The Automatic Generation of Children's Stories from a Multi-Scene Input Picture

Karen Ang[1], Janine Antonio[2], Dan Sanchez[3], Sherie Yu[4], and Ethel ChuaJoy Ong[5]

College of Computer Studies, De La Salle University
Manila, Philippines

karenang0903@yahoo.com[1], janine.antonio@yahoo.com[2], dansanchez.dlsu@yahoo.com[3], sherie_yu@yahoo.com[4], ethel.ong@delasalle.ph[5]

ABSTRACT

For children, the act of telling stories develops their language and literacy skills. With this in mind, the Picture Books story generation system was developed to produce children's stories from a single-scene input picture comprising of the elements (i.e, characters and objects) that were provided by a child, to impart that picture stories have textual counterparts. However, this may result in the production of stories wherein the flow of events do not necessarily match the actual idea of the child when he/she conceptualized the picture story, since the system randomly selects from a predefined list of themes associated with the selected background. Because people tell stories as a sequence of events to share their daily activities and experiences, Picture Books 2 has been developed to provide an environment where a child can creatively define a sequence of scenes depicting the flow of major events in his/her target story. The task of the story planner then is to figure out how to connect this sequence of scenes to form a single coherent and interesting storyline of the fable form for children age 6-8 years old.

Keywords

Story Generation, Semantic Ontology, Storytelling

1. INTRODUCTION

Stories are an essential part of human life. When people tell stories to one another, they are able to share their beliefs, their experiences, their culture, their knowledge and many more. In children, storytelling develops language and literacy skills and provides an enjoyable activity where they can exercise their imagination to creatively produce their own stories.

The popularity of children's storybooks is their inherent ability to relate stories using pictures and words. In their book "Let's Begin Reading Right", Fields and Spangler [1] studied how storytelling and literacy development of children begin through picture-telling, or narrating a story from a drawn picture. This is the motivation behind Picture Books [2], which automatically generates stories of the fable form for children 4-6 years old. The input to the system is a picture containing components that were selected by the child from a library of background images, and character and object stickers. Based on these story elements, the story planner then selects a theme and generates a story reflecting the flow of events from negative to positive, where a child violates a stated rule and learns the consequence. The stories are set in settings common for young children, such as different rooms in the house, school and park, and involve plots that promote good values, like being honest and being brave in new situations. All stories involve a main child character who performs various actions and experiences various emotions and events as the story unfolds. A secondary adult character is also present to provide guidance and support to the child through his/her process of acquiring a moral lesson.

One limitation of Picture Books is that the generated story may not match the intended story of the child when he/she created the input picture-story. This is because the major scenes in the story, subdivided into problem, rising action, resolution and climax [4] are defined according to the theme that has been randomly assigned by the theme planner based on the elements of the story found in the input picture.

Another concern with Picture Books is that the characters are not embodied with traits which may affect the flow of the story. Having animals as the system's characters, child educators believe that enabling these characters to possess traits would help kids relate to the story better because they can convey themselves with the characters. Riedl [7] further notes that character believability is an essential property of narratives because the events that occur in the story are motivated by the beliefs, desires and the goals of the characters.

Picture Books 2 is an extension to Picture Books and provides an environment where a child can creatively define an input picture consisting of a sequence of scenes that comprise his/her intended story. The target age group consists of older children age 6-8 years old, who are more independent and learn about life's lessons on their own. They are also more creative in their storytelling and can develop three or more scenes depicting various stages of a story. Because embodying the story characters with traits can help the children to relate to the story better, each character in the system has been assigned three character traits which are also utilized by the story planner in developing the story plan.

A major task of the story planner is to figure out how to connect the input sequence of scenes to form a single coherent and interesting storyline that still flows from negative to positive, and still promotes moral values, this time set in more adventurous places like the camp, park or street. A number of issues have to be considered to facilitate the flow of the story from one scene to the next, specifically on character and object existence and movement across two adjacent scenes. Existence deals with the question of whether or not a particular object that is present in one scene is also present in the next. Movement, on the other hand, deals with the change in position of either a character or an object from one scene to the next.

