Introducing Multimodal Character Agents into Existing Web Applications

Kimihito Ito
Meme Media Laboratory, Hokkaido University
Sapporo 060-8628, JAPAN
itok@meme.hokudai.ac.jp

ABSTRACT

This paper proposes a framework in which end-users can instantaneously modify existing Web applications by introducing multimodal user-interface. The authors use the IntelligentPad architecture and MPML as the basis of the framework. Example applications include character agents that read the latest news on a news Web site. The framework does not require users to write any program codes or scripts to introduce multimodal user-interface to existing Web applications.

Categories and Subject Descriptors

H.5.2 [**Information Systems**]: *User Interfaces*

General Terms

Human Factors

Keywords

Web application, IntelligentPad, Multimodal user interface, MPML

1. INTRODUCTION

Web applications, which are computer programs ported to the Web, allow us to use various remote services through our Web browsers. There are a huge number of Web applications on the Web, and they are becoming the basic infrastructure of everyday life. Meanwhile, multimodal character agents [4], which interact with human users through both verbal and nonverbal behaviors, are also well-developed recently. The use of multimodal character agents is one of the hot topics in the Web development communities [1].

Despite their fundamental potential to present information to users, multimodal character agents are not used in many Web applications. There were two problems that need to be solved:

1. the lack of a scripting scheme to support authoring of multi-modal presentations in a reasonable time, and
2. the lack of a framework to change the user-interface of an existing Web application into another form.

WWW 2005, May 10–14, 2005, Chiba, Japan.
ACM 1-59593-051-5/05/0005.

Recently, M. Ishizuka's group in University of Tokyo designed a markup language called MPML (Multimodal Presentation Markup Language), and succeeded to solve the first problem[2]. IntelligentPad architecture[6][5], which was proposed by Y. Tanaka in Hokkaido University in 1989, enables application programs to be functionally linked to other programs, and gives a solution to the second problem[3].

This paper proposes a framework where even end-users can introduce multimodal user-interface with character agents to existing Web applications (Figure 1). We focus on how instantaneously end-users can perform this process of extending existing Web applications.

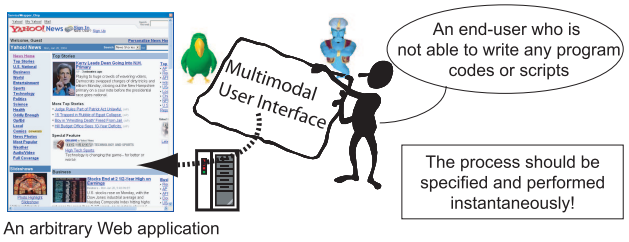

Figure 1: The goal of this research

2. INTELLIGENTPAD AND MPML

We employ the IntelligentPad architecture [6] to reuse the functions of existing Web applications.

IntelligentPad is a meme media system [5] that allows users to functionally combine media objects (called *pad*s) through direct manipulation of them. Each pad has data I/O ports called *slot*s to input/output data to/from the pad. Through the use of IntelligentPad, users may easily change the user-interface of application programs (including Web applications[3]) by connecting their data slot to a slot in another application.

We also employ MPML(Multimodal Presentation Markup Language)[2] to give users multimodal presentations with character agents. MPML has been developed by M. Ishizuka, in order to provide a meta-level description language commonly applicable to various multimodal character systems. MPML provides a simple scripting scheme for users to create a multimodal presentation in the XML form.

3. FRAMEWORK

For the collaboration of IntelligentPad and MPML, we introduce two pads that allow users to interact with Web

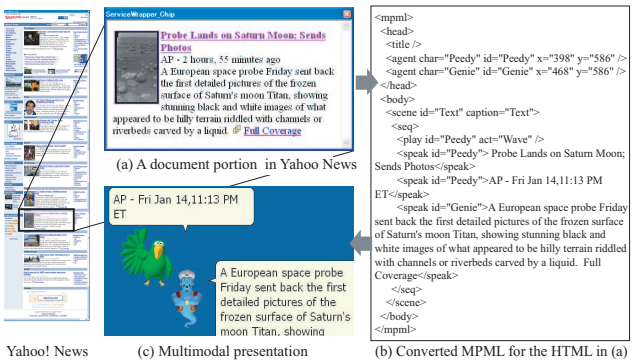

Figure 2: Conversion from HTML to MPML

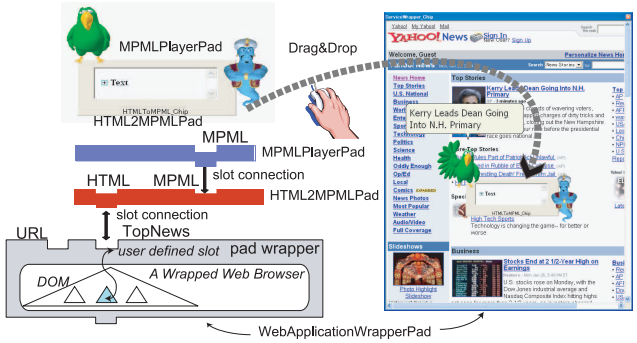

Figure 3: A sample application of the framework

applications through multimodal character agents. These pads are 'MPMLPlayerPad' and 'HTML2MPMLPad'.

