Hitch Haiku: An Interactive Supporting System for Composing Haiku Poem

Naoko Tosa, Hideto Obara, and Michihiko Minoh

Academic Center for Computing and Media Studies, Kyoto University
Yoshida-Nihonmatsu-cho, Sakyo-ku, Kyoto, 606-8501 Japan
{Tosa,Obara,Minoh}@media.kyoto-u.ac.jp

Abstract. Human communication is fostered in environments of regional communities and cultures and in different languages. Cultures are rooted in their unique histories. Communication media have been developed to circulate these cultural characteristics. The theme of our research is "Cultural Computing", which means the translation of cultures using scientific methods representing essential aspects of Japanese culture [1]. We study the reproduction of a traditional Japanese Haiku by computer. Our system can abstract an essence of human emotions and thoughts into a Haiku, a Japanese minimal poem form. A user chooses arbitrary phrases from a chapter of the essay "1000 Books and 1000 Nights" [2]. Using the phrases chosen by the user, our system generates the Haiku which includes the essence of these words.

Keywords: Poem, Haiku, Haiku generation, Art, Interactive art.

1 Introduction

Haiku is a Japanese classical poem style with minimal length of five-seven-five characters including a seasonal word called *"Kigo."* The original form of Haiku was called Hokku and in the late 19th century Shiki Masaoka revised it and finally established the present form of Haiku [3]. Haiku include various imaginative expressions and thus has been applauded by many people. Haiku is a story that generates context - the shortest story in the world. Known as the first great Haiku poet in the Japanese history, Matsuo Basho is responsible for *"Oku No Hosomichi"*, a prime example of his work [4].

We developed a new interactive system, "Hitch Haiku", which supports a user for composing a Haiku. The user inputs some words into the system, and the system composes phrases consisting of five-seven-five characters which most fit with the user inputs. The system is called Hitch Haiku as it generates a Haiku "hitching" the phrases chosen based on the user inputs. If the user does not like the composed Haiku, the user can modify the Haiku and make the system to learn the composition of better Haiku. Hitch Haiku is one of the automatic poem generation systems. In 1959, Theo Lutz developed a system of poem generation for the first time [5]. The system only showed words at random based on grammatical rule, and could not generate a poem in its real meaning. In 1971, for the first time Masterman developed the generation system of a Haiku [6]. By rearranging the words which users chose from the pull

down menu in the interaction process the system generates a Haiku. However, in these interactions, users could input only a few limited words into the system.

In the field of Interactive Art or Game, the quality of contents is important [7]. But in these cases only simple techniques have been used. On the other hands, in the field of AI, many researchers have been using various kinds of technologies to find some relations among input words/phrases by users and to compose answers in relation to these inputs [8]. These techniques have been often used, because using one of these techniques they can develop an interactive system that can achieve relatively interesting interactions. But the relations they try to find out and they try to use in their systems are static, and the quality of their interactions have been mostly dependent on the quality of the relations given beforehand.

Based on his long carrier as an editor and a philosopher, Matsuoka fund four several basic forms called *"Thoughtforms,"* that exist as basic forms of relations among things [9]. By tracing and re-constructing some relations indicated by *"thoughtforms,"* Tosa and Matsuoka created an art work called *"i.plot"*, which displays dynamically hidden relations and contextual emergences of English, Chinese Character, and so on [10]. As this technique can re-construct interesting relations and enable to generate Haiku poems, we apply the technique to our system. Furthermore in our system we included the learning function. If the users do not like the generated Haiku based on the users inputs they can modify the Haiku. The whole process is observed by the system and it learns method of generating better Haiku using these examples. The more the users play with our system, the better relations among words/phrases our system learns, and the better Haiku our system can generate.

2   Process of Generation

Our system generates a Haiku according to the following process (Fig. 1). We describe the detail of the process below.

(1) A user chooses arbitrary phrases from a chapter of a famous Japanese essay called "1000 Books and 1000 Nights", which introduces more than 1000 books covering many genres from all over the world [2].

(2) The system carries out a syntactic analysis for each of the phrase and detects a basic form of noun or a verb from each phrase.

(3) Then the system composes a phrase of a Haiku by adding a special propositional particle called *"Kireji"*, which not only separates a Haiku into three phrases but also gives each phrase a Haiku-like feeling.

(4) In addition to these phrases the system tries to generate new phrases so that the combination of the phrases would expand the imagination of a reader/listener. For this the system uses six types of databases; Haiku thesaurus, Kigo thesaurus, idiom thesaurus, case frame of onomatopoeia, thesaurus, and case frame.

(5) From these databases, the system searches the phrase which is most related to the user inputs.

