Computational Modeling of Narrative Structure: A Hierarchical Graph Model for Multidimensional Narrative Structure

Taisuke Akimoto
Kyushu Institute of Technology
Fukuoka, Japan
akimoto@ai.kyutech.ac.jp

ABSTRACT: *This paper proposes a basic and general model of narrative structure representation for computer processing in order to provide a common basis for both generative and analytic processing of narratives. In particular, we address the following two aspects. First, we formalize a multidimensional framework to integrate the three structural aspects of a narrative: the story world, which is the background world structure; the story, which represents chronologically organized events as the content of a narrative; and discourse, which represents the structure used for the expression of a narrative. Second, we propose a hierarchical graph model as a common mode of representing each structural dimension. This model organizes various types of structural elements (e.g., entities, events, and discourse units) by hierarchical grouping and network-like linkage. Then, we discuss the advantages and future challenges of the proposed model from the perspective of representational capacity by presenting several examples based on the experimental implementation of the model.*

Keywords: Artificial Intelligence, Intelligent Narrative Technology, Computational Modeling of Narrative, Narrative Structure, Story and Discourse, Hierarchical Graph

Received: 12 February 2017, Revised 1 April 2017, Accepted 5 May 2017

1. Introduction

Narrative is a universal form of human information and communication. Computer technologies for narrative information processing have a wide range of applications. For example, the narrative is the most important mode for the communication of knowledge or information among humans. This will continue to be true in the context of human-computer communications. Thus, narrative abilities including both the generation and interpretation processes are important in the implementation of human-friendly computers or intelligent agents. In addition, narrative generation technologies will be applied to automatization, or to the aid of narrative content creation for various purposes, e.g., entertainment, education, advertisement, and news reports.

The modeling of a narrative structure is an important but difficult task in the design of both generative and analytic narrative processing. In particular, the core task of narrative generation is to compose a narrative structure based on some type of creative or communicative goal. Conversely, the task of narrative interpretation or understanding includes the composition of a semantic structure from an expressed narrative.

In the field of artificial intelligence (AI), many studies on narrative processing have been conducted; these studies range from basic theories (e.g., [11, 34, 35]) to comprehensive or applied systems (e.g., [4, 19, 23]). However, the common basis of narrative structures has not been sufficiently discussed at the computational level.

This study adopts a structure-centered approach to intelligent narrative technology. Thus, the purpose of this study is to design a basic model for computational representation of narrative structure according to the following three design directions:

- *Representational capacity* for a flexible and comprehensive treatment of the fundamental narrative elements and structures.
- *Theoretical and formal simplicity* for computer implementation.
- *Extensibility* for its application based on various purposes and thoughts.

Although the design of the proposed model is directed toward a general-purpose framework for the computation of narrative structures, the detailed requirements of a narrative structure vary depending on application-specific issues, purposes, concepts, or ideas in the system design. Hence, the proposed model provides only a minimal and abstract framework for the comprehensive representation of diverse narratives. The model can be used with an adaptational extension for various application purposes. In addition, a narrative is a complex object involving ill-defined natures; therefore, simplification is necessary for computer-based modeling. A model that has a good balance between representational capacity and computational simplicity should be designed. From this above standpoint, we propose a comprehensive model of narrative structure representation based on object-oriented modeling.

In Section 2, previous computational models of narratives are reviewed from a structural perspective and the key features of the proposed model is clarified; then, in Sections 3 and 4 the theories of multidimensional framework for a narrative and hierarchical graph model for a narrative, respectively, are described. In Section 5, several representation examples based on the proposed model are presented. In Section 6, difficult challenges for future expansion of the representational capacity are considered. Finally, the conclusion is stated in Section 7.

2. Background

In this section, we review the various approaches used to model narratives in the field of AI. First, we describe two different aspects—i.e., story and discourse—of a narrative based on narratological terminology. Second, we categorize previous models of narrative structure representation into tree- and graph-based structures, and discuss the advantages and disadvantages of each type.

2.1 Story and Discourse

The fundamental nature or structure of narrative has been explored in the field of narratology. According to Prince’s “The Dictionary of Narratology” [29], narrative is defined as “[t]he representation (…) of one or more real or fictive [events] communicated by one, two, or several (…) [narrators] to one, two, or several (…) [narratees].” In narratology, the two planes of narrative—content (*what*) and expression (*how*)—are distinguished by the terms “story” and “discourse.” Thus, the content plane corresponds to a chronological sequence(s) of events and the expression plane corresponds to the manner of narrating the events, such as the “order” and “perspective” (or point of view) in/from which the events are narrated.

By using the above distinction, the main issues addressed in most of the early AI studies on narrative can be positioned as the content plane (story). The composition of a coherent and dramatic sequence of events is a significant issue in computational narrative generation [19, 27, 32, 37]. The main challenge in narrative understanding is the parsing and inferring of the story—i.e., events, situations, and their higher-level meaning—from a narrative text [22, 36]. Knowledge models on events, e.g., “conceptual dependency,” “script,” and “goal” [35], provide the theoretical basis for such studies.

Computational models relevant to the expression plane (discourse) also exist. The knowledge model of a story grammar [34] provides a cognitive framework to understand and generate structured narrative texts. Discourse structure models in terms of natural language processing—e.g., rhetorical structure theory (RST) by Mann and Thompson [17] and the concept of coherence relations by Hobbs [11]—provide methods to represent a logical organization among sentences. These models or theories have been applied to computational models for both generation and understanding of narratives [4, 12, 25, 28].

Since the early 2000s, the aspect of narrative discourse has attracted an increasing amount of attention in computational studies on narratives, with the expansion of interdisciplinary approaches with narratological knowledge [9, 16, 23]. Several recent studies on computational narrative generation apply the narrative discourse theory developed by Genette [7] to the structural modeling of narrative discourse generation [2, 14, 20, 23, 26]. The structural notions on narrative discourse based on Genette's theory [7] are also incorporated into the computational studies on narrative analysis or understanding. For example, Mani [16] proposed NarrativeML, a scheme of structural annotation for narrative texts.

A full-fledged system of narrative generation or understanding must unify the processes of the story and of the discourse. Several unifying studies in computational narrative generation have been conducted. For example, an integrated narrative generation system by Ogata and his colleagues [23] is based on a three-phased architecture consisting of the generative mechanisms for story structure, discourse structure, and surface expression including natural language, image, and music. In addition, several researchers have been discussing system frameworks for the integration of multiple narrative aspects including story and discourse [8, 15, 33].

2.2 Tree and Graph

Previous structural models of narrative can be broadly categorized into tree- and graph-based models (Figure 1). We consider the advantages and disadvantages of both the models.

