ABSTRACT
Picture Books generates stories for children age four to six based on a set of picture elements selected by the user, namely background, characters and objects. The system creates an abstract story representation of the theme for the given picture and uses NLG techniques to generate the sentences in the story. Manual evaluations on 15 generated stories were conducted to validate the correctness of their grammar and appropriateness of their content. The generated stories received an average score of 91.5% for grammar correctness and 88% for appropriateness of content. Automated evaluation was also performed by comparing the generated stories against reference stories, with the system receiving a low word error rate of 9.71% and a sentence error rate of 41.96%. The test results demonstrated that computers can generate stories from pictures, and the system’s knowledge representation, using the combination of story pattern data structure and semantic ontology, is feasible and usable to all story generation systems regardless of the domain.

Keywords
Natural language generation, story generation, story planning operators, story tree, semantic ontology

1. INTRODUCTION
Stories serve as tools in helping children develop their knowledge because children focus more on pictures/images of objects compared to words [1], thus the basic motivation of published children’s story books is to relate stories using pictures and words. Combining this motivation with natural language processing technologies, it is of a significant interest to research on the possibility of automatically generating stories given a picture as its starting point and using an ontology for knowledge representation.

An automated story generator consists of story planning goals combined with Natural Language Generation (NLG) techniques to convert story plans to human understandable text. For all NLG systems, the knowledge resource dictates the amount of information the system can output, thus highlighting its importance.

Several attempts have been made to create a story generator that would closely resemble “human made” stories using NLG techniques. However, several factors have to be considered in creating a story, including the following: (1) availability of NLG techniques to suit the narrative goals of a story; (2) the creativity that should be included in the story, for it would be erroneous to generate the same story over and over again; (3) the length of the story produced, whether sentential or multi-paragraph; and (4) the ambiguities of the English language, or any other language that will be used in the story generator systems.

This paper is subdivided as follows. Section 2 presents an overview of storytelling for children, which is the basis in designing the story planner of Picture Books. Section 3 discusses Picture Books in detail, its processes and knowledge sources. Section 4 presents the various tests that were conducted to validate the stories generated by Picture Books. The paper ends with a summary of further works that can be done to extend the capabilities of the story generator system.

2. STORYTELLING FOR CHILDREN
Various aspects have to be considered in generating stories appropriate for a target age group, namely story patterns, themes, and sentence structures. Interviews with child educators and literature review were conducted to determine these components.

A common story plot for children age four to six flows from negative to positive as follows.

i. The main character wants something.
ii. The main character is informed of the rules and/or restrictions.
iii. Nevertheless, the main character disobeys.
iv. The main character is either caught or experiences natural consequences of disobedience (e.g. a tooth ache from eating too much candy).
v. The main character learns a lesson.

This is depicted in Figure 1.

Fables are often used in storytelling. Fables are short, simple stories with animals as characters. In fables, the characters convey morals or lessons which are usually contained in the last line of the story. Common lessons taught in children’s story books
include being honest, being brave, value of sharing, being cheerful, eating fruits and vegetables, being polite, keeping one’s toys, and helping at home.

Stories for four year olds have simple sentence structures and contain line redundancy. This is lessened as the child grows older. Words used should also be simple and easy to understand. Positive adjectives describing what the child must do are used. The words used to describe things vary depending on the age of the child. For example, it is appropriate to use the word happy for a four-year old child, merry for five-year old and delightful for six-year old.

Another aspect to consider in storytelling is the story’s title. Story titles are short and often contain the story’s theme as a hint to what the story is about. The main character’s name should also be included in the title [8].

3. THE PICTURE BOOKS SYSTEM

Picture Books is an automated story generation system intended for children age four to six. It derives the story elements from a given input picture with components selected by the user from a library of background images, character stickers and object stickers. The genre (fables) and story goal (moral lesson) are applied as part of the system’s domain knowledge.

3.1 Knowledge Representation

In a story, the knowledge required is represented by the story elements, namely the settings, characters and events; and the goal of the story. The design of the knowledge representation of Picture Books is divided into two functionalities: (1) the story patterns to direct the story goal (moral lesson) and to ensure the completeness of the plot; and (2) a semantic ontology to provide domain knowledge or to constitute concepts that describe the target genre of the system, in this case, fables.

