

A Novel Approach for Intelligent Reasoning of Stories using Ontology

A.Jaya

Research scholar, Anna University
Department of CSE

G.V.Uma

Assistant Professor, Anna University
Department of CSE

ABSTRACT

Stories are really an interesting one to read, especially kids love to listen to stories. It grabs the reader's attention and gives them great pleasure or excitement or thrill or suspense while reading it. Readers go along with stories and may have the expectation and predict the flow of story based on the previous situations narrated by the authors. Predicting the flow of story requires reasoning capacity to analyze the same. Human beings can easily reason the story based on their cognitive process whereas reasoning the stories by the system is not as easy and it requires a lot of intelligence to perform the same. This paper concentrates on to provide an environment for analyzing the stories on the basis of characters, events and the situations. It aims for reasoning the stories sentence by sentence based on the real world description using ontology. Ontology helps to investigate the stories by extracting the characters and events from the given story and provides the semantic relation among them. Ontology is formal explicit shared conceptualization. Ontology provides the domain knowledge which can be utilized for reasoning the stories semantically. Reasoning the stories based on the characters acts as a lead for the construction of the new different variety of stories with change in the characters, their nature and the events.

Categories and Subject Descriptors

I.2.7 [Natural Language Processing]: Discourse, *Language generation, language models, Language parsing and understanding, machine translation, speech recognition and synthesis and text analysis*

General Terms

Design of a Reasoner

Keywords

Story reasoning, ontology, knowledge transfer, semantics, and concepts

1. INTRODUCTION

A story or Narrative is temporally an ordered sequence of events to describe the incident or interaction between two characters or self description. Story is one of the methods in transferring knowledge within society and its subsequent generations. The traditional oral story telling that has evolved into our contemporary modes of narrative have been recognized as core to the transfer of knowledge within society [1]. Children learn their moral value and social responsibilities in the form of stories narrated to them by their parents' guardians and peers [2].

Stories are mainly used to teach good moral values and entertain them with different emotions. They are used to express the human's thoughts about the events, or interaction among two characters, or the objects in a creative way and they are the outgrowth of imagination and creativity. There are mainly three ways to express the stories such as text, audio, animation. Usually Stories can be expressed in the natural language in the form of text. Incorporating the story writing concepts into the computer system leads to automatic story generation which makes the computer as an 'Artificial author' to generate the story based on the user's desire. Automatic story generation system facilitates the users to generate story by selecting the different characters, suitable environment or settings, different objects, according to their wish and desire. The story will be generated based on the above parameter value. Change in parameter value, makes the author to be more curious about the generation of new story. Generally, the normal flow of the story is represented in Freytag triangle which is depicted in Figure-1.

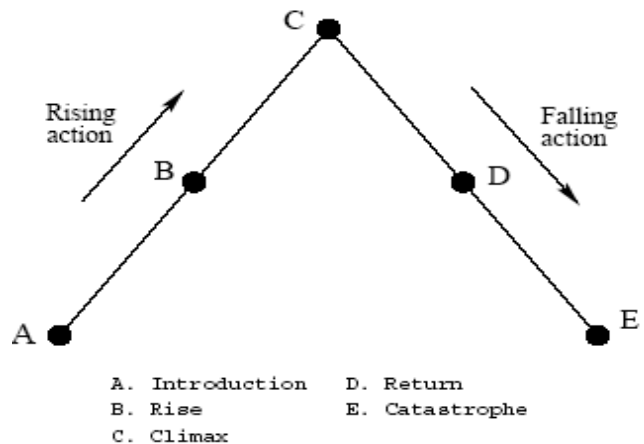


Figure. 1 Freytag triangle story

Freytag triangle depicts the normal flow of story with five nodes such as Introduction, Rise, Climax, Return and Catastrophe. Generally, the preliminary setup of story has been introduced with selected settings, environment and the situation in the introduction phase. Based on the situation, the main character has to react with situation which gives the rising action or the reaction. Because of rising action the activities related to the character comes to a climatic end which is known as the Climax. It is a moment of great or culminating intensity in a narrative or

drama, especially the conclusion of a crisis. It is also the turning point in a plot or dramatic action. After the peak situation, definitely, there will be a falling action which may have catastrophe end or it may be the good end [9].

The flow of the story depends on the theme conception or narrative sequences of events. Set of related order of events are responsible for the flow of story. Ontology possess different kinds of events which constitute the theme for a story.