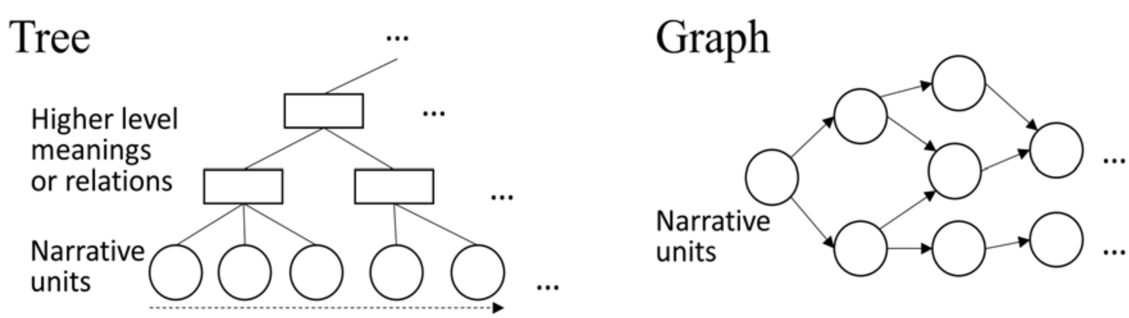


Figure 1. Tree- and graph-based models

2.2.1 Tree-Based Models

In tree or hierarchical structures of narratives, a terminal node generally corresponds to a narrative unit such as an event or discourse segment. An internal node provides a higher-level meaning or relation binding the child nodes, e.g., the main or subgoals in planning-based story generators [19], functional or compositional categories in schematic models for narrative [12, 28, 34], and various types of relations in discourse structure models for natural language processing [11, 17] and "narrative tree" models for narrative generation [23, 24].

Tree-based models can organically operate a story or discourse structure based on a whole-part structure. This approach is advantageous in integrating the macroscopic and microscopic processing of narrative structures. However, a tree structure basically assumes a linear and clearly organized structure; therefore, its representational capacity and operational flexibility are restricted. In particular, tree-based modeling is unsuitable in the handling of parallel or branched structures, fragmentary or complexly organized structures, and multiple relations among nodes.

2.2.2 Graph-Based Models

In a graph structure of a narrative, a node generally corresponds to a narrative unit and an edge defines a pathway or relation between any two nodes. A typical usage of a graph structure is to represent the potential course of events. In particular, a graph structure is widely used in interactive narrative systems [30, 31]. Typically, an interactive narrative includes multiple potential storylines that are branched based on the actions of a user-controlled character in the virtual world. Such branched storylines are represented as a directed graph in which the events or scenes (nodes) are connected based on the actions (edges) selected by a user-character. A graph structure is also used in the knowledge representation of a script-like chain of events as a potential course of events in a specific situation [5, 18].

León [13] proposed a knowledge representation model for narrative-form memories including episodic and procedural memories. His knowledge model represents a narrative (more precisely, the primary focus of this model is on the story aspect in terms of narratology) as a network-like graph structure associating narrative units with various relations including causality, relative time relation, and abstraction.

Graph-based models have an advantage in terms of their structural flexibility. Graphs can represent branched or parallel structures and multiple relations among nodes. However, planar graph structures have no higher-level structures; therefore, the flexible operation of large narrative structures is difficult.

2.3 Key Features of the Proposed Model

In this study, we design a narrative structure model from the perspective of narrative generation in contrast to most of the previous narrative structure models that were designed from an analytical perspective. Hence, the primary focus of the proposed model is to provide a computationally manipulatable representation of a narrative structure. Although NarrativeML [16] provides a systematic scheme of a narrative structure, the proposed model adopts an essentially different approach. While León's model of narrative memory [13] focuses on the mental structure of narrative-form memories in a cognitive architecture, the proposed model focuses on the complex structure of the narrative itself.

This study introduces the following two novel theories of the proposed computational model:

• **Unification of multidimensional narrative structure:** The model introduces a nesting structure to consistently unify the four dimensions of the narrative, including the three structural dimensions (story world, story, and discourse) and a surface dimension (expression).

• **Hierarchical graph structure:** The model merges the advantages of the previous tree- and graph-based models, i.e., the potential of a hierarchical organization and flexible relationships between various narrative elements, in the form of a hierarchical graph structure.

Overall, this study develops a framework for unifying multiple aspects of a narrative structure. This framework provides the capacity for treating complex narrative structures in contrast to previous one-dimensional models based on tree- or graphbased structures. The framework is particularly useful for creative tasks, e.g., generation, editing, manipulation, and recomposition of narrative structures. However, this framework may lead to the complication of programs requiring simplicity. In such cases, a tree- or graph-based representation may be suitable.

3. Multidimensional Framework for a Narrative Structure

Based on the background described in Section 2, we design a comprehensive model of narrative structure representation for computer processing. The first step is to formalize a multidimensional framework for a narrative structure by dividing and combining the different structural aspects of a narrative. From a computational perspective, the two planes of narrative—i.e., story and discourse in narratological terminology—can be further divided as shown in Figure 2 (this division is partially similar to Chatman's [6] classification of narrative elements). First, the content plane consists of two structural aspects—events and the background world. We use the term Story (in a narrow sense) for event data and StoryWorld for world data Second, the expression plane can be divided into a surface expression and its structural representation. We use the term Discourse (in a narrow sense) for a structural representation and Expression for a surface expression. These four dimensions are defined below:

• **StoryWorld:** The background world structure containing information about the attributes and states of the entities (characters, objects, places, etc.). It is similar to an ontology of a story.

• **Story:** The structure of the chronologically organized events in a StoryWorld.

• **Discourse:** The structure of how a Story is organized into an Expression.

• **Expression:** The surface expression using natural language, still and/or moving images, sounds, or other forms of media.

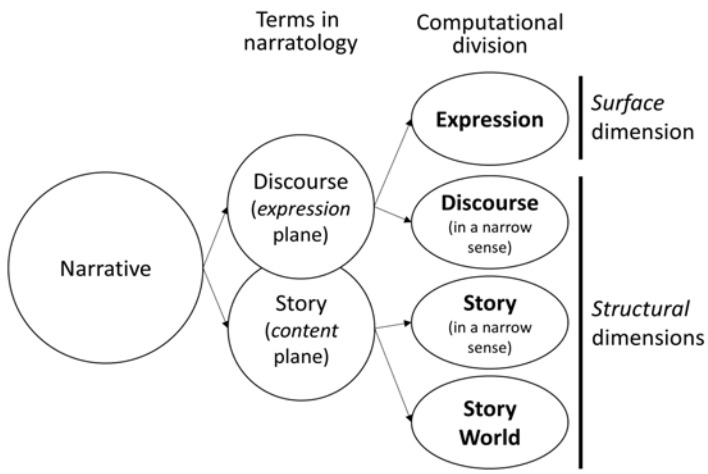


Figure 2. Division of narrative dimensions

In this model, the three structural dimensions are combined by using a nested structure (Figure 3), i.e., a Story contains a StoryWorld as the background information, and a Discourse includes a Story as the content information. An Expression is generated based on the Discourse. This framework assumes the following two principles for the process of narrative structure composition:

- The included dimension must be available in order to compose the including dimension, i.e., a StoryWorld (or Story) must be composed before the Story (or Discourse) composition process.
- The composition process of the including dimension can optionally operate the structure of the included dimension, i.e., the structure of the StoryWorld (or Story) may be operated during the Story (or Discourse) composition process.

