouraged to articulate their thoughts and feelings, and to express themselves, their preferences, aspirations and reflections (Mantei & Kervin, 2010), through their stories.

In group storytelling sessions, the narrator and the listeners share their personal recollections and interpretations (Lawrence & Thomas, 1999), and the storytelling session evolves through the exchange of viewpoints about character motivations, interpretations of events occurring in the story, and the rationale for character responses. This collaborative nature of storytelling can be achieved in computational systems through virtual agents.

From merely telling stories to children, the virtual agent can become a more intelligent partner by engaging in conversations to define the characters and the setting, rationalize a character's actions in relation to his/her motives, imagine the varying sequences of events that may come next, and even relate one's personal experiences to that of the character. Collaboration with a virtual peer further extends opportunities for children to acquire language skills needed for literacy, and helps develop them to be critical listeners of others' stories (Ryokai, Vaucelle, & Cassell, 2002).

Sam is a childlike androgynous conversational agent that employs a turn-taking model to share and listen to children's oral storytelling (Cassell et al., 2005). Storyfighter (Lieberman et al., 2004) uses a virtual peer to co-author stories with children, where the peer proposes the start and end state of the story to be told, and the child contributes to the story by selecting a sentence template form a list generated by the system, then filling this with words to form the story text.

For collaboration to take place, communication among the entities in the form of dialogues is necessary. Dialogue exchange is characterized in terms of communicative goals or speech acts, such as *inquiry* or *direct question* to solicit additional information, *informing* to respond to a request for information, *elaboration* to provide additional definitions or descriptions, *justification* to explain actions, *motivation* to persuade someone to carry out an action, *exemplification* to model how the task can be carried out, and even *repair* to resolve misunderstanding.

Conversational agents (Milne et al., 2011) are equipped with the capability to facilitate a conversation by holding a dialogue with the user. In the study of (Doering et al., 2008), they found out that initial conversations between the students and the conversational agents revolve around

topics outside the scope of the learning task. This shows that the students perceive the agents both as a learning companion and a social agent. While the agents are not always accurate in addressing complex content-related questions posed by the learners, they were able to provide assistance that encouraged the students to continually consult with them throughout the learning experience.

In this paper, we describe the design of the dialogue model of our conversational storytelling agent, Orsen. Orsen is designed to engage children in conversation to elicit the elements of a story, namely the characters involved, the setting where the story takes place, and the sequence of events. It is built on top of Google Home, to support oral rather than written communication.

2. Dialogue Model

A collaborative system requires a flexible dialogue model that can support the flow of conversation between the virtual storytelling agent, Orsen, and the human user. Orsen should be able to adapt and to respond to a variety of inputs when engaging the user in storytelling, such as acknowledging the story text (Mandelbaum, 2013), and asking for more details about the story elements.

The types of input provided by the user during its dialogue turn were categorized into four, namely answer, silence, command and story text. This is illustrated in Figure 1. Children's responses to Orsen's questions, which include "*yes*", "*no*" or "*ok*", are classified as answer. Any user input that contains the keyword "*Orsen*" or "*Orson*" is flagged as a command for the agent to perform. When Orsen does not hear the child's input, this dialogue turn is flagged as silence. All other text input is categorized as story text and will be processed as such.

