The Smart Personal Assistant: An Overview

Wayne Wobcke, Anh Nguyen, Van Ho and Alfred Krzywicki
School of Computer Science and Engineering
University of New South Wales
Sydney NSW 2052, Australia
{wobcke|anht|vanho|alfredk}@cse.unsw.edu.au

Abstract

The Smart Personal Assistant (SPA) enables users to access e-mail and calendar information using natural language dialogue through a PDA platform. The user interface to the SPA must present the system as a single unified set of back-end task assistants, enabling the user to conduct a dialogue in which it is easy to switch between these domains. The SPA is implemented using an agent platform and includes a special BDI Coordinator agent with plans both for coordinating the actions of the task assistants and for encoding the system's dialogue model. The agent-based dialogue model is at a high level of abstraction, enabling the domain-independent plans in the dialogue model to be reused in different SPA systems.

Introduction

Our research is focused on natural language interaction with personal assistant systems for use on mobile devices such as PDAs and mobile phones. Our current Smart Personal Assistant (SPA) is a personal information management assistant that provides users with integrated access to e-mail and calendar information. The e-mail assistant, EMMA, described in Ho, Wobcke and Compton (2003), enables users to easily define a rule base that enables messages to be associated with a sorting folder for display, a priority and/or an action to be taken. The calendar assistant enables users to define rules for suggesting various attributes of appointments. The SPA must present a unified interface to users so that they can switch seamlessly between the different domains. In addition, though speech recognition systems have greatly improved in recent years, performance is far from perfect, so the dialogue model must be built to include mechanisms for recovery from speech recognition errors.

SPA Agent Architecture

The SPA is implemented as a multi-agent system using JACK Intelligent Agents™, as shown in Figure 1, and includes "wrapper" agents for the e-mail and calendar task assistants, a User Interaction Agent, and a special Coordinator agent that mediates communication between the user and the specialist task assistants, similar to its use in the Intelligent Assistant, Azvine *et al.* (2000). The Coordinator is built using a BDI (Belief, Desire, Intention) agent architecture in which both dialogue management and coordination of the task assistants are encoded in the agent's plans.

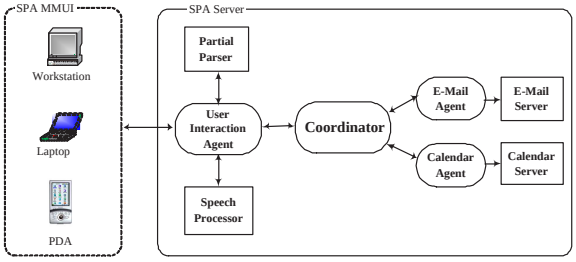

Figure 1: SPA Agent Architecture

The role of the Coordinator, further described in Wobcke *et al.* (2005), is to present a single point of contact for the user to interact with the SPA, to maintain the dialogue context (both information about the physical context of the user and the dialogue history), to delegate tasks to the e-mail and calendar assistants to fulfil user requests, to notify the user of any important events, and to learn the user's preferences for interaction on particular device types.

The current version of the SPA uses Dragon NaturallySpeaking in dictation mode for speech recognition, and Lernout and Hauspie TTS for speech synthesis. The User Interaction Agent uses the ProBot scripting language of Sammut (2001) for computing a partial syntactic analysis of the speech input. Due to the modularity of the architecture, it is straightforward to replace both the speech processor and partial parser by other equivalent systems.

Dialogue Management and Coordination

A BDI agent architecture based on the PRS system, as implemented in JACK, has been employed for the development of the SPA's Coordinator agent. The dialogue model is encoded in the plans of the agent, as described in Nguyen and Wobcke (2005). There is a separation between domain-independent dialogue plans, handling discourse-level goals such as recognizing the user's intention, and domain-dependent plans, handling domain-level dialogue aspects and task delegation to the back-end assistants. The discourse-level plans are generic, though make

use of domain-specific knowledge. The modularity of this approach enables the reuse of such discourse-level plans in other SPA-type applications.

Figure 2 shows the overall structure of our agent-based dialogue model, in which the Coordinator's plans are roughly arranged into four groups according to their purpose: *semantic analysis*, *pragmatic analysis*, *task processing* and *response and clarification generation*. Each group itself contains several plans. The system contains roughly 40 plans, including 20 discourse-level plans and six domain-specific plans for each task domain (plus some auxiliary plans). The main point is that dialogue processing is performed automatically as the result of the BDI interpreter selecting and executing plans according to the current context.