3.1.1 Themes and Story Patterns

Themes dictate the plot of the story to be generated and are composed of four story stages or story patterns (see Figure 2) namely, the problem, rising action, solution and climax which are the four fundamental stages of the main plot of any story [4]. In the take bath theme for example, these four stages would contain the following story patterns:

- **Problem**: Defy - not do rule
- **Rising Action**: Experience consequence
- **Solution**: Do the lesson
- **Climax**: Learn the benefit

The theme is executed through the story plot assigned to each story stage. A story plot (see Figure 3) represents the events that happen per story stage. It contains at least two author goals representing the scenes that would make up the event. An author goal (see Figure 4) is composed of a goal of the scene and the corresponding consequence of the goal to ensure consistency in the scene.
the target as the location or object of the action verb, and the instrument as the object used to perform the verb. This structure makes all fields nullable except for the action and agents fields.

Figure 5 Character Goal Data Structure

The character goal is generic so that it only contains default values for action, agents and nullable patients. These character goals are customized to be used for a particular scene through parameter passing of attribute values at the author goal level. For example, the generic character goal adult tells main character (CGL01) has the following default attributes:

- **action**: tell
- **agents**: adult
- **patients**: main character

When a scene requires the adult to inform the main character of the lesson, the target attribute would then be assigned with the lesson value to denote that the adult is talking to the main character about the lesson. The invocation of the character goal in the author goal level would look like this:

CGL01(target:lesson)

This customizes the character goal to “adult tells the lesson to the main character” to fit the scene.

Parameters for character goal attributes include not only the story element variables (i.e. object, lesson, background), but invocations to inner character goals and ontology accesses as well. An inner character goal is a character goal assigned as an attribute of an outer character goal. It represents a clause in a sentence and is usually assigned as a value of the target attribute in the outer character goal, for example:

CGL03(target:CGL05(target:lesson))

Picture Books would interpret the inner character goal main character is not doing the lesson (CGL05) first before appending it to the outer character goal secondary character told the adult character that (CGL03) the main character is not doing the lesson (CGL05).

The ontology access gets the value(s) returned by the ontology and depending on the nature of access may either fill up an assigned attribute or create dynamic character goals which contribute to the length and variation of the generated story.

3.1.2 Ontology

Picture Books uses an ontology to have a flexible knowledge resource based on the target domain. The ontology used serves as the knowledge domain model to provide relevant concepts applicable to the story being planned. Its design was adapted from ConceptNet [6], an ontology of common sense knowledge, and populated with concepts familiar to children aged four to six. The idea of semantic relationship, which is the relationship connecting two concepts, and semantic category, which serves as the classification of semantic relationships, are adapted from ConceptNet’s design.

The ontology is accessed to search for a related concept or for a path of relationships. The search for concepts is used when filling up an attribute in a character goal. For example, the instrument attribute in a character goal may have the following value that triggers an ontology search for a concept:

- **Instrument**: onto<semantic category>%(object%)%

Searching for a concept requires a single concept input (%object%) and a semantic category where the desired semantic relationship falls under. The ontology narrows down the search coverage based on the semantic category constraint and returns the concept directly connected to the input concept. For example, ontoSpatial(play) denotes an input concept of play and spatial semantic category. This invocation would make the ontology search through all the concepts connected to play, choosing only from “locationOf” and “oftenNear” semantic relationships that are under the spatial semantic category.

The search for relationship paths is used when a relationship between two concepts is describable through a series of connected relationships. Ontology paths are usually expected to generate dynamic character goals.

Searching for path of relationships requires two input concepts – the source and destination concepts, and the semantic category that constrains the search coverage. The search is facilitated by an ontology tree to store the relationship paths that are already visited as the ontology search proceeds. The searching and filtering mechanism for this type of ontology search is the same as searching for a concept, except that the search continues until the destination concept is located and until all concepts in the same tree level with the destination concept is in the ontology tree. This algorithm enables the ontology to possibly return more than one applicable path. The path returned by the ontology access is the shortest path connecting the source to the destination concept.

As an example search for path, ontoEvent(break object, punishment) denotes that a search for a series of semantic relationships must be done to relate break object to punishment. The event semantic category narrows down the search by focusing only on the semantic relationships under it (i.e. “firstSubeventOf”, “lastSubeventOf”). A path returned by performing the search on ontoEvent is as follows:

- **break object**: lastSubeventOf : get punished
- **get punished**: lastSubeventOf : not allowed to play today
- **not allowed to play today**: isA : punishment

When a path is returned, hypernym relationships (i.e. “isA” and “conceptuallyRelatedTo”) are filtered as these will not be useful when converted into sentences. Thus, the last relationship above will be disregarded. The system accommodates the returned relationship paths through semantic relation rules that map each relationship to dynamic character goals. For example, the first relationship above maps the lastSubeventOf relationship by assigning the second concept (get punished) as the action attribute of the dynamic character goal. Therefore the first relationship would be converted into a character goal whose output sentence is “main character gets punished.”
3.2 Architectural Design

The Picture Books system has three main modules: the Picture Editor module, the Story Planner module, and the Sentence Generator module, as depicted in Figure 6.