This paper is subdivided as follows. Section 2 presents an overview of the necessary storytelling knowledge needed by Picture Books 2. Section 3 then presents the system's architectural design and discusses the story planning process. Section 4 presents an evaluation on the system's output. The paper ends with a summary of what we have achieved thus far and further work that can be done in the field of story generation.

2. KNOWLEDGE SOURCES

The process of telling a story requires a large body of knowledge including the narrative concepts on how a story is to be told, as well as knowledge about the story world (the various settings and objects in the story world), character representation (traits, emotions, behavior), and a causal chain of actions and events. Actions are performed directly by characters, while events occur as a result of some actions, another event, or as a natural occurring phenomenon.

The events in a story are the result of the complex interplay of the characters using the objects in the story world environment. Thus, the knowledge representation should model the concepts that exist in the real world as defined by the properties and relationships that exist among them. This knowledge base is then utilized by a story planner to perform its tasks.

Picture Books has a semantic network of commonsense concepts about objects and activities in a child's daily life [6]. These concepts are connected by semantic relations that are patterned after those defined in ConceptNet [3], but the actual contents have been derived in order to provide concepts that are not only familiar to children but are also relevant to the system's themes.

Similar to its predecessor, Picture Books 2 also models the following basic storytelling concepts in its semantic ontology – i) the objects that may exist in the story world and which the character may use in his/her pursuit of his/her goal; ii) the attributes of and the locations where objects can be found in the environment; iii) the actions that a character may perform in order to reach a desired goal or state, or as a response to the occurrence of some events; iv) the events that may occur in the story world as a result of some character actions (explicit events) or as a naturally occurring phenomenon (implicit events); and v) abstract concepts such as conflict, discomfort, and resolution.

Aside from these, the ontology of Picture Books 2 has also been populated with additional concepts based on new storytelling requirements included in the scope of the system. New semantic relations were defined in order to model concepts relating to existence, movement, and orientation or positioning of characters and objects in a scene. While maintaining the binary structure of the semantic relations found in ConceptNet, the new relations include *isTransition* which indicates whether an action implies an appearance, disappearance or movement of an object between the scenes created by the user; *hasResolution* which specifies the goal to be achieved or the solution of the conflict that arose in the story; *leadsTo* which indicates the possible story conflicts that may arise due to the object(s) present in the scenes; and *causesConflictOf* which denotes an action carried out by the main character that will spark the conflict in the story and is used to define the plot of the whole story.

Some relations were also enhanced for clarity, such as the *feels* relation used for conveying a character's emotional state, and the *receivesAction* relation used for specifying the action that can be applied on an object. The *receivesAction* complements the existing *capableOf* action which is used to model the actions that a character can perform. The causal chain of events is now modeled by a single relation, *effectOf*. An additional *hasSubevent* relation has been added to specify the optional event which may happen before another event can take place.

Because the input picture is now comprised of at least three scenes, the story planner has to determine changes that may have happened across two adjacent scenes. These transitions are of two types, the movement and the appearance or disappearance of characters and/or objects (collectively referred to as *stamps*).

To track movements of a stamp across two adjacent scenes, each of the four available backgrounds has been subdivided into 36 grids. *Grid* concepts are then used to represent the different sections of the background. For example, the *camp* background, shown in Figure 1, is subdivided into three *grid* concepts – forest, bonfire and tent.

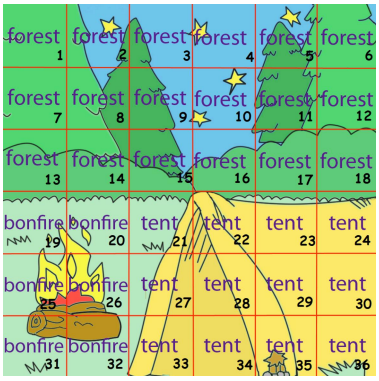

Figure 1. Camp Background subdivided into Grids

Six personality *trait* concepts – brave, honest, persevering, helpful, obedient, and responsible – were defined to convey a character's trait, and are used to relate to the actions that can only be performed by a character possessing a specific trait.