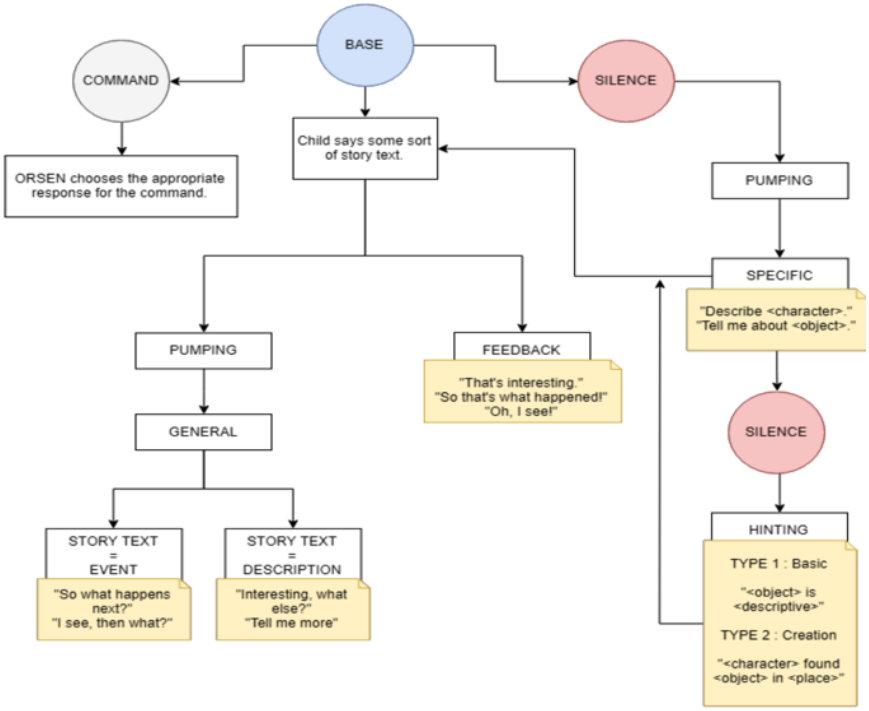

Figure 1. Dialogue Model for four Input Categories: *Answer*, *Silence*, *Command* and *Story Text*.

2.1 Dialogue Moves

Orsen can employ different speech acts during its dialogue with the users. Taxonomy of children's collaborative speech acts has been identified in the study of Wang and Cassell (2003). This includes suggesting an event, seeking clarification, asking for missing information, acknowledging a suggested storyline, playing the role of a story character, and even competing for a turn in the conversation.

Certain dialogue strategies can also promote a collaborative environment for children. These include feedback, pumps and hints, as identified in (Graesser et al., 1999). Pumps are used

when the user provided a story text that is lacking in details, e.g., the location where the event takes place, the objects used in performing an action, and even descriptions of characters and objects. Pumps work by encouraging the user to express additional information that can expand the story, such as describing the character's attributes or the consequential event. Hints, on the other hand, are used when the user suffers from writer's block. They work by stating a fact or adding details related to the previous story text that the user has already shared to help trigger ideas.

Orsen depends on the child's input to decide on its dialogue move. If the input is a story text, Orsen may choose to do either a feedback or a general pump to acknowledge and prompt the child to continue with his/her story. Statements such as "*Interesting*" and "*Okay*" are examples of feedback, while "*Tell me more.*" and "*What happens next?*" are considered to be general pumps. On the other hand, if Orsen encounters silence, that is, no input was given, such as the case when the child could be encountering some difficulty in continuing with his/her story, a specific pump or a hint may be employed. A specific pump is a type of pump that asks the child to expand on a certain part of the story. Listing 1 illustrates a sample conversation.

Listing 1. Example conversation that uses feedback, pump and hint.

User:	There was a princess.
Orsen:	I see. (*feedback)*
User:	Her name was Maddie.
Orsen:	And then what? (*general pump*)
User:	**silent**
Orsen:	Tell me more about the princess. (*specific pump*)
User:	**silent**
Orsen:	How about "The princess went to the castle."? (*hint*)

2.2 User Validation

To validate the dialogue model, five (5) children age 7 to 11 years old interacted with Orsen through the Google Home platform. Orsen's responses were limited to acknowledging the child's input. The following are the findings. First, children tend to introduce themselves to Orsen before starting their story, as a response to Orsen's self-introduction. Since the agent has not been designed to process an introduction, it would treat introductory text as regular story text and responds with moves like the general pump, i.e., "*What happens next?*" This leads to the child's confusion or amusement at Orsen's response, often ending the storytelling session before they even started.

Second, children tend to ignore the agent's generic and neutral replies as they find these to be predictable. Responses such as "*Okay*", *"I see*" and "*All right*" were not very effective in keeping the children engaged since their inputs were not really acknowledged by these generic responses.

Third, some children tend to speak continuously without relinquishing the dialogue turn back to Orsen. This is considered as a non-silence type of input, triggering Orsen to generate feedback and general pumps. Specific pumps and hints, which give Orsen the ability to be a collaborative storytelling partner, are never generated in these instances.

