

Capturing Common Sense Knowledge via Story Generation

Kaoru Sumi

National Institute of Information and Communications Technology
Seika-cho, Soraku-gun, Kyoto, 619-0289 Japan

kaoru@nict.go.jp

ABSTRACT

This paper introduces an animation story generation system called WordsAnime using the animation database provided by animation CGM (consumer generated media). A user generates a story by repeating the steps of inputting scenarios and viewing animations. The system presents manual, semiautomatic, and fully automatic modes of creating scenarios using a common knowledge that links different scenarios with each other. The system automatically extracts the common knowledge of the IF-THEN rules via scenario generation. This paper reports attempt to capture common sense knowledge through the workshops on WordsAnime for children.

Author Keywords

Common sense knowledge, story generation, consumer generated media.

ACM Classification Keywords

I.2.6. Learning: Knowledge acquisition.

INTRODUCTION

WordsAnime was developed for automatically generating content of a three dimensional animation story by inputting or selecting a user's simple scenario; thus, this system can be operated even by children. The animation is created by combining animation parts corresponding to the words of scenario. The user generates the story by repeating the steps of inputting scenarios and viewing animations. Data on animation parts in an animation database referred to as Animebase correspond to words provided by the Animation CGM (consumer generated media): Anime de Blog [1]. WordsAnime is an application system using Animebase. Anime de Blog can collect 3D animation and image data. Users have indicated the need for a more appealing way of

creating animation in blogs. In this system, animation or image data are searched for and selected from a shared consumer-generated database by using simple words. If users cannot locate the appropriate data, they can easily create new data by using an animation editor, and they can upload the new data by entering simple, related words. Once the user uploads a new animation of a particular character, its motion can also be applied to all the other characters through a process called motion retargeting [2], and thus, new animations are created by reusing the uploaded motion data.

Drawing from scratch is the traditional way of generating animation. For example, animation can be created in this fashion by using professional three dimensional computer graphics (3DCG) tools. Users require much effort, however, to master professional 3DCG tools, and because they are expensive, they are not for novice users. This limits their creativity because novice users need easier trial and error practice and learning for creating content.

Our concept is that by providing animation data corresponding to the natural language, we can enable novice users to create content. We believe that this is highly significant because of its potential for wide applications.

WordsAnime has rules of a common knowledge base for automatic scenario generation or scenario suggestion. The system automatically extracted the common knowledge of the IF-THEN rules via scenario generation. Cyc [3] is an artificial intelligence project that attempts to assemble comprehensive ontology and database of everyday common sense knowledge, with the goal of enabling AI applications to perform humanlike reasoning. Cyc needs an elaborate logical language that precisely specifies the meaning. Our system does not need description of strict knowledge by users. Simply creating stories (or describing easy rules) will collect knowledge of the scenarios. StoryNet [4] is a very large narrative-based knowledge database of story scripts that can be used for common sense reasoning. ComicKit [5] is an interface for acquiring a StoryNet script over the Internet. ComicKit can also extract rules of scenario automatically when the user is creating comics using arranged graphics.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

CHI 2009, April 4–9, 2009, Boston, MA, USA.

Copyright 2009 ACM 978-1-60558-246-7/08/04...\$5.00



Figure 1: Snapshot of “Introduction”.

The characteristic of our system is that it has both 3D animation consumer generated database based on words and common knowledge database based on simple subject, predicate, and object, etc. These databases can grow up by users and the data is based on words.

ANIMATION STORY GENERATION SYSTEM

WordsAnime leads users to input the story line through the method of paragraph writing. If a user inputs a scenario, an animation corresponding to the scenario appears on the display using animation database (Animebase) collected by Anime de Blog. WordsAnime displays GIF animation data for displaying on the screen from Animebase, stored on the Anime de Blog server. WordsAnime includes functions for story generation support, automatic scenario generation using rules, rule extraction and story browsing. It is based on Servlet/JSP and Ajax (JavaScript) and requires Windows XP and Internet Explorer 6.0 SP2 or Firefox 2.0.0.9.