3. ARCHITECTURAL DESIGN

The process of generating stories in Picture Books 2 proceeds in four main stages, commencing with the specification of an input picture with at least three scenes in the Story Editor environment. The Story Planner then generates a story plan based on this input multi-scenes picture. A story plan contains the chronological enumeration of actions and events that occur in the story world. The Sentence Planner maps the story plan concepts into their corresponding character goals and applies the rhetorical structure theory to relate two or more character goals together to form coherent sentences. Finally, the Surface Realizer transforms these character goals into actual English text.

3.1 Story Editor

The Story Editor, shown in Figure 2, provides an environment for the user to specify the basic elements of his/her story. Once a background representing the setting of the story world has been selected (Figure 2a), a sequence of at least three scenes must be specified by the user through the placement of character and object stamps onto the background (Figure 2b). Because a background has 36 grids, the system can track the location of each

character and each object stamp within a scene. Movement is recognized when a stamp is located at a grid position that is different from the previous scene.

The Story Editor performs the following validation to ensure that certain restrictions are met in the input picture:

- The first scene should have at least one character stamp.
- The first scene should have at least one object stamp.
- There must be at least one character stamp in each of the scenes at all times.

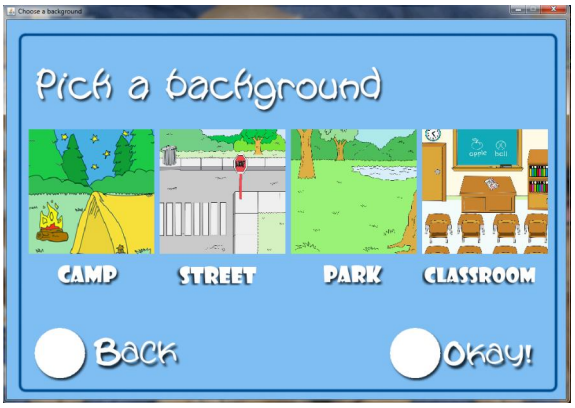

Figure 2a. Story Editor – Background Selection

Figure 2b. Story Editor – Scene Creation

When an input picture has been validated, relevant information namely the characters, the objects and their positions, are extracted from each scene and stored in an abstract story representation, as shown in Table 1. Each scene contains facts that model the state of the world, including the name of the main character and his/her grid position in the background, and the name of the object and its grid position in the background. Each scene also includes the kind of transition that has been detected by the Story Editor. The main character is the first character that the user places on the first scene.

3.2 Story Planner

The Story Planner uses the abstract story representation to generate a story plan. A story plan contains specifications of the actions and events that take place in the story world, beginning with an introduction of the setting, followed by the conflict that a character faces and must resolve as the story progresses, and ending with a resolution of how the character achieves the goal of the story. In the case of Picture Books 2, the solution to the identified conflict is through the acquisition of a new moral lesson either in the form of acquiring a new trait or reinforcing an existing trait.

Table 1. Abstract Story Representation

Scene Number: **0**
Name: ***Danny***; gridNumber: ***25***; trans: ***appearance***
Name: ***marshmallow***; gridNumber: ***25***; trans: ***appearance***

Scene Number: **1**
Name: ***Danny***; gridNumber: ***28***; trans: ***movement***
Name: ***flashlight***; gridNumber: ***28***; trans: ***appearance***

Scene Number: **2**
Name: ***Danny***; gridNumber: ***28***; trans: ***null***

The Story Planner performs three main activities, as shown in Figure 3, to derive the story plan, namely theme formulation, setting formulation, and events generation. The theme of the story is identified based on the background, character, and objects present in the conflict or middle scene. Each theme defines a character trait that the main character lacks, and focuses on developing this trait as the story progresses.