Fourth, the vague wordings used in Orsen's dialogue would lead children to misunderstand the agent's request for them to repeat their statements. The children would respond by repeating their story from the beginning. This is a problem because Orsen did not realize that the story was being retold. Orsen maintains a story world model to track elements of the story that the child has already shared, such as the story setting (location, time), the different story characters, and the sequence of events or character actions. These elements are extracted from the user's input and stored as assertions. If the child ended up re-telling the story in a slightly different manner, it could lead Orsen to extract duplicate assertions from the user's input that already exist in the story world model.

Lastly, the children would sometimes panic when Orsen states that it did not hear the statement. The child would respond in one of two ways: he/she either tries to speed through the narration of the story to end the session faster, or to shut down and not speak to Orsen until the latter automatically ends the session.

3. Response Generation

The dialogue model was revised to handle the different concerns identified during user validation. The input category, *introduction*, was added to enable Orsen to include a greeting and a question to start the story. The dialogue move to handle inputs in the *silence* category is also expanded as shown in Figure 2. If the previous response is a general pump, it triggers a specific pump. On the other hand, if the previous response is a specific pump, a hint is triggered.

To enable Orsen to provide relevant feedback to a child's input, it makes use of an ontology that has been populated with concepts and their semantic relations retrieved from ConceptNet (Speer & Havasi, 2012), vocabulary lists of common nouns (ESLDesk, 2010) and verbs (Rundell, 2003), and fables. These are stored as assertions of the form *[concept1 relation concept2*, and are used to define real-world entities, such as *[bread IsA food]* and *[king IsA ruler*]; describe their attributes, such as *[bread HasProperty soft]* and *[king HasProperty brave]*; and depict event relations, such as *[study AtLocation library]*. Orsen uses the assertions in the ontology and story world knowledge extracted from the child's inputs to create responses for its dialogue moves. Examples are provided in Table 1.

Stories contain characters, objects and locations where events take place. Orsen tries to extract these elements from the user input by utilizing the knowledge it gained to formulate a response. As shown in examples #1 and #2 in Table 1, Orsen identified and utilized two story world objects, i.e., *bear* and *forest*, as the main subject of its *pump* dialogue move. It then generates pumps that can encourage the user to provide further details regarding these objects.

In example #1, Orsen wants the user to supply missing details to complete the assertion *[character CapableOf event-verb]*. The agent can determine that the character is sentience, and as such, has the ability to perform some action. In example #2, Orsen needs information to complete a descriptive assertion *[object IsA description]* so that it can increase its knowledge about the object.

Table 1. *Example responses generated by Orsen.*

User Input	Orsen's Response	Dialogue Move
#1 The bear ran from the hunter.	What happens next to the bear?	Pump
#2 The bear lived in a forest.	Tell me more about a forest.	Pump
#3 The bear like to sing.	Then the bear is eating honey.	Hint
#4 Ben is a student.	Then, Ben went to school.	Hint
#5 My mom likes movies.	My mom is a good person.	Hint

Stories also contain sequence of events, represented in the story world model as an event chain. Example #3 illustrates Orsen's ability as a storytelling partner by utilizing the *hint* dialogue move that suggests story text to add onto the story's event chain. Since previous user input already established the *bear*'s sentience, the agent uses the semantic ontology to find assertions that depict what a sentient character is capable of doing, using the *CapableOf* relation.

Example #4 shows another function that storytelling partners should be able to do, that is, to relate one concept with another. Because the user provided story text that classifies the character, *Ben*, as a *student*, Orsen's task is to find assertions containing concepts related to the concept *student*, following this assertion chain:

[Ben IsA student] ➔ *[student AtLocation school*]
assertion from user input assertion from the semantic ontology

Given this, Orsen can generate the response *Ben went to school*. Orsen uses the same principle to generate attributes describing a character or object, as illustrated in Example #5.

4. Further Work

Orsen is a collaborative storytelling agent that is capable of processing user input in order to generate responses that encourage the user to continue with his/her story and to suggest story text. At this stage, the agent is able to understand a variety of sentence structures and perform coreference

resolution. These structures include simple active and passive sentence forms, prepositional phrases, conjugated attributes or clauses, clausal complements, and possessive cases.