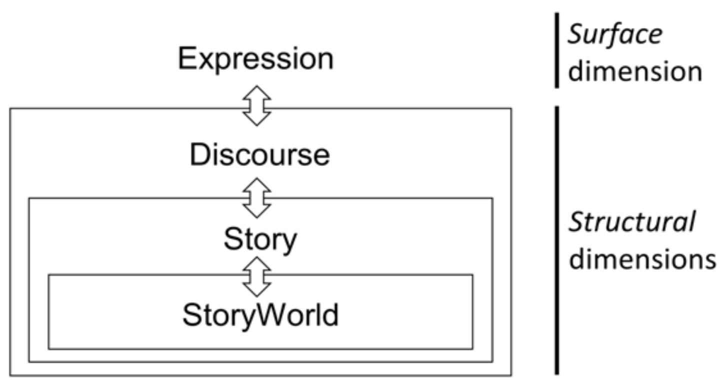


Figure 3. Relationship among StoryWorld, Story, Discourse, and Expression

In the above framework, a simple architecture of the narrative generation process is a pipeline method from StoryWorld to Discourse; this architecture ignores the second principle. However, the second principle provides a bidirectional architecture that allows a more flexible generation process, e.g., the story world generator composes an outline of StoryWorld, and then, the story generator details or changes the StoryWorld when it generates a Story.

We assume that the mechanism of ideal narrative interpretation will also be modeled as a process to compose the structural representations of StoryWorld, Story, and Discourse by analyzing an input narrative text. In particular, a StoryWorld is composed from signs representing objects and situations in an input narrative text. Similarly, the structures of Story and Discourse are composed from signs representing events and discourse-level structures, respectively. We will present an example of this concept in Section 5.3.5.

4. Hierarchical Graph Model

Next, we develop a common model to represent each of the aforementioned three structural dimensions: StoryWorld, Story, and Discourse. Our approach is to merge the advantages of the previous tree- and graph-based models. Based on this perspective, we propose a hierarchical graph model in which various types of narrative elements are organized by hierarchical grouping and network-like linkage. Each structural dimension is qualitatively different; the hierarchical graph model is applied to each dimension. A formal unification among the dimensions will contribute to the theoretical and mechanical simplification of full-fledged systems. For a computer implementation of this model, we adopt object-oriented modeling. The following basic design corresponds to an organization of classes for a hierarchical graph-based narrative representation.

4.1 Basic Framework

The hierarchical graph model consists of three types of structural elements: SE (semantic element), Group, and Link. Figure 4 shows a structural diagram of this model. SE/Group and Link correspond to a node and edge, respectively. Link includes both directed and undirected links. The basic concepts and functions of each element are defined below:

- SE contains semantic information for the basic unit in each structural dimension of a narrative.
- Group integrates any number of nodes (SEs and/or Groups) into a broader unit. This function is similar to the linguistic abstraction or categorization of the member nodes. A Group itself also has semantic information as a comprehensive meaning that binds or abstracts the member nodes.
- Link connects any two nodes with a relation or a set of connective attributes.

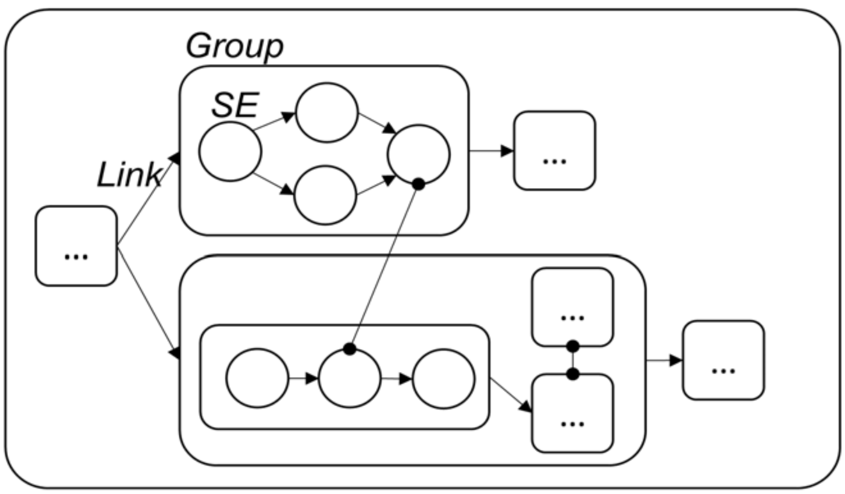


Figure 4. Structural diagram of the hierarchical graph model

The model provides the subclasses specialized to represent each structural dimension, as listed in Table 1. In addition, for structural simplicity, the model sets basic constraints on multiple grouping of a node and linkage beyond the grouping hierarchy, as shown in Table 2. The following subsections define each class specified in Table 1.

	SE	**Group**	**Link**
StoryWorld	Entity	wGroup	wLink
Story	Event	sGroup	*sLink (tLink, cLink)
Discourse	DU	dGroup	dLink

*sLink is further divided into the two subclasses, tLink and cLink.

Table 1. Subclasses of SE, Group, and Link for each dimension

	Multiple grouping	Linkage beyond grouping hierarchy
StoryWorld	permitted	permitted
Story	not permitted	tLink: not permitted cLink: permitted
Discourse	not permitted	not permitted

Table 2. Basic constraints on structuring

4.2 StoryWorld Structure

4.2.1 Entity

An Entity refers to an individual existence in a StoryWorld. Entity can be divided into several subclasses, e.g., Human, Animal, Plant, Artifact, NaturalObject, and Place. An Entity is represented by a set of attributes in a manner similar to the "frame" model. For instance, Human attributes may include "name," "age," "sex," "job," "physical feature," "clothes," and "personality."

4.2.2 wGroup

A wGroup contains Entities and/or wGroups. The grouping or categorization perspectives include a human group (e.g., family, community, and social organization), spatial or geographical unit (e.g., country, city, building, and island), and functional kinship (e.g., instruments of a festival). For example, a human organization such as the International Olympic Committee (IOC) is represented as a set of Human members. Here, "IOC" corresponds to the higher-level meaning. In addition, a member of the IOC can be a member of other wGroups such as "their family."

4.2.3 wLink

A wLink relates any two nodes in a StoryWorld. The types of relations indicated by wLink may include interpersonal relations, spatial relations, possessive relations, and functional relations. wLink includes both directed and undirected links. For instance, "X is a teacher of Y" is defined as a directed link, and "X is a friend of Y"—which is equal to "Y is a friend of X" from an objective perspective—is defined as an undirected link.

4.3 Story Structure

4.3.1 Event

An Event refers to a semantic representation of an action by an Entity or a happening in the StoryWorld. An Event is basically represented as a "case frame." For instance, the Event "Taro eats an apple" is represented as "(Eat (agent Taro) (object Apple))."

4.3.2 sGroup

An sGroup contains Events and/or sGroups. The grouping or abstraction perspectives include scene or temporal and/or spatial continuity (e.g., events at a park on a Sunday), key person (e.g., events around Taro), a functional role in the story (e.g., departure, struggle, or solution), and other types of abstractions. For example, the sequence of Events in "On a Sunday at a park, Taro played with a ball. Next, Taro played with sand…" can be grouped with the abstracted event, "Taro played at a park on a Sunday."

4.3.3 tLink

A tLink defines the relative temporal relation between any two nodes for the chronological organization of Events and sGroups. By applying Allen's classification of temporal relations between events [3], we provide the following relation types for tLink: "before," "equal," "meets," "overlaps," "during," "starts," and "finishes." "Before" and "equal" are the most fundamental types for a temporal organization. In particular, a sequential chain of nodes is basically composed of "before" relations, and a parallel or simultaneous structure consists of "equal" relations. The absolute time information (e.g., the date and hour) of a node is optionally defined as an attribute of the node.

4.3.4 cLink

A cLink defines a causal relation between any two nodes. Causal relations can be classified into several types. For example, Abe and Muramoto [1] categorized causal relations as "result," "enablement," and "motivation."

