Implementing Story Units of Japanese Folktales with Conceptual Dictionaries

Takuya Ito

Graduate School of Software and Information Science, Iwate Prefectural University, Sugo 152-52, Takizawa, Iwate, Japan

Jumpei Ono

Faculty of Software and Information Technology, Aomori University, Seishincho 2-10-1, Edogawaku, Tokyo, Japan

Takashi Ogata

Faculty of Software and Information Science, Iwate Prefectural University, Sugo 152-52, Takizawa, Iwate, Japan
E-mail: g231r003@s.iwate-pu.ac.jp, j.ono@aomori-u.ac.jp, t-ogata@iwate-pu.ac.jp

Abstract

This paper reports the development of story units that are knowledge representing story structure for story generation. A story structure in our narrative generation study is formed by a tree that consists of events connected based on the relationships, such as the causal and temporal relations of events. We have been implemented story units based on the "Type of Folktale" in the *Nihon Mukashi-banashi Taisei* (*The Concluded Compilation of Japanese Folktales*). The current main theme is to use our conceptual dictionaries, including verb and noun conceptual dictionaries, to combine the current version of story units to our "integrated narrative generation system," which is a general framework of our narrative generation study. In this paper, we explain the above mechanism.

Keywords: List four to six keywords which characterize the article.

1. Introduction

In a study of folk tales and folktales from around the world, Propp collected a genre of Russian folk tales called magical folktales, and analyzed the tales structurally in terms of commonalities among the collected stories[1]. Propp called the structural commonalities "functions," which are related to the classification of characters, and in this respect are like the main action of the main characters in the definition of motifs.

This research is underpinned by narratology. We analyze existing stories and propose story units for generating a story using the results of the analysis in the Integrated Narrative Generation System (INGS), which is being developed by the authors[2,3].

Story units are in the prototype stage, and the verbs representing events in the Story Unit were created according to the descriptions of the folktales to be analyzed, and do not correspond to the descriptions in the verb concept dictionary that INGS refers to when generating stories.

In this paper, we aim to map the verbs in the story unit to the verb concept lexicon of INGS in order to generate a story using the story unit. In this paper, we aim to map the verbs in story unit to the verb concept dictionary of INGS. Specifically, we examine the descriptions of the verbs in story unit, and if the verbs in story unit can be mapped to the existing verb concepts in the verb concept dictionary, we register the verbs in story unit as the verb concepts of INGS. This correspondence enables us to use the rich sentence generation mechanism

and about 110,000 noun concepts in INGS, and to generate a variety of stories from a story unit.

2. Motif and Form in the study of folktales

A motif is a unit that contains one major action of a major character in the composition of a story, as well as actions that directly correspond to that action [3].

In addition to the research on motifs mentioned in section 1, Aarne and Thompson collected folk tales mainly from various parts of Europe and produced the International Folk Tale Abstracts [4,5] (compiled by Aarne and expanded and revised by Thompson, so it is called the Aarne-Thompson type index (AT index)).

To analyze Aarne-Thompson's International Folktale Abstracts, it was noted that there were differences in the characters and actions of the different categories of folktales in the International Folktale Abstracts.

There are several efforts by Seki. The folktales used in this study are the efforts of Keigo Seki. The details are described below.

3. Seki's Study of Folktales

In *Nihon Mukashi-banashi Shusei* (*Collection of Japanese Folktales*), Seki classified about 8,700 folktales collected in various parts of Japan. Later, Seki and his colleagues expanded the number of tales to about 35,000 and updated the classification in *Nihon Mukashi-bahashi Taisei* (*The Complete Japanese Folktales*)[7-18]. Seki summarized the contents of 827 "type of folktale," which were the result of the classification. The 827 types of folktales based on the AT index[5,6]. The types of folktale is a classification of Japanese folktales as defined by Seki and a summary of their contents.

[Seki, 78a-e, 79a-d, 80a] has collected and classified about 35,000 folktales from various parts of Japan in *Nihon Mukashi-banashi Taisei*. There are 12 volumes of *Nihon mukashi-banashi Taisei* in total, and volumes 1 to 10 [7-16] contain the collected folktale themselves, volume 11 [17] contains materials on the type of folktale, an index, and a bibliography, and volume 12 contains contributed articles on folktales [18].