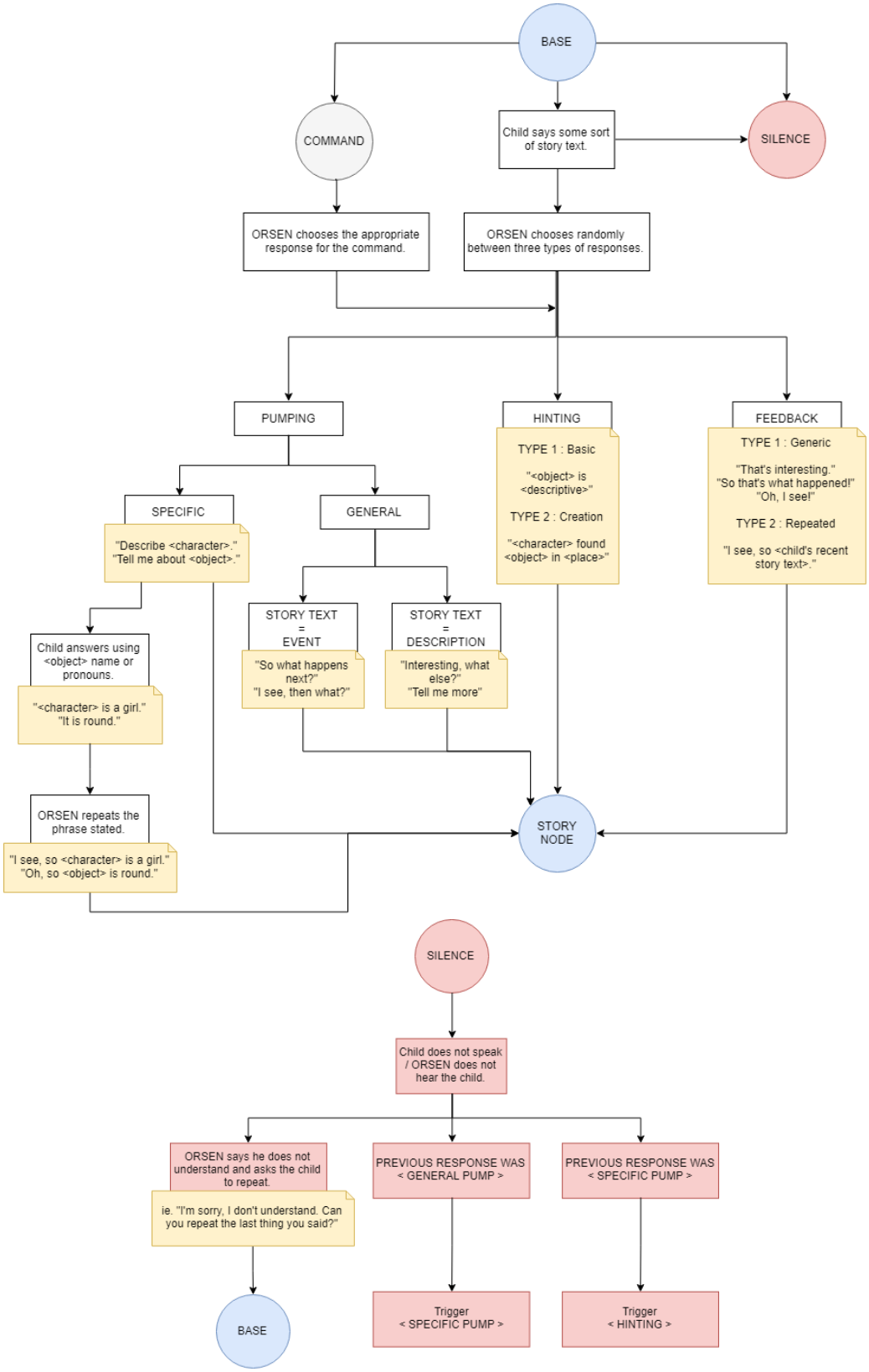

Figure 2. Expanded Dialogue Model to handle User's *Silence*.