4.4 Discourse Structure

4.4.1 DU

A discourse unit (DU) is a minimum segment of a Discourse. Based on the terms of narratology [29], a DU is divided into three subclasses: Narration, Description, and Commentary. A Narration is the most basic unit that recounts an Event or sGroup. It has attributes regarding “voice” [7], i.e., the situation of a narration by a narrator and narratee, with its content (i.e., Event or sGroup). Description refers to a static expression for an Entity or wGroup, and Commentary refers to a commentarial excursus by the narrator about an Entity, Event, or another aspect.

4.4.2 dGroup

A dGroup contains DUs and/or dGroups. It forms a semantic or formal segment of a Discourse, similar to a paragraph, section, or chapter in terms of narrative prose. The grouping structure of a Story can be directly reflected in the grouping structure of a Discourse (e.g., the Events in an sGroup corresponding to a scene can be packaged into a dGroup that narrates that scene); however, a dGroup can freely compose a part of a discourse based on various compositional perspectives that are not bound by the Story structure.

4.4.3 dLink

A dLink provides a pathway between two nodes with their relation or connective logic. For example, coherence relations [11] or rhetorical relations [17]—such as elaboration, background, and parallel—will be used for this element. The thread(s) of a narrative is (are) composed of dLinks by defining the order in which the DUs are expressed.

4.5 Implementation

We implemented a prototype of the above model by using Java. The classes of the structural components are organized as shown in Figure 5 (where each colon indicates an inheritance). Methods to represent the semantic information of the subclasses of SE, Group, and Link are specified in a simple manner.

```
Node:
  SE:
    Entity:  Human, Animal, Plant, Artifact, NaturalObject, Food, Place
    Event
    DU:  Narration
  Group:
    wGroup
    sGroup
    dGroup
Link:
  wLink
  sLink:
    tLink
    cLink
  dLink
```

Figure 5. Class hierarchy of structural components

StoryWorld, Story, and Discourse are each defined as a class consisting of the above components in the field variables of each dimension class. The StoryWorld and Story classes include a method to compose a class instance from hand-coded text data that define the information related to the instances of Entity/Event, wGroup/sGroup, and wLink/sLink. The Discourse class includes three types of simple algorithms that compose different Discourse structures by using path searches of the Story structure. These algorithms are exemplified in Section 5.1. (Because the main focus of this study is to develop a representation model for narrative structure, we omit the detailed explanation of these algorithms in this paper.)

5. Validation of Representational Capacity based on Examples

We analyze the representational capacity of the proposed model by presenting several examples of narrative structure representation based on the proposed model; further, we discuss the advantages and future challenges related to representational capacity.

5.1 Example Representation (1)

We present narrative structure representations based on the proposed implementation. We created by hand a StoryWorld and Story for a simplified fictional version of the Olympic Games, as illustrated in Figures 6 and 7. The semantic information of the Entities and Events is annotated on the right side of the figures.

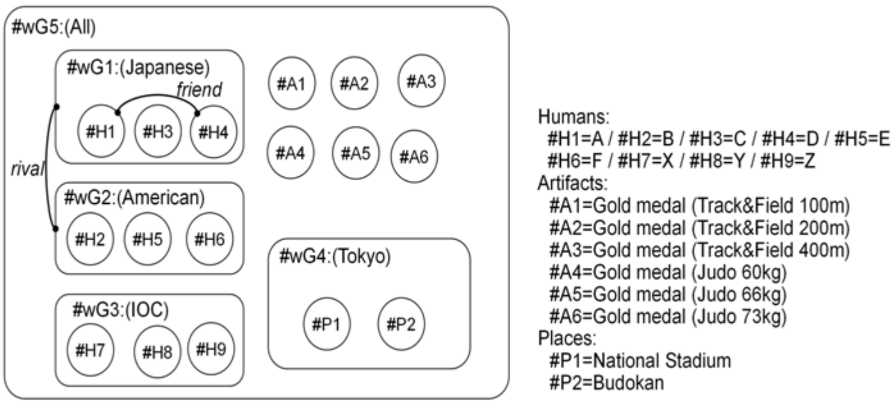


Figure 6. Diagram of a hand-coded StoryWorld structure[1]

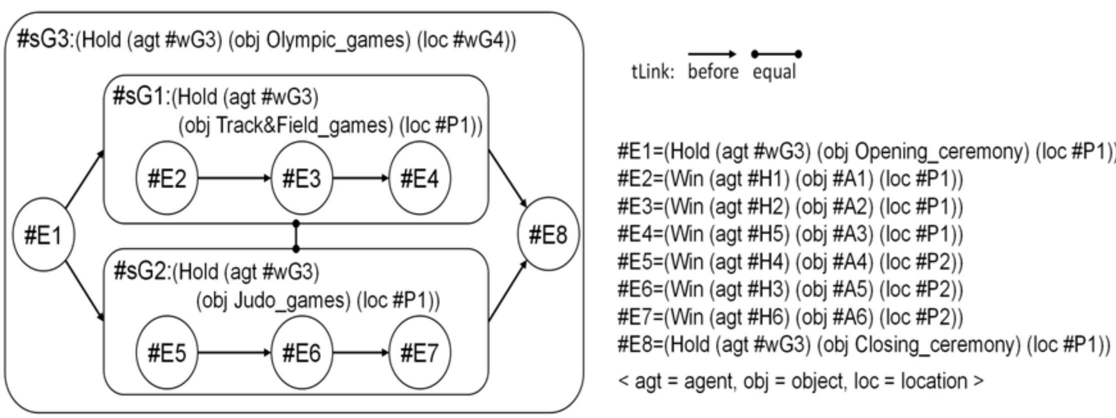


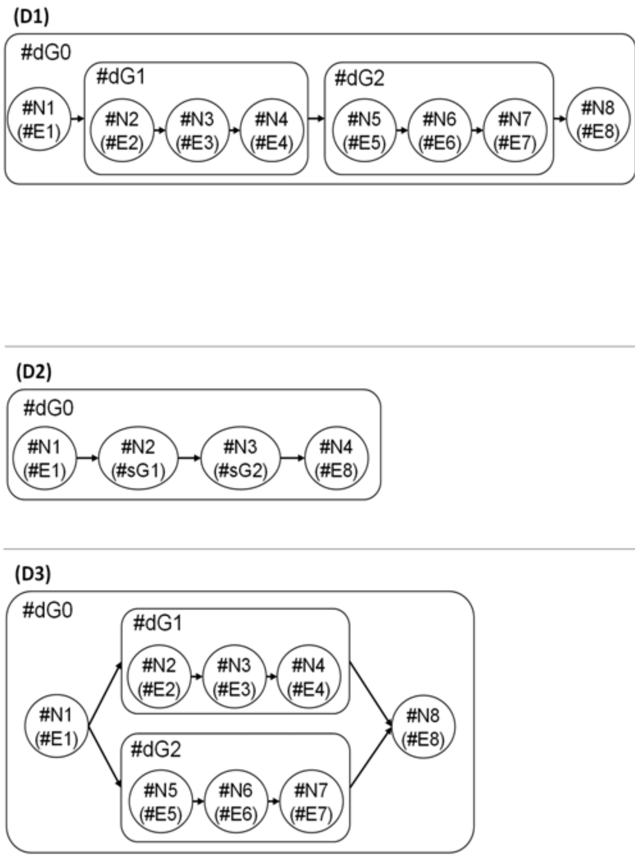
Figure 7. Diagram of a hand-coded Story structure

In the StoryWorld (Figure 6), 17 Entities including Humans (#H1–9), Artifacts (#A1–6), and Places (#P1–2) are organized into five wGroups (#wG1–5) and two wLinks. For example, "IOC," the host of the Olympic Games, is represented as a set of three Humans (#wG3).