![Figure 6 Picture Books' Architectural Design](image)

Storytelling begins when the user selects one of the available backgrounds in the Picture Editor module of Picture Books. A background is needed to identify the corresponding objects that a user can place in his/her picture. Objects, represented as stickers, dictate the possible themes of the stories that can be generated. Themes represent moral lessons of the stories, and contain plots for the problem, rising action, solution, and climax, following the classic story pattern presented in Section 2.

The selected background, the selected main character and supporting character stickers, the list of the object stickers placed in the background, and the user’s age are converted into an abstract input content representation and forwarded to the Story Planner module.

The Story Planner module is subdivided into the Story Content Planner and the Story Organizer. The Story Content Planner selects the appropriate theme based on the combination of the background and sticker objects in the abstract input content representation. The theme selected creates a theme specification plan consisting of a problem, a rising action, an answer or solution and a climax that defines the story events that must be included in order to effectively portray them.

Story patterns selected under each plot of the theme specification plan are then executed, utilizing the system’s ontology together with the semantic relation rules. As discussed in section 3.1.2, the ontology is used whenever a story pattern triggers its invocation so as to add story details other than the details provided by the story pattern and reduce the predictability of the story generated. The semantic relation rules are then used in order to convert the ontology outputs into newly created story patterns, i.e. the character goals. The generated story events are stored in an abstract plot representation.

The last activity performed by the Story Content Planner involves adding the title, the introduction and the appropriate ending of the story to the abstract plot representation, to complete the story plan. The story plan containing the story content is represented as an abstract story tree that will be transformed into sentences in later stages of the generation process.

The Story Organizer arranges the story events in a pre-order manner and organizes the story as it is supposed to be presented to the user, resulting in a structured abstract story tree that is passed to the Sentence Generator module.

The Sentence Generator module is subdivided into the Sentence Planner and the Realizer that correspond to the microplanner and realisation modules, respectively, of a general NLG architecture [2]. The Sentence Planner converts the structured abstract story tree to sentence specification form that can be processed by the Realizer. This conversion embodies three pipelined tasks, namely Referring Expression Generation, Lexicalisation, and Phrase Specification Mapping.

In Referring Expression Generation, redundant noun phrases are replaced with appropriate pronouns, for example, “Daniel did not want to take bath. He wanted to play.” Descriptors are also inserted to character names, for example, “Simon the sheep”.

Lexicalisation converts each concept into its text form. Phrase Specification Mapping converts the character goals to sentence specifications. Grammatical rules are applied to each character goal to convert them to phrase specifications and compound them to produce the sentence specifications that serve as input to the Realizer. One character goal is equivalent to one sentence.

The Realizer utilizes the simplenlg realiser [3], an open-source Java class library, to complete the story generation process by converting sentence specifications from the Sentence Planner to actual sentences that comprise the story. The Realizer outputs both the story title and the text of the final story.

4. Test Results

Several tests were conducted to help in the analysis of the effectiveness and performance of the system. 15 generated stories were evaluated to determine whether they conform to the standards set by the linguist, the child educators and the common and classical story pattern. The 15 stories consist of 5 themes varied according to user age. These 5 themes are selected, among the 10 themes that the system currently can generate, based on the variation of their story plots and the differences in the story length per age. Two child educators validated the appropriateness of the content of the stories for the target age group. A linguist validated the correctness of the generated stories in terms of grammar.

Automatic evaluation of the generated stories was also performed using the General Text Matcher [7] tool to compare the generated stories with the corresponding reference stories. The automatic evaluation returned quantitative values corresponding to the Word Error Rate (WER) and Sentence Error Rate (SER).

The knowledge base of Picture Books currently contains 9 backgrounds, 37 objects, 40 characters, 15 themes, 61 character goals and 77 author goals. The lexicon has been populated with 419 words appropriate for the target age group, while the ontology has 240 concepts and 369 semantic relationships.
4.1 Linguist Evaluation
During manual evaluation, the linguist and the child educators were asked to evaluate 15 stories in all, 5 themes for each age level: 4, 5, and 6. Each story was rated per category from 1-4, with 4 being the highest. A rate of 4 means that the criterion was completely present in the story, 3 means the criterion is present but incomplete, 2 means the criterion is partially present, and 1 means the criterion is not present.

