

**Emotion Expression Functions attached to
Multimodal Presentation Markup Language MPML
Yuan ZONG Hiroshi DOHI Mitsuru ISHIZUKA**

Department of Information and Communication Engineering
School of Engineering, University of Tokyo
E-mail:{yzong, dohi, ishizuka} @ miv.t.u-tokyo.ac.jp

Abstract

With the increase of multimedia contents in the WWW, multimodal presentation using interactive lifelike agents is attractive and becoming important. However, it is not easy for many people to write such multimodal presentations. In order to allow many people to write attractive multimodal presentations easily, MPML (Multimodal Presentation Markup Language) has been developed. In this paper, we present a new emotion function attached to MPML. With this function, we are able to express emotion-rich behaviors of the character agent in MPML. Some multimodal presentation contents are produced in the new version of MPML to show the effectiveness of the new emotion expression function.

Keyword: multimodal presentation, lifelike agent, emotion expression

1. Introduction

In a sense, an interface is a necessary evil. The ideal user interface would let us perform our tasks without being aware of the interface as the intermediary. The longevity and ubiquity of the now two-decades old graphical user interface should not mislead us into thinking that it is an ideal interface.

Among many possible Post-GUI interfaces, multimodal interface is supposed to be the most potential one. Multimodal interface uses the character agent as the middle layer between user and computer,

interacting with user and controlling the device. Character agent recognizes user's command and run the task as user requests. After the task is completed, character reports the result from verbal words or actions. With the favor of character agent, user can get information from many information channels (ex. verbal words, intonation, emotion, actions and so forth) .

One important implement of utilizing multimodal interface is multimodal presentation.

Presentation is the process of showing others achievement or approach. With the development of the multimedia technology, presentation technology developed. Centuries ago, people use text to make appeal. Because text only conveys information through single channel, it was not the effective presentation method. Recently people use presentation tools to make presentation. (ex. OHP, PowerPoint and so forth) As shown in Fig. 1, with the help of this method, images, movies, text, presenters' sounds, intonation convey different information to the audience through different channels. Audience integrates all information from different channels, then understand what presenter wants to say completely. Because this method conveys much information through different channels, it is effective and became the most popular presentation method at present.

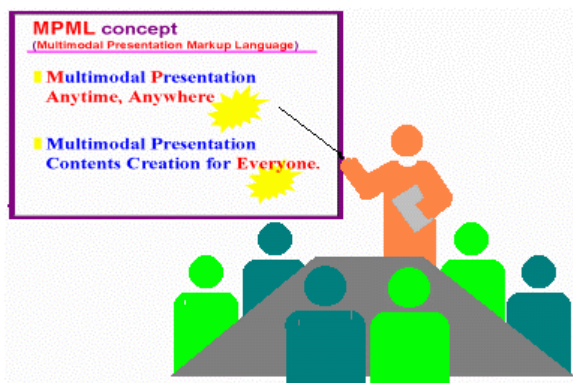


Fig. 1 Present Presentation

However, its disadvantage is that the presenter has to be at the meeting hall, which means time and place is restricted.

The solution is Multimodal presentation. It is a new presentation method to make presentation without the restriction of time and place. Fig. 2 illustrates this kind of presentation. The character agents make the presentation instead of human. You can download the presentation contents from WWW, then ask character agent make the presentation according the contents.

However, this attractive presentation method didn't replace the current popular PowerPoint presentation tools yet. The reason is that it is too difficult to write the multimodal presentation contents. There are many character agents, and different script language specifications are defined to control different character agents respectively. Most of these script language are individual (often low-level).

In order to overcome the complexity of describing various behaviors of character agents, and write attractive presentation contents easily, we developed MPML (Multimodal Presentation Markup Language).