Ontology is essential in conception of themes and generation of stories. They are now being recognized as important components of information systems and information processing. Ontology is a formal explicit specification of shared conceptualization [17]. It is expressed as the basic terms and relations comprising the vocabulary as concepts and the rules for combining terms and relations to define extensions to the vocabulary for a specific domain. The ontology is constructed based on the components of story domain. It contains characters, location settings like forest, palace, home and etc. Ontology helps in theme conception for generating new stories. Even though constructed ontology is domain specific, it can be updated, modified, reengineered, reused for other purposes also.

The section 2 offered an idea about the evolution of the story generation system which describes about the earlier system since 1968 to till date. The section 3 discussed semantic reasoning of stories. The section 4 insisted the role of ontology in semantic reasoning. The section 5 and 6 discussed about the results and the conclusion.

2. RELATED WORKS

According to Propp (1968), a cohesive story can be formed by connecting a series of any set of the thirty-one functions in order. This project explores this component of Propp's argument by randomly generating a fairy tale from selected functions. Each function has several passages written specifically to express that function, and the generator will randomly select one passage for each selected function. While each passage appropriately expresses its respective plot element, the tone, characters, and settings may vary [3]. Dynamic selection of Propp's function yields the text message but no guarantee of making the story more interesting. Semantic reasoning of the story helps to improve the interest and the characters originality.

Klein (1973, 1975) used real-world story structures such as canned story sequences and story grammars to generate the stories. They introduced the automated model for propp function for the two Russian fairy tales by utilizing the text grammar. The text grammar is a structural description of a linguistic performance to present the story [4]. Here, they generated the story sequences for two Russian fairy tales. But the problem here is possibility for generating variety of stories is less.

Meehan's Tale-Spin (1976) described the story telling in three modes. Two modes are interactive, asking the audience to make decisions about features of the story world, while one mode "fixes" the world to assure the production of particular stories. In the interactive mode, the user can introduce new character with the description and can be represented in conceptual dependency [5]. Here, Semantic reasoning of the generated story helps to

enhance the performance of dynamic change of characters in the existing model.

MINSTREL (1993, 1994) was a complex computer program which evolved in mid 80's that wrote short stories about King Arthur and his Knights of the Round Table. The system is a case-based problem-solver where past cases are stored in an episodic memory [7][8]. The above system retrieves existing stories as such; it requires the reasoning to produce the new kinds of story to grab the attention of readers.

MEXICA is a computer model released by Pérez y Pérez (1999) which was based on the engagement-reflection cognitive account of creative writing that produces stories about the Mexicas. MEXICA processes the Previous Stories to build structures in memory representing content and rhetorical knowledge. Story-actions and Previous Stories are defined in a text file through a set of syntactic rules known as definition languages designed for that purpose [10]. The above said system does not guarantee about meaning of stories as well as does not promotes user interest on stories. Semantic reasoning of the story helps to produce stories in a meaningful way.

Ferrucci's BRUTUS (2000) story generation system utilized the story-grammars to represent the structure of the stories it produces. It includes the descriptions of how to create phrases and sentences called paragraph-grammars and sentence-grammars. The outputs produced during the instantiation of the theme and the development of a plot processes are employed by the grammars during its expansion to generate the sentences and words that form the final output of the system. It contains the story elements like characters, characters' goals, events, etc. and are represented by structures that can be referred as story-frames [11]. Brutus proposed the story generation model with story grammar which helps to generate story automatically.

Hugo Liu (2002) proposed the MAKEBELIEVE story generation method which utilized the common sense knowledge for story generation in the year 2002 which inherits from both the structuralist and transformationalist traditions [6]. The above system follows the cause and effect pattern for story generation method. By introducing the semantic reasoning, the story could be more effective.

Peinado's (2004) Knowledge-Intensive Interactive Digital Storytelling (KIIDS) generates stories step by step, adding narrative episodes and simulation elements when they are needed to continue the storytelling process. KIIDS system mainly tried to resolve the problem of "Computational Creativity" i.e.: how to create something new and useful at the same time [12]. But the introduction of semantic reasoning can improve the performance of computational creativity.

In the existing models, the reasoning was not focused for the construction of stories. By integrating the reasoning with the story generation, definitely the reader will get the good satisfaction while reading the story.

3. REASONING THE STORIES

Reasoning comprises of all cognitive process activities that involve making or testing inferences to provide a solution. It is also closely related to problem-solving and creativity. Reasoning

the story gives an idea about concepts in semantic way. As per our story generation system, there is a provision for generation of short stories by selection of characters, settings, environments, objects, events and etc based on the user desire. But the problem is there is a chance of automatically generated stories to be meaningless or provide non-real world objects.

For example,

A tuple Event(x, y) represents for sentence generation. Where Event is an action performed by the actor 'x' in the destination 'y'.

