Supplemental Material: Generation of Small Groups with Rich Behaviors from Natural Language Interface

ABSTRACT

This supplemental material contains group behavior models, the keyword list, the algorithm for computation of reference objects, the user study results, the user interface, and some examples.

1 GROUP BEHAVIOR MODELS

A group behavior model contains the information about a behavior of a small group. It consists of several attributes including a region required for the characters of the group, interpersonal social distance between characters, and orientation of each character. These attributes are not easy to be learned. Figure 1 shows two examples.

(a) (b)

Figure 1: Animation of a group of characters that is created based on group behavior models. There are two examples. (a) Seven active characters are dancing and three passive characters are watching. (b) Eight active characters are chatting.

2 SENTENCE ANALYSIS

The character-related keywords are listed as follows:

(1) Quantity: the quantity of people, e.g., "a", "one", "a group of", "some", "all".
(2) Character: e.g., "people", "man", "woman", "boy", "police officer", "soldier".
(3) Non-character object: e.g., "chair", "waterfall", "advertising pillar", "campfire", "car".
(4) Character behavior: e.g., "cheer", "walk", "run", "laugh", "dance".
(5) Preposition: e.g., "to", "with", "from", "beside", "near", "at".
(6) Emotion: e.g., "happy", "sad", "angry", "excited".
(7) Age: e.g., "young", "old", "elderly", "adult".
(8) Skin (or race): e.g., "Asian", "white skin", "black skin".
(9) Body posture: e.g., "tall", "short", "fat", "skinny".
(10) Wearing type: e.g., "short pants", "jeans", "business suit", "dress", "swimming shorts".

Figure 2 shows the structure of a situation node. Figure 3 shows a parsed result for a situation node. Figure 4 shows the results for tagging.

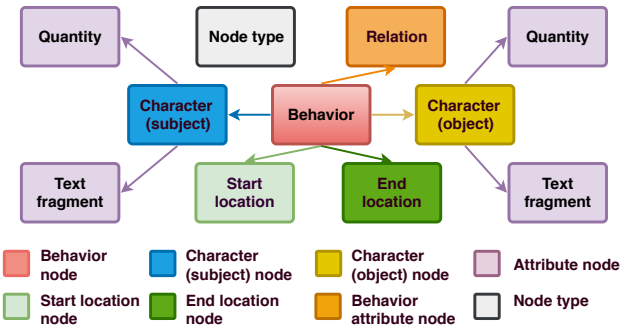

Figure 2: The data structure of a situation node. A situation node is created for each simple sentence.

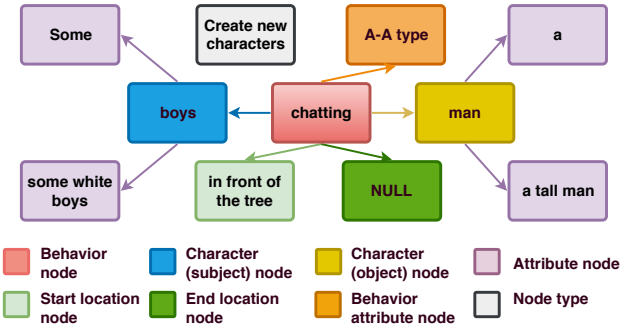

Figure 3: The content of the situation node for the input: "some white boys are chatting with a tall man in front of the tree.

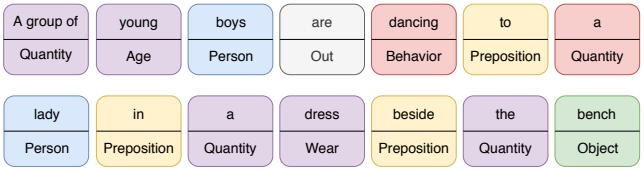

Figure 4: A result of the keyword extraction: *A group of young boys are dancing to a lady with a dress in the station.*

3 USER INTERFACE

Figure 6 illustrates that we can use entity retrieval to obtain an existing group of characters. Figure 7 shows the user interface.

