Journal of Robotics, Networking and Artificial Life
Vol. 5(3); December (2018), pp. 212–216
DOI: 10.2991/jrnal.2018.5.3.15; ISSN 2405-9021 print; 2352-6386 online
https://www.atlantis-press.com/journals/jrnal

Research Article

Acquiring Short Scripts and Setting a Case Frame in Each Acquired Script: Toward Random Story Generation

Jumpei Ono[1,*], Takashi Ogata[2]

[1]*Department of IT Solution, Vocational School of Digital Arts Sendai, Sendai, Miyagi 980-0003, Japan*
[2]*Department of IT Solution, Faculty of Software and Information Science, Iwate Prefectural University, Takizawa, Iwate 020-0693, Japan*

ARTICLE INFO

Article History

Received 3 June 2018
Accepted 30 September 2018

Keywords

Case frame
integrated narrative generation system
knowledge acquisition
random story generation
script
verb concept

ABSTRACT

The integrated narrative generation system (INGS), that the authors have developed, generates a story to translate the story into the surface representation. In the story generation process, the integrated narrative generation system uses narrative knowledge that was automatically acquired from existing narrative works. This paper presents a method to acquire short scripts, which are a kind of narrative knowledge, from existing works in Aozora Bunko for the story generation. This paper presents a mechanism to generate random story-like event sequences by using 23,751,142 bi-gram scripts acquired based on the method proposed below. The authors aim to use the scripts generated by the method as a first set to be revised through the next learning process.

1. INTRODUCTION

The integrated narrative generation system (INGS) [1] is a system that generates a narrative automatically. The system not only generates a story but translates the story into the surface representation. In story generation, the INGS uses knowledge about story structure and dictionaries [2] that systematically store nouns, verbs, and so on. The authors call story content knowledge, which is stored knowledge bases of the INGS. The dictionaries are called conceptual dictionaries. A variety of generated stories depend on the scale of the knowledge base. If the knowledge base is small scale, story generation by the INGS fails frequently. However, expanding the knowledge base manually has limitations. Therefore, the authors have attempted to acquire knowledge automatically [3]. This paper details the attempt to acquire short scripts based on bigrams acquired from existing narrative works for story generation.

A script that Schank [4] has mentioned involves knowledge about human actions in a particular situation. For example, a restaurant script describes the procedure procedures for ordering and eating dinner at a restaurant. We attempt to not only acquire Schank's scripts but to also acquire the temporal sequences of events.

The INGS consists of one-part mechanisms and one-part knowledge. Knowledge refers to dictionaries for concepts and language notation [1], and to knowledge bases that build partial structures of a story (e.g., the narrative content knowledge base). By using this knowledge, the mechanisms generate various aspects in a narrative, such as a story, discourse, and surface representation (i.e., sentences, music, or images).

* *Corresponding author. Email: j.ono@sugawara.ac.jp*

2. METHOD FOR STORY GENERATION IN THE INGS

Figure 1 shows a story structure that is generated in the INGS. The structure has a hierarchical structure that consists of states, events, and relations. A state describes static information regarding characters, things, locations, and times in a story. An event describes dynamic information that represents the difference between two states. A relation links several events by some relationship such as Causal relationship.

Story generation in the INGS is the expansion and transformation of a story structure using story techniques. A story technique is a formal procedure for expanding the structure of a story and draws on a knowledge base that corresponds to the particular technique. Figure 2 shows an example of using a story technique; the technique selects applicable knowledge and expands the structure of the story by using that knowledge.

Figure 3 shows an example of automatically-acquired knowledge. This knowledge is applied to an event that includes "eat," and the story technique adds an event that includes "drink" to the structure of the story. (&v age1) or (&v obj1) are variables described by the cases. These variables denote the position or location where the generated character, thing, or location is inserted based on the restrictions. Cases that have the same variables also have the same character, thing, or location.