Each type has a summary of the contents of one or more folktales corresponding to the type. These types of folktales are systematized, with the largest categories being "animal folktales," "authentic folktales," and "funny stories," which are based on the AT index. In addition, these three classifications have "divisions" that group folktales according to their motifs, with 11 divisions for "animal folktales," 16 for "authentic folktales," and 12 for "funny stories" (a total of 39 divisions). Under each category, 827 types of folktales are assigned. The three major categories are further divided into 39 divisions. 827 types of folktale belong to one of the divisions, and each division contains a minimum of 2 and a maximum of 66 types of folktale.
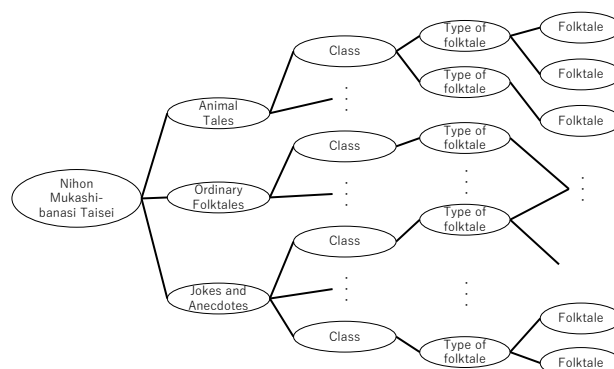

Fig. 1. The hierarchy of folktales in the "Nihon Mukashi-banashi Taisei"

Figure 2 shows an example of a type of folktale. type of folktale has four items.

- Serial number: The serial number for type of folktale (e.g., 一四二 (One hundred forty-two))
- Title: Motifs representing the characteristics of folktales (e.g., 蛇息子 (The snake son))
- AT Classification Number: These are the numbers corresponding to the motifs in the AT index. Some types of folktales do not have corresponding numbers. (e.g., AT 四三三 B)
- Content: The structure of a type of folktale. It consists of one or more sentences. The sentences are numbered roughly by scene. It may also contain branching patterns of events and patterns of elements. For example, the structure is as shown in Figure 3.

一四二 A 蛇息子 (AT 四三三 B)
1、子どものない夫婦が(a)神に祈願して蛇の子を生む。(b)卵を拾い、それから蛇の子が生まれる。2、あまりに大きくなりすぎて村人に恐れられるので山に捨てる。3、この蛇の援助によって(a)雨を降らせる。(b)田に雨がかかるようにする。(c)蛇が禍をなして爺婆に仇を討たれる、4、そのために爺婆は富を得る。

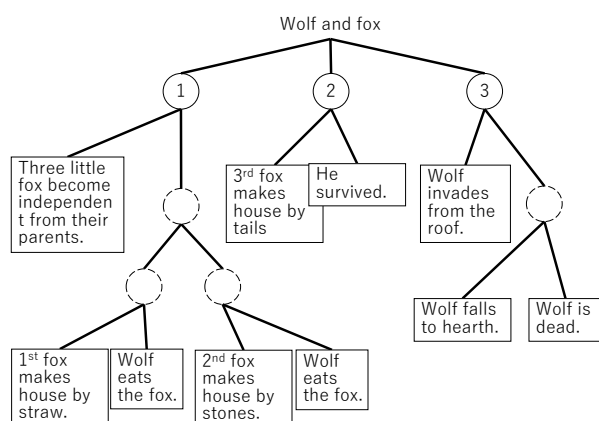Fig. 2. An example of a type of a folktale

Wolf and fox

Fig. 3. An example of structure of a type of a folktale

The type of folktale is a structure that integrates the actual collected folktales. Figure 4 shows an example of collected folktales.

「昔、三匹の子狐がおった。一番兄狐が、もう一匹前になったからといって、親狐に藁から家をつくってもらった。ところが、ある日、・・・」[Once upon a time, There are three little foxes. An eldest brother fox is presented a house that makes by straw from their parent, because the eldest brother fox becomes an adult. But, one day, …]

Fig. 4. An example of motif variation

Each type of folktale has a serial number called a taisei number and a title that represents the most characteristic motif of the content (the bold Chinese numerals in the figure are taisei numbers).

4. First Step for Making Story Units

4.1 Story Units

A story unit is the knowledge that makes up a story. There are many kinds of story units. For example, there is a story unit that shows the structure of a large story. On the other hand, there is a story unit that shows the local structure of a story. We used "Type of folk tale" to create

story units [19]. The created story units can be used to create folktale-like story. The story units can be used to create folktale-like story, and by combining the story units, it is possible to create a story that is different from the original story.

