Multi-Modal Seamless Messaging

John F. Meech and Suhayya Abu-Hakima

Seamless Personal Information Networking Group,
Institute for Information Technology,
National Research Council,
Building M50, Montreal Road, Ottawa, Canada.
{John.Meech,Suhayya.Abu-Hakima}@iit.nrc.ca

Abstract

This paper provides an overview of a framework for multi-modal interface design in the field of Seamless Messaging™[1]. The objective of seamless messaging is to intelligently tailor message handling based on user preferences and device capabilities. Messages can include voice mail, email, fax, video-mail and pager messages. Incoming messages are dealt with in a manner independent of the user's physical location. Because of the varying message modalities and the differing capabilities of the access devices the system must be capable of representing numerous modalities and intelligently converting messages of one modality to another. This paper presents an agent-based framework that is currently under development that can provide this functionality.

Introduction

Seamless Messaging™ is a concept in which any modality of message may be forwarded to a user automatically, taking into account user preferences, access device capability and network capability (e.g. permissions and bandwidth). The objective is to provide the user with a seamless personal information network, in which messages ('phone calls, faxes, emails, etc.) can be received, created, stored and delivered independent of physical location. Thus, messages are delivered based on the preferences of the user and the capabilities of their access devices.

The message types in such a system are necessarily multi-modal, including sound, speech, text, static images and dynamic graphics. The access devices include desktop PC's, mobile PC's, PDA's, cellular phones, pagers, etc. In addition to the modalities contained within the messages, the access devices can provide haptic and tactile modalities for interacting with the devices (cellular phones and pagers that vibrate being obvious examples).

The challenge in such a system is to manage the presentation of these modalities in an intelligent manner. Essentially this message management revolves around converting one message type to another in order to display the message type on the user's selected access device. This might include performing key phrase extraction on an email message in order to display it on a pager, or performing optical character recognition on a fax in order to display it on a digital cellular phone. In addition the user's preferences might automatically switch a pager from audible to vibrate when the user's calendar indicates a meeting is underway.

[1] Seamless Messaging and SPIN are registered trademarks of NRC.

Intelligent Multi-Modal Messaging

As part of current research we are developing a Seamless Personal Information Networking (SPIN) testbed based on cooperating agents. The SPIN testbed is intended to serve as a development and prototyping environment for seamless messaging and intelligent network management. A prototype seamless messaging system has been operating as of 1996 and is currently undergoing redevelopment using an agent-based software paradigm. The first prototype is essentially based on a blackboard paradigm where various knowledge sources act on a message after it is posted. The second generation system is decoupling the knowledge sources and providing them added functionality in the agent paradigm. The ability to provide seamless messaging implies that the system has the ability to change the content and modality of messages in a way which exhibits some degree of intelligence or adaptivity in structuring both the data and the interface to that data (Meech 1994).

Agent-Based Architecture

Agents can be viewed as software entities that can act autonomously (or with some guidance) on the user or system's behalf. They represent active computational entities that are persistent, perceptive, can reason and act within their environments and can communicate with other agents (including users) (Abu-Hakima et al., 1998). They are particularly well suited for applications that require some form of distributed intelligent cooperation, such as that necessary in telecommunications networks.

Here we are focussing particularly on agents which interact with the user at the user interface and consequently we adopt a general definition (after Lieberman, 1997) of an

Interface Agent as *a program that can affect the interface without explicit instruction from the user.*

The function of the interface agent in seamless messaging may be thought of as reducing the work and information overload of the user (Maes, 1994). Essentially it is an agent which automates based on a set of known user preferences, converting messages to a format which the preferred user device can receive in the current operating context, without user intervention. In adapting interaction to user preferences, a seamless messaging agent has some features in common with other agent implementations, but is not visible to the user as an agent, bypassing many of the possible interaction problems (Norman, 1994). Essentially such an agent provides the radical tailorability proposed by Malone et al (1997).

Framework

A high-level agent framework of the system currently under development is shown in figure 1.

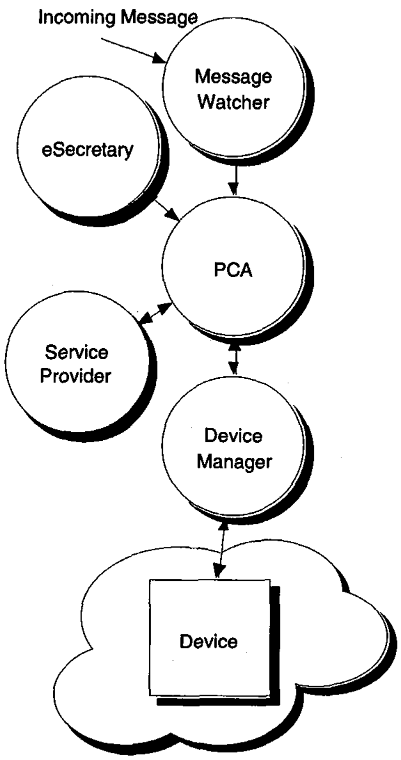

Figure 1. A high-level view of the agent design.