Potential selection points in story generation involve mainly the "application point of the story technique," "applying a story technique," and "using knowledge." The "application point of story technique" and "applying a story technique" steps can be controlled

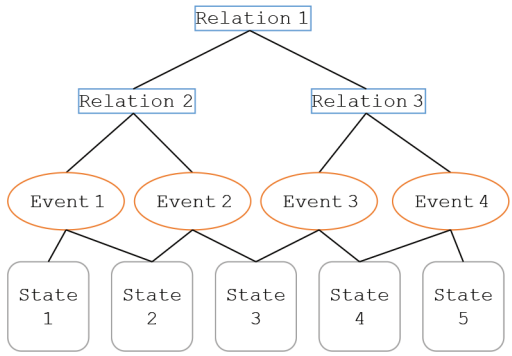

Figure 1 | An example of a story structure

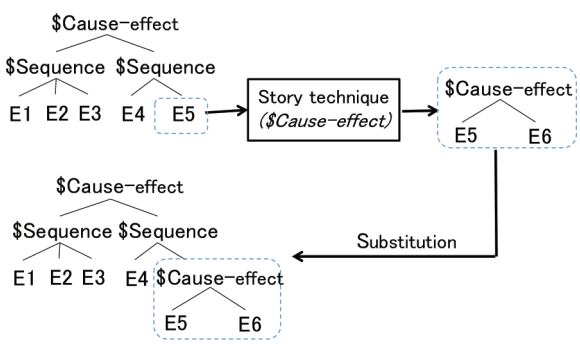

Figure 2 | A method for expanding a story structure (Figure 8 in Ogata [1])

```
(script (食べる 2[eat] (食べる 2[eat] 飲む
2[drink]))
 ((event 食べる 2[eat] (agent (&v age1))
   (counter-agent (&v age2)))
  ((1 (event 食べる 2[eat] (agent (&v
   age1))(counter-agent (&v age2))))
  (2 (event 飲む 2[drink] (agent (&v
   age1)) (object (&v obj1))))) ))
```

Figure 3 | Example of a short script

based on a parameter that is inputted into the INGS. For example, if "parameter length" is increased, a story technique is selected that increases the number of events within the story.

3. ACQUIRING SHORT SCRIPTS FROM EXISTING WORKS

The mechanism for acquiring short scripts has three steps: acquiring bigrams, making short scripts, and setting a case frame. This section explains each of the three steps.

3.1. Acquiring Bigrams from Existing Text

Bigrams that consist of two verbs are acquired by applying morphological analysis to existing text. The acquired bigram is called a verb bigram. For example, verb bigrams are acquired as shown in Figure 4. The procedure for acquiring bigrams is as follows:

(i) Morphological analysis: The mechanism analyzes a text using a morphological analyzer.

(ii) Extracting verbs from results: The mechanism extracts verbs from the results of morphological analysis.

(iii) Making bigrams: The mechanism creates pairs of verbs in the order of the results of extraction. For example, if the result of extraction is $(A \times B \times C)$, then the mechanism combines two pairs of verbs $(A \times B)$ and $(B \times C)$. $N - 1$ verb bigrams are created from N verbs.

(iv) Removing a part of the bigram: The mechanism removes bigrams that include verbs that are not stored in the conceptual dictionary, because the INGS cannot use such bigrams in story generation.

3.2. Making Short Scripts

The acquired verb bigrams are made into short scripts. This step connects the verb concepts to verbs that are included in verb bigrams. Verb concepts are assigned a number in the INGS because the INGS provides the plural of the verb. For example, the verb concept "食べる1" has a meaning like "earn," while the verb concept "食べる2" has a meaning like "eat".

Short scripts $(X_n \times Y_k)$ were created from verb bigram $(X \times Y)$ [$(X_n \times Y_k)$ consists of verb concepts X_n and Y_k; $(X \times Y)$ consists of verb X and verb Y]. In this case, verb X_n has i kinds of meaning, verb Y_k has j kinds of meaning, and the number of short scripts is equal to i multiplied by j. In making short scripts, we need to consider the meaning of a verb. However, in this paper, the authors created all patterns of $(X_n \times Y_k)$ with the procedure shown in Figure 5.

