PeopleFinder: A Multimodal Multimedia Communications Tool for Interconnecting Office Staff
Ferguson, Innes; Davlouros, J.D.

Publisher's version / Version de l'éditeur:

National Research Council Canada   Conseil national de recherches Canada

Canada

PeopleFinder: a Multimodal Multimedia Communications Tool for Interconnecting Office Staff

Innes A. Ferguson and James D. Davlouros
Knowledge Systems Laboratory (IIT)
National Research Council
Ottawa, ON, Canada
innes@ai.iit.nrc.ca; jams@ais.dsmsp-agency.ca

Abstract

The PeopleFinder is a knowledge-based tool to assist users in determining the whereabouts of other staff located in an office or network environment. The tool makes use of several modes of input and output, as well as employing a number of interface and communications media with which to present information and interconnect geographically distributed system users. The accompanying video contains example uses of the tool which help illustrate some of its functionality.

PeopleFinder

The PeopleFinder employs agent-oriented design techniques as a way of integrating a variety of heterogeneous applications running on a number of