(6) The system scores all phrases using the following weights: Haiku thesaurus is 3, Kigo thesaurus is 3, idiom thesaurus is 3, case frame of onomatopoeia is 3, thesaurus is 3, case frame is 1, user's relation is 5. If a phrase includes two or more related

words, the system sums the score respectively. The system chooses one of the phrases with the highest score, and by presuming the season of the Haiku from the user inputs choose a Kigo using the Kigo database (Fig. 2).
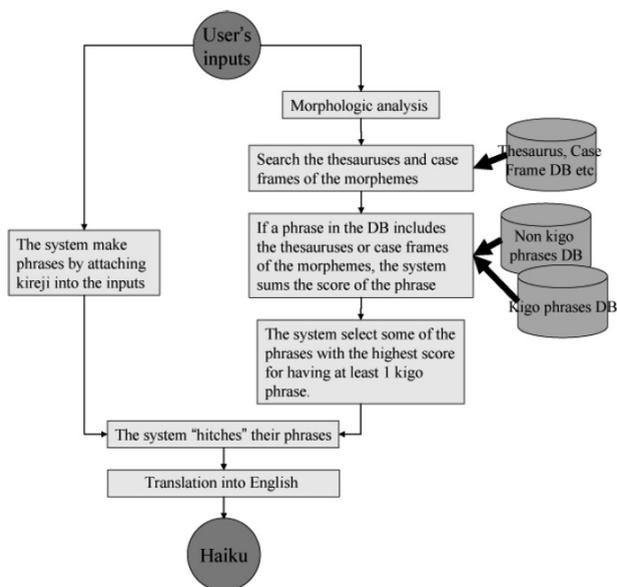
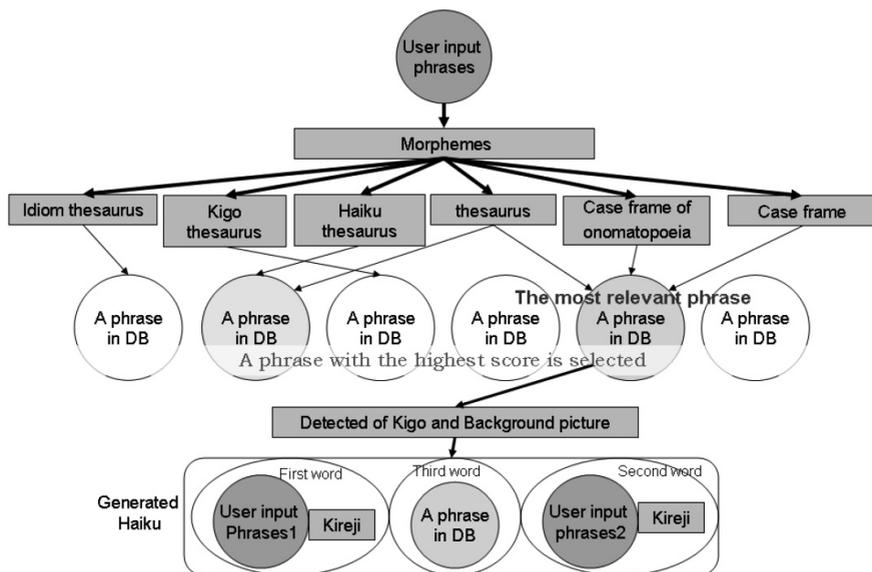**Fig. 1.** Haiku generation process

Fig. 2. Selection of the phrase with the highest score

(7) The system translates Japanese Haiku into English Haiku using the translation system, *"Language Grid"* developed at NICT [11]. Language Grid has many dictionaries of translation in many communities. By choosing and using a suitable dictionary of a Haiku community in all dictionaries and by using it, Language Grid translates phrases.

(8) If the user does not like the generated Haiku, the user can modify the Haiku phrases and register these new phrases in the system. We assume that the user inputs have strong relativity with the phrases modified by the use. The system adds the relations between the user inputs and the morphemes of the modified phrases into the database, thus the system learns the new relativity.
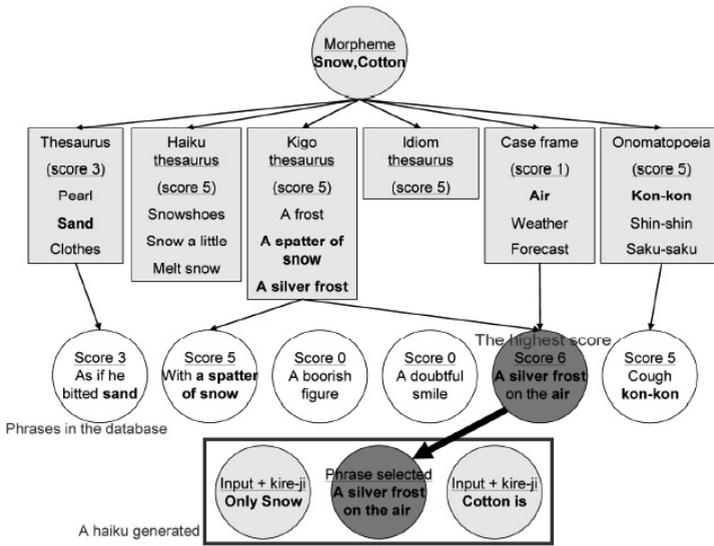