Case1:

Input: Sleep (Lion, den) –Past tense;

Output: Lion was sleeping in the Den.

Lion slept in the Den

The above tuple generates the sentences in the form of Simple past tense and past continuous sentence. Output of the sentence generation in case 1 shows the simple concept that wild animal 'Lion' was doing the action 'sleep' in the location 'Den'. The real world concept is clear.

Case 2:

Input: fly (Crow, sky) – past

Output: Crow was flying in the sky.

Crow flew in the sky.

Output of the case 2 shows the simple concept that bird 'Crow' was doing the action 'fly' in the location 'sky'. The real world concept is clear and meaningful.

Case 3:

Input: sleep (Crow, nest) – past

Output: Crow was sleeping in the nest.

Crow slept in the nest.

Output of the case 3 shows the simple concept that bird 'Crow' was doing the action 'sleep' in the location 'nest'. The real world concept here is obvious.

Case 4:

Input: sleep (Lion, nest) – past

Output: Lion was sleeping in the nest.

Lion slept in the nest.

Case 5:

Input: fly (Lion, sky) – past

Output: Lion was flying in the sky.

Lion flew in the sky.

In the last two cases 4 and 5 shows the conflict in the real world concepts. Case 4 showing the statement that 'Lion was sleeping in the nest' is impossible and in the case 5 shows the sentence as

'Lion was flying in the sky'. It produces the semantically incorrect concept.

Semantic reasoner helps for identifying the real world concept and makes the story as meaningful one. Domain knowledge of the birds and animal world is required for performing the semantic reasoning. Ontology is used to provide the shared conceptualization of any kind of real world objects and their properties.

4. ROLE OF ONTOLOGY IN REASONING THE STORY

4.1 Ontology

Ontology is the science of what is, the kinds and structures of objects, properties, events and relations in every area of reality. For an information system, ontology is a representation of some pre-existing domain of reality which [13]:

- (1) Reflects the properties of the objects within its domain in such a way that there is a systematic correlation between reality and the representation itself
- (2) Is intelligible to a domain expert
- (3) Is formalized in a way that allows it to support automatic information processing

It is formally defined precondition for a computer to interpret the concepts and properties. Rules are used to infer the knowledge from the ontology. Typical elements of Ontologies are:

- Concepts and its attributes.
- Taxonomies to categorize concepts by generalization and specifications.
- Axioms to define statement, which are always true. They are used to prove the consistency of the knowledge modeled by an ontology and to deduce further facts
- Individuals are instances of concepts in relation

The criteria for high quality ontology are:

- Meaningful—all named classes can have instances
- Correct—captured intuitions of domain experts
- Minimally redundant—no unintended synonyms
- Richly axiomatised—(sufficiently) detailed descriptions
- Answer queries over ontology classes and instances, e.g.:
- Find more general/specific classes
- Retrieve individuals/tuple matching a given query
- Integrate and align multiple Ontologies

4.2 Ontology in reasoning

The basic building blocks are concepts, roles and individuals. Description logics are knowledge representation languages tailored for expressing knowledge about concepts and concept hierarchies [14]. Concepts describe the common properties of a collection of individuals and can be considered as unary predicates, which are interpreted as sets of objects. Roles are

interpreted as binary relations between objects. [15]. some of the language constructs in description logic are union, intersection, role quantification, etc which are helpful for defining the new concepts and roles. Since OWL was derived from DAML + OIL, it possesses the existing reasoning algorithms in Description Logics (DL). The main reasoning tasks are [16].

- Satisfiability of a concept - Checks about the description of concept exists or contradictory.
- Subsumption of concepts – It determines whether concept c subsumes the concept D.
- Consistency of ABox with respect to TBox - determine whether individuals in ABox do not violate descriptions and axioms described by TBox.
- Check an individual - check whether the individual is an instance of a concept
- Retrieval of individuals - find all individuals that are instances of a concept
- Realization of an individual - find all concepts, which the individual belongs to, especially the most specific ones

To reason the concepts, it needs the matching with existing ones in the ontology. There are three types of matches:

Exact match - the concept to be found is found,

Plug-in match - the concept to be found is more specific than the Concept in ontology, and

Subsume match - the concept to be found is more general than the concept in ontology.