2. MPML 1.0

The goal of MPML (Multimodal

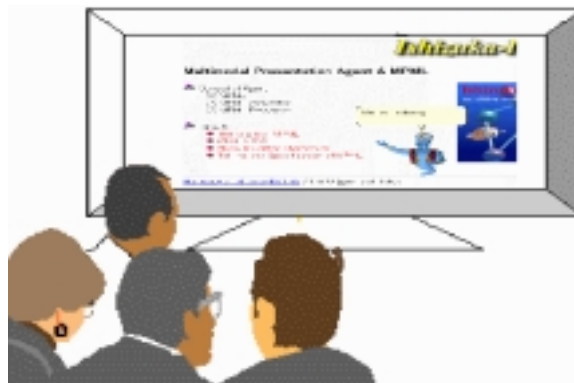


Fig. 2 Presentation Using Life-Like Agent

Presentation Markup Language) is to make everyone to write attractive multimodal presentation easily.

Most current multimodal presentation contents is written for the particular character system. In most cases, we have to write the detailed description to control the particular agent system.

We hope people can write multimodal presentation easily, just as people can write homepage easily using HTML. So MPML is designed to write multimodal presentation contents independent of the character agents.

Some feathers for MPML Version 1.0 are showed as below:

- *Independent of the Character agent system.*
- *Easy to describe*, anyone who understands HTML, should have no problem to learn MPML in short period.
- *Media synchronization supported*, because MPML conforms to SMIL.
- *Easy control to character*
- *Interactive presentation supported*

3. Emotion and MPML 2.0e

As the interface layer between the computer and user, character agent should have not only the communication functions, but also the personality which

lets user feel the affection. Audience will feel boring after communicate with the agent for a while, if character agent who has the face, body, however, can only perform machinelike reactions.

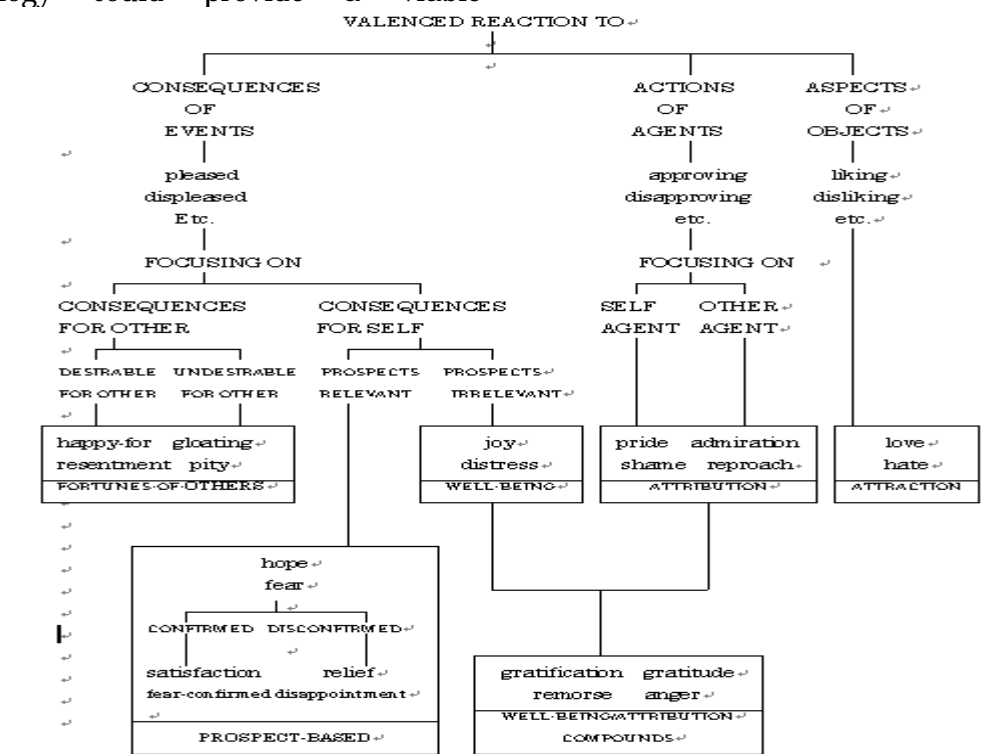
Considering how to attach the personality and sociality to character agent, we focus on the emotion expression functions.