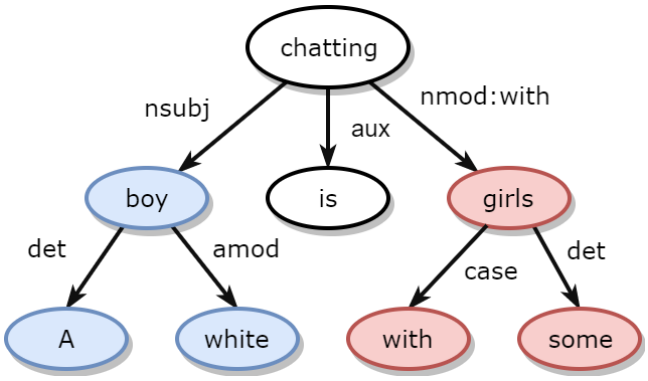

Figure 5: A parsing result of the sentence: A white boy is chatting with some girls. The blue and red subtrees represent two human subtrees, which are subject and object type, respectively.

Figure 6: Small groups are created based on the sentences: *Some boys and girls are dancing in the station. Some people are chatting. Some men with business suits are chatting. Some elders are cheering for Group1.* Our system automatically labels the first group as Group1. Thus, Group1 can be used in the second sentence.

4 REFERENCE OBJECTS FOR CHARACTER LOCATIONS

When a character or a group performs a behavior, we need to know the start location and the end location. These locations may be referred by scene objects. For example, given the sentence "Some people are dancing near a campfire", in this case, the end location is near to the campfire. Algorithm 1 shows the psuedo code which computes the reference objects for character locations.

Algorithm 1: Computation of reference objects for character locations.

```
Input  : The movement node, the node type, the start
         location object, and the end location object
Output : Reference location objects
if Both location objects != NULL then
    return location objects;
if The movement mode = large movement mode then
    if The start location object = NULL then
        if The node type = "control main group" then
            The main group is the start location object itself.
        else
            Random select a scene object as the start
             location object;
        end
    if The end location object = NULL then
        Randomly select a scene object as the end location
         object;
    return location objects;
else
    if The start location object = NULL then
        if The end location object != NULL then
            Set the end location object as the start location
             object;
        else if character node (object) != NULL then
            Set the character(object) as the start location
             object.
        end
        else if The node type = "control main group" then
            The main group is the start location object itself.
        end
        else
            Randomly select a scene object as the start
             location object;
        end
    return the start location object;
end
```

5 USER STUDY

We conducted a user study to evaluate our system. There were 3 female and 11 male participants whose average age was 24, ranging between 22 and 31. The procedure of the user study was as follows:

(1) Overview: Our researcher introduced our system to each participant. Then the participant watched a tutorial video which showed the tools and the procedure of generating a crowd animation.
(2) Crowd generation and control: The participant learned to perform basic tasks to get familiar with the system. The participant could refer to example sentences to finish several tasks including 1) generation of characters with behaviors in a certain location, 2) controlling existing characters to perform specific behaviors, 3) generation of characters in locations specified by hot regions, and finally 4) producing a simple animation.
(3) Reproduction of crowd animation in a given animation video and a video clip of a real crowd: The participant was shown a subway video and then a zombie attack story animation generated by our system. The associated story scripts were provided. The participant needed to reproduce crowd animations that were similar to the content in video and the story script.
(4) Exploration and feedback: The participant freely input different sentences to produce animations and used different tools to explore the system.

Figure 7: User interface. The animation graph on the left side is composed of several time nodes, and situation nodes. Situation nodes have four types and they are represented by different icons. The right side shows the input interface. The parsed result is also shown. A unique name is assigned to a character or a group after creation.

(5) Questionnaires: Participants filled a demographic and background questionnaire, a post experiment questionnaire with 10 questions in 5-points Likert scale, and System Usability Scale (SUS) [**?**]. The post-experiment questionnaire has ten questions which asked for whether our system could generate desirable small groups of characters at proper locations. Some questions were about the behaviors of small groups and animation results.

Table 1 shows the questions in the background questionnaire and the average score of each question. Table 2 shows the average score of each question in the post-experiment questionnaire. The overall average score of all the questions was 4.37 (SD = 0.22). Figure 8 shows the subway and forest scenes that were reproduced by P6, P12, P1, and P7.

