Keywords: character connections, comics, manga, Dragon Ball, frames

Abstract: We presented a method to create character connections from manga using the frequencies of characters and their co-occurrences by referring to frames. First, we input characters and frames with a data input tool. Second, we calculated the frequencies of characters and the relationships among characters and group-related characters. Third, we created character connections. Preliminary experiments using Dragon Ball vol. 32 suggest the usefulness of our approach.

1 INTRODUCTION

In recent years, the popularity of manga (Japanese comics) has increased worldwide, and a large amount of manga is being published. As of 2008, the U.S. and Canadian manga market generated $175 million in annual sales (Reid 2009).

Manga is usually first serialized in magazines and later compiled in books. Some manga is a long epic. For example, Dragon Ball, a well-known manga written and illustrated by Akira Toriyama, was originally serialized from 1984 through 1995; later its 509 individual chapters were published as 42 book volumes.

Since finding a chapter, a specific volume, or a manga itself from a large amount of manga collections is difficult, we investigate how to find a chapter or a particular manga volume.

Character connections are often created to introduce the contents of multimedia such as movies, anime, and TV dramas that help people understand the characters and the complicated stories. We believe that character connections are also useful for finding and understanding manga. However, creating character connections from manga is unclear, time-consuming, and expensive.

This research creates character connections from manga to help users find and understand its contents. Below, in Section 2 we explain our approach’s overview. Algorithms and preliminary experiments are described in Sections 3 and 4. Related work is shown in Section 5.

2 OUR APPROACH

This research creates character connections from manga.

First, we input characters and frames with a data input tool. Second, we calculate the frequencies of characters and relationships among characters and group-related characters. Third, we create character connections.

The main feature of this research is creating character connections from manga using the frequencies of characters and their co-occurrences by referring to frames. No current work creates character connections from manga or comics.

Figure 1 shows created character connections from an experiment using Dragon Ball vol. 32.
3 ALGORITHMS

3.1 Data Input
Data input is time-consuming. We prepared a data input tool using Excel macro to reduce the burden of the task. First, we manually extracted the character names contained in the manga and input them into the tool. Next, a user read the manga and pressed a button for the registered character names when they appeared. Finally, a data file was produced. It takes approximately one hour to input data from one volume (about 130 pages).

Figure 2: Example of data input.

Figure 2 shows an example of creating a file from a page, beginning at chapter 374 of Dragon Ball vol. 32. Manga flows from top to bottom and right to left. In the first frame of the first page, Cell (character name) appears. In the next frame, Android #16 (hereafter #16) and Android #18 (hereafter #18) appear.

3.2 Judging Relations

3.2.1 Relations inside Frames
We assume that characters who appear in the same frame are related. When more than one character appears in a frame, we give related scores to each character and assume that when the number of characters inside frames is smaller, the relation is stronger. Related scores inside frames between characters \( C_i \) and \( C_j \) are calculated as:

\[
R_{\text{in}}(c_i, c_j) = \begin{cases} 
\frac{1}{n} & \text{if } n = 1, 2, 3, 4 \\
0.25 & \text{if } n \geq 5 
\end{cases}
\]  

Here \( n \) is the number of different characters from the designated character inside a frame.

3.2.2 Relations outside Frames
We assume that characters who appear in the next frame are related. When different characters appear in subsequent frames, related scores are added to each character. Related scores outside frames between characters \( C_i \) and \( C_j \) are calculated as:

\[
R_{\text{out}}(c_i, c_j) = \begin{cases} 
\frac{0.5}{n} & \text{if } n = 1, 2, 3, 4 \\
0.25 & \text{if } n \geq 5 
\end{cases}
\]  

Here \( n \) is the number of different characters from the designated character outside (next) the frame.

3.2.3 Unifying Relations
Finally, related scores between \( c_i \) and \( c_j \) are calculated as the summation of the related scores of relations inside and outside frames, as shown in Equation (3).

\[
R(c_i, c_j) = \sum_{k=1}^{n} R_{\text{in}}(c_i, c_j) + \sum_{k=1}^{n-1} R_{\text{out}}(c_i, c_j)
\]  

Here \( n \) is the number of frames in designated units (e.g., chapters, volumes, etc.).