The scoring function of matching degree is given below:

Exact Match > *Plug in Match* > *Subsume Match*

Based on the level of matching, the information will be retrieved and processed accordingly. For example, the part of the ontology contains

Living being → animals → wild → Lion (king, legs, anger, roar, yellow, Den)

→ Domestic → Rat (small, legs, frightens, gray, holes)

→ Bird → crow (friendly, legs, caws, fly, black, nest)

Tuple in the ontology for living beings concept is

“Noun (speciality, physical feature, emotions, color, locateAt)”

In the case 4, the generated sentence is ‘Lion was sleeping in the nest’. In the concept of Lion, one of the attributes is “locateAt – Den “. In the above generated sentence, the Lion was sleeping in the nest. It is conflicts with concept Lion. The sentence can be changed in anyone of the following options to provide the real world meaning.

Crow was sleeping in the nest.

Lion was sleeping in the Den.

The User can select anyone of the option to generate a story for the readers. The Figure 2 shows the semantically corrected story for the sentence “Rat Killed the Lion” has been corrected as “Rat killed by the Lion “.

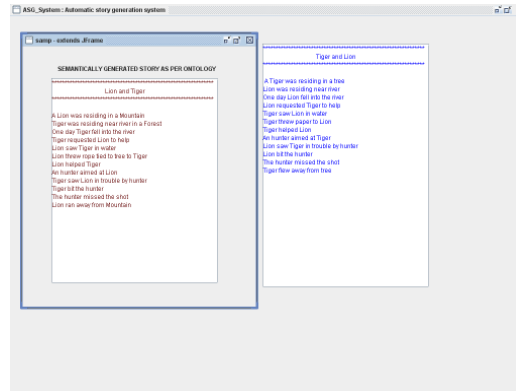


Figure. 2 Semantically corrected story

4.3 Ontology development tools

Protégé tool is used for construction of ontology for the story domain. It provides extensible knowledge model to enable users to redefine the representational primitives. Using Protégé tool the ontology is constructed and the owl file is generated. This file is then parsed using Jena parser for owl parsing which helps in identifying the class subclass relationships. The extracted knowledge from ontology is the input for story generation phase where the generated story will be checked for semantic correctness of each sentences based on domain ontology concepts.

5. RESULTS

The experimental results obtained for the effective retrieval of concepts and sub concepts in the ontology. The retrieval and matching with the existing concepts in the ontology helps to reason the sentences. The ontology has more number of concepts and attributes to represent the domain of the story. The required concepts can be obtained appropriately from the ontology. The unavailability of concepts needs to be upgraded through the ‘update module’ of ontology. Checking for the unavailability of terms or concepts from the base is mandatory.

Test 1:

Out of 120 number of testing, the details about the retrieval of terms are listed below: 54 attributes are retrieved appropriately (exact match), 42 are approximately retrieved and 24 are retrieved not relevantly or saying that the search is unsuccessful.

Table.1 results of the terms retrieved from ontology

S.No	Nature	No of terms attempts	No of retrieved terms	% of retrieval
1.	Exact	120	54	45
2.	Relevant	120	42	35
3.	Unrelated	120	24	20

The Exact match retrieval was improved by concentrating the retrieval of terms with their synonyms and their semantics. The semantic retrieval of terms or concepts from the ontology helps to improve the accuracy of semanticness in the story generation system. The second test result was given in table 2.

Test 2:

Out of 150 number of testing, the details about the retrieval of terms are listed below: 96 attributes are retrieved appropriately (exact match), 35 are approximately retrieved and 19 are retrieved not relevantly or saying that the search is unsuccessful.

Table.2 results of the terms retrieved from ontology

S.No	Nature	No of terms attempts	No of retrieved terms	% of retrieval
1.	Exact	150	96	64
2.	Relevant	150	35	23
3.	Unrelated	150	19	13

The Figure 3 depicts the comparative results of both the test 1 and test2. The test 2 results have been improved because of the enhancement of ontology with the synonyms of the terms.

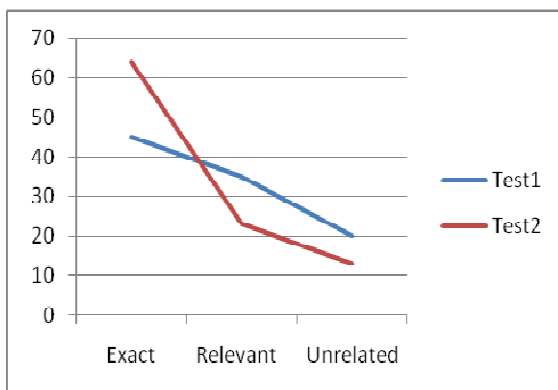


Figure 3 results of the terms retrieved from ontology

6. CONCLUSION

This is the first step towards the semantic reasoning sentences to produce the meaningful story. Semantics of story can be validated by the semantic checker in the automatic story generation system using ontology. Performance can be enhanced by further

extending the ontology with more number of concepts and the attributes. This project lacks in the ontology updation. Even though Ontology has the provision for extending with new concepts, Ontology alignment is also crucial for reusing the existing Ontologies and for facilitating their interoperability. In future, the domain ontology has to be enhanced with widespread information and their meanings. The reasoning should be concentrated on all types of sentences. This project can be further enhanced for deriving tacit knowledge from the axioms which helps to improve the semanticness of the story. By concentrating on these three factors, automatic story generation system can be updated with interesting ones.