In the Story (Figure 7), eight Events (#E1–8) are organized into three sGroups (#sG1–3) and nine tLinks. No causal relation (cLink) exists among the nodes. The semantic information of an Event or sGroup is represented with reference to the nodes of the StoryWorld. wGroup can also be a component of an Event, e.g., #wG3 ("IOC") is referred to as the agent in #E1, #E8, and #sG1–3. The sGroups #sG1 and #sG2 bind #E2–4 and #E5–7, respectively, from the perspective of the types of sport. These two parts are temporally organized as parallel courses.

[1] *This model permits the presence of isolated nodes that do not have a Link (edge) to other nodes.*

Next, we show different Discourses that are composed from the Story by using the three types of simple algorithms. Figure 8 illustrates the structures of the three Discourses with their English versions generated by hand. In the figure, each small circle indicates a Narration—e.g., #N1(#E1) represents a narration of #E1—and each arrow denotes a dLink. In this prototype version, dLinks only define the pathways between the nodes and do not provide a connective logic.



IOC holds the opening ceremony at a national stadium.

A wins a gold medal (100 m track and field) at the national stadium. B wins a gold medal (200 m track and field) at the national stadium. E wins a gold medal (400 m track and field) at the national stadium.

D wins a gold medal (60 kg judo) at Budokan. C wins a gold medal (66 kg judo) at Budokan. F wins a gold medal (73 kg judo) at Budokan.

IOC holds the closing ceremony at the national stadium.

IOC holds the opening ceremony at the national stadium. IOC holds the track and field games at the national stadium. IOC holds the judo games at Budokan. IOC holds the closing ceremony at the national stadium.

IOC holds the opening ceremony at the national stadium.

<#dG1 (track and field games) or #dG2 (judo games) is inserted here.>

IOC holds the closing ceremony at the national stadium.

Figure 8. Diagrams of three Discourse structures

In Figure 8, D1 narrates all Events based on a breadth-first order search including recursive searching inside sGroup type nodes. The start node is #E1. D2 is also based on a breadth-first order search; however, it narrates at an abstracted level by referring to the semantic information of each sGroup (#sG1 and #sG2). This type of discourse variation corresponds to the "duration" [7], the narratological notion about the relation between the story time (length of time) and discourse time (length of the text). D3 has a branched structure, and the reader can choose #dG1 or #dG2 as the next path for #N1.

5.2 Example Representation (2)

We demonstrate another example using Aesop's fairy tale "The Fox and the Crow" (Figure 9). This example consists of two phases. First, we show a representation of the narrative structure of this tale. Second, because the proposed model aims to provide a basis for structural manipulation of a narrative, we present a recomposed version of that structure. These example structures were manually developed. Developing mechanisms for such interpretive and generative processes based on the proposed model is outside the scope of this study.

In Figure 10, we show a set of the structural representations of StoryWorld, Story, and Discourse for this tale. These structures were composed based on our interpretation of the text in Figure 9, including several supplementations of implicit elements. For example, because the place in which the events occur is not recounted in the text, we arbitrarily added "forest" (#P1) as a Place in the StoryWorld. The causal relations between Events in the Story, i.e., #E5–#E6, #E5–#E7, and #E7–#E8, were also interpretatively complemented. This kind of task corresponds to causal inference in natural language processing. We interpreted the first phrase “A Crow was sitting on a branch of a tree with a piece of cheese in her beak” not as a Narration of an Event, but as a Description of the crow. #De1 in the Discourse structure corresponds to this part. The Description was represented in a string format, because the framework for symbolic representation of Description (and Commentary) is not yet defined.

A Crow was sitting on a branch of a tree with a piece of cheese in her beak when a Fox observed her and set his wits to work to discover some way of getting the cheese. Coming and standing under the tree he looked up and said, “What a noble bird I see above me! Her beauty is without equal, the hue of her plumage exquisite. If only her voice is as sweet as her looks are fair, she ought without doubt to be Queen of the Birds.” The Crow was hugely flattered by this, and just to show the Fox that she could sing she gave a loud caw. Down came the cheese, of course, and the Fox, snatching it up, said, “You have a voice, madam, I see: what you want is wits.”

Figure 9. “The Fox and the Crow” (English translation by V. S. Vernon Jones)[2]

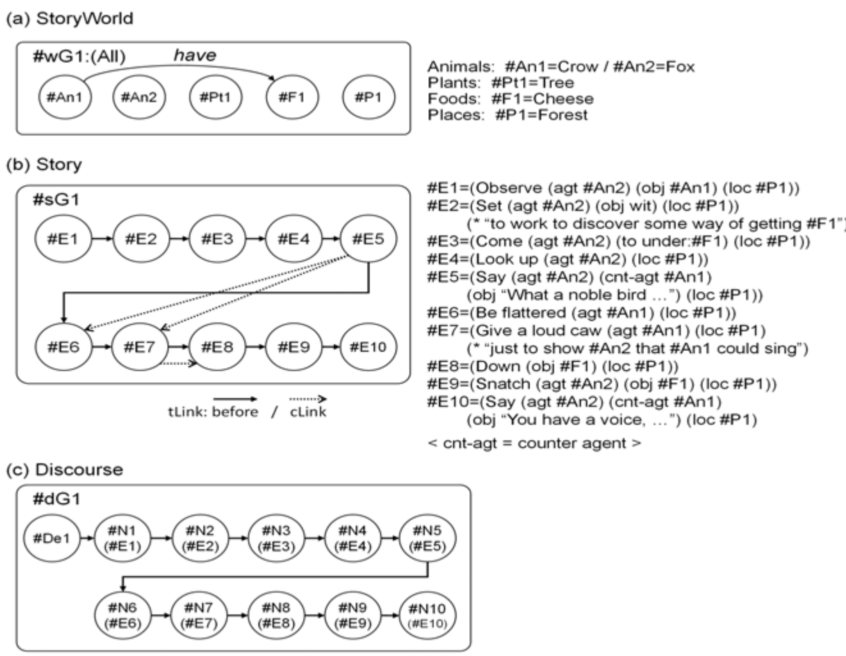


Figure 10. A structural representation of “The Fox and the Crow”

However, there are several issues with regard to the representational capacity. First, the StoryWorld model cannot represent the change in the world structure. For example, in this tale, the cheese is in the possession of the crow at the initial state. The StoryWorld in Figure 10 represents that structure. However, the cheese falls into the fox’s hand in the final state of this tale. This change is not represented in the StoryWorld. Second, the Story model cannot represent an objective or intention of a character’s action in phrases such as “set his wits *to work to discover some way of getting the cheese*” and *“just to show the Fox that she could sing she gave a loud caw.”* We will consider solutions for these issues in Section 6.

