5N−10

Internet Shiritori using Java*

Tiemi Christine Sakata Masahide Sugiyama
Computer Science & Eng The University of Aizu

1 Introduction

Word Play Game is an introduction to language learning to the young and help young children learn the letters and sounds of the alphabet and develop first reading and spelling skills. Shiritori Game is one typical word play game in Japan and this game also works just as well with non-native learners of Japanese. Java is a flexible language, provides resources for the Internet and allows the creation of a user-friendly interface. The propose of this paper is development of Shiritori game in Java language [1] [2] to execute on the Internet using a large vocabulary and to discover intelligent strategies to obtain real-time answers to the players.

2 Shiritori Game

Shiritori game is a Japanese word play game in which one player has to say a word that must not end with the /N/ sound and starting with the last syllable of the word given by the previous player. Computer follows the rules of game and verifies if user inputed a correct word.

2.1 Vocabulary

The vocabulary used was obtained by getting the column with words in Hiragana characters of jddic dictionary and contains 9500 words. In this paper, the Vocabulary File is represented by VF. In Table 1, NWI is number of words that initiates with a given Hiragana and NWF is the number of words that finishes with a given Hiragana.

Computer's weakness is when NWI < NWF.

Table 1: Analisys of Vocabulary

Character	NWI	NWF
あ	249	2
い	298	1109
...		
ぴ	0	9
...		
Total	9500	9500

2.2 Strategy

Branching Number (BN) is defined as the number of connectable words of a given word. Computer supposes that its weak point is the same weak point of the user. In other words, the system based on itself vocabulary to get its best answer. Accordingly this purpose, computer checks all possible answer words and returns the word with a minimal branching number. For example, if user input the word "えいが", computer will answer "がんばる" because its BN = 11 is the minimal BN of all words of possible answers, demonstrated in Figure 1.

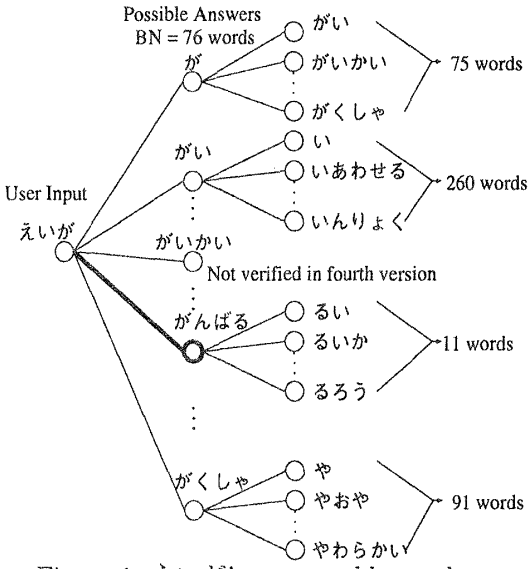

Figure 1: えいが's connectable word tree

2.3 Development

Four versions of Shiritori game were implemented providing better time response.

2.3.1 First Version - Sequencial Search

Initially, the VF was not sorted. For this reason, to get all possible answer words it is necessary to read all words of VF and to calculate the BN of each word it is also necessary to read all VF's word. For example, if the number of possible answers of a given word is 3, it is necessary to read the VF 4 times shown in Figure 2. The memory used in this version is the VF size. As result, answers can be obtained from 15 seconds to 8 minutes.

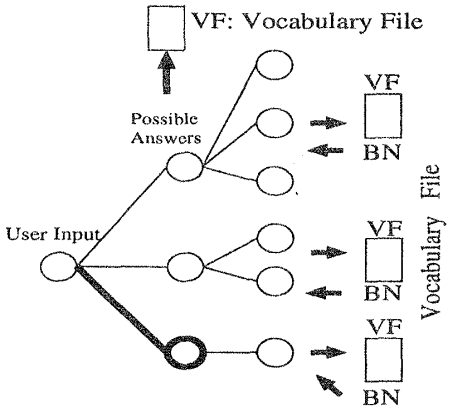

Figure 2: First Version

* Java 言語を用いた Internet 対応しりとりゲーム, 坂田 知恵美クリスチーネ, 杉山 雅英 (会津大学 コンピュータ理工学部)

2.3.2 Second Version - Sorted Sequencial Search

The great difference of this version is that the VF is sorted. Therefore, to get possible answers words and to calculate the BN of each word it is necessary to read just an appropriate peace of VF, shown in Figure 3. The size of memory used in this verion is equal to the size of VF. The average of answer time is half of first version but until delayed to the fact that it is a game.