We created a table whose rows are characters sorted by their frequencies and whose columns are related characters sorted by their related scores.

Table 1 shows the five most frequent characters and the five most related characters. For example, Cell is the most frequent character and appears 266 times in vol. 32. The most related character to Cell is Vegeta whose related score with Cell is 172.5.
Table 1: Five most frequent characters with five most related characters.

<table>
<thead>
<tr>
<th>Character</th>
<th>Related characters</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cell</td>
<td>Ve</td>
<td>172.5</td>
<td>43.1</td>
<td>24.0</td>
<td>23.7</td>
<td>15.7</td>
</tr>
<tr>
<td>Vegeta</td>
<td>Ce</td>
<td>172.5</td>
<td>28.2</td>
<td>10.4</td>
<td>10.2</td>
<td>8.3</td>
</tr>
<tr>
<td>Krillin</td>
<td>Tr</td>
<td>52.6</td>
<td>43.1</td>
<td>20.1</td>
<td>14.0</td>
<td>11.4</td>
</tr>
<tr>
<td>Trunks</td>
<td>Kr</td>
<td>52.6</td>
<td>43.1</td>
<td>28.2</td>
<td>13.2</td>
<td>8.1</td>
</tr>
<tr>
<td>Goku</td>
<td>Gh</td>
<td>43.8</td>
<td>15.0</td>
<td>14.3</td>
<td>13.2</td>
<td>10.4</td>
</tr>
</tbody>
</table>

Note: [ ] indicates appearance frequency of characters; numbers on right hand are related scores.
Note2: Ce: Cell; Ve: Vegeta; Kl: Krillin; Tr: Trunks; Go: Goku; #18: #18; #16: #16; Gh: Gohan; Bu: Bulma; Ti: Tien Shinhann; Pi: Piccolo; Po: Mr. Popo

3.3 Grouping
We group strongly related characters using the table generated by the previous step. The grouping algorithm is shown in Figure 3.

Step 1. We verify the differences between the related scores in a raw and set the border using the biggest difference. We make an initial group from the left-most characters to the characters next to the border.
Step 2. We check all initial groups and merge those include same characters in them.

Figure 3: Grouping algorithm.

For example, for the first raw of Cell, 129.4 (172.5-43.1) is the biggest among (172.5-43.1), (43.1-24.0), ...; Cell and Vegeta are grouped into an initial group. Likewise, Vegeta and Cell are grouped from the second raw. These two initial groups are merged into one group.

Table 2: Grouping results.

<table>
<thead>
<tr>
<th>Group</th>
<th>Evaluation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cell, Vegeta</td>
<td>5.0</td>
</tr>
<tr>
<td>Krillin, Trunks</td>
<td>3.8</td>
</tr>
<tr>
<td>Goku, Gohan</td>
<td>5.0</td>
</tr>
<tr>
<td>Tien Shinhann, Piccolo</td>
<td>3.0</td>
</tr>
<tr>
<td>#18, #16</td>
<td>4.8</td>
</tr>
<tr>
<td>average</td>
<td>4.3</td>
</tr>
</tbody>
</table>

3.4 Generating Character Connections
We generated character connections using the frequencies of characters, their related scores, and groups. The algorithm is shown in Figure 4.
We created a related table (Table 4) for each method. Characters ranked under 7 were omitted because there was no correct answer. We counted the rank of the correct dataset: smaller is better. Table 4 shows the table for our method. For example, for the first raw Cell, 1 (Vegeta) + 5 (Trunks) + 6 (#16) + 4 (Krillin) + 2 (#18) + 3 (Goku) = 21. The totals for our method, inside only, and outside only were 62, 66, and 63, respectively. Our method was the best.

Table 4: Dataset for our method.