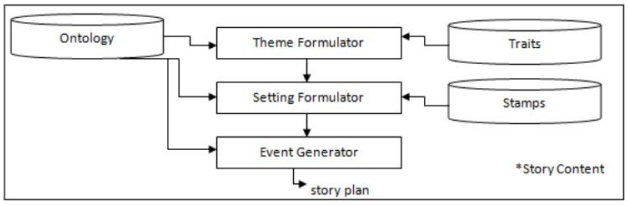

Figure 3. Architectural Design – Story Planner Module

The setting of the story is constructed based on the first scene. The setting describes the location of the story and the time the story takes place. In the case that the first scene contains an object located in the same grid as the character, the setting formulator also introduces that object. The description of the location, i.e., the activity that can be performed in that location and an adjective describing the location, as well as the time that the story should take place are determined by randomly selecting from the available knowledge in the ontology, such as those shown in Tables 2 and 3.

Table 2. Ontology Concepts for Background and its Purpose

Concept 1	Relation	Concept 2
Camp	UsedFor	Camping
Park	UsedFor	Picnic
Class	UsedFor	Learn

Table 3. Ontology Concepts for Background and its Property

Concept 1	Relation	Concept 2
Camp	HasProperty	Far
Camp	HasProperty	Crowded
Park	HasProperty	Clean
Class	HasProperty	Quiet

Events generation involves searching the ontology for a story path comprising of conceptual relations starting from the identified theme (or conflict) to the selected background property. This path of cause-and-effect relations comprises the possible sequence of actions and events that may take place in the story. Nodes included in the story path are validated to be considered as effects. Effects may or may not require a character to perform the action. If the event requires a character action, then the corresponding action concept must also be present as a node in the story path. The same condition holds when validating the object. The object present in the scene should be related to the effect, that is, some action must be done to the object to achieve the effect. There are also certain effects that do not require an object, such as a character *singing*.

The action associated with the effect is also validated if it can be performed on a specific grid and/or background. For example, in the camp background shown in Figure 1, the *sleep* action can only be performed by a character who is in the *tent* portion of the grid, and not anywhere else.

3.3 Sentence Planner

The Sentence Planner transforms the sequence of conceptual relations in the story plan into a set of character goals. One character goal represents one action or one event. The design of the character goal is based from the action operators of Uijlings [9] comprising of five fields – the *action*, the *agens* or doer of the action, the *patiens* or receiver of the action, the *target* or the location or object of the action, and the *instrument* used to perform the action. This design allows a character goal to be easily transformed into a declarative sentence in active voice with the use of the surface realizer simpleNLG [10].

Because the main premise of Picture Books 2 is that a story is simply a chain of cause-and-effect relations between character actions and events, there are two general types of character goals defined in the system, namely the *capableOf* character goals and the *feels* character goals. A *capableOf* character goal represents an action that a character can perform and contains the following:

Agent:	noun/CHARACTER
Verb:	the associated verb concept in the *capableOf* relation
Patien:	optional receiver of the action (from the *receivesAction* relation)
Location:	optional place where the action is to be performed (either the specific grid or the background)

A *feels* character goal represents an emotional response of a character as a result of the occurrence of an event in the story world. It contains the following:

Agent:	noun/CHARACTER
Verb:	'*feels*'
Patien:	the associated emotion concept in the *feels* relation
Location:	Null

The Sentence Planner also takes care of aggregating two or more character goals to form more coherent sentences with the use of the Rhetorical Structure Theory (RST) of Mann and Thompson [5]. RST states that every sentence in a paragraph is connected with each other to define the logical flow of transition from one sentence or phrase to another and to make the coherence between sentences possible. This connection can be made explicit with the use of discourse markers, which can elaborate a sentence, give a time element to the sentence, signal a conclusion and many more. Taylor [8] provided a list of common discourse makers. Table 4 lists those that are appropriate for elementary age kids.