An MPMLPlayerPad gives a multimodal presentation, according to the given script written in MPML. This pad provides a slot named 'MPML' as a place holder of MPML script. When the MPML script stored in the 'MPML' slot has been changed into a new one, the pad plays the presentation, according to the definition of the presentation written in MPML format.

An HTML2MPMLPad translates text string given in HTML format into MPML format. The HTML string to be translated may be a subtree of the Web page represented as a DOM tree. An HTML2MPMLPad provides two slots named 'HTML' and 'MPML'. When the text string in the 'HTML' slot has been updated, the pad generates an MPML from the HTML. The conversion is an XSL-translation from HTML to MPML.

In a possible implementation of this pad, the translation may generate a presentation in which two agents read the given texts alternatively, in the same way as two casters reads a news in TV news programs. The conversion algorithm divides the given HTML text into two parts by detecting the paragraph changes. We use a set of tags to detect paragraph changes. This set of tags includes ⟨br⟩, ⟨p⟩, ⟨div⟩, ⟨h1⟩,. . ., ⟨h6⟩, and ⟨tr⟩. Then it alternatively assigns one agent to each new paragraph.

Figure 2 shows an example conversion from HTML to MPML. The document (a) is a news summary in Yahoo! News. The MPML code (b) is generated from HTML code of (a) by HTML2MPMLPad. Figure (c) shows the multimodal presentation defined in the MPML (b).

Figure 3 shows an example application in which two character agents read news posted on an external news site. The user has defined a slot named 'TopNews' with the WebApplicationWrapperPad [3] at the bottom. This slot is specified by the user through mouse operations, and it holds the latest top news on the Yahoo! news. The HTML slot in the WebApplicationWrapperPad is connected to the HTML slot in the HTML2MPMLPad, in which MPML slot is connected to the MPML slot in the MPMLPlayerPad. These slot connections are all instantaneously defined by users through the direct manipulation of pads. [1]

Whenever the top news is updated, the character agents automatically start to read the latest news. The update is propagated from WebApplicationWrapperPad to HTML2MPMLPad and MPMLPlayerPad through the slot connection mechanism in the IntelligentPad architecture.

This framework does not require users to write any programming code either for wrapping an existing Web application or for attaching multimodal character agents to the wrapped Web application.

In case other types of behaviour of agents are required, we must modify the XSL-translation. Supporting users to modify XSL-translation remains for future work.

[1] A video demonstration is available at http://km.meme.hokudai.ac.jp/people/itok/CHIP/movies/

4. CONCLUDING REMARKS

In this paper, we have proposed a meme media framework that enables end-users to instantaneously introduce a multimodal user-interface to existing Web applications.

Recently many news portal sites use RSS(RDF Site Summary) to provide summary of their news. Our framework can be directly applicable to such RSS for multimodal agents to present the latest news with verbal and non-verbal behaviors. This would be a practical application of our framework.

Acknowledgments

I would like to express sincere thanks to Prof. M. Ishizuka and his MPML development team at University of Tokyo. I had very fruitful discussions with them during my visit to their laboratory. I also wish to express thanks to Prof. Y. Tanaka in Hokkaido University, for his kind help and appropriate advice on this research.

5. REFERENCES

[1] E. André and T. Rist. From adaptive hypertext to personalized web companions. *Commun. ACM*, 45(5):43–46, 2002.

[2] M. Ishizuka, T. Tsutsui, S. Saeyor, H. Dohi, Y. Zong, and H. Prendinger. MPML: A multimodal presentation markup language with character control functions. In *Proc. of Agents'2000 Workshop on Achieving Human-like Behavior in Interactive Animated Agents*, pages 50–54, 2000.

[3] K. Ito and Y. Tanaka. A visual environment for web application composition. In *Proc. of 14th ACM Conference on Hypertext and Hypermedia*, pages 184–193, 2003.

[4] H. Prendinger and M. Ishizuka, editors. *Life-Like Characters*. Springer-Verlag, 2003.

[5] Y. Tanaka. *Meme Media and Meme Market Architectures:Knowledge Media for Editing, Distributing, and Managing Intellectual Resources*. IEEE Press, John Wiley& Sons, 2003.

[6] Y. Tanaka and T. Imataki. Intelligentpad: A hypermedia system allowing functional composition of active media objects through direct manipulations. In *Proc. of IFIP'89*, pages 541–546, 1989.