Animation Story Generation by scenario inputting

The user generates the complete story by repeating the steps of inputting scenarios and viewing animations. WordsAnime uses a method of paragraph writing derived from a Finnish method [6] enabling users to easily create stories. The system leads a user to create scenes for “Introduction”, “Problems”, “Challenges”, “Problem Resolution” and “Conclusion”. The sequences from “Problems” to “Problem Resolution” and from “Problems” to “Challenges” are repeatable.

To create the complete story, the user accesses the WordsAnime URL, and the system shows the scene for “Introduction” on the scenario input page (Figure 1).

For each scene, the system leads users to input “Who?” and “Doing what?” as necessary input items. All the input items

can be selected by clicking alternatives in the scrolling list or typing the user’s own words.

As unnecessary items, the system provides “When?”, “Where?”, “What?”, “Whom?”

“When?”, and “Where?” corresponding to the background image, “What?” corresponding to an item, and “Whom?” corresponding to a character as an object of a sentence. These unnecessary items can be chosen by clicking the balloons in the menu, and they can be removed by re-clicking the balloons.

If the user types the “Who?” text or chooses from the scrolling list, the system shows the information dialogue window of the character, along with its thumbnail. In this window, the user can choose the kind of character, the height (tall/short), weight (heavy/light), strength (strong/weak), and mood (happy/gloomy). If the characters or background images include more than one data file, the system show the choices in a new window and leads the user to select one.

After the user selects the character, if the motions of the character corresponding to “Doing what?” include more than one data file, the system also shows the motion choices in a new window and leads the user to select. These choices are sorted by frequency of usage in Anime de Blog.

Next, if the characters corresponding to “Whom?” include more than one data file, the system again shows the choices in a new window and leads the user to select. Here, note that the character for “Whom?” is performing the passive mode action of “Doing what?”, because it is an object of the scenario.

Once the user has input all the input items, including all necessary items, the scenario is completed and it

corresponds to one sentence. When the user clicks the “Finish the sentence” button, the animation appears with overlapping animations and images according to the scenario.

If the user wants to connect additional scenarios in the same scene, he or she can connect the sentences by clicking “Why?”, “How?”, “Then what happened?” or “Then how do they think?” in place of a conjunction. Then, the user can input next scenario in the same scene.

The user can choose suitable music or sound effects in every scenario. The system has dozens of sounds, such as busy music, joyful music, sad music, footsteps, snoring, and so on. One story is completed by creating all the scenes from “Introduction” to “Conclusion”.

Rule Execution and Extraction

WordsAnime has rules of a common knowledge base for automatic scenario generation as follows.

Rule 19: IF Who “?a” Whom or What “?b” Doing what “call”
THEN Who “?b” Doing what “come”

Rule 19 signifies “If a calls b, then b comes”. These rules are shown as lists. Authoritative users can register rules which are shared by users via the rule control pages on the Web. As more rules are collected and refined, the function of automatic scenario generation becomes more intelligent.

These users can describe rules by writing IF and THEN sentences directly. An IF sentence represents the condition of the scene. A THEN sentence represents the condition of the next scene. These users can select scenes for rules, including “Introduction-Problem”, “Problem-Challenges”, “Challenges-Problem resolution”, and “Problem resolution-Conclusion”. “Introduction-Problem” signifies that the condition of the IF sentence occurs in the Introduction scene and the condition of the THEN sentence occurs in the Problem scene. The user can also input the character’s feature in a rule’s condition, such as height, weight, strength, and mood.

Each rule has a weight for conflict resolution, indicating how much the rule has been used. Rule candidates are shown sorted by their weights.

When the user creates a scenario and clicks the button “Finish the scene”, then some scenario candidates appear as hints for the user. The user can select one of the candidates or create his or her own scenario when creating a scene. If the user selects one scenario, then the corresponding animation appears. There are semi-automatic and fully automatic modes.