Table 4. Common Discourse Markers

Used to signal	Transition Word
Addition	Also, again, and, besides
Time	After, before, during, later, now, then
Cause or Reason	Because, since
Effect	Because, hence, so, thus
Direction	Above, behind, below, between, near
Summary	So, thus

Sentence planning also involves performing lexicalization and generating referring expressions. Most of these algorithms were adapted from Picture Books [2]. Picture Books 2 recognizes three different types of articles namely, definite, indefinite and personal. An article type is assigned for every noun or character concept in the character goal. Definite articles, '*the*', are used to refer to definite nouns or specific objects. Indefinite articles, '*a*' or '*an*', are used to refer to non-specific nouns. Lastly, personal articles denote pronouns such as '*his*' or '*her*' for nouns that are owned by a specific person.

During lexicalization, character goals are mapped into their equivalent English language words. These mapping of words is based on the manually-built lexicon of the system, which contains three lexical terms for each concept to differentiate the words to be used for children with different ages. For example, the concept of *delicious* can be lexicalized as '*yummy*' for 6-year olds, '*tasty*' for 7-year olds, and '*delicious*' for 8-year olds.

Character concepts are mapped based on the names found in the knowledge base. However, in order to make the references more natural, pronouns such as '*he*' or '*she*' are also generated. When a character is first introduced, the actual name of the character is used. Pronouns are used for the succeeding references to the character until the fourth reference, after which, the actual name of the character is used again in order to reduce ambiguity.

4. TESTING AND RESULTS

Three evaluators (two linguists and one writer) assessed the quality and the correctness of the generated stories. Each evaluator was given 10 stories to evaluate. These stories were generated from an ontology that has been manually populated with 1,002 relevant concepts and 1,442 semantic relations, and a lexicon that has been populated with 769 words.

The Story Editor repository currently contains four background images (camp, street, park, and classroom), four character stamps (Danny the dog, Hammy the hippo, Peter the pig, and Rizzy the rabbit), and four object stamps for each of the background. Each of the characters has also been assigned with three traits.

The linguists provided quantitative evaluation to each of the 10 stories using four criteria: *language*, *coherence and cohesion*, *character objects and background*, and *content*. Each of these criteria has a set of associated questions that are rated with the

following scores: 5 strongly agree, 4 agree, 3 neutral, 2 disagree, and 1 strongly disagree.

The linguists also provided qualitative comments and suggestions some of which were implemented to improve the quality of the generated stories, specifically in the areas of grammar, flow and content. Ten stories were again generated from the revised software and submitted to the evaluators for review. Table 5 shows the results of both rounds of evaluation.

Table 5. Results of Quantitative Evaluation

Criterion	First Eval	Second Eval	Average
Language	3.87	4.43	4.01
Coherence and Cohesion	3.52	3.66	3.28
Characters, Objects and Background	3.78	4.02	4.02
Content	3.50	3.76	3.64
Overall Evaluation	**3.67**	**3.96**	**3.74**

The *language* criterion evaluates the correctness of the generated sentences' structure and the appropriateness of the words used, including the correct usage of articles, prepositions and pronouns. Although the system received an average score of 4.01, there are still cases of incorrect usage of these grammar constructs. SimpleNLG does not provide a function that generates articles automatically. Thus, Picture Books 2 implemented its own set of simple rules for article generation. However, a standard rule does not always apply to all words, for example, the articles *a* and *an* cannot be used with the word *juice*, as shown below.

Incorrect usage of article
Output: She spilled a juice.
Correct: She spilled juice.

Missing article
Output: She played a game in park.
Correct: She played a game in the park.

Incorrect usage and missing article
Output: Hammy the hippo saw a money, therefore she felt tempted to use money.
Correct: Hammy the hippo saw money, therefore she felt tempted to use the money.

Incorrect preposition
Output: Peter the pig was in the camp for relax.
Correct: Peter the pig was in the camp to relax.

Another issue is the absence and the misuse of pronouns in sentences. References to characters usually appear consecutively in the generated story, making the sentences sound redundant, as shown in the excerpt below.

Output: Danny the dog learned that when Danny the dog is scared, he should search sound.
Correct: Danny the dog learned that when he is scared, he should search sound.