Emotion can be expressed as *joy, sadness, anger, surprise, hate, fear* and so forth. The classification category is not defined clearly yet. So we focus to the related research about emotions in cognitive psychology domain.

Cognitive psychology was established in the 1970s, In 1981, Donald Norman identified the topic of emotion as one of twelve major challenges in cognitive science. It was at about this time that three authors decided to collaborate to explore the extent to which cognitive psychology could provide a viable

foundation for the analysis of emotions. Their names are Andrew Ortony, Gerald Clore, Allan Collins. In 1988, they published the book called *The Cognitive Structure of Emotion*, in which they explain the emotion system provided. We took the first letter of their names, and dubbed their emotion structure model as OCC model.

According OCC's theory, All emotions can be divided into terms according to the emotion-eliciting situation. Emotion-eliciting situations can be divided roughly into 3 types. The first type of emotion-eliciting situation is consequences of events. The second type of emotion-eliciting situation is actions of agents. The third type of emotion-eliciting situation is aspect of objects. According to the classification of emotion-eliciting situations, All emotions can be divided into 3 classes, 6 groups and 22 types of emotion.(Fig.3)



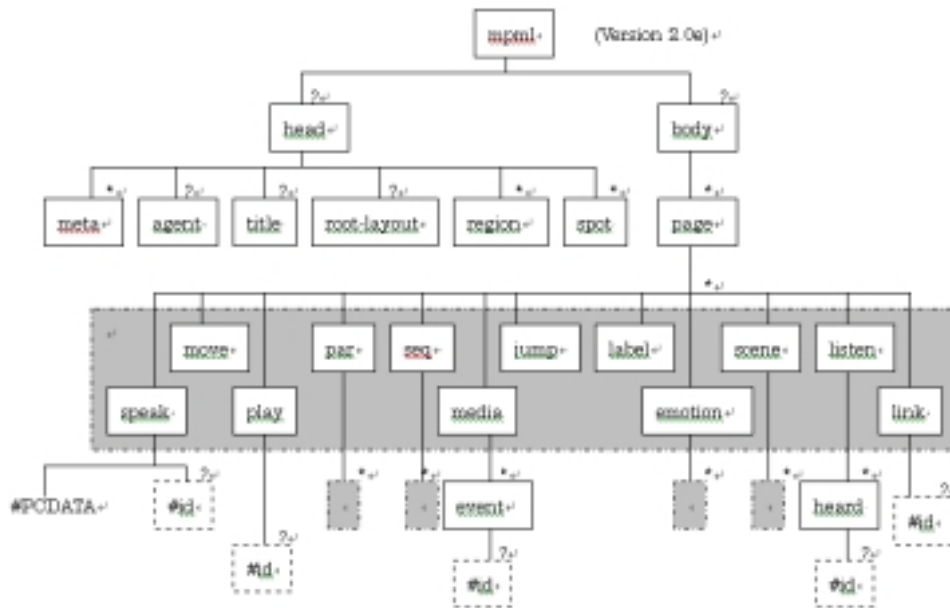


Fig. 4 Tag Structure for MPML 2.0e

In MPML Version 2.0e, we provide the emotion expression function to control agents' emotion more conveniently. Contents provider can specify 22 types of emotion defined in OCC emotion model, to modify the action performed by character agent. Character agent express the emotion with performing different action, changing the pitch, volume, speed, emphasis of the voice. For example, when the emotion type is specified as "pride", the character agent would wave his hands, then speak loudly with the emphasis at the beginning of the sentence.

Except for the emotion expression functions, some new functions are added in Version 2.0e:

- *Page*
Every presentation is divided into pages. Contents provider has to describe contents page by page.
- *Fast-forward*
Audience can request from page to next or previous page when watching the presentation.
- *Presentation-macro*
Some templates are prepared for particular presentation purpose.

Fig. 4 illustrates the tag structure for MPML Version 2.0e.