Linguist evaluation involved checking the grammar, completeness of the plot, and coherence of the sentences in the generated stories. Table 1 shows the average score of the sample stories in each criterion.

Picture Books received the highest score in the criterion “completeness of the plot”. The system’s knowledge base ensures that each story generated has all the essential story elements (problem, rising action, solution, and climax) as well as the introduction of the time and setting of the story.

The description of objects received the lowest score, because although objects are present in the stories, for example, “She played near a lamp.”, they were not described.

The generated stories also lacked of transitional devices, as shown below (with the underlined words as the identified missing transitional devices), resulting in an average score of 3.0 for the criterion “The story has transition”.

He apologized to her. Mommy Patricia then helped Porky to clean up. Soon he found the lost toys.

The linguist also noted that the sentences were simple, thus they were less prone to grammatical errors and are also very appropriate for the target age group. Over-all the evaluation yielded a score of 91.36 % (3.66 out of 4) in terms of correctness of grammar.

<table>
<thead>
<tr>
<th>CRITERIA</th>
<th>AVG</th>
</tr>
</thead>
<tbody>
<tr>
<td>The words used are contextually correct</td>
<td>4.0</td>
</tr>
<tr>
<td>Sentences are grammatically correct</td>
<td>3.2</td>
</tr>
<tr>
<td>Sentences are coherent</td>
<td>3.6</td>
</tr>
<tr>
<td>The story has transition</td>
<td>3.0</td>
</tr>
<tr>
<td>If there are pronouns, they are used correctly</td>
<td>3.9</td>
</tr>
<tr>
<td>Articles are used correctly</td>
<td>3.2</td>
</tr>
<tr>
<td>The characters in the story were described</td>
<td>3.3</td>
</tr>
<tr>
<td>The settings of the story were described</td>
<td>4.0</td>
</tr>
<tr>
<td>Objects in the story were described</td>
<td>2.8</td>
</tr>
<tr>
<td>The actions of the characters make sense</td>
<td>4.0</td>
</tr>
<tr>
<td>There is a problem</td>
<td>4.0</td>
</tr>
<tr>
<td>There is a rising action</td>
<td>4.0</td>
</tr>
<tr>
<td>There is a climax</td>
<td>4.0</td>
</tr>
<tr>
<td>There is a solution to the problem</td>
<td>4.0</td>
</tr>
<tr>
<td>Story is appropriate to target age(Rate them 1 being lowest and 4 highest)</td>
<td>3.8</td>
</tr>
<tr>
<td>Story is understandable(Rate them 1 being lowest and 4 highest)</td>
<td>3.47</td>
</tr>
<tr>
<td>Words used are comprehensible(Rate them 1 being lowest and 4 highest)</td>
<td>3.53</td>
</tr>
<tr>
<td>Based from your knowledge of previous stories, how far is the generated story from them(Rate them 1 being farthest and 4 nearest)</td>
<td>3.00</td>
</tr>
</tbody>
</table>

4.2 Child Educator Evaluation
Evaluation by the child educators involved checking the generated stories in terms of grammar, completeness of the plot, coherence and comprehensibility of sentences, and appropriateness to the target age group. Table 2 shows the average score of the sample stories in each criterion.

The completeness of the plot was the strong point of the system, while the criteria regarding the grammar, coherency of sentences and the pronouns usage were the weak points. The quality of some stories could be improved if additional information is provided. For example, the story below is missing a sentence to depict that Ellen did something to show she is trying to be brave.

She wanted to be brave. Ellen was brave. She wanted to play with others. <more information could be provided here> Ellen made friends.

Since the ontology did not contain any other information of what to do in order to be brave, the story content planner did not place any detail describing the action of the main character depicting her attempt to be brave. This can be remedied by simply adding knowledge regarding this in the ontology.

The child educators also noted that although the stories were relatively good, they could have been better if they contain dialogues. Over-all the evaluation yielded a score of 88% (3.52 out of 4) in terms of correctness.