Section 4.2 and 4.3 show the procedure for creating story units.

4.2 The Structure of Type of Folktales

We first reconstructed the folktale type into a form suitable for the program. We split the content of the type of folktale so that an item contains a verb. Table 1 shows a form of splitting.

- Serial number: Referring Figure.2 in the section3.
- Title: Referring Figure.2 in the section3.
- Content number: The specific content of the folktale type is summarized by scene, with the events contained therein.
- Sentence number: Indicates a sentence break in type of folktale.
- Event: Events segmented by verb units

4.3 Making Story Units

To make a story unit, we create a case structure for each of the "event" descriptions in Table 1.

This story unit consists of three scenes, each of which is represented by A, B, and C. The events in each scene are numbered, and there are six events in this story unit. Each event is composed of a verb and its case information based on the way events are described in INGS. The descriptions of verbs and case elements in the Story Unit are based on the descriptions in type of folktale, and do not correspond to the verb and noun concepts in INGS. However, in the case of verbs, a temporary number 1 is

Table 1. The example of a structure form of type of folktale

大成番号	モチーフ（昔話）		内容番号			文番号	事象（内容）
474	鬼の面					①	ある女(男)が山(化け物屋敷)で鬼の面をかぶっている。
						②	化け物がそれを見て逃げる。
						③	女は
				(a)		①	宝物をとって
							帰る、
							&または&
				(b)		①	化け物屋敷の主人になる。

written after the verb in order to correspond to the verb concept.

The branching structure of the Type of folktale is represented by a description starting with "or", as shown in Figures 5.4a and 5.4b. In addition, there are branching structures that affect the choice of subsequent branching structures. In order to preserve the pairing relationship, there are examples where several different story units were created from the same type of folktale. Therefore, there are 912 story units that we have created.

The Story Unit is created based on the description of the type of Folk. In the case structure of verbs involving a dialogue such as "say," the content of the dialogue is inserted directly into the object case. In this case, the elements of the object case cannot be changed or expanded without changing the content of the description. Therefore, we consider a more flexible use of the story unit by creating a nested structure of events.

Among the case elements of the story units, there are 761 cases that can be nested. Table 2 shows the classification of nested structures.

Table 2. The type of nested structure in the story units

Category	Amount	Definition
An action	217	Contents consisting of a single action
Actions	50	Contents consisting of multiple actions
Avoid	1	Contents that attempt to avoid a certain action.
Prohibition	15	Contents that prohibit certain actions
Propose	10	Contents of a proposal for an action
Desire	33	Contents that want to an action
Speak	4	Contents that characters speak
Denial	29	Contents that negate an action.
Impossibility	3	The impossibility of an action.
Command	68	To command an action.
Rewording	3	The action is difficult to understand, so it is replaced by a different verb expression.
合計	433	

5. Combining Story Units and a Verb Conceptual Dictionary

We combine the story unit shown in section 4 with a verb concept dictionary.

5.1 Goal of Combining Story Units and a Verb Conceptual Dictionary

The combination of verb concept dictionaries provides two advantages to story units. (1). Through the sentence pattern of the verb concept, single sentence generation and sentence generation with conjunction in INGS will be available, and the following functions will be available--"Ordering words," "Transforming inflected form," "Transforming notation," "Multi-sentence." (2). Story units can use the system structure of the verb concept dictionary. For example, the story units can manipulate events using the systematic structure of the verb concept dictionary, and to use the co-occurrence information of the verb concepts associated with the verb concept dictionary.

5.2 Verb Conceptual Dictionary

A concept here is data that shows the meaning and classification of a word. A concept dictionary is a kind of ontology of words belonging to a certain part of speech. In this section, two types of concept dictionaries are explained. One is the verb concept dictionary, and the other is the noun concept dictionary. The verb concept dictionary is a dictionary that systematically registers verb concepts. The verb concept dictionary has about 10,000 verb concepts. A noun concept dictionary is a systematic dictionary of noun concepts. The Noun Concept Dictionary has about 110,000 noun concepts.