3.3. Setting a Case Frame

By setting a case frame in a short script, the INGS can generate a story using the short script. Figure 6 shows the procedure for

<Target text>

男がウィスキーを運ぶと、女はパスタを食べていた。彼はウィスキーを机に置き、椅子に座った。

[When a man carried whisky, a woman was eating pasta. He puts the whisky on the table and sits down on the chair.]

<Acquired bi-grams>

(運ぶ 食べる) (食べる 置く) (置く 座る)

[(carry eat) (eat put) (put sit)]

Figure 4 | Example of acquired bigrams

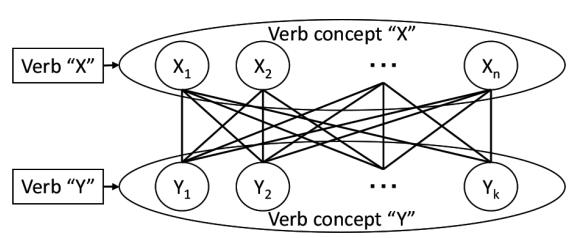

Figure 5 | An image of making short scripts by a bigram

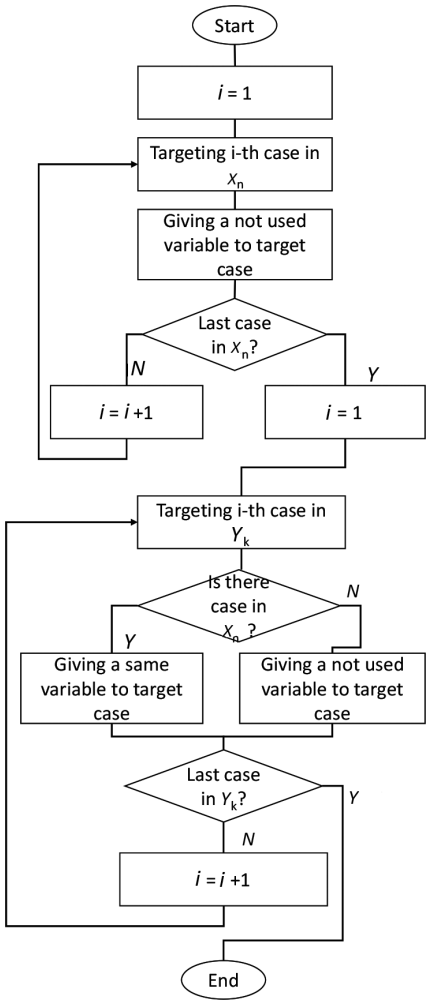

Figure 6 | A flowchart for setting a case frame

setting the case frame in short scripts ($X_n \times Y_k$). A case frame consists of cases that require a verb. For example, a verb concept has a description as seen in Figure 7. Here, the elements of the "case-frame" are the case frame. A case frame has restrictions based on the noun conceptual dictionary. The restrictions define the probability of cases that are included in a verb concept.

4. RESULT

In this paper, the authors acquired short scripts from 13,331 works in the Aozora Bunko [5]. The acquisition mechanism used MeCab in the morphological analysis. We acquired short scripts from each work. In analyzing each work, we did not consider structures such as paragraphs in each work. Table 1 shows the number of acquired verb bigrams and short scripts (MeCab was downloaded from "http://taku910.github.io/mecab/").

5. USING SHORT SCRIPTS FOR STORY GENERATION

The INGS uses a technique that allows the generated story to converge on one result based on several parameters that are set beforehand. Specifically, the INGS performs story generation by using selection rules at several points of the generation process. The authors have adopted random generation in order to gradually change a narrative. A summary regarding random narrative generation is presented in Ogata et al. [6]. In this section, a generated story is presented.

