Theme-Based Cause-Effect Planning for Multiple-Scene Story Generation

Karen Ang, Sherie Yu and Ethel Ong

Center for Language Technologies
College of Computer Studies, De La Salle University
Manila, 1004 Philippines
karenang0903@yahoo.com, sherie_yu@yahoo.com, ethel.ong@delasalle.ph

Abstract

Early literacy in children begins through picture drawing and the subsequent sharing of an orally narrated story out of the drawn picture. This is the basis for the Picture Books story generation system whose motivation is to produce a textual counterpart of the input picture in order for the child to associate words with images. However, stories are comprised of sequences of events that occur in a cause-effect loop, and the single-scene input picture approach may lead to a story whose event flow may not match the child's original intended story. In this paper, we present Picture Books 2, which provides an environment for a child to creatively define a sequence of scenes for his input picture and then uses a theme-based cause-effect planner to generate a fable story narrating the flow of events across the scenes for children age 6-8 years old.

Introduction

Storytelling is an important aspect of human life. People use stories to share knowledge, experiences and ideas. Researches have shown that in their early years, children would draw pictures and then tell stories out of these afterwards. This helps develop their literacy skills and creativity, as one form of measuring creative thinking abilities is to assess the articulatenes of children in telling stories through drawings (Torrance, 1977). Furthermore, it has been found that children recognize pictures more easily than words (Fields and Spangler 2003).

Picture Books (Hong et al 2009) is an existing system that generates fable-form stories for children age 4-6 years old based on a given single-scene input. This input picture contains the basic story elements – background, characters and objects – that are selected by the child from a predefined list of stickers in the system's Picture Editor. The generated stories embody a moral lesson or theme encapsulated in a plot structure that flows from negative to positive, where a child violates a stated lesson, experiences the consequences of such violation, and learns the required value at the end of the story. The themes are randomly selected from a list of pre-defined themes associated with the specified background, while the plot structure follows the classic story pattern presented by Machado (2003) that flows from problem, rising action, resolution, to climax.

Although Picture Books showed the potential for computers to exhibit creativity in the form of literary art, there are a number of factors in storytelling that are currently not supported by the system (Ong 2009). Stories are sequences of events or scenes, and the single-scene structure of the system limits the planner on the events that it may generate, which may not necessarily match the original intent of the child whilst defining the scene. This makes the generated story less interesting as it may not adequately capture the story that was originally conceptualized by the child.

Picture Books 2 (PB2) extends the first system (from here-on referred to as PB1) by allowing the children, this time age 6-8 years old, to define multiple scenes which serve as the input picture to the story planner. Enabling the children to input several scenes can lead to stories that are longer and have more complex plot. Computational storytelling can then be used to enhance the creative abilities of children, as they fluently elaborate their stories through connecting sequences of scenes to form a single storyline. Fluency and elaboration are two measures of creative thinking abilities as defined in (Torrance, 1977).

Following PB1, PB2 also provides a set of background that the child can select, and a library of character and object images (called stamps) that can be pasted onto the selected background in order to create the scenes. It uses a theme-based cause-effect planning algorithm to generate the content of the stories that still promote moral values, this time set in more adventurous places like the camp or the street to allow older children to learn to explore the world and learn life's lessons on their own.

In order to facilitate the flow of the story from one scene to the next, two types of transitions are identified. The existence transition refers to the presence of a stamp in a particular scene. The movement transition refers to the movement or change in position of a particular stamp between adjacent scenes.

Another factor is the addition of traits to the characters. Using common animal characters in fables, child educators believe that embodying these characters with traits (e.g., a monkey is mischievous, a panda is loyal, and a rabbit is hardworking) would help children to relate to the story better. Riedl and Young (2004) noted 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. Thus, each of the characters in PB2 also possesses traits that comprise one of the factors affecting the flow of the story to be generated.

The rest of this paper is organized as follows. First, the knowledge base used by the story planner is presented. This is followed by a discussion of PB2's architecture, with emphasis on the planning process. The paper then presents the results of qualitative and quantitative analysis performed by linguists on the generated stories, and ends with a summary of research findings and further work that can be done to improve the system.

Storytelling Ontology

Storytelling relies on a large body of knowledge about 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 performing some actions, the occurrence of another event, or as a natural occurring phenomenon.

The ontology of PB2 stores storytelling knowledge comprising of world concepts and events common in a child's everyday life. It contains a network of binary semantic relations patterned after ConceptNet (Liu and Singh 2004), a free, machine-usable lexical and commonsense ontological knowledge representation. The PB2 ontology was then populated with concepts that are suitable for the target users and relevant to the identified themes.