Orsen is also able to generate responses by utilizing an ontology populated with relevant concepts and assertions from ConceptNet, basic vocabulary lists and fables. Three dialogue moves

are utilized, specifically feedback, pump and hint, to process user inputs that include introductory text, commands, silence and actual story text. These dialogue moves aim to encourage collaboration at varying levels of detail, from the very basic pump or feedback which can work even with minimal story information, to the more specific pumps that ask for additional information about story elements, to full hints where the agent suggests its own story text to help the user move the story forward.

Orsen's platform and dialogue model have undergone a series of preliminary testing to validate the children's responses to a virtual peer in a collaborative setting. Our results showed that most children are able to accept the virtual peer after few trials. Further evaluation has to be performed on the latest iteration of the dialogue model, to assess Orsen's input understanding and text generation capabilities. Speech patterns of children derived from this evaluation can also be incorporated to enhance the agent's dialogue turns. Additionally, adjusting the scope of the semantic ontology into a more specific domain of children's stories will be explored, so that Orsen's dialogue turns would contain responses that are less general, and more attuned to the user inputs.

References

Basic ESL Vocabulary, http://www.esldesk.com/vocabulary/basic, last accessed 2018/04/10.

Cassell, J., Tartaro, A., Rankin, Y., Oza, V., & Tse, C. (2005). Virtual Peers for Literacy Learning. *Educational Technology special issue on Pedagogical Agents*, 47(1), pp. 39–43.

Doering, A., Veletsianos, G., & Yerasimou, T. (2008). Conversational agents and their longitudinal affordances on communication and interaction. *Journal of Interactive Learning Research*, 19(2), 251–270. Waynesville, NC: Association for the Advancement of Computing in Education (AACE).

Graesser, A. C., Wiemer-Hastings, K., Wiemer-Hastings, P., Kreuz R., & Group, T. R. (1999). Autotutor: A simulation of a human tutor. *Cognitive Systems Research*, 1(1), pp. 35–51.

Isbell, R., Sobol, J., & Lindauer, L. (2004). The effects of storytelling and story reading on the oral language complexity and story comprehension of young children. *Early Childhood Education Journal*, 32(3), p. 157–163. Springer Science + Business Media, Inc.

Lieberman, H., Liu, H., Singh, P., & Barry, B. (2004). Beating some common sense into interactive applications. *AI Magazine*. AAAI Press.

Mandelbaum, J. (2013). Storytelling in conversation. In Sidnell, J., Stivers, T. (Eds.), *The Handbook of Conversation Analysis*, pp. 492–507. Blackwell Publishing Ltd.

Mantei, J., & Kervin, L. (2010). This is me! Empowering children to talk about their learning through digital storytelling. In Kanganas, A. (Eds.), *aWAy with Words*, pp. 1–15. Adelaide: South Australia: ALEA/AATE.

Milne, M., Luerssen, M., Lewis, T., Leibbrandt, R., & Powers, D. (2011). Embodied conversational agents for education in autism, Chapter in *A Comprehensive Book on Autism Spectrum Disorders*, pp. 387–412. InTech Open.

Mokhtar, N. H., Halim, M. F. A., & Kamarulzaman, S. Z. S. (2011). The effectiveness of storytelling in enhancing communicative skills. *Procedia Social and Behavioral Sciences*, 18, 163–169. Elsevier Ltd.

Rundell, M. (2003). Macmillan essential dictionary: For learners of English. Macmillan.

Ryokai, K., Vaucelle, C., & Cassell, J. (2002). Literacy learning by storytelling with a virtual peer. In *Proceedings of the Conference on Computer Support for Collaborative Learning: Foundations for a CSCL Community*, pp. 352–356, Colorado.

Speer, R., & Havasi, C. (2012). Representing general relational knowledge in ConceptNet 5. In Choukri, K., Declerck, T., Dogan, M.U., Maegaard, B., Mariani, J., Odijk, J., Piperidis, S. (Eds.), *Proceedings of the 8th International Conference on Language Resources and Evaluation* (LREC'12), Istanbul, Turkey (2012).

Wang, A., & Cassell, J. (2003). *Co-authoring, corroborating, criticizing: Collaborative storytelling for literacy learning.* In *Proceedings of the Vienna Workshop '03: Educational Agents - More than Virtual Tutors*, Vienna, Austria.