Fig. 3. An example of the phase selection

Figure 3 shows an example of the selection process where two words *"snow,"* and *"cotton"* are selected by an user. First to each input word, the most relevant Kire-ji is selected so that two phrases composed from these words would have five or seven characters. As Haiku consists of three phrases, it is necessary to generate one more phrase. For this, the system searches the words related to *"snow"* and *"cotton,"* from the database. Form the thesauruses, *"pearl, sand, cotton"* are found, and from Haiku thesauruses, *"snowshoes, snow a little, melt snow,"* and so on. Then again using the database the system searches Haiku phrases which include these words scores each of these obtained phrases. For example, the phrase, *"As if he bitted sand"* is scored 3 because the word "*sand*" is including in the thesaurus whose score is 3, and *"A silver frost on the air"* is scored 6 because *"a silver frost"* is included in the Kigo thesaurus whose score is 5 and *"air"* is included in the case frame whose score is 1. The system chooses the phrase, *"A silver frost on the air"*, as it has the highest score. Thus the system generates a Haiku by "hitching" the user inputs. If the user does not like the generated phrase "A silver frost on the air," he/she can modify the phrase to new

phrase, for example "Like the wind on the air." Then system adds the relations between *"snow"* and *"air,"* *"snow"* and *"wind,"* into the system. (Here, snow and air are the user inputs, wind is a morpheme of the modified.)

Table 1. Examples of the database contents

Database	A Example of Category	Expression
Case frame	fade	sound, taste, part, focus ...
Thesaurus	Expression	Enlightenment, Crying, A facial expression showing familiarity ...
Haiku thesaurus	Love	Relics, Courting, Affection, Loving one another ...
Kigo	Spring	Lunar New Year, First day of spring ...
Kigo	Summer	Summer, Early summer, April, May, First day of summer, Summer-like, Slight heat ...
Kigo	Autumn	Autumn, Early autumn, July of the lunar calendar, August, First day of fall, Lingering summer heat. Autumn-like ...
Kigo	Winter	Winter, November of the lunar calendar, December, Winter solstice, The end of a year, The days of year-end
Kigo	New Year	New year, Early spring, The New Year of the lunar calendar, This year, Last year ...
Idiom thesaurus	Love	The whip of love. Autumn breeze blows, Parents often cannot judge the right way to love their children, ...
Idiom thesaurus	Gather	hang out. Many faces, In a crowd, ...
Idiom thesaurus	Walk	turn legs toward, Leave stranded, Extend one's trip ...
Case frame of onomatopoeia	Oh	Mouth, Large, Open, Cry ...
Case frame of onomatopoeia	Frank	Taste, Color, Deep, Style, Form, ...
Case frame of onomatopoeia	Hesitant	Attitude, Action Decision, Condition, ...

3   Database

We have prepared six types of databases. The case frame database (with about 31,000 records) describes the relation between verbs and nouns. The thesaurus database (with about 32,000 records) is constructed by classifying words with similar meaning into one category. It is constructed based on a general thesaurus dictionary. The Haiku thesaurus database (with about 2,500 records) is constructed by finding relationship among entries in a general Haiku dictionary. The Kigo thesaurus database (with about 13,000 records) is constructed by setting a Kigo (a seasonal word) as its entity and various kinds of expressions as its contents. The Idiom thesaurus database (with about 1,300 records) contains words as its entries and various kinds of idiom phrases frequently used in Haiku poems. The Case frame of onomatopoeia (with about 8,800 records) contains the relation between an entry word and various kinds of onomatopoeias that are used in relation to the word. The database of user's relation contains the relation between the user input and the morpheme extracted from a phrase modified by the user. The records in this database increase whenever users make any revision to a generated. Some of the examples of the records in this databases are shown in Table 1.

4   Interaction Example

We show an example of Haiku generation by our system.