Among the 20+ relations in ConceptNet, only the *CapableOf, UsedFor, ReceivesAction, EffectOf, HasSubevent, HasProperty* and *IsA* relations were relevant to PB2. Additional semantic relations, namely *Feels, CausesConflictOf, LeadsTo, IsTransition*, and *HasResolution* were defined to support concepts related to scene transitions, character traits, and theme-based planning. Table 1 describes these relations and provides examples defined in PB2.

In order to facilitate the flow of the story from one scene to the next, the story planner must be able to identify changes that have occurred between two adjacent scenes. These transitions are classified into two, stamp (character or object) appearance/disappearance and movement.

Appearance and disappearance is easily determined by checking if a stamp that is present in one scene is still present in the subsequent scene. Concepts related to this type of transition are then modelled in the ontology using the semantic relation *IsTransition*. For example, *eat – IsTransition – disappearance* associates the action that can cause an object, such as *marshmallow*, to disappear across two adjacent scenes. Similarly, if the *marshmallow* that was absent in the first scene appears in the second scene, the relation *buy – IsTransition - appearance* is used to model a possible action necessary for this.

To model stamp movements, each background image is divided into 6x6 grids, as shown in Figure 1. Each grid is labelled to track the position of a stamp in the background. A stamp is considered to have moved between two adjacent scenes if its position label in the first scene is different in the next scene. An example concept for this type of transition is *walk - IsTransition – movement*.

PB2 also recognizes six traits – responsible, honest, brave, helpful, obedient, and persevering. Each character has been assigned to possess three positive and three negative traits.

Relation	Definition	Example
Feels	Denotes the emotional response of the character to an event.	Character – *Feels* – Sad
CausesConflictOf	Used for selecting a story theme (conflict) based on the character's negative trait.	Brave – *CausesConflictOf* – Scared
LeadsTo	Used to associate an object to a theme (or conflict)	Flashlight – *LeadsTo* – Scared
IsTransition	Used to associate an action to a transition.	Eat – *IsTransition* – Disappearance; Bring - *IsTransition* – Appearance
HasResolution	Used to determine the appropriate resolution for a conflict.	Scared – *HasResolution* – Search
CapableOf	Represents an action that a character can execute in the story.	Character – *CapableOf* – Eat
UsedFor	Represents an activity that can take place in a specified location.	Camp – *UsedFor* – Camping
ReceivesAction	Relates an action that can be performed on an object.	Marshmallow – *ReceivesAction* – Eat
EffectOf	Provides a causal chain relationship between two events.	Tired – *EffectOf* – Sleep
HasSubevent	Specified an event that can occur before another event.	Eat – *HasSubevent* – Cook
HasProperty	Specifies an adjective to describe a noun.	Camp – *HasProperty* – Far
IsA	A generalized concept of a noun.	Marshmallow – *IsA* – Food

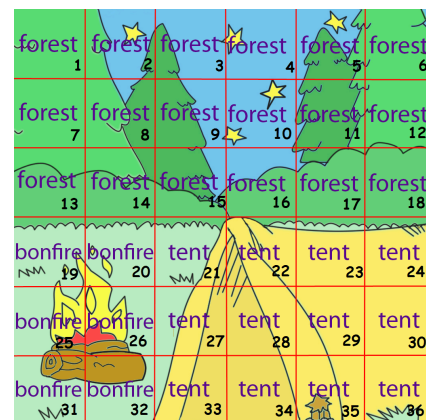Table 1. Semantic Relations in the Ontology of PB2

Figure 1. Grids in the Camp Background

System Architecture

Picture Books 2 has four main modules, namely, the Story Editor, the Story Planner, the Sentence Planner, and the Story Generator. This is depicted in Figure 2.
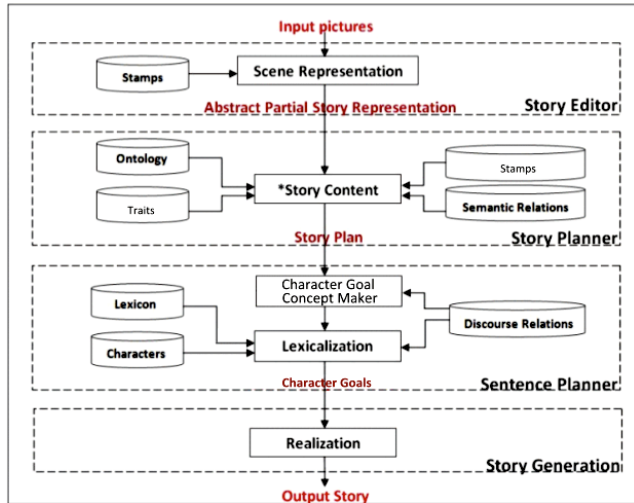