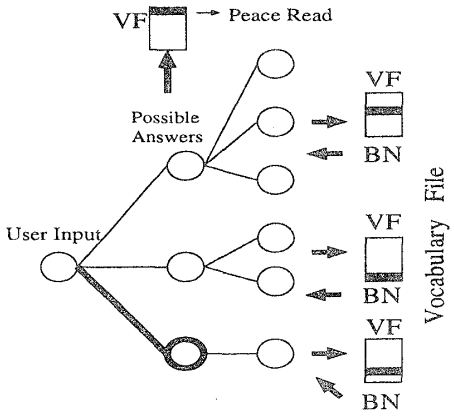

Figure 3: Second Version

2.3.3 Third Version - Index and BN based Search

In the initialization, the VF is read and the object with some characteristics of this File is created shown in Figure 4. Therefore, to get the BN of a given word, it has just get it in this object. In this version, it is necessary to use not only the size of VF, but the size of Vector that contains Hiragana characters and the size of vector with characteristics of each Hiragana character. In spite of the fact that the memory size increased, it is not necessary to read a file many times, but just once and good response time coud be obtained, being the average of a answer time about 105 milliseconds.

Vector Hiragana

	あ	い	う	え	お	...	わ
BN	232	260	143	108	238	...	63
Index of VF	0	1223	2714	3402	3998	...	48499
check	false	false	false	false	false	...	false

Vector IdentChar

Figure 4: Characteristics of each Hiragana Character

2.3.4 Fourth Version - Third Version + Black Board Search

It is a customization of the third version. In this version, the BN of words with the same final character is verified just once. Furthermore, if the BN of a given word is zero, this word is returned immediately. Consequently, this version increased mainly the performance of word with large number of possible answer words.

3 Result

This four version was create to improve the time answer to the user. Figure 5 shows the average of answer time in the four versions and the memory used can be compared in Table 2.

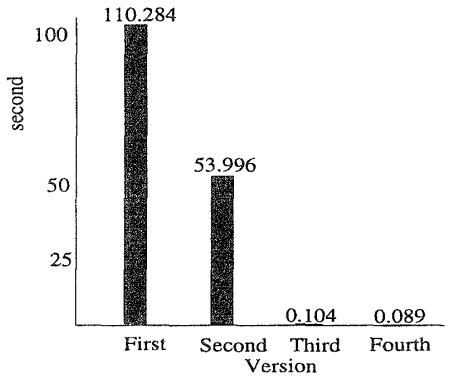

Figure 5: Response Time in Four Versions

Table 2: Memory used in Four Versions

Version	Memory Size
First Version	VS
Second Version	VS
Third Version	VS + 3 * NH
Fourth Version	VS + 4 * NH

VS: Vocabulary Size
NH: Number of Hiragana Character

4 Conclusion

It is possible to get answer word in this game quickly, in spite of the fact that computer uses an vocabulary with 9500 words. Although increase the vocabulary, the time of answer will not have a great difference. The most difficult point is that the user's vocabulary is open and computer works just with suppositions. After get a great time answer, it is possible to develop more intelligent strategy and use voice to return word to the player. Furthermore, it is interesting to adopt it to multi player game.

References

[1] John Withers, "*Developing Java Entertainment Applets*", Wiley Computer Publishing, 1997.

[2] Ken Arnold & James Gosling, "*The Java Programming Language*", Addison-Wesley Publishing Company, 1996.