Picture Books 2 introduced the *Personal* concepts to represent nouns that imply possession of the character, such as *hair*, *teeth* and *sister*. These are classified as such to allow the system to recognize and use the appropriate articles. For example, it is more appropriate to use "*CHARACTER brushed his hair*", rather than "*CHARACTER brushed the hair*". However, because of the size of the lexicon, majority of the concepts are not regarded as words that require a possessive pronoun, thereby producing sentences that lack pronouns, such as the following excerpts:

Missing possessive pronoun
Output: He explored using flashlight.
Correct: He explored using his flashlight.

Missing possessive pronoun and incorrect pronoun replacement
Output: Friend burned the marshmallow. He ignored. Peter the pig felt sad.
Correct: His friend burned the marshmallow. Peter the pig ignored. He felt sad.

The *coherence and cohesion* criterion evaluates the presence of a smooth and logical flow from one sentence to another and if the generated sentences make sense and can be understood by the target age group. This includes the proper use of discourse markers and conjunctions to join two sentences together. Although Picture Books 2 made attempts to address issues on coherency, the system still received the lowest average score of 3.28 in this criterion. The simple conjunction generation algorithm that was implemented resulted in the inappropriate usage of conjunctions that are not applicable to the context of certain sentences as shown in the excerpts below:

Incorrect discourse marker
Output: Danny the dog wanted to eat, since he felt hungry.
Correct: Danny the dog wanted to eat, because he felt hungry.

Missing discourse marker
Output: He slept in tent. He heard a sound.
Correct: He slept in tent. While he was sleeping, he heard a sound.

Insufficient discourse marker
Output: Danny the dog walked, as a result he felt tired.
Correct: Danny the dog walked and as a result he felt tired.

The third criterion examines the *characters, objects and background* found in the story, specifically the character trait, the moral lesson, the objects used, and the background. The system received the highest average score of 4.02 because the generated stories reflect the story elements selected by the user in the Story Editor. There are, however, instances wherein the objects placed in the scenes are not included in the generated story. For example, when the input scene contains a *marshmallow* and a *flashlight* in the *camp* background, the system is able to introduce both of the objects, as shown in lines 2 and 11 below, but the *marshmallow* does not affect the plot of the story.

[1] *One dark evening, Danny the dog was in the camp for camping.*
[2] *He gets a white marshmallow.*
[3] *The camp is very far.*
[4] *He feels tired, since he walks.*
[5] *Danny the dog feels thirsty.*
[6] *He feels refreshed, because he drinks a water.*
[7] *He walks.*
[8] *Danny the dog sees a shadow.*
[9] *He feels scared.*
[10] *He does not know what to do.*
[11] *He turns on a flashlight.*

[12] *Danny the dog searches the shadow.*
[13] *He is not scared anymore.*
[14] *He learns that when He is scared, He should search shadow.*
[15] *Since then, Danny the dog learns to be brave.*

It is also evident in the generated stories that the moral lesson is based on the character's trait, the objects present in the scenes, and the selected background. In the sample story above, Danny the dog's predefined traits are honest, obedient, and persevering. Therefore, he is not brave, responsible, and helpful. As can be seen in the story above, the story ended with the character developing a trait he lacked, which is bravery.

The *content* criterion is used to assess the story as a whole, including the correctness of the structure of the story, the appropriateness of the story, and if the story contains enough detail. The system received an average score of 3.64 because the evaluators found the generated stories to follow the basic structure of a children's story, and they catch the attention of the reader. A conflict is always present and is resolved at the end. However, there are also cases where the generated story is too short containing few details which affect the quality of the story, as shown in the excerpt below:

[1] *She felt thirsty.*
[2] *Rizzy the rabbit drank using water jug.*
[3] *She felt refreshed.*
[4] *She hopped.*
[5] *She saw shadow.*
[6] *Rizzy the rabbit felt scared.*
[7] *She did not know what to do.*
[8] *She cried.*
[9] *She searched the shadow.*
[10] *It was a tree.*
[11] *Rizzy the rabbit was not scared anymore.*
[12] *Rizzy the rabbit learned that when Rizzy the rabbit is scared, She should search shadow.*
[13] *Since then, Rizzy the rabbit learned to be brave.*

Another linguistic concern involves aesthetics, specifically, observing proper capitalization and the use of punctuation marks. Errors in grammar are very common. Some of the words used in the system are grammatically and contextually incorrect due to following. First, because the ontology contains limited concepts, the system may not find the appropriate words to use in a sentence, as shown below.

Output: He fixed his bedsheet.
Correct: He fixed his sleeping bag.

Second, because of the binary nature of the semantic relations, certain events are related to each other even though they are not logically related, for example, the event that may cause a character to feel dizzy in the excerpt below is the crowd and not the marshmallow.

[1] *The camp is very crowded.*
[2] *He feels dizzy, because he sees the marshmallow.*
[3] *Danny the dog fixes his bedsheet.*

Line 2 above was generated from the following relations:

propertyOf camp crowded
feels CHARACTER dizzy
effectOf dizzy see
receivesAction see crowd
receivesAction see marshmallow

While a character seeing a crowd can make him feel dizzy, seeing a marshmallow would not have the same effect. However, because of the binary relations employed, the planner was not able to distinguish the difference. This is because there is not enough constraint that tells the planner that the marshmallow cannot be used to make the character feel dizzy. Thus, the planner may choose to use the marshmallow in the event.

Lastly, simpleNLG can only realize simple structured sentences. Some sentences that may need to use a different tense, such as a gerund, cannot be generated properly, as shown below:

Output: He slept in tent. He heard a sound.
Correct: He slept in tent. While Danny the dog was sleeping, he heard a sound.

Some of the recommendations provided by the evaluators with regards to improving the linguistic quality of the generated stories were implemented prior to subjecting the stories to a second round of evaluation. In terms of the usage of articles and prepositions, the entries for nouns as "*money*" and "*juice*" in which articles are not applicable were annotated in the knowledge base to inform the article generation algorithm that articles should not be attached to these nouns. Nouns which refer to places such as "*park*" have been modified to accommodate the usage of the definite article "*the*".

The proper capitalizations of words in cases where conjunctions are used have also been revised.

Previous Output: Danny the dog learned that when he is scared, He should search sound.
Revised Output: Danny the dog learned that when he is scared, he should search sound.

Personal pronouns have been used in cases where the noun being referred to entails possession, such as "*hair*" below:

Previous Output: She combed a hair.
Revised Output: She combed her hair.

Finally, the sequence of two character goals in the story plan may be revised depending on the conjunction used, in order to produce more coherent sentence structure, as depicted below where the resulting sentence begins with the effect followed by the cause:

Previous Output: Danny the dog ate a marshmallow, thus he felt sleepy.
Revised Output: Danny the dog felt sleepy because he ate a marshmallow.

5. CONCLUSION

The work presented in this paper, Picture Books 2, demonstrated that grammatically correct and coherent stories comprising the four basic elements of a story (problem, rising action, solution, and climax) can be generated from a given picture, provided that the appropriate domain knowledge is present. The input picture is comprised of multiple scenes and the story planner is tasked with finding a logical story path to form a coherent story based on scene transitions, specifically character and object movement and existence across two adjacent scenes. A semantic ontology populated with binary relations of concepts relevant to a child's

daily activities is used to provide the storytelling domain knowledge and the knowledge about narrative structures needed by the planner.

Picture Books 2 addressed some of the concerns present in the previous system, including the use of discourse markers to combine two or more character goals in order to generate more complex sentence structure, and assigning character traits to each of the system's characters which in turn is one of the main factors that affect the selection of a theme for a given input picture.