Next, we show a recomposed version of the structure mentioned above. Here, we focus only on the Story structure. Because “The Fox and the Crow” has a simple linear structure, the hierarchical and parallel structuring features are not utilized in the Story structure in Figure 10. These potential features are useful when composing more large and complex structures. Figure 11 shows a derivative Story by an extensional recomposition of the original Story. In this derivative Story, the original version is transformed into a revenge of the crow. The original Story is incorporated as an sGroup (#sG1) into the larger structure including three new sGroups (#sG2–4). Detailed Events in each sGroup are omitted in the figure.

[2] *Quoted from: http://www.gutenberg.org/ebooks/11339*

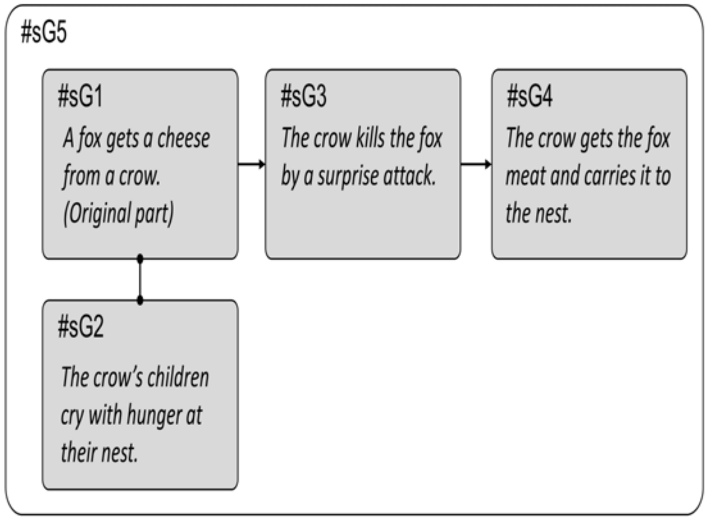


Figure 11. A derivative Story from "The Fox and the Crow"

5.3 Advantages

We discuss the characteristics and advantages of the proposed hierarchical graph model from several perspectives.

5.3.1 Integration of the Multiple Structural Dimensions of a Narrative

The multidimensional framework for StoryWorld, Story, and Discourse provides a basis for the systematic processing of the structure of a narrative, a complex object involving multiple aspects. Although most of the previous computational models for narrative structure can be positioned as models for either the story or the discourse aspect in a narrative, they have been designed as unidimensional structures without a clear division between the story and discourse. For example, as described in Section 2, the goal-based hierarchy in a planning-based story generator [19] and the graph structure of script-like chains of events [5, 18] focus on the story aspect. On the other hand, story grammar [34] and discourse structure models for natural language processing [11, 17] primarily focus on the discourse aspect.

5.3.2 Hierarchical Segmentation of Entities in a StoryWorld

The structural unit of wGroup in StoryWorld is useful for flexible reference to a StoryWorld object in a Story or Discourse. The inherent characteristic of language allows humans to refer to objects at various levels of segmentation according to the context or situation in which the language is used. For example, humans can use different levels of words to refer to the space where an event is held—such as "the Olympic Games are held at Tokyo" and "the judo games are held at Budokan (in Tokyo)"—and wGroup enables this type of flexible reference in the representation of a Story or Discourse.

5.3.3 Flexible Organization of Events in a Story

The hierarchical graph model can directly represent the parallel courses of events in a Story of the Olympic Games. Treebased models are unsuitable for the handling of this type of structure. The proposed model can also systematically divide and integrate the tasks in composing a Story structure based on its hierarchical structure, e.g., the composition of the overall structure, the games in each type of sport, and more detailed events focusing on a specific player. This type of systematic processing is difficult for general graph-based models.

5.3.4 Discourse Structuring with Grouping and Linkage

Tree-based modeling of a discourse structure such as RST [17] is also a simple and strong method to represent a linear discourse. However, tree-based models assume that the discourses are clearly and logically organized, and they seem unsuitable for complex or loosely organized discourses in many narrative texts. Hobbs [10] also indicated that incoherent discourses such as dialogues are not represented as normal tree structures. One difference between a hierarchical graph model and a tree-based discourse model is that the hierarchical graph model separates the structurization principles of relating (dLink) and grouping (dGroup), whereas a tree-based model forms a hierarchical structure by binding the discourse segments based on their relations. In our opinion, the hierarchy (grouping) and relation are different concepts in the structures of the narrative discourses and cannot be unified into the concept of relation, except in the case of clearly written texts in formal documents.

In addition, in the Discourse dimension, the proposed model can represent both types of linear narratives in many traditional narrative genres—e.g., novel, film, and manga—as well as branched narratives in the digital age. This is a unique feature of the proposed model. However, a branched narrative based on this model does not provide the readers with interactive changes in the dimension of the Story—as available in many interactive storytelling systems—but it provides the possibility of multiple readings for a fixed Story. In order to permit some interactive change in a Story, an extended mechanism must rewrite the Story in response to the actions of the reader through the reading process of the narrative.

5.3.5 Natural Language Expression based on Multidimensional Narrative Structure

The example texts in Figure 8 consist of only simple sentences based on the plain expression of each Event or sGroup. However, the proposed model will allow the representation or production of more sophisticated sentences by combining multidimensional information with the Event, i.e., attributes of the Entities, relations between Events (cLink and tLink), the attitude or modality of a narrator toward an Event (voice attributes involved in DU), and the connecting logic between Events (dLink).

We can explain this concept by using an example sentence: “Because a bear with red eyes bit a child, a hunter cruelly shot the bear.” In this sentence, the words “bear,” “child,” and “hunter” correspond to Entities in the StoryWorld and the phrase “with red eyes” corresponds to an attribute of “bear.” In the Story, this sentence contains the two Events, “a bear bit a child” and “a hunter shot the bear.” The causality between these Events—which is expressed by the word “because”—is represented by a cLink. In the Discourse dimension, this sentence recounts the two Events in chronological order and the information corresponding to “cruelly” is placed as a “voice” attribute of the Narration for the second Event.

6. Future Works

We consider several difficult challenges with regard to the future expansion of the representational capacity of the proposed model.

6.1 Nested Structure in Discourse

A narrative structure has various complexities. The proposed model integrates the three structural dimensions as a nested structure (Figure 3). However, we must provide a capacity for the representation of nested structures in each structural dimension. We consider two types of nested structures in the dimensions of Discourse and Story.