It also automatically extracted the IF–THEN rules in which the sentence of the scenario in the before scene as the IF rule and the sentence of the scenario in the after scene as the THEN rule were stored in the knowledge base at the request. When the rule is extracted for the first time, the weight value is set to the default. If the extracted rule is the

same as the rule in the rule base, the weight increases by one.

EXPERIMENT

The story generation system—WordsAnime—was evaluated by conducting twice of children’s workshops after analyzing the questionnaire data. At the first workshop, the examinees included 39 family and children groups including children from 5 to 14 years old. At the second workshop, the examinees included 68 family and children groups including children from 5 to 14 years old. One group uses one notebook PC. Because some of these children might have been using a PC for the first time, if a child required help in inputting, the adults were made to use the PC. First, the experimenters make the examinees think about characters and continuity of the story, followed by making them use the system. After they finish making the story, their work was displayed on the front screen. Finally, they completed the questionnaires. The examinees score the question according to five levels.

As a result, 96.2% of the children enjoyed and 84.6% would love to make another story. Further, 51.0% felt that making animation stories by words is easy. However, they did not always feel like making animation stories by words because 47.1% felt that it is difficult. The users pointed out that some animations corresponding to their words were not always provided by the system. It was considered that the users input the word through a trial and error process because the operations are provided only inputting or selecting words and they were not taught how to make animations using AnimeEditor. 66.3% felt that introducing continuity in the story is easy. However, 30.8% felt that this is difficult. It made us feel that introducing continuity in the story required training even when using the paragraph writing method. The provision of certain alternative scenarios as hints was helpful for 45.1% of the users. According to their comments, there were some children that did not need hints, whereas some children needed hints. This depended on the child. The feeling whether the rendered animation compared to the animation they imagined before using the system is different or same also depends on the child. We think increasing the number of characters, motions, and background pictures can help achieve animations similar to the user’s imagination. 80.1% of the respondents answered that children can use this system without problems. 67.3% of the respondents also answered that elderly people can use this system without problems.

We have received the comments such as “I’d love to create another one. I can create better than what I expected” and “It is easier than expected, and it is useful to make the imagination come true.”

A total of 54 stories were made at the first workshop and a total of 78 stories were made at the second workshop. The predicates are “eat,” “play,” “go,” “hear,” “come,” “think,” “laugh,” “say,” “look,” and “run” in the order of usage