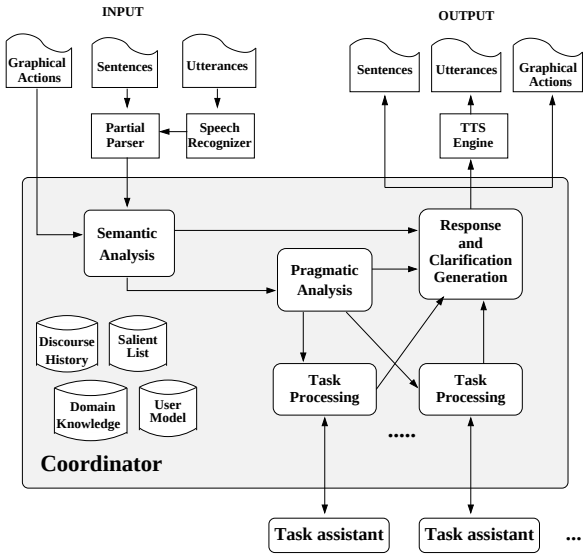

Figure 2: Agent-Based Dialogue Model

The Coordinator maintains the dialogue model, including the conversational context and other domain-specific knowledge as its internal beliefs, as follows:

- *Discourse History*: for maintaining the conversational context such as information about the current and past dialogue states.
- *Salient List*: for maintaining a list of objects which have been mentioned previously in the conversation, i.e. the objects that are in the focus of attention.
- *Domain-Specific Knowledge*: includes domain-specific vocabulary and information of the tasks that are supported, used in interpreting the user's requests.
- *User Model*: for maintaining information about the user such as current device, preferred modality of interaction, physical context, preferences, etc.

As we have shown in Nguyen and Wobcke (2006), learning can be incorporated into the plan selection process of the Coordinator so that the agent can select the most suitable plan amongst those applicable, enabling the SPA to tailor its responses according to the conversational context and the user's physical context, device and preferences.

Scenario

The following shows a simple scenario in which the user checks e-mail and creates an appointment using speech interaction over a PDA. Actions of the SPA corresponding to points in the dialogue are written in italics.[1]

```
User  Is there any new mail from John?
SPA   You have two new messages from John Lloyd.
User  Show me the one about slides please.
SPA   Displays correct message from John Lloyd.
User  I need to see him at 5pm tomorrow about the slides.
SPA   Do you want to enter that appointment to calendar?
User  Yes, enter it to the Research category.
SPA   Where are you going to meet him?
User  My office.
SPA   Appointment has been created.
      Displays calendar showing new appointment.
```

Conclusion

The Smart Personal Assistant provides natural language interaction tailored to the user's device, and coordination of a range of specialist task assistants. The architecture is a BDI architecture with plans for coordination and dialogue actions, and point-to-point communication between agents. The agent-based approach provides a high degree of modularity, making the domain-independent aspects of the dialogue model reusable across applications. The agent-based approach also makes it possible to integrate learning into the system so that the user's preferences for interaction can be acquired automatically.

References

Azvine, B., Djian, D., Tsui, K. C. and Wobcke, W. R. 2000. The Intelligent Assistant: An Overview. In Azvine, B., Azarmi, N. and Nauck, D. D., eds., *Intelligent Systems and Soft Computing*. Berlin: Springer-Verlag.

Ho, V. H., Wobcke, W. R. and Compton, P. J. 2003. EMMA: An E-Mail Management Assistant. In *Proceedings of the 2003 IEEE/WIC International Conference on Intelligent Agent Technology*, 67–74.

Nguyen, A. and Wobcke, W. R. 2005. An Agent-Based Approach to Dialogue Management in Personal Assistants. In *Proceedings of the 2005 International Conference on Intelligent User Interfaces*, 137–144.

Nguyen, A. and Wobcke, W. R. 2006. An Adaptive Plan-Based Dialogue Agent: Integrating Learning into a BDI Architecture. In *Proceedings of the Fifth International Joint Conference on Autonomous Agents and Multiagent Systems*, 786–788.

Sammut, C. A. 2001. Managing Context in a Conversational Agent. *Electronic Transactions on Artificial Intelligence* 5(B):189–202.

Wobcke, W. R., Ho, V. H., Nguyen, A. and Krzywicki, A. 2005. A BDI Agent Architecture for Dialogue Modelling and Coordination in a Smart Personal Assistant. In *Proceedings of the 2005 IEEE/WIC/ACM International Conference on Intelligent Agent Technology*, 323–329.

[1] A more sophisticated scenario is shown in the video at `http://www.cse.unsw.edu.au/~wobcke/spa.mov`.