The below is a sample for MPML script:

```

<mpml>
<head>
  <title> MPML Presentation </title>
  <agent id="PD" character="peedy" >
</head>
<body>
  <page id="first" ref="self_intro.html">
    <emotion type="pride">
      <speak>
        My name is Zong Yuan,
        I am from Tokyo University.
      </speak>
    </emotion>
  </page>
</body>
</mpml>
  
```

4. MPML concerning tools

In order to be accepted by most people, authoring tools and audience tools should be provided for MPML.

As the authoring tools, 2 types are considerable. One is plain text editor, MPML is easy to learn and write, so it

should be easy to write with the plain text editor. The other authoring tools is visual editor. Just as people use Homepage Builder to built homepage, with the help of visual editor for MPML, people can built the multimodal presentation content without the knowledge of MPML. The visual editor for MPML is under construct.

Audience tools is also necessary for audience to watch the multimodal presentation. 3 types of audience tools are considerable and have been developed already. One type is the MPML player. One player called ViewMpml was developed for MPML 2.0e already. The second tool type is converter which converts MPML to the script that are understood by particular agent system. At present 2 kinds of converters are already developed for MPML 1.0(an old version of MPML). And the third tool type is XML browser with plug-in. Because MPML conforms to XML, it can be understood by XML browser. At present, one plug-in program written by XSL is already developed.

Fig. 5 displays ViewMpml, a MPML player developed for MPML Version 2.0e. It supports all tags defined in MPML Version2.0e's specification. It is free and can be download from site: <http://gem.miv.t.u-tokyo.ac.jp/MPML/en/2.0e/>

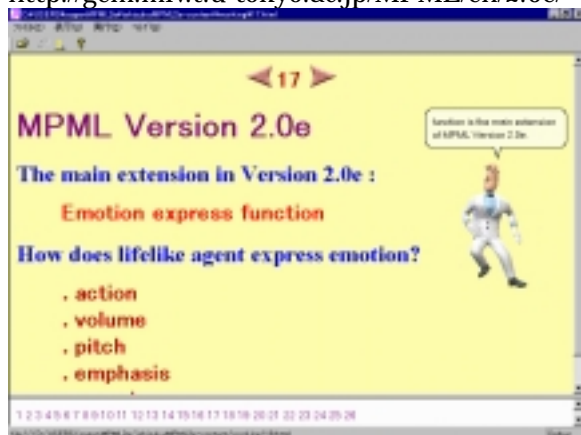


Fig. 5 MPML Player

5. Conclusion

The goal of MPML is to let most people to provide multimodal presentation contents easily.

In MPML Version 2.0e, we keep the features of Version 1.0 and applied some new functions to MPML. The main extension is emotion expression functions which conforms to OCC emotion models. For future potential change of the emotion models, this system didn't restrict emotion expression functions to OCC models. We can change the emotion model by changing the text setting files, which means we give enough scalability to emotion expression functions attached to MPML.

As the future work, because the current existing character agents were not designed for emotion expression, we are going to use the customized 3D character agent to express emotion more freely and naturally. Also, the eliciting-emotion situation's input is considerable.

References

1. A. Ortony, G. L. Clore and A. Collines: *The cognitive Structure of Emotions*, Cambridge Univ. Press, 1988.
2. *Proc. Workshop on Recognition, Analysis, and Tracking of Faces and Gestures in Real-Time Systems*, IEEE Computer Society Press, Los Alamitos, Calif. , 1999
3. M. Elson, "The Evolution of Digital Characters," *Computer Graphics World*, Vol. 22, No. 9, Sept. 1999, pp.23-24
4. F. Thomas and O. Johnson, *Disney Animation: The Illusion of Life*, Abbeville Press, New York, 1981
5. Nagao, K. and Takeuchi, A.: *Speech Dialogue with Facial Displays*, Multimodal Human-Computer Conversation 32nd Annual Conf. Of ASSOC of Computational Linguistics, 102-109, 1994.
6. <http://msdn.microsoft.com/workshop/imedia/agent/>
7. <http://www.miv.t.u-tokyo.ac.jp/MPML>
8. <http://www.w3.org/TR/REC-xml/>