Figure 2. System Architecture

The Story Editor allows the child to choose from a pre-defined list of background, character and object stamps. Currently, there are four backgrounds (camp, street, park and classroom), four characters (dog, pig, hippo and rabbit), and 16 objects to choose from. The available objects vary depending on the selected background. Each story can have only one background for all its scenes, one character, and up to four objects. An input picture is required to have a minimum of three scenes to depict the initial setting, the problem phase, and the resolution phase of the story.

An abstract representation of the input picture that is forwarded to the Story Planner is shown in Figure 3. It shows that the input picture contains three scenes. The first scene (scene 0) contains a character stamp named *Danny* who appeared in this scene and is at grid 25, and an object stamp named *marshmallow* which also appeared in this scene and is also at grid 25. The second scene showed a movement of the character stamp from grid 25 to 28, and the appearance of another object, the *flashlight*. The *marshmallow* is assumed to have disappeared as it is not in the scene anymore. The value *null* signifies no transition.

Story Planner

The Story Planner produces a story plan comprising of semantic relations retrieved from the ontology. These semantic relations represent the progression of the story through a causal chain of character actions and events that will lead the main character to overcome one of his/her negative traits. The planner works by considering the traits of the character, the objects present in the scenes, and the scene transitions.

A theme depicting the conflict of the story is selected based on the main character's non-traits and objects present in the conflict (or middle) scene. Candidate conflicts

are retrieved from the ontology using the relation *CausesConflictOf*. Table 2 presents some *character trait* concepts and their associated *conflict* concepts. Note that each binary relation means that the absence of the *trait* concept in the character (e.g., *not brave*) leads to a story with the stated conflict (*scared*).

```
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
```

Figure 3. Abstract Story Representation

Concept 1	Relation	Concept 2
Brave	CausesConflictOf	Scared
Responsible	CausesConflictOf	Lose
Obedient	CausesConflictOf	Disobey

Table 2. *CausesConflictOf* Relations between Non-Character-Trait and *Conflict* concepts

The setting of the story is based on the background and the selected theme. This includes the time when the story takes place and the adjective to be used in describing the background. For instance, given the theme of a *character learning to be brave* that is set in the *camp* background, the most likely story will be that the *character is scared of the dark*, and thus, the time should be set to *evening*.

The planner generates the chain of events in the story by taking the background adjective as the root node and finding a path in the ontology to connect this to the identified conflict. Table 3 presents some *adjective* concepts associated to backgrounds through the *HasProperty* relation.

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

Table 3. *HasProperty* Relations for Backgrounds

Events generation also considers the possible events that may happen given the transitions between scenes. Furthermore, an event can be considered in the story if the character is capable of doing the associated action and the object required for its performance is present in the scenes. Other events may also require a specific location.

Once these preconditions are met, the planner finds a causal chain of events from the root node (background adjective) to the target node (conflict) using *EffectOf* relations. Table 4 presents *EffectOf* relations showing the causal link between concepts, starting from the background adjective *far*, leading to its effect, e.g., *tired*. *Tired*, in turn, may necessitate the character to *eat*. The chain continues until the concept node matching the identified conflict, e.g., *scared*, has been reached.

Concept 1	Relation	Concept 2
Tired	EffectOf	Far
Eat	EffectOf	Tired
Sleepy	EffectOf	Eat
Sleep	EffectOf	Sleepy
Hear	EffectOf	Sleep
Scared	EffectOf	Hear

Table 4. Chain of *EffectOf* Relations

If a path is found, the planner also checks for candidate sub-events that can occur in order to increase the length of the story. Currently, only one sub-event is included in the story plan.