Figure 6 shows an example of a verb concept. A verb concept has a sentence pattern, a case frame, and a constraint. Constraint describes the constraints for


```
(motif0669 (鬼の面 (被る1 見る1 逃げる1 (or (取る1 帰る1) なる1)))
    (
    (継起
        (A
        (1 (event 被る1 (agent (&sc 女)) (object (&sc 鬼の面)) (bcation (&sc 山))))
        (2 (event 見る1 (agent (&sc 化け物)) (counter-agent (&sc 女))))
        (3 (event 逃げる1 (agent (&sc 化け物))))
        (or
            (4a (event 取る1 (agent (&sc 女)) (object (&sc 宝物)))
                (event 帰る1 (agent (&sc 女))))
            (4b (event なる1 (agent (&sc 女)) (to (&sc 化け物屋敷の主人))))
        ))))
```

Fig. 5. An example of a story unit

selecting a case frame from the noun concept dictionary, and is-a information for systematically classifying verbs.

5.3 The Method for Combining Story Units and a Verb Conceptual Dictionary

This paper describes the binding of verbs contained in story units to a verb concept dictionary. We organized the types of verbs contained in story units and implemented methods according to the classification.

5.4.1 Counting kinds of verbs in the story units

The 912 story units we created have 1107 different verbs. These verbs can be divided into those that can be easily combined with the verb concept dictionary and those that require some work to be done in order to be combined with the verb concept dictionary. First, we classified the verbs in the story unit in terms of the challenges in binding. Table 3 shows the breakdown of the classification, the number of units corresponding to the classification, and an example.

Table 3. The category of task in the story units

Task	Amount	Example
(A). Direct connect	615	食べる
(B). New registration	363	あさる
(C). Notation	121	あきらめる
(D). Passive form	13	追われる
(E). Negative form	5	承知しない
Total	1117	

5.4.2 (A). Direct connect

(A) is a verb that can be easily combined with the verb concept dictionary. For verbs that have the same notation, once the description of the Story Unit is used as is, the sentence generation mechanism of INGS is used to expand the sentence into a natural sentence, and if a natural sentence that matches the content of the Story Unit is expanded, it is used as is. In case of no case, we selected an appropriate verb concept number from the verb concept dictionary and changed the number of the Story Unit verb to that number.

5.4.3 (B). New registration

For the verbs in (B), we first checked whether there were any verb concepts with the same meaning but different notation in the verb concept dictionary, and if so, we processed them as the task in (C). If there was no verb with different notation, we registered it as a new verb concept in the verb concept dictionary.

For each new verb, we created a sentence-pattern and a case-frame based on the case structure of the target verb in the story unit. case-frame refers to the case structure of the verb in the story unit, and case-frame refers to the case of the target verb in the story unit. The sentence-pattern was created by specifying the particle corresponding to each case used in the case-frame, and combining the "case number" and the "particle corresponding to the case". In the tentative data, we did not give any constraints on the cases to be used, and all is-a to classify verb types were set to "physical action". In this example, the case registered in the original story unit is the agent case, which represents the subject, and the object case, which represents the object.

5.4.4 (C). Notation

This problem stems from the linguistic nature of Japanese. The Japanese language can express homophonic words with different notations. For example, "食べる (taberu)" which means "eat," can be expressed as "たべる," "タベル," etc. In story units, "気づく (notice)" can be expressed as "notice," "notice," "notice," etc.

In the story unit, there is a verb "気づく (kiduku)" which means "notice." However, the verb concept

```
(set ’通行する 1
    ’((name 通行する 1)
      (sentence-pattern ″N1 が N2 を 通行する″)
      (case-cons-set
        ((case-frame ((agent N1) (location N2)))
         (constraint ((人 -死人 -人間〈人称〉 -準人間)
                      (場所 -世界 -景 -宇宙 建造物 -屋根 -柱・梁〔はしら・はり〕 -壁 -家屋{部分〈要素{その他}〉}[アン
テナ] -家屋{部分〈要素{その他}〉}[避雷針] -建具 -塀)
             ))))
      (is-a (v 物理的移動))))
```

Fig. 6. An example of a verb concept

dictionary registers it as "気づく." Therefore, we changed the notation of the verb in the story unit side to that of the verb concept dictionary.

5.5.5 (D). Passive form and (E). Negative form

(D) and (E) are issues that are addressed by revising the structure of the story unit. (D) is dealt with by using the original verb concept of the verb and placing the person who is the target of the action in the counter-agent case. (E) is expressed by adding a not structure to the case structure of the verb.

6. Conclusion

In this paper, we combined story units and the verb conceptual dictionary. Therefore, the story units can use sentence generation in the INGS. And the system structure of the verb concept dictionary.