Picture Books 2 also generated stories that revolve around the internal conflict and personality development of only a single main character. One improvement involves providing a story world where at least two characters interact (or compete) to achieve a common (or competing) goal(s). Allowing and considering the presence of multiple characters will enable the system to generate more complex stories. This will also require the application of intelligent agents, each embodying its own behavior, capabilities, states and goals. The story planner will not only be responsible for identifying which of the agents will perform a certain action, but is also tasked with resolving conflicting character goals and behaviors if necessary. This is currently being explored in Picture Books 3.

While Picture Books 2 showed the potential for generating stories for multiple scenes, the scenes take place in the same setting or background. Thus, another area for exploration is to allow the user to define different backgrounds for each scene. This will require modeling the set of adjacent scenes, as well as knowledge (actions that can take place and events that can occur) about how a character or an object can move from one background to a different background across two adjacent scenes.

Similar to its predecessor, Picture Books 2 also relies heavily on the available concepts and semantic relations present in its knowledge repository. Currently, this has been manually populated to match the identified themes, character traits, objects and backgrounds. But various tests conducted showed that over-population of the ontology led to the retrieval of illogical story paths, while under-population led to dead ends (that is, no story can be generated). Aside from addressing the issue on providing an automated or semi-automated approach to building the knowledge base, the development of a reasoning engine to perform inferencing on the set of candidate story paths should also be explored. Still another direction is to utilize the latest version of ConceptNet which has its own knowledge acquisition tool (the Open Mind Common Sense project) and may be a good resource of common sense knowledge to provide story generation systems with the common sense needed to produce sensible stories and provide large enough data for generating numerous story content.

Finally, a more comprehensive representation model to reflect the current state of the story world and the changes that had already taken place, such as previous actions of the character and the resulting consequences, and changes in the objects that are in the character's possession or in the story world, should be developed in order for the system to consistently generate story events that are logical and believable.

6. ACKNOWLEDGMENTS

The Picture Books 2 project has been funded by the Department of Science and Technology - Philippine Council for Advanced Science and Technology Research and Development (PCASTRD).

7. REFERENCES

[1] Fields, M. and Spangler, K. 2000. *Let's Begin Reading Right: A Developmental Approach to Emergent Literacy.* Upper Saddle River, N.J: Merrill.

[2] Hong, A.J., Solis, C., Siy, J.T., Tabirao, E. and Ong, E. 2008. Picture Books: An Automated Story Generator. *Proceedings of the 5th National Natural Language Processing Research Symposium* (Manila, Philippines, November 2008). 5NNLPRS, DLSU, Manila, Philippines.

[3] Liu, H. and Singh, P. 2004. ConceptNet – A Practical Commonsense Reasoning Tool-Kit. *BT Technology Journal*, 22, 4 (Oct. 2004), 211-226. Netherlands, Springer. DOI= http://dx.doi.org/10.1023/B:BTTJ.0000047600.45421.6d

[4] Machado, J. 2003. Storytelling. *Early Childhood Experiences in Language Arts: Emerging Literacy*, 304-319. Clifton Park, N.Y. Thomson/Delmar Learning.

[5] Mann, W. and Thompson, S. 1987. Rhetorical Structure Theory: Toward a Functional Theory of Text Organization. *Text*, 8,3, 243-281.

[6] Ong, E. 2010. A Commonsense Knowledge Base for Generating Children's Stories. *Proceedings of the 2010 AAAI Fall Symposium Series on Common Sense Knowledge* (Virginia, USA, November 11-13, 2010). CSK '10, AAAI, USA, 82-87.

[7] Riedl, M. 2004. *Narrative Generation: Balancing Plot and Character*. PhD Dissertation, North Carolina State University.

[8] Taylor. 2009. Elementary transition words. http://www.greenville.k12.sc.us/taylorse/Taylorsy.

[9] Uijlings, J. 2006. *Designing a Virtual Environment for Story Generation*. MSc Thesis. University of Amsterdam, Amsterdam.

[10] Venour, C. and Reiter, E. 2008. Tutorial for Simplenlg (version 3.7). http://www.csd.abdn.ac.uk/~ereiter/simplenlg/