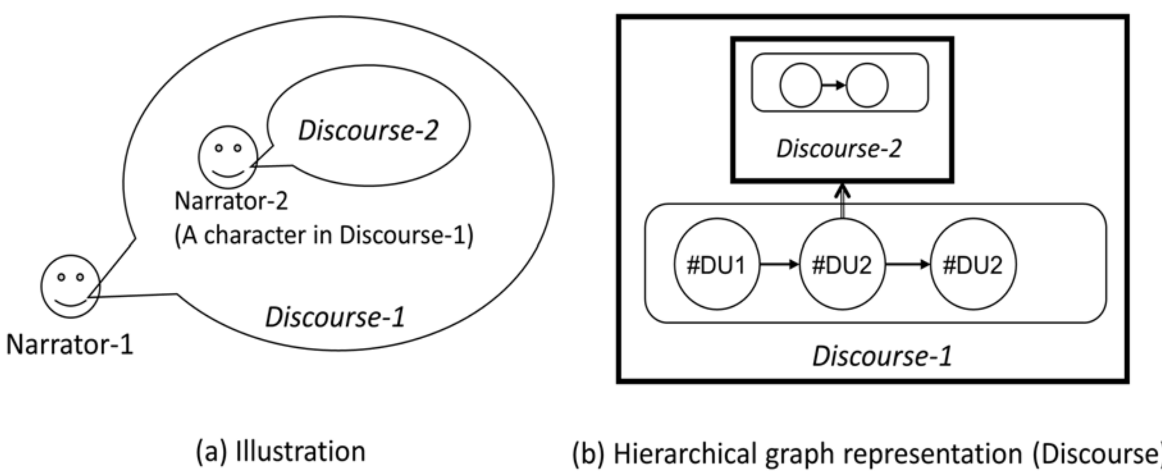


Figure 12. Narrative level: Nested structure in a Discourse

In narratology, Genette [7] conceptualized the notion of “narrative level” (or “diegetic level”) in a narrative discourse. It is a subcategory of “voice” (the situation of a narration by a narrator and narratee). The narrative level considers a type of nested structure explained as “a (diegetic) narrative narrated within a narrative.” This concept is illustrated in Figure 12(a). In the proposed model, this type of nested structure will be positioned as a nested structure in a Discourse. (Lönneker [14] also formalized the structure of narrative level for natural language generation systems.) In particular, as shown in Figure 12(b), a diegetic Discourse (Discourse-2) is connected from a node (#DU2) of the outer Discourse (Discourse-1). Here, the source node #DU2 corresponds to the Narration for an Event; the meaning is that “a diegetic narrator (Narrator-2) narrates a diegetic Discourse (Discourse-2).”

6.2 Nested Structure in Story

Next, we consider a nested structure in the Story dimension by using an example sentence: “Risa plans to rob a bank to get rich, but she gets arrested before executing that plan.” In this sentence, “rob a bank” and “get rich” as Risa’s plan correspond to unexecuted actions or events. Although these types of elements—i.e., goal, intention, desire, or plan—have a crucial role in narrative, the proposed model does not have the capacity to represent such elements. In considerations for future expansion, we assume that these types of imagined events are also represented as an independent Story that is nested within the main Story. When we assume that the main Story is composed from the narrator’s standpoint, a nested Story is associated with a character in the main Story. Figure 13(a) illustrates this idea. This type of nested structure in the Story dimension will be represented as nesting of multiple Stories (Figure 13(b)), in a manner similar to nested Discourses. For example, in the example sentence, “Risa gets arrested” is an Event in the main Story and the Events corresponding to “rob a bank” and “get rich” are placed inside a nested Story.

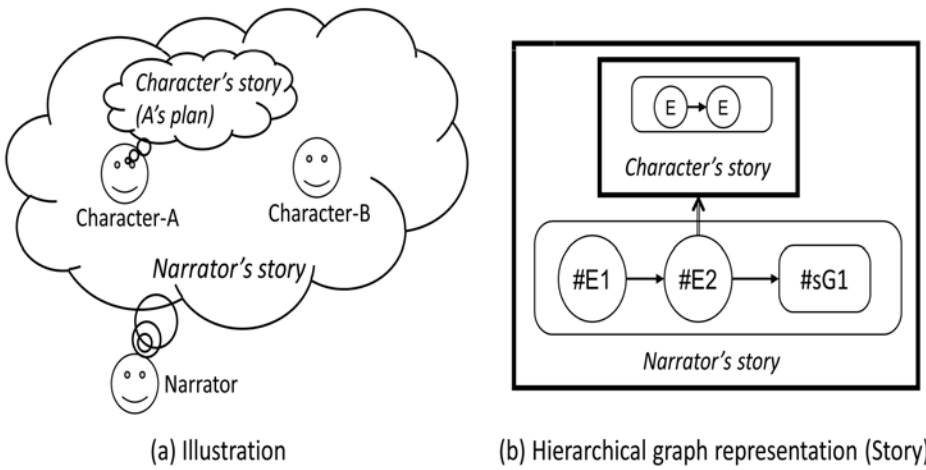


Figure 13. Character’s goal, intention, desire, or plan: Nested structure in a Story

6.3 Temporal Changes in StoryWorld

A world is a dynamic object that undergoes changes in state with temporal progress. For example, a human will be removed from the world when an event “the human died” or “the human is killed by someone” occurs, and an interpersonal relation will be added or removed by events such as “marry” and “divorce.” However, the proposed StoryWorld model provides only static data as a basic world structure (Figure 14(a)). We consider an expanded StoryWorld model that includes dynamic changes in its structure according to the development of Events in the Story dimension.

A straightforward approach to address this issue is to provide multiple StoryWorld instances as before and after states of each Event (or sGroup) in the Story. This approach requires a mechanism to simulate or infer changes in state (i.e., StoryWorld structure) caused by each Event. Here, StoryWorld(t_i) is derived from StoryWorld(t_{i-1}) by applying a change operation corresponding to the intermediate Event(t_{i-1}–t_i). However, this processing encounters complex and difficult issues that require considerable commonsense knowledge (in [21], a computational method for this type of inference is described). In addition, such a complete simulation of changes in the StoryWorld will not always be required in generative and interpretative processing of narratives.

Another realistic approach is to provide only A) a basic StoryWorld instance and B) additional StoryWorld instances as states for the key time points in the Story (Figure 14(b)). These key time points are determined depending on the objective of narrative processing. For example, when processing a folk tale Story, Events such as “die” and “marry” affecting the main characters will result in a crucial change in the StoryWorld structure, whereas Events such as “a farmer mows a yard” may not lead to significant changes. Our future plan is to design such an ad hoc StoryWorld model that has a capacity to represent the incomplete information of the temporal changes in the world structure.

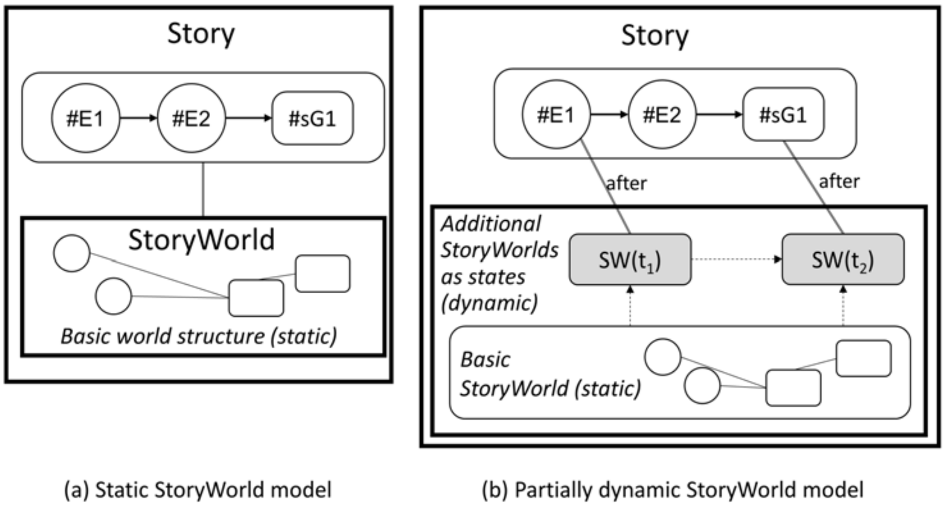


Figure 14. Static and partially dynamic models for StoryWorld representation

7. Conclusion

This paper presented the basic theory for a computational model of narrative structure representation. The proposed model has two characteristics. First, we formulated a multidimensional framework to systematically integrate multiple structural aspects of a narrative, i.e., the story world, story, and discourse. Second, we designed a hierarchical graph-based representation for each structural aspect. This hierarchical graph model combines the advantages of the previous two major models of narrative representation, i.e., tree- and graph-based models. In particular, the hierarchical graph model integrates the hierarchy-based systematic narrative operation and the flexible linkage among the narrative units. In addition, we implemented a prototype version of the proposed model based on object-oriented modeling and discussed its representational capacity through examples. Several future challenges in extending the representational capacity are also considered.