Currently, parameters have not been applied by the authors in order to evaluate whether acquired knowledge alone suffices for random story generation. This method accomplishes generation by applying various story techniques that INGS has available, and uses scripts that were acquired automatically which are stored

```
((name 食べる2[eat])
 (sentence-pattern "N1 が N2 を 食べる"["N1 eat N2"]))
 (case-cons-set
  ((case-frame
   ((agent N1) (counter-agent N2) (location N3) (object nil) (instrument nil) (from
       nil) (to nil) (adverb nil) (possessive nil) (situation nil) (purpose nil)
       (experiencer nil) (source nil) (idiom nil) (information nil) (as nil)))
   (constraint
    ((人[human] -死人[corpse] -人間〈人称〉[human<person>]-準人間[semi-human])
     (鯨[whale] 馬[horse] 牛[cattle] 豚[hog] 山羊[goat] 羊[lamb] 鹿[deer] 猪[wild boar]
            兎[rabbit] 家禽[poultry] 猟鳥[game bird] 魚[fish] たこ・いか・えび・かに
            [octopus/squid/shrimp/crab])
     (場所[locaiton] -交通路[traffic route] -公共施設{その他}[public building{etcetera}]
             -地域[local] -崖[cliff] 山{部分}[mountain{parts}] 平地[flatland] 岸
             [bank]))))
  ((case-frame
   ((agent N1) (counter-agent nil) (location N3) (object N2) (instrument nil) (from
       nil) (to nil) (adverb nil) (possessive nil) (situation nil) (purpose nil)
       (experiencer nil) (source nil) (idiom nil) (information nil) (as nil)))
   (constraint
    ((人[human] -死人[corpse] -人間〈人称〉[human<person>]-準人間[semi-human])
     (食料[food] -調味料[flavoring] -飲物・たばこ[drink/cigarette] -汁[soup])
     (場所[locaiton] -交通路[traffic route] -公共施設{その他}[public building{etcetera}]
            -地域[local] -崖[cliff] 山{部分}[mountain{parts}] 平地[flatland] 岸
            [bank])))))
  (is-a (v 身体動作[physical action]))))
```

Figure 7 | Example of a verb concept

Table 1 | Amount of short scripts

Type	Amount
Verb bigram	5,52,773
Short script	23,751,142

若主人は溝で酩酊しました。若主人は溝で粛清の夢を見ました。特急が溝で滑りました。若主人は溝でわれに帰りました。引力がトレーシングペーパーより起こりました。若主人は特急をトレーシングペーパーと反対にしました。若主人はトレーシングペーパーで地政学へ駄目を押しました。若主人はトレーシングペーパーでカメレオンを抜きました。若主人は派出所で酩酊しました。若主人は派出所で派出所を放浪しました。若主人は派出所で酩酊しました。若主人は派出所で酩酊しました。若主人は派出所で酩酊しました。若主人は派出所で寝ました。若主人は上で上を放浪しました。若主人は上で専攻に音を上げました。若主人は上でカメレオンを捕獲しました。審理に上で競争力がありました。カメレオンは上で稲刈りに干渉しました。若主人に上でカメレオンは居ました。審理は上で審理との矛盾がありました。若主人は上で審理を発音しました。若主人は上で働きました。若主人は鉄塔で働きました。若主人は鉄塔で働きました。若主人は鉄塔で働きました。若主人は鉄塔で居残り手当を得ました。若主人は溝で働きました。若主人は溝で働きました。若主人は溝で居残り手当を得ました。若主人は縄暖簾でビアを飲みました。[A young host is extremely drunk in a ditch. The young host dream of liquidation in the ditch. A super - express slips by a ditch. The young host comes to the young host' self in a ditch. A tracing paper is generated by attraction. The young host reverses the super - express against a ditch. The young host makes sure of the tracing paper. The young host passes a Chamaeleon. The young host is extremely drunk in a police box. The young host wanders in a police box. The young host is extremely drunk in the police box. The young host is extremely drunk in the police box. The young host is extremely drunk in the police box. The young host goes to bed in the police box. The young host wanders through the upper part. The young host is at the young host's wits' end for a special study. The young host catches the Chamaeleon. A judge is competitive. The Chamaeleon interferes in the upper part. The young host has the Chamaeleon. The judge contradicts the judge. The young host pronounces the judge. The young host works in the upper part. The young host works in a pylon. The young host works in the pylon. The young host works in the pylon. The young host earns in the pylon. The young host works in the ditch. The young host works in the ditch. The young host earns in the ditch. The young host drinks a beer.]