Figure 4 illustrates a sample causal chain of events based on the relation *EffectOf*. The orange-colored nodes denote concepts found through the *HasSubevent* relation. For example, the *sleep* concept has the sub-events *pray*, *comb*, and *brush*. The dark colored nodes are the root node and the target node representing the background adjective and the theme or conflict of the story, respectively. A possible story path is: "*crowded-bump-dizzy-pray-sleep-hear_sound-scared*".
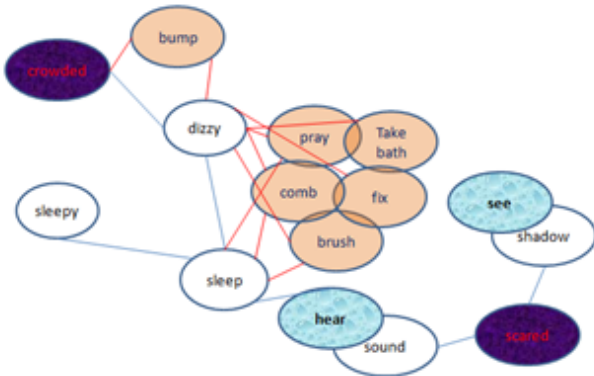

Figure 4. Sample Causal Chain of Events

In order to achieve variances in the generated stories based on the same input picture, the planning algorithm uses a simple random selection approach to identify the nodes to be included in the chain of events. Future work on PB2 should consider having a scoring function to guide the planning process.

Events generation is repeated in order to find a path from conflict to its possible resolutions. Table 5 presents the relationships between *conflict* concepts and *resolution* concepts. For example, if a character is *scared*, then the resolution phase of the story should involve actions requiring the character to *search for the causes* that lead to his being scared, e.g., *what is making the sound in the night*?

Concept 1	Relation	Concept 2
Scared	HasResolution	Search
Lose	HasResolution	Admit
Disobey	HasResolution	Apologize

Table 5. *HasResolution* Relations between *Conflict* and *Resolution* concepts

The Sentence Planner produces character goals (Uijlings 2006) by aggregating two or more consecutive semantic relations using discourse markers. A character goal represents one sentence in the final story. Figure 5 shows a sample output of character goals for the first scene in the abstract story representation in Figure 3. For each character goal entry, *agent* is the actor, *art(n)* is the article to be used, *verb* is the action to be performed, *patient* is the receiver of the action, *rst:n* specifies the type of discourse marker to be used, *type* signifies if the sentence will be joined with another sentence, and *tense* is the verb's tense. Based on feedback from the linguist, children's stories are usually written in past tense.

```
0agent: art(0) CHARACTER; verb: is; patien: art(2) camping;
  location: art(1) camp; time: fine evening;
  rst: 0type: 0; tense: 0;
1agent: art(0) CHARACTER; verb: see; patien: art(2) marshmallow;
  rst: 0type: 0; tense: 0;
2agent: art(1) camp; verb: is; patien: art(0) crowded;
  rst: 0type: 0; tense: 0;
3agent: art(0) CHARACTER; verb: spot; patien: art(1) marshmallow;
  rst: 0type: 0; tense: 0;
4agent: art(0) CHARACTER; verb: want; patien: art(0) eat;
  rst: 0type: 0; tense: 0;
5agent: art(0) CHARACTER; verb: cook; patien: art(1) marshmallow;
  location: art(1) bonfire; rst: 0type: 0; tense: 0;
6agent: art(0) CHARACTER; verb: eat; patien: art(1) marshmallow;
  rst: 3type: 1; tense: 0;
7agent: art(0) CHARACTER; verb: feel; patien: art(0) sleepy;
  rst: 3type: 1; tense: 0;
```

Figure 5. Sample Output Character Goals

The Sentence Planner also lexicalizes concepts and generates a set of sentence specifications, which is then forwarded to the Story Generator to produce the surface text with the use of an external surface realizer, simpleNLG (Venour and Reiter 2008).

Results and Analysis

10 sets of stories were given to the evaluators comprising of two linguists and one storywriter. These evaluators were chosen to provide expert judgement on the quality of the generated stories in terms of linguistic structure, narrative content, and appropriateness for the target audience. No feedback was solicited from the intended audience themselves at the time of this writing, because as the results below will show, the system still needs major work in its planning algorithm and knowledge representation in order to produce stories that the children may truly appreciate.

The stories were generated from an ontology that contains 1,002 concepts and 1,442 semantic relations. The lexicon has been populated with 769 terms. The evaluation was performed twice; after the first evaluation, PB2 was revised to address some of the feedback, then ten stories were regenerated to undergo a second evaluation.

Four criteria were used: *language*; *coherence and cohesion*; *character, objects and background*, and *content*, each of which has a set of associated questions that are rated by the following scores: 5-strongly agree, 4-agree, 3-neutral, 2-disagree, 1-strongly disagree. Table 6 shows the results.