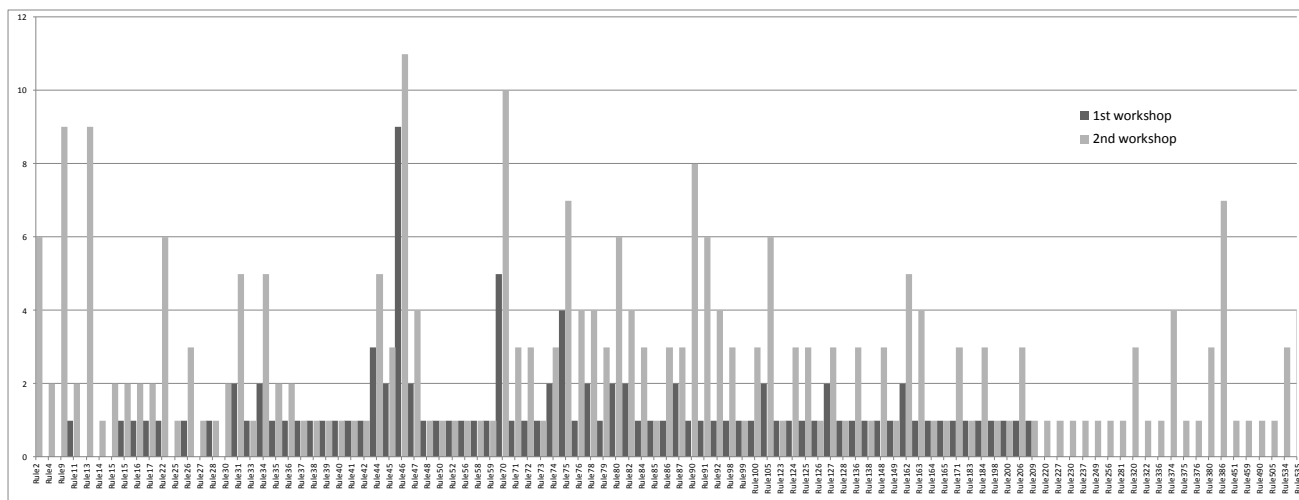


Figure 2: Extracted rules and its increased weights.

frequency. The words that indicate the behavior encountered several times by children in their daily lives are ranked higher. The words indicating emotions such as “laugh,” “angry,” and “cry” and playing such as “play ball,” “play football” and “dance” were also selected by the users.

Accordingly, 32 rules set as the default increased to 199 rules at the first workshop, and then increased to 207 rules at the second workshop. The rules that frequent use are such as IF someone “play” THEN someone “laugh” (Rule22 and 105), IF someone “eat” THEN someone “eat” (Rule46), IF someone “play” THEN someone “get rest” (Rule31 and Rule34), IF someone “play” THEN someone “play” (Rule70) , IF someone “go” THEN someone “go” (Rule75), and IF someone “run” THEN someone “go home” (Rule91) (Figure 2). For refinement of the rule bases, more user trials and better quality stories would be needed. However, the system stores and learns scenarios and rules as common sense knowledge automatically according to their usage. Since the rules of popular sequences are weighted and refined, these rules can be used as common sense knowledge.

DISCUSSION

We have introduced a story generation system called WordsAnime as an application of Animebase, enabling even children to create animation content by selecting words. We have shown that users can create animation content by combining animation materials (animations and images), and that the content can be searched by using keywords. According to the workshops, users enjoyed and wanted to try the application. Almost of the users create animation easily by words, but some felt difficult for inputting. Therefore, the system needs data enhancement such as characters, motions, background pictures, and knowledge rules, and easier interface.

The mental images that people have when they create something vary from person to person. In our system, users

can select parts of animation content and revise them through the reuse function. Although they cannot always create content that precisely matches their own imaginations, they can come close by using our approach and large animation databases.

According to the trials of rule extraction at the workshops, the rules of the system are increased. We will investigate the learning and refinement of the rules.

CONCLUSION

In this paper, we have introduced a story generation system, WordsAnime, as an application of the Animebase database. By using these systems, users can easily create animation content by simple familiar user interfaces involving inputting of words. This paper reports the attempt to collect rules of the common knowledge via WordsAnime.

REFERENCES

1. Kaoru Sumi: Anime Blog for collecting Animation Data, published in Mark Cavazaza and Stephane Donikian ed., Virtual Storytelling, LNKS 4871, Springer Lecture Note in Computer Science, pp.139-149, Springer (2007).
2. Michael Gleicher: Retargetting Motion to New Characters, in Proc. of SIGGRAPH '98, pp.33-42(1998).
3. Lenat, Douglas B. “CYC: A Large-Scale Investment in Knowledge Infrastructure.” Communications of the ACM, November 1995 (Vol.38, No.11).
4. P. Singh, B. Barry, and H. Liu, "Teaching Machines about Everyday Life," BT Technology J., vol. 22, no. 4, 2004, pp. 227–240.
5. Williams, R., Barry, B. and Singh, P.: ComicKit: acquiring story scripts using common sense feedback, Proceedings of the 10th international conference on Intelligent user interfaces: 302-304 (2005).
6. Koskipää, Ritva Töllinen, Markku Wäre, Mervi: Uusi salainen kerho 3, WSOY, Finland, 2004.