Web2Talkshow:Transforming Web Content into TV-program-like Content Based on the Creation of Dialogue

Akiyo Nadamoto
National Institute of
Information and
Communications Technology
Seikachyo, Kyoto, Japan
nadamoto@nict.go.jp

Masaki Hayashi
NHK Science & Technical
Research Laboratories
Kinuta, Setagaya, Tokyo,
Japan
hayashi.m-hk@nhk.or.jp

Katsumi Tanaka
Kyoto University
Yoshida Honmachi, Sakyo-ku,
Kyoto, Japan
ktanaka@i.kyoto-u.ac.jp

ABSTRACT

We propose a new browsing system called "Web2Talkshow". It transforms declarative-based web content into humorous dialog-based TV-program-like content that is presented through cartoon animation and synthesized speech. The system does this based on keywords in the original web content. Web2Talkshow enable users to get desired web content easily, pleasantly, and in a user-friendly way while being able to continue working on other tasks. Thus, using it will be much like watching TV.

Categories and Subject Descriptors: H.5.2[User Interface]:Prototyping I.7.m[Document and Text Processing]:Miscellaneous

**General Terms:** Design, Documentation

Keywords: Web Content, TV-program-like Content, Dialogue, Humor

1. INTRODUCTION

The current web browsing environment typically demands that users engage in active operations such as reading, scrolling, and clicking. We have to concentrate on these activities needed to reach sources of web content, and we cannot get to web content while also performing other tasks. Furthermore, it is difficult for many older people and young children, who cannot operate computers well, to get web content. In contrast, we can passively obtain information from TV by simply watching and listening. This ease of getting information means we can work on other tasks while obtaining information from TV. Furthermore, older people and young children can easily obtain information by watching and listening to TV. We believe it would be beneficial if users could get web content in a manner similar to watching TV. We are now developing a passive-user (audio-visual) browser that will provide web content without requiring the user to actively engage in a content search. We call this system Web2Talkshow. Our goal is that users will be able to get desired web content easily, pleasantly, and in a user-friendly way. Web2Talkshow transforms web content into TV-program-like content by using cartoon animation and synthesized speech, causing the web content to resemble a TV program. Furthermore, it automatically transforms declarative sentences on a web into dialogue sentences with humor. There are many researches about Dialogue analysis. Ishizaki et al.[1] is a good summary about in this area. Their approaches are analysis about real world dialogues and extract intentional from the dialogue. Our research, that transforms
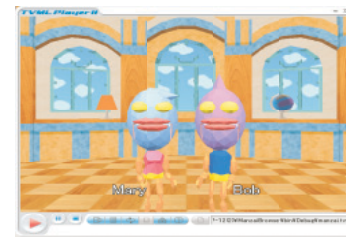
(a) Image Displayed by Web2Talkshow
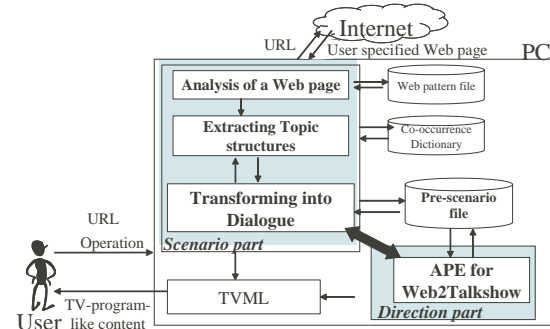
(b) Web2Talkshow System Overview

Figure1: System for Web2Talkshow

web content into dialogue-based TV-program like content, is opposite to their approaches. Moreover, we believe humor provides the easiest way to ensure that users of any age understand the substance of content. In Japan, we have a traditional form of comedy called "manzai". "Manzai" typically consists of two or three comedians participating in a humorous dialogue. It is a bit like "stand-up comedy" in the U.S., or "xiang shen" in China. Web2Talkshow uses manzai-metaphor.

Web2Talkshow consists of a scenario part and a direction part. In the scenario part, the system transforms into dialogue sentences based on keywords in the original web content. Our transformation method is only a partially automatic operation. We write a dialogue framework in XML, and create a scenario based on the original web content's story-flow and keywords along with this dialogue framework. In the direction part, we create the direction for each dialogue framework, and combine the directions based on the story. Figure 1 shows a typical Web2Talkshow display and system overview. Two cartoon characters engage in humorous dialogues, and act in a comedic fashion.

2. SCENARIO

2.1 Extracting Keywords

To extract the keywords of a page, we use a topic-structure model based on the method proposed by Matsukura et al.[2]. For a given page P, its topic t_i, $i \in \{1, \dots, n\}$ is simply represented as a pair consisting of *a subject term* s_i and a set C_i of *content terms*. C_i consists of multiple *content terms* c_{im}, $m \in \{1, \dots, k\}$. s_i is a noun and its term frequency is more than the threshold α. c_{im} in a given page is intuitively the term that has a high cooccurrence relationship with s_i in the page. A web page P may have more than one topic, and so, s_i is associated with multiple c_{im}.