<table>
<thead>
<tr>
<th>CRITERIA</th>
<th>AVG</th>
</tr>
</thead>
<tbody>
<tr>
<td>The words used are contextually correct</td>
<td>2.80</td>
</tr>
<tr>
<td>Sentences are grammatically correct</td>
<td>2.60</td>
</tr>
<tr>
<td>Sentences are coherent</td>
<td>2.67</td>
</tr>
<tr>
<td>The story has transition</td>
<td>3.47</td>
</tr>
<tr>
<td>If there are pronouns, they are used correctly</td>
<td>3.00</td>
</tr>
<tr>
<td>Articles are used correctly</td>
<td>3.20</td>
</tr>
<tr>
<td>The characters in the story were described</td>
<td>3.67</td>
</tr>
<tr>
<td>The settings of the story were described</td>
<td>3.86</td>
</tr>
<tr>
<td>Objects in the story were described</td>
<td>4.00</td>
</tr>
<tr>
<td>The actions of the characters make sense</td>
<td>3.93</td>
</tr>
<tr>
<td>There is a problem</td>
<td>4.00</td>
</tr>
<tr>
<td>There is a rising action</td>
<td>4.00</td>
</tr>
<tr>
<td>There is a climax</td>
<td>4.00</td>
</tr>
<tr>
<td>There is a solution to the problem</td>
<td>4.00</td>
</tr>
<tr>
<td>Story is appropriate to target age(Rate them 1 being lowest and 4 highest)</td>
<td>3.80</td>
</tr>
<tr>
<td>Story is understandable(Rate them 1 being lowest and 4 highest)</td>
<td>3.47</td>
</tr>
<tr>
<td>Words used are comprehensible(Rate them 1 being lowest and 4 highest)</td>
<td>3.53</td>
</tr>
<tr>
<td>Based from your knowledge of previous stories, how far is the generated story from them(Rate them 1 being farthest and 4 nearest)</td>
<td>3.00</td>
</tr>
</tbody>
</table>

4.3 Automated Evaluation
The automated evaluation was conducted by the proponents using the General Text Matcher (GTM) version 1.4 tool created by [7] for machine translation systems. GTM computes for the Word Error Rate (WER) and the Sentence Error Rate (SER) of a computer-generated translation of an input document against a manual translation of the same input done by a linguist, to validate the correctness of machine translators. The proponents utilized this same tool to determine the correctness of the generated story in terms of sentence structure and grammar by comparing the generated stories with their corresponding reference stories.

WER is the percentage of words that must be added, modified or deleted to the generated story to make it similar to the
reference story. SER refers the frequency of sentence mismatch in the actual generated story against the reference story.

The reference stories were derived from the stories generated by the system and were corrected or edited by a linguist. These stories were used as the reference stories because the linguist evaluator focused on the structure and grammar aspect of the stories' sentences. The child educator evaluators, on the other hand, focused more on the coherency and content of the story that are currently not measurable through automated evaluation.

Automated evaluation yielded an average Word Error Rate of 9.71% and an average Sentence Error Rate of 41.96%. Both WER and SER were caused by lack of possessive pronouns, lack of transitional devices and lack or inappropriate use of articles. Furthermore, stories with low WER may still incur a high SER because of the high distribution of word errors across sentences in the stories.

5. Conclusion
The work presented in this paper explored the use of natural language generation techniques to the generation of children's stories. Picture Books 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 design of the system models the way humans perform storytelling through story patterns that utilize author goals and character goals. Furthermore, the ontology design that has been adapted from ConceptNet modelled concepts and their semantic relationships for the target age group. The flexibility of the story pattern design and the ontology design of Picture Books makes them suitable to be adapted to other story generation systems, regardless of the domain.

Based on the results of the manual evaluations performed, Picture Books can be further improved by extending the design of the character goal in order to indicate a connection or the passage of time between sentences through the use of transitional devices. Character goals can also be designed to generate dialogues, which will make the stories more interesting for the target age group. Applying Rhetorical Structure Theory to the author and character goals can also be performed for a more effective discourse structure that would provide smoother story flow.

Picture Books generates a story based on only one picture. However, in real life, storybooks written for children have different illustrations per page. The Picture Editor module could be expanded to allow users to create a sequence of pictures. Each picture in the sequence can be considered as a story event and the story plot would then be planned around them in order to make the stories more complex and more dynamic.

The knowledge base of Picture Books is a very crucial resource in generating appropriate stories. Populating the lexicon, ontology, and picture editor library is very time-consuming and places a limit on the possible stories that the system can generate. There are two ways to address this concern. The first is through the use of an authoring tool that would allow system administrators to add and modify the contents of the knowledge base of Picture Books. A second approach is by exploring the use of machine learning techniques to automatically extract different types of knowledge from sample stories fed to the system. This knowledge can be story patterns and word relationships.

6. REFERENCES