The operation of the system is based around the concept of a Personal Communications Agent™[2] (PCA™) which coordinates message handling, including the interface management. The PCA contains a set of stored user preferences in the form of rules to classify messages, actions to take in given circumstances and a calendar of user activity. The operation of the architecture is as follows:

1. The Message Watcher detects an incoming message and passes it to the PCA.
2. The PCA classifies the message based on the Classification Rules encapsulated in the User Preferences. Here, the message can be classified as "urgent" or "non-urgent" based on a keywords and header information.
3. The PCA then checks the Action Rules that are also encapsulated in the User Preferences. If the action the PCA needs to take is "contact me" versus "forward to my Unified Message Box", the PCA needs to find the user.
4. The PCA uses the eSecretary Agent™ to locate the user.
5. The PCA obtains stored User Preferences on being contacted based on the current context (i.e., the user is in a meeting, the user cannot be interrupted, etc.) and reasons about what it needs to do next in order to deliver the message.
6. If necessary (e.g. the user preferences dictate it) the PCA can use a Service Provider agent to change the modality of a message.
7. The PCA then contacts the appropriate Device Manager Agent™ which determines the required means to deliver the message to the specified device.
8. The Device Manager negotiates with the device (possibly by moving through the network to the device) in order to determine actual device settings (e.g., ringer is off).
9. The PCA may use the Service Provider agents to further change message modality.
10. The message is then delivered.

The system handles synchronous and asynchronous messages (e.g. telephone calls, emails) and can provide indication in one modality that a message has arrived in another modality (e.g. an email message that a fax has arrived). In some message types which contain more than one modality (video mail, email, etc.) it may be necessary to change (or remove) one of the modalities in order to allow the message to be displayed on a particular device, or over restricted network bandwidth.

Scenario

To illustrate the concept of Seamless Messaging, let us consider an example scenario. Assume two users: Tom and Jane. Tom urgently emails Jane a 3 page multimedia message which includes some text and an annotated attachment.

1. The Message Watcher detects an incoming message and passes it to Jane's PCA.

---

[2] Personal Communication Agent, PCA, eSecretary Agent, and Device Manager Agents are registered trademarks of NRC.

2. Jane's PCA calls on some of the Service Provider agents of the system to help interpret the message. The header information is examined indicating that the message is from Tom. Tom is Jane's boss. The text in the front of the message is condensed into a set of key phrases. The classification rules are used to check if the message can be classified based on the key phrases as urgent or not. The message is classified as urgent and the key phrase "business plan and opportunity" are captured as the most representative.
3. Jane's PCA checks her Action Rules in her User Preferences. According to Jane's rules, if a message is received from her boss (Tom), her PCA is to contact her.
4. Jane's PCA asks the eSecretary Agent™ to locate Jane. Through Jane's electronic calendar, the eSecretary determines that she may be at a client site ABC Corp. at a meeting.
5. Jane's PCA checks her User Preferences on being contacted based on the current context. Jane has indicated that she cannot be noisily interrupted at ABC Corp. (i.e., implying that her cellular phone is likely to be turned off).
6. Jane's PCA asks a Service Provider agent to change the modality of the abbreviated urgent message. The text is converted to 30 seconds of voice (as Jane's User Preferences indicate) and this voice mail indicates to Jane that she has an urgent message from Tom and tells her the key phrases she received.
7. Jane's PCA then contacts the cellular phone Device Manager Agent™ to deliver the message.
8. The Device Manager negotiates with Jane's device through the network to make sure the phone is at least in message notification mode since the ringer is off.
9. The voice mail is then delivered to Jane's cellular phone as a 30 second condensed voice clip. She is also told she can access the original message through a remote login or she can have it faxed to her at a secured number at ABC Corp. (again, through her User Preferences and Personal Directories she had specified a contact address at ABC Corp.).
10. Jane replies to Tom's message through voice mail.

From this example it can be seen that the architecture of the system is designed to handle the representation of multi-modal messages for both human-human communication (where messages originate from and are transmitted to a person). This messaging may also take place between a human user and a computer, as in the case of database query in the field of sales automation. As such the capacity of the system to intelligently transform and filter modalities is a key capability in maintaining a messaging capability, particularly for mobile users. In our system this capability is provided using an agent-architecture which acts on messages using user preferences and device capabilities to manage the modalities. This agent-based framework allows for delivery of multi-modal messages, but also provides for the intelligent transform and combination of modalities to form new messages. This transformation of message modality may be thought of as a form of multi-modal dialogue management between multiple devices, and therefore fits into the wider arena of multi-modal human-computer interaction in a more general sense.

Summary

We have presented a framework and a scenario for representing and managing multi-modal interaction in a seamless messaging application. Part of this framework has been realized in a prototype Seamless Messaging system that provided text-to-speech and various text messaging transforms. Research is currently focussed on the second generation Seamless Messaging system which makes better use of the agent metaphor, and allows a greater range of message and device modalities to be managed (OCR-speech, voice-to-text, MIME-based messaging, etc.) We expect to have a system that is fully demonstrable in early 1999.

References

Abu-Hakima, S, Liscano, R., Impey, R. (1998) A Common Multi-Agent Testbed for Diverse Seamless Personal Information Networking Applications. To appear in *IEEE Communications Magazine*, July 1998.

Lieberman, H. (1997) Autonomous Interface Agents Proceedings of CHI'97. ACM press

Maes, P. 1994 Agents that reduce work and information overload *Communications of the ACM* Vol.37, No.7. July 1994

Malone, T.W., Grant, K.R., Lai, K-Y. (1997) Agents for Information Sharing and Coordination: A History and Some Reflections in *Software Agents*, J.M. Bradshaw (ed) AAAI Press/MIT Press

Meech, J.F. 1994. The Intelligent Management of Information for Human Interaction with Complex, Real-Time Systems. In Proceedings of AAAI Spring Symposium on Intelligent Multi-Media Multi-Modal Systems. Stanford University, USA. March 21-23. AAAI Press

Norman, D.A. 1994 How might people interact with agents? *Communications of the ACM*, Vol.37, No.7. July 1994