Criterion	First	Second
Language	4.01	4.43
Coherence and Cohesion	3.28	3.66
Character, Objects and Background	4.02	4.02
Content	3.64	3.76
Overall Rating	**3.74**	**3.96**

Table 6: Summary of Quantitative Evaluation

The *language* criterion deals with the correctness of the sentence structure and appropriateness of the words used. Also included are the proper usage of articles, pronouns and punctuation marks. Here, PB2 received an average score of 4.43 after the second evaluation since the English grammar rules and the lexicon used during sentence generation were defined specifically with the target users in mind. During revision, rules related to correct usage of pronouns and articles provided by the linguists were implemented. However, there are still cases of incorrect usage as shown in the examples below where there is an incorrect usage of the article and a missing article.

> *Output: She spilled <u>a juice</u>.*
> *Correct: She spilled <u>juice</u>.*

> *Output: She played a game in park.*
> *Correct: She played a game in <u>the</u> park.*

The *coherence and cohesion* criterion is concerned with the transition of events and the flow of the sentences, to evaluate if the generated story makes sense and is easy to understand. Coherence between sentences can be enhanced through the use of discourse markers (Mann and Thompson 1987). Taylor (2009) provided a list of common discourse markers, and those appropriate for elementary age kids are presented in Table 7.

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

Table 7: Common Discourse Markers for Children

PB2 received the lowest average score of 3.66 in this criterion because although the stories contained discourse markers, these are sometimes used inappropriately with respect to the context of the sentence, as shown below.

> *Output: Danny the dog ate a marshmallow, <u>thus</u> he felt sleepy.*
> *Correct: Danny the dog ate a marshmallow, <u>and thus</u> he felt sleepy.*

On the other hand, the absence of discourse markers resulted in the generation of choppy sentences.

> *Output: He slept in tent. He heard a sound.*

> *Correct: He slept in tent. <u>While he was sleeping</u>, he heard a sound.*

Because the planner utilizes a random selection approach and does not perform reasoning over the resulting path of semantic relations, there are also cases in which the generated story is not logical.

> *He brought a blue water jug.*
> *The camp was very far.*
> *He felt tired.*
> *He felt thirsty.*

The *characters, objects and background* criterion examines the appropriate interplay between the character, object and background elements of the story with the story itself. This includes checking for the incidence of character traits and moral lesson, and the appropriateness of the objects with respect to the chosen background. The system received an average score of 4.02 for both rounds of evaluation. There are instances wherein the objects placed in the input scenes are not included in the generated story. There are also instances where the object, although introduced in the story, does not play any role in the story. The excerpt below illustrates this.

> *[1] It was a fine evening.*
> *[2] Danny the dog was in the camp for a trip.*
> *[3] He buys a packed marshmallow.*
> *[4] The camp is very big.*
> *…*
> *[5] He sees a shadow.*
> *[6] Danny the dog feels scared.*
> *[7] He does not know what to do.*
> *…*
> *[8] Since then, He learns to be brave.*

The given story excerpt was generated from an input picture comprising of three scenes. In the first scene, a marshmallow object has been included and introduced in line [3]. However, it plays no part towards the development of the theme where the main character has to learn to be brave. In fact, aside from line [3], no other text in the story mentioned the *marshmallow* again.

The overall *content* of the story includes evaluating the appropriateness of the story to the target age group, the adequacy of details provided, as well as the believability of the events in the story. The evaluators noted that the generated stories follow the basic structure of a children's story. They also found these stories to be quite interesting due to the interplay of the conflict and resolution to the theme of the story as well as the chain of events.

Conclusion

Picture Books 2 demonstrated that a coherent story with the four basic classic story subplots of Machado (2003) can be generated from a given input picture with at least three scenes. This is achieved by a theme-based cause-effect planner that utilizes an ontology of semantic domain and narrative knowledge, and a sentence planner that utilizes discourse markers to connect two or more events together.

The story planner provided a mechanism to control the sequencing of events to adhere to the basic story plot suitable for children's stories, but also allowed for flexibilities and variances in the generated stories. This is done by manually populating the semantic ontology with the relevant binary relations and concepts. However the population should be done with caution. Based on tests conducted, over population can lead to illogical story paths and under population can lead to not being able to generate stories.