Table 1: Background questionnaire

No	Question	Average
1	How good is your written English?	3.07
2	How good are you using animation software?	2.86

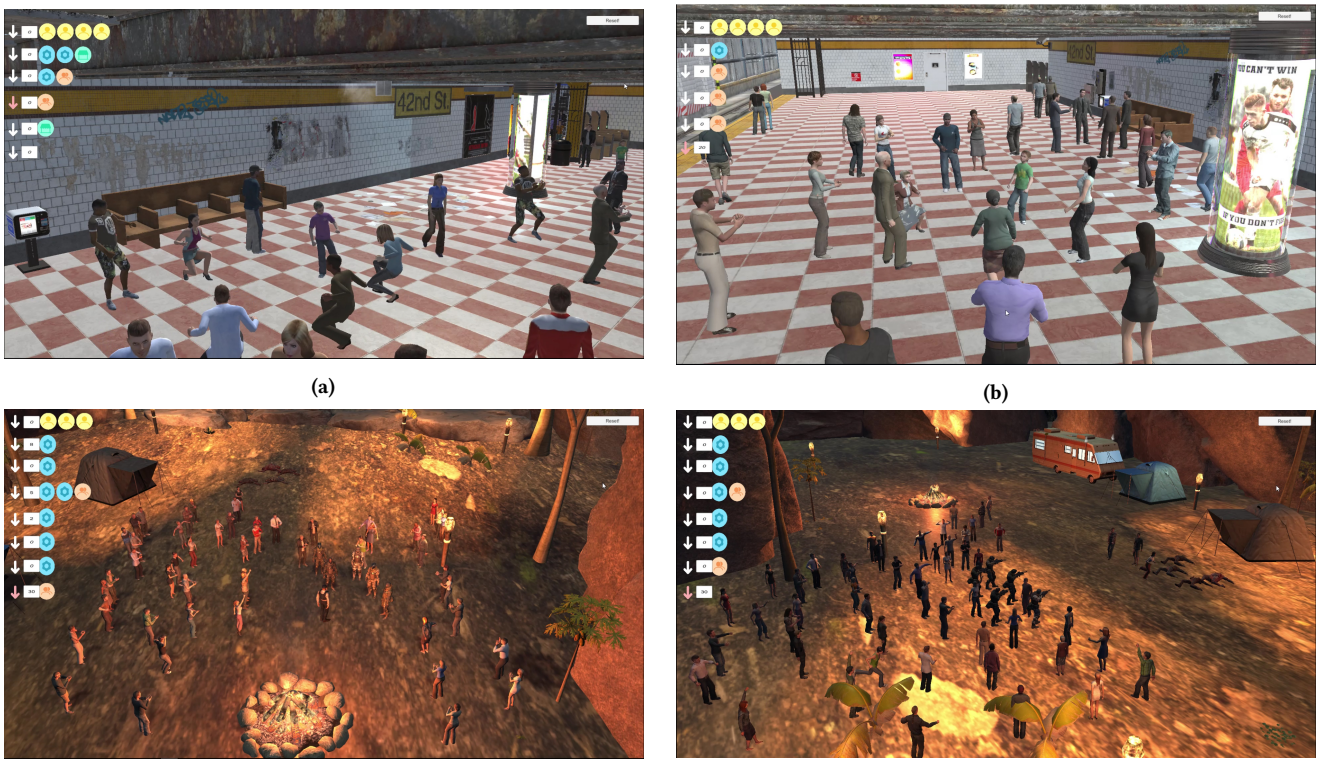

Figure 8: (a) and (b) were the reproduction scenes of a subway video by P6 and P12. (c) and (d) were the reproduction scenes of the zombie attack animation by P1 and P7. The small groups were created in different ways in the two examples. An animation graph is shown at the upper left corner of each image.

Table 2: Post experiment questionnaire. How much do you agree each of the following items? 5-point Likert scale (1=very disagree, 3=neutral, 5=very agree).

No	Question	Scores
1	The appearances of characters are generated based on the input sentences?	4.50
2	The behaviors of characters are generated based on the input sentences?	4.29
3	The interaction between characters is generated based on the input sentences?	4.29
4	The locations of characters are generated based on the input sentences?	4.29
5	The animation graph assists in producing a desirable animation?	4.29
6	Hot regions are useful when characters are created?	4.71
7	Hot regions are useful in specifying the destinations of characters?	4.64
8	The interaction between characters is nature?	3.93
9	The ambient crowd increases the richness of the animation?	4.34
10	The scene events increase the richness of the animation?	4.43
	The average score of the questionnaire	4.37