In the future works, we have two tasks. At first, we have to combine story units with the noun conceptual dictionary. By combining with the noun conceptual dictionary, story units can manipulate the case of story units based on the systematic structure of noun concepts. At second, we revise the new verb concepts. The newly registered verb concepts have been created according to the structure of the folktale. In particular, we have to revise the constraints of the verbal concepts in order to make them more generic.

References

1. Propp
2. T. Ogata, *Toward an Integrated Approach to Narrative Generation: Emerging Research and Opportunities*, IGI Global, PA, USA, 2020.
3. T. Ogata, *Internal and External Narrative Generation Based on Post-Narratology: Emerging Research and Opportunities*, IGI Global, PA, USA, 2020.
4. [小澤 97] 小澤俊夫：昔話入門，ぎょうせい，1997.
5. [Aarne 61]Aarne, A.: The types of the folktale: a classification and bibliography, "FF communications," 184, Helsinki: Academia Scientarum Fennica, 1961.
6. [Aarne 69]Aarne, A.: 関啓吾訳. 『昔話の比較研究』. 東京: 岩崎美術社, 1969. (Antti A.: Leitfaden der vergleichenden. Marchenforshung, 1913.)
7. [Seki 1979] Seki, K., Nomura, J., & Ōshima, H. (Eds.): Nihon mukashibanashi taisei: vol. 12 kenkyū hen [Complete collection of Japanese folktales: Vol. 12, researches]. Tokyo, Japan: Kadokawa Shoten, 1979.
8. [Seki 1980] Seki, K., Nomura, J., & Ōshima, H (Eds.): Nihon mukashibanashi taisei: Vol. 11 shiryō hen [Complete collection of Japanese folktales: Vol. 11, data]. Tokyo, Japan: Kadokawa Shoten. 1980.
9. [Seki 1979a] Seki, K.: Nihon mukashibanashi taisei vol.1: dobutsu monogatari [The Concluded Compilation of Folktales]. Tokyo, Japan: Kadokawa Shoten, 1979.
10. [Seki 1978a] Seki, K.: Nihon mukashibanashi taisei vol.2: honkaku mukashibanashi (1) [The Concluded Compilation of Folktales]. Tokyo, Japan: Kadokawa Shoten, 1978.
11. [Seki 1978b] Seki, K.: Nihon mukashibanashi taisei vol.3: honkaku mukashibanashi (2) [The Concluded Compilation of Folktales]. Tokyo, Japan: Kadokawa Shoten, 1978.
12. [Seki 1978c] Seki, K.: Nihon mukashibanashi taisei vol.4: honkaku mukashibanashi (3) [The Concluded Compilation of Folktales]. Tokyo, Japan: Kadokawa Shoten, 1978.
13. [Seki 1978d] Seki, K.: Nihon mukashibanashi taisei vol.5: honkaku mukashibanashi (4) [The Concluded Compilation of Folktales]. Tokyo, Japan: Kadokawa Shoten, 1978.
14. [Seki 1978e] Seki, K.: Nihon mukashibanashi taisei vol.6: honkaku mukashibanashi (5) [The Concluded Compilation of Folktales]. Tokyo, Japan: Kadokawa Shoten, 1978.
15. [Seki 1979b] Seki, K.: Nihon mukashibanashi taisei vol.7: honkaku mukashibanashi (6) [The Concluded Compilation of Folktales]. Tokyo, Japan: Kadokawa Shoten, 1979.
16. [Seki 1979c] Seki, K.: Nihon mukashibanashi taisei vol.8: warai banashi (1) [The Concluded Compilation of Folktales]. Tokyo, Japan: Kadokawa Shoten, 1979.
17. [Seki 1979d] Seki, K.: Nihon mukashibanashi taisei vol.9: warai banashi (2) [The Concluded Compilation of Folktales]. Tokyo, Japan: Kadokawa Shoten, 1979.
18. [Seki 1980] Seki, K.: Nihon mukashibanashi taisei vol.10: warai banashi (3) [The Concluded Compilation of Folktales]. Tokyo, Japan: Kadokawa Shoten, 1980.
19. [Ono, 2019] Ono, J., A., Ito, T., & Ogata, T.: Mukahi banashi no motif no program ka to motif kozo no hikaku. In Proceedings of the 61th Special Interest Group on Language Sense Processing Engineering, pp. 51-62, 2019.