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ce</td>
<td>Ve *</td>
<td>Tr *</td>
<td>16*</td>
<td>Kl *</td>
<td>18*</td>
<td>Go *</td>
</tr>
<tr>
<td>Ve</td>
<td>Ce *</td>
<td>Tr *</td>
<td>Go *</td>
<td>Kl</td>
<td>16</td>
<td>18</td>
</tr>
<tr>
<td>Kl</td>
<td>Tr *</td>
<td>Ce *</td>
<td>18*</td>
<td>Bu *</td>
<td>16</td>
<td>Ve</td>
</tr>
<tr>
<td>Tr</td>
<td>Kl *</td>
<td>Ce *</td>
<td>Ve *</td>
<td>Go</td>
<td>Bu</td>
<td>16</td>
</tr>
<tr>
<td>Go</td>
<td>Gh *</td>
<td>Ti</td>
<td>Ce *</td>
<td>Tr</td>
<td>Ve *</td>
<td>Pi</td>
</tr>
<tr>
<td>18</td>
<td>16*</td>
<td>Kl *</td>
<td>Ce *</td>
<td>Ve</td>
<td>Tr</td>
<td>Ti</td>
</tr>
<tr>
<td>16</td>
<td>18*</td>
<td>Ce *</td>
<td>Kl</td>
<td>Ve</td>
<td>Tr</td>
<td>Ti</td>
</tr>
<tr>
<td>Gh</td>
<td>Gh *</td>
<td>Bu</td>
<td>Tr</td>
<td>Ti</td>
<td>Po</td>
<td></td>
</tr>
<tr>
<td>Bu</td>
<td>Kl *</td>
<td>Tr</td>
<td>Pi</td>
<td>Go</td>
<td>Gh</td>
<td>Ti</td>
</tr>
<tr>
<td>Ti</td>
<td>Go</td>
<td>Pi</td>
<td>Ce</td>
<td>Bu</td>
<td>Gh</td>
<td>Tr</td>
</tr>
<tr>
<td>Pi</td>
<td>Ti</td>
<td>Go</td>
<td>Bu</td>
<td>Tr</td>
<td>Gh</td>
<td>Ve</td>
</tr>
<tr>
<td>Po</td>
<td>Go</td>
<td>Gh</td>
<td>Tr</td>
<td>Ve</td>
<td>Pi</td>
<td>Ti</td>
</tr>
</tbody>
</table>

Note: *: correct

4.3 Grouping
Five subjects evaluated whether the generated groups were appropriate by five values (5: very appropriate; 4: appropriate; 3: intermediate; 2: not very appropriate; 1: inappropriate).

The average value was 4.3 (See Table 2), and the result shows the grouping algorithm is satisfactory.

4.4 Character Connections
Five subjects evaluated the created character connections by five values as grouping. Finally, we conducted interviews about the above experiments.

The average evaluation value for the character connection was 4.4. All subjects believed that Cell was the most important character, although Goku is the hero of the Dragon Ball. They felt that the positions of Cell (center) and Goku (upper left) were appropriate.

These results suggest that the created character connections expressed the contents of the volume well and the frequency of the characters indicates the importance.

5 RELATED WORK

No previous research creates character connections from manga or comics.

Some research create character connections from other media. Goto et al. (2008) create character charts from EPG texts that introduce movies. Both studies use natural language understanding techniques to identify relations between characters. They do not deal with manga or comics.

Spysee, whose algorithm is based on Matsuo et al. (2006) extracts person information from the Web and displays social networks. Some famous character names were input as person names. For example, Cell is connected not only with such other characters as Krillin but also a voice actor of Cell in animation based on manga. This method is not adequate to express the character connections of designated units such as chapters or volumes.

Ogasawara et al. (2008) extracted persons from broadcast videos to construct a human correlation graph and examined both text and image processing; they didn’t create graphs.

6 SUMMARY

We presented a method to create character connections from manga using the frequencies of characters and their co-occurrences by referring to frames. Preliminary experiments using Dragon Ball vol. 32 suggest the usefulness of our approach. Since this is merely the first step of our research, we need to improve our algorithms and conduct further experiments using different manga. We believe our approach is applicable to other types of comics and should be investigated in the future.

REFERENCES