References

[1] Abe, J., Muramoto, T. (1985). A proposal of a notation system for representing the meaning structure of narrative story. *The Japanese Journal of Psychonomic Science*, 4 (2) 67-74. (in Japanese)

[2] Akimoto, T., Ogata, T. (2015). Experimental development of a focalization mechanism in an integrated narrative generation system. *Journal of Artificial Intelligence and Soft Computing Research*, 5 (3) 177-188.

[3] Allen, J. F. (1984). Towards a general theory of action and time. *Artificial Intelligence*, 23 (2) 123-154.

[4] Bringsjord, S., Ferrucci, D. A. (1999). *Artificial Intelligence and Literary Creativity: Inside the Mind of BRUTUS, a Storytelling Machine*, Lawrence Erlbaum.

[5] Chambers, N., Jurafsky, D. (2008). Unsupervised learning of narrative event chains. *In*: Proceedings of 46th Annual Meeting of the Association for Computational Linguistics, pp. 789-797.

[6] Chatman, S. (1978). *Story and Discourse: Narrative Structure in Fiction and Film*, Cornell University Press.

[7] Genette, G. (1980). *Narrative Discourse: An Essay in Method*, J. E. Lewin Trans., Cornell University Press. (Original work published 1972)

[8] Gervás, P., León, C. (2014). The need for multi-aspectual representation of narratives in modelling their creative process. *In*: Proceedings of the 5th Workshop on Computational Models of Narrative, pp. 61-76.

[9] Gervás, P., Lönneker-Rodman, B., Meister, J. C., Peinado, F. (2006). Narrative models: Narratology meets artificial intelligence.

In: Proceedings of Satellite Workshop: Toward Computational Models of Literary Analysis, 5th International Conference on Language Resources and Evaluation, pp. 44-51.

[10] Hobbs, J. R. (1985). The coherence of incoherent discourse. *Journal of Language and Social Psychology*, 4 (3-4) 213-232.

[11] Hobbs, J. R. (1990). *Literature and Cognition*, CSLI Lecture Notes Number 21.

[12] Lang, R. R. (1999). A declarative model for simple narratives. *In*: Narrative Intelligence: Papers from the 1999 AAAI Fall Symposium (Technical Report FS-99-01), pp. 134-141.

[13] León, C. (2016). An architecture of narrative memory. *Biologically Inspired Cognitive Architectures*, 16 19-33.

[14] Lönneker, B. (2005). Narratological knowledge for natural language generation. *In*: Proceedings of the 10th European Workshop on Natural Language Generation, pp. 91-100.

[15] Lönneker, B., Meister, J. C., Gervás, P., Peinado, F., Mateas, M. (2005). Story generators: Models and approaches for the generation of literary artefacts. *In*: Proceedings of the 17th Joint International Conference of the Association for Computers and the Humanities and the Association for Literary and Linguistic Computing, pp. 126-133.

[16] Mani, I. (2013). *Computational Modeling of Narrative*, Morgan & Claypool Publishers.

[17] Mann, W. C., Thompson, S. A. (1988). Rhetorical structure theory: Toward a functional theory of text organization. *Text*, 8 (3) 243-281.

[18] McIntyre, N., Lapata, M. (2009). Learning to tell tales: A data-driven approach to story generation. *In*: Proceedings of Joint Conference of the 47th ACL and the 4th IJCNLP, pp. 217-225.

[19] Meehan, J. R. (1980). *The Metanovel: Writing Stories by Computer*, Garland Publishing.

[20] Montfort, N. (2009). Curveship: An interactive fiction system for interactive narrating. *In*: Proceedings of the NAACL HLT Workshop on Computational Approaches to Linguistic Creativity, pp. 55-62.

[21] Mueller, E. T. (2006). *Commonsense Reasoning*, Morgan Kaufmann Publishers.

[22] Mueller, E. T. (2007). Understanding goal-based stories through model finding and planning. *In*: Intelligent Narrative Technologies, Papers from the 2007 AAAI Fall Symposium, pp. 95-101.

[23] Ogata, T. (2016). Computational and cognitive approaches to narratology from the perspective of narrative generation. *In*: T. Ogata and T. Akimoto Eds., *Computational and Cognitive Approaches to Narratology*, Chapter 1, IGI Global.

[24] Ogata, T., Hori, K., Ohsuga, S. (1996). A basic framework for narrative conceptual structure generation based on narrative techniques and strategies. *Journal of the Japanese Society for Artificial Intelligence*, 11 (1) 148-159. (in Japanese)

[25] Ogata, T., Terano, T. (1991). Explanation-based narrative generation using semiotic theory. *In*: Proceedings of Natural Language Processing Pacific Rim Symposium 91, pp. 321-328.

[26] Ogata, T., Umehara, S., Yamakage, S., Ueda, K., Hosaka, Y. (2004). Aspects of narrative discourse process and their integration by computer simulation. *In*: Cognition, Text and Society of Narrative: Notes of Intensive Course at University of Yamanashi 2002 (Ogata, T. Ed., Tech. Rep. JCSS-TR-52, pp. 136-143).

[27] Okada, N., Endo, T. (1992). Story generation based on dynamics of the mind. *Computational Intelligence*, 8 (1) 123-160.

[28] Pemberton, L. (1989). A modular approach to story generation. *In*: Proceedings of the Fourth Conference on European Chapter of the Association for Computational Linguistics, pp. 217-224.

[29] Prince, G. (2003). *A Dictionary of Narratology (Revised Ed.)*, University of Nebraska Press.

[30] Riedl, M. O., Bulitko, V. (2013). Interactive narrative: An intelligent systems approach. *AI Magazine*, 34 (1) 67-77.

[31] Riedl, M. O., Young, R. M. (2006). From linear story generation to branching story graphs. *IEEE Computer Graphics and Applications*, 26 (3) 23-31.

[32] Riedl, M. O., Young, R. M. (2010). Narrative planning: Balancing plot and character. *Journal of Artificial Intelligence Research*, 39 217-267.

[33] Ronfard, R., Szilas, N. (2014). Where story and media meet: computer generation of narrative discourse. *In*: Proceedings of the 5th Workshop on Computational Models of Narrative, pp. 164-176.

[34] Rumelhart, D. E. (1975). Notes on a schema for stories. *In*: D. G. Bobrow & A. Collins Eds., *Representation and Understanding: Studies in Cognitive Science*, Academic Press.

[35] Schank, R. C., Abelson, R. P. (1977). *Scripts, Plans, Goals, and Understanding: An Inquiry into Human Knowledge Structures*, Lawrence Erlbaum.

[36] Schank, R. C., the Yale A.I. Project (1975). SAM: A story understander, Research Report #43, Yale University, Department of Computer Science.

[37] Turner, S. R. (1994). *The Creative Process: A Computer Model of Storytelling and Creativity*, Lawrence Erlbaum.