Figure 8 | Generated sentence using a story structure

hierarchically, based on verbs. When the INGS selects an event that randomly, and if the event can be expanded by any story technique, the INGS selects a story technique in those story techniques randomly. However, if the event cannot be expanded by any story technique, the INGS then selects another event randomly. Figure 8 shows how a generated story is transformed into a natural language sentence by a sentence generation mechanism in the INGS. Although the authors have experience with the topic of random generation, previous attempts to generate a story through random methods failed owing to the lack of a knowledge base.

6. CONCLUSION

This paper showed short script acquisition from existing works. The authors acquired 23,751,142 short scripts from existing works that were stored the Aozora Bunko. The acquired short scripts were used for story generation in the INGS. In future works, the authors will acquire longer scripts. In addition, they will attempt to acquire scripts semantically, because the results in this paper did not consider the semantic structure of a story.

The authors' goal is the random generation of a story. We presented story generation by using knowledge, which was obtained automatically. Based on the discussion above, a method to change narratives and to set parameters for a narrative generation was presented.

For future research, the aim is to increase automation by using learning mechanisms. For example, we propose a generation strategy composed of a preliminary story generation, a valuation of the story, and elements conjecture where operating parameters are not input manually, and instead are selected automatically. Through the learning operations at some selection points, generation results improve using a bottom-up process. A second option would be the construction of a system although maintaining the consistency of a story in a narrative generation process is very difficult, and maintaining consistency would be our goal.

REFERENCES

[1] T. Ogata, Computational and cognitive approaches to narratology from the perspective of narrative generation, in: T. Ogata, T. Akimoto (Eds.), Computational and Cognitive Approaches to Narratology, IGI Global, Pennsylvania, 2016, pp. 1–74.
[2] T. Ogata, Building conceptual dictionaries for an integrated narrative generation system, J. Robot. Netw. Artif. Life 1 (2015), 270–284.
[3] T. Arai, T. Ogata, Acquiring and using event sequences based on verbs from novels, Proceedings of the 31st Annual Conference of the Japanese Society for Artificial Intelligence, 2017, 1D2-OS-29a-5in2.
[4] R.C. Schank, Tell Me a Story: A New Look at Real and Artificial Memory, John Brockman Associates, New York, 1990.
[5] Aozora-Bunko, Available from: http://www.aozora.gr.jp, 1998 (accessed December 2016).
[6] T. Ogata, T. Arai, J. Ono, Random Narrative Generation in an Integrated Narrative Generation System, Proceedings of the Forum on Information Technology, Information Processing Society of Japan, 2016, pp. 133–134.

Authors Introduction

Dr. Jumpei Ono

He received his Ph.D. from Iwate Prefectural University, Japan in 2018. He is now working as an instructor in Vocational School of Digital Arts Sendai in Japan. His research interest includes artificial intelligence, cognitive science, natural language processing, narratology, narrative generation, and game.

Dr. Takashi Ogata

He received his Ph.D. in the University of Tokyo in 1995. He has industrial experience since 1983 at software development companies. He is a Professor of the Faculty of Software and Information Science at Iwate Prefectural University since 2005.

His major research interests include artificial intelligence, cognitive science, natural language processing, narratology and literary theories, an interdisciplinary approach to the development of narrative generation systems based on AI and narratology, and the application to narrative creation and business.