(1) First if an user logs in the system, he sees a map of the City of Book (Fig.4).
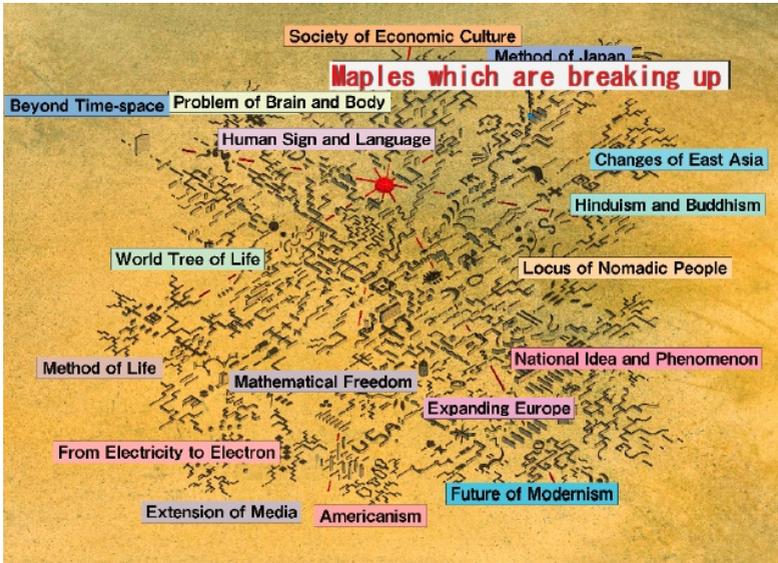(2) By manipulating a mouse, he can walk around in the city.

Fig. 4. An example of the map of the City of Book

Fig. 5. Selection of words/phrases

(3) If he finds an interesting keyword of a book at particular points of the city, and clicks the keyword, he sees the title and the author of the book. By clicking again, he sees the essay of the book contained in *"1000 Books and 1000 Nights."*

(4) When the user marks some words on the essay using a pen or an ink brush, the Haiku system generates a Haiku based on the marked words (Fig.5).

(5) If he does not like the generated Haiku, then he can modify and save it.

Figure 6 shows an example of generated Haiku. In Fig. 5, an user selected *"Zen"* and *"ink."* The system chooses the word *"god"* that comes from "*Zen*" using the thesaurus database then chooses the phrase *"The one of the seven gods of Good Luck"* based on the algorithm described above Finally, the system generates the Haiku, *"Even Zen, The one of the seven gods of Good Luck, It's in Ink."*

Fig. 6. An example of a generated Haiku

5   Conclusion

In this paper we proposed an interactive system which supports an user for composing a Haiku. Haiku is a short poem with an abbreviation. Atmosphere and emotion of Haiku is also part of the abbreviation. Based on the algorithm we have developed, we found that sensitive atmosphere and emotion which even we could not expect emerges. We have exhibited this system at ACM SIGGRAPH 2007 and obtained fairly good responses from the visitors [12].

There are several issue for further studies. To increase the contents of each database is one of the important issues. As this system was originally developed for the purpose of Japanese Haiku generation, the process of the translation into English needs a further improvement. At the same time our future target is a system that would support the generation of Haiku based on an interaction between the system and user. So far, what users can do is only to select several phrases/words from texts they like. Although this would be relevant to give a beginner a feeling of Haiku generation, experts would need a system that would really help their creation process. To achieve this target starting from our present system would be a big challenge.

References

1. Tosa, N., Matsuoka, S.: ZENetic Computer: Exploring Japanese Culture, vol. 39(3), pp. 205–211. MIT Press, Leonardo (2006)
2. Matsuoka, S.: 1000 Books and 1000 Nights. Editorial Engineering Laboratory, `http://www.isis.ne.jp/mnn/senya/senya.html`
3. Wikipedia: Haiku, `http://en.wikipedia.org/wiki/Haiku`
4. Matsuo, B.: A Haiku Journey -Basho's Narrow Road to a Far Province-. Kodansha International Ltd. Translated by Dorothy Britton (1974)
5. Lutz, T.: Stochastische text. Augenblick 4(1), 3–9 (1959)
6. Masterman, M.: Computerrized Haiku. In: Cybernetics, pp. 175–183 (1971)
7. `http://digitalmedia.risd.edu/billseaman/workMajor.php`
8. Weizenbaum, J.: Eliza - a computer program for the study of natural language communication between man and machine. Communications of the ACM 9(1), 36–45 (1966)
9. Matsuoka, S.: Intelligent Editorial Engineering. Asahi Shinbun Publisher (2001)
10. Tosa, N., Matsuoka, S.: i.plot. ACM SIGGRAPH Emerging technologies. In: Proc. the 32nd International Conference on Computer Graphics and Interactive Techniques, p. 154 (2005)
11. Ishida, T.: Language Grid: An Infrastructure for Intercultural Collaboration. In: IEEE/IPSJ Symposium on Applications and the Internet (SAINT 2006), pp. 96–100 (2006)
12. Tosa, N., et al.: Hitch Haiku. ACM SIGGRAPH 2008 Art Show (2007)