2.2 Transforming Basic Dialogue

Long sentences are often used in web pages, especially news pages. Web2Talkshow divides compound sentences and complex sentences into short sentences. After dividing a long sentence, Web2Talkshow transforms the divided sentence into dialogue sentences based on the topic structure. We focus on two types of sentence, which is the subject term includes a sentence or not, and the transformation into dialogue is based on the sentence type.

1. A Sentence Including Subject Term(s)
We transform into a question-and-answer-based dialogue. We focus on a subject and a object in a sentence as follows:

<**Subject term is a subject in a sentence**>

* Content terms depend on the subject
The subject term and content terms are strongly related. We transform these sentences into a question-answer dialogue regarding the relationship between the subject term and content terms.

* Content terms are independent of the subject
The subject term is not strongly related to the content terms. We transform a sentence into dialogue about an answer which is a verb or object co-occurring with the subject terms in a sentence.

<**Subject term is an object in a sentence**>
The answer from the dialogue is a term that depends on the subject term.

2. Sentence Not Including a Subject Term
When a sentence includes a date or place, we expect these to be important to the sentence meaning. In this case, the system transforms the sentence into a when-type or where-type question.

When a sentence does not include a date or place, we presume the sentence does not contain important terms. The system also transforms into yes/no questions or tag questions.

2.3 Transforming into Dialogue with Humor

A humorous dialogue is often based on a strange or unexpected point of view regarding common situations. This can be expressed through exaggeration, deliberate mistakes, or misunderstandings.

Mistakes and misunderstandings
In our system, the topic structure of content terms co-occur with the subject term; that is, content terms are terms ordinarily used with the subject term. The system deliberately uses mistaken topic structure sets consisting of incorrect content term(s) and a subject term, to transform dialogue. Thus, it can transform into humorous dialogue based on mistakes.

Exaggeration
The first step to create dialogues using exaggeration is to use bigger numbers. When a sentence includes numbers, the system increases the numbers by a substantial factor.

2.4 Pre-scenario

We create dialogue frameworks in a XML pre-scenario file. The pre-scenario consists of structure tags, content tags, and direction

Table 1: Pre-scenario Direction Tags

setup	Pre-processing part. CG studio set, CG character, camera setup, lighting setup
d_intro	Introduction part. Background music, entrance of cartoons.
ending	Conclusion part. Background music, cartoons move from stage.
LookAtCamera	Viewpoint of cameras.
LookAtFellow	A cartoon character looks at another one.
hit	A cartoon character hits another one.
nod	A cartoon character nods.

tags. An example of a pre-scenario would be as follows:

$< question\, type = "1", num = "1" >$

$< line\, chara = 1 >$ Do you know \$S1? $< line >$

$< line\, chara = 2 >$ I know, \$C1 of \$S1. $< line >$

$< line\, chara = 1 >$ That's right! \$news. $< line >$

$< question >$

Here, 'type="1"' represents the type of sentence regarding the topic structure, and 'num="1"' represents the variation number for the question type. We can create many kinds of framework by creating many variations of each sentence. In this example, the question type is the case where content terms depend on the subject terms. \$S1 is a subject term, \$C1 is a content term, and \$news is a sentence in the web page without subject.

3. DIRECTIONS

In the TV-program-like content, directions are as important as the scenario because the TV-program-like content is very much like audio-visual content. Creating directions for each content item is troublesome, though, so we separate the directions from the content by using APE[3]. We describe here the APE we use to create the directions for Web2Talkshow. At present, the directions for cartoons depend on the lines in Web2Talkshow. We plan, however, to use XML tags written in the pre-scenario for Web2Talkshow. The direction tags of the pre-scenario will be transformed into cartoon animation by the APE for Web2TV. Table 1 shows an example of direction tags for Web2Talkshow. When users develop other APE and XML tags, they can create various other types of cartoon animation. The benefit of using the APE for Web2TV is that the system can easily transform either the same directions and different content or different directions and the same content into many types of TV-program-like content.

4. CONCLUSION

Web2Talkshow is designed to automatically transform web content into humorous dialog-based TV-program-like content. In this paper, as the first step towards automatically transforming dialogue, we examined dialogue transformation based on keywords in the web content. Our keyword-based dialogue is a surface-like dialogue. In our future work, we will consider the use of deeper dialogue as enabled by natural language technology.

Coding to change the references to 8 pt:

5. REFERENCES

[1] M.Ishizaki and Y.Den, "Computation and Language", Volume3:Discourse and Dialogue, University of Tokyo Press, 2001.

[2] T.Matsukura, H.Kondo, Y.Hirata, and K.Tanaka, "Discovery of semantic relationship among web pages based on web topic structures", Proc. of 9th IFIP 2.6 Working Conference on Database Semantics, 2001.

[3] M.Hayashi, M.Douke, and N.Hamaguchi, "Automatic TV Program Production with APEs", The 2nd International Conference on Creating, Connecting and Collaborating through Computer(C5), IEEE Press, pp20-25, 2004. Jan.