Because the ontology representation makes use of binary relations, this can lead to logical errors in the resulting stories. An instance of this is the relation "*dizzy – EffectOf – see*" and "*people – ReceivesAction – see*", which logically means that *if the character sees many people he or she feels dizzy*. However, given that there is also a relation "*marshmallow – receivesAction – see*", this resulted to a story text where *a character feels dizzy because he or she saw a marshmallow*, which makes no sense at all.

Even though both PB1 and PB2 utilize a semantic ontology, the story planner of PB1 has a set of predefined planning operators in the form of author goals (Hong et al 2009; Cua et al 2010) similar to Minstrel (Turner 1992) that represent high-level tasks to guide the construction process by focusing on the narrative structure of the story. The author goals in turn are divided into two or more character goals (Uijlings 2006), again predefined, and these two types of goals are used to constrain the stories being generated. The ontology is used only to provide information needed to fill in the attributes in the character goals in a theme-driven story plot template (Ong 2010). This generated stories that are of good quality, coinciding with findings of Peinado and Gervas (2006) that "ontology-based stories obtain good results on coherence because the ontology forces explicit links between events".

PB2, on the other hand, did away with predefined author goals while its character goals are generated dynamically based on the semantic relations found in the story path retrieved by the planner. The development of a reasoning engine that contains rules for checking the logical inconsistencies and performing inferencing on the set of candidate story paths should be explored to address the issue on resolving ambiguities and generating logical stories.

The development of a more comprehensive model for representing the current state of the world and the changes that had already taken place, such as previous actions of the character, previous events that have taken place, and changes in the objects that are in the character's possession or in the story world, should also be explored to address these issues and enable the system to consistently generate story events that are logical and believable.

Finally, feedback should be solicited from the intended users (the children) to provide an assessment on the effectiveness and usability of the system in supporting their creative expression, the degree of consistency of the actual story to the target story conceived through the input picture, and the ability of the generated story to capture and retain their attention.

References

Cua, J., Manurung, R., Ong, E., and Pease, A., 2010. Representing Story Plans in SUMO. *Proceedings of the NAACL-HLT 2010 Workshop on Computational Approaches to Linguistic Creativity,* 40-48, June 2010, Los Angeles, California: ACL.

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

Hong, A.J., Solis, C., Siy, J.T., Tabirao, E., and Ong, E., 2009. Planning Author and Character Goals for Story Generation. *Proceedings of the NAACL Human Language Technology 2009 Workshop on Computational Approaches to Linguistic Creativity*, 63-70, June 2009, Colorado, ACL.

Liu, H., and Singh, P., 2004. ConceptNet — A Practical Commonsense Reasoning Tool-Kit. *BT Technology Journal*, 22(4): 211-226. Netherlands: Springer.

Machado, J., 2003. Storytelling. *Early Childhood Experiences in Language Arts: Emerging Literacy*, 304-319. Clifton Park, New York: Thomson/Delmar Learning.

Mann, W., and Thompson, S., 1988. *Rhetorical Structure Theory: Towards a Functional Theory of Text Organization. TEXT*, 8(3):243-281.

Ong, E., 2009. Prospects in Creative Natural Language Processing. *Proceedings of the 6th National Natural Language Processing Research Symposium*, Center for Language Technologies, September 2009, Manila, Philippines: De La Salle University.

Ong, E., 2010. A Commonsense Knowledge Base for Generating Children's Stories. *Proceedings of the 2010 AAAI Fall Symposium Series on Common Sense Knowledge*, 82-87, November 2010, Virginia, USA: AAAI.

Peinado, F., and Gervas, P., 2006. Evaluation of Automatic Generation of Basic Stories. *New Generation Computing*, 24(3):289-302, Ohmsha, Ltd. and Springer.

Riedl, M. O. and Young, R. M., 2004. An Intent-Driven Planner for Multi-Agent Story Generation. *Proceedings of the Third International Joint Conference on Autonomous Agents and Multi-Agent Systems*, 186-193, Washington, DC, USA: IEEE Computer Society.

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

Torrance, E.P., 2007. *Creativity in the Classroom: What Research Says to the Teacher*. National Education Institution, Washington, D.C.

Turner, S.R., 1992. *Minstrel: A Computer Model of Creativity and Storytelling*. University of California, Technical Report CSD920057.

Uijlings, J.R.R., 2006. *Designing a Virtual Environment for Story Generation*. Master's Thesis, University of Amsterdam, The Netherlands.

Venour, C. and Reiter, E., 2008. A Tutorial for Simplenlg. http://www.csd.abdn.ac.uk/~ereiter/simplenlg