7. ACKNOWLEDGEMENT

The authors would like to express their thanks to Mr. Abdul Qadir A. Rahman Bubari, the Correspondent, Dr. V.M.Periasamy, the Principal and Dr. P.Sheik Abdul khader , HOD / Department of CA, B.S.A.Crescent Engineering College Chennai, Tamilnadu, India and Dr.Chellappan, HOD, Department of Computer science engineering, Anna Univerisity , Chennai , Tamilnadu, India for the environment provided.

8. REFERENCES

- [1] Alani, H., et al.: Automatic Ontology-based Knowledge Extraction from Web Documents. IEEE Intelligent Systems, 18 (1) (Jan-Feb 2003), 14-21.
- [2] Tuffield, M. M., Shadbolt, N. R. and Millard, D. E. (2005) Narrative as a Form of Knowledge Transfer: Narrative Theory and Semantics. In: 1st AKT Doctoral Symposium, June 2005, Milton Keynes, UK.
- [3] Propp, Vladimir. "Introduction Theory and History of Folklore". Ed. Anatoly Liberman. University of Minnesota: University of Minnesota Press, 1984. pg ix.
- [4] Klein, S., L.A. Price, J.F. Aeschlimann, D.A. Balsiger and E.J. Curtis, (April 1975). "A Meta-symbolic Simulation Model for Five Myths from Levi-Strauss" The Raw and The Cooked University of Wisc. Computer Science Department Tech Report Presented at 2nd International Conference on Computers in the Humanities, Chicago.
- [5] Meehan, James R. (1976). "The Metsnovel: Writing Stories By Computer". Ph. D. Thesis. Reserch Report #74, Department of Computer Science. Yale Univeraity, Mew Haven, CT.
- [6] Hugo Liu, Push Singh. (2002). "MAKEBELIEVE: Using Common sense to Generate Stories". Proceedings of the Eighteenth National Conference on Artificial Intelligence, AAAI 2002, July 28 - August 1, 2002, Edmonton, Alberta, Canada. AAAI Press, pp. 957-958.
- [7] Turner S R (1994), "The Creative Process: A Computer Model of Storytelling", Erlbaum Associates, Hillsdale NJ
- [8] Turner S R (1993.), "MINSTREL: A computer model of creativity and storytelling", PhD Dissertation, University of California LA.
- [9] <http://www2.cnr.edu/home/bmcmansu/freytag.html>
- [10] Pérez y Pérez R and Sharples M (2001), "MEXICA: A computer model of a Cognitive account of creative writing,

Journal of Experimental and Theoretical Artificial Intelligence vol 13 pg: 119-139.

- [11] Bringsjord S, and Ferrucci D A (2000.), “Artificial Intelligence and Literary Creativity Inside the Mind of BRUTUS, a Storytelling Machine”, Lawrence Erlbaum Associates, Hillsdale NJ.
- [12] Federico PEINADO and Pablo GERVAZ (2006), “Evaluation of Automatic Generation of Basic Stories”, New Generation Computing 24(3):
- [13] http://www.ontologyworks.com/what_is_ontology.php
- [14] Bosch, Mela.”Ontologies, Different Reasoning Strategies, Different Logics, Different Kinds of Knowledge Representation: Working Together”. Knowledge Organization, Ergon Verlag. 33(3) 153-159, ISSN 0943-7444 Knowl.Org. 2006
- [15] Lambrix, Patrick. “Description Logic. Document of: Intelligent Information Systems” Laboratory, Department of Computer and Information Science, Linköping University, S-581 83 Linköping, Sweden 2003
- [16] Ong Siew Kin, Tang Enya Kong “Conceptual Modeling and Reasoning using Ontology” National Computer Science Postgraduate Colloquium 2005 (NaCSPC’05) School of Computer Science, Universiti Sains Malaysia, Penang, Malaysia. June 2005.
- [17] Natalya F. Noy and Deborah L. McGuinness. “Ontology Development 101: A Guide to Creating Your First Ontology”. Stanford Knowledge Systems Laboratory Technical Report KSL-01-05 and Stanford Medical Informatics Technical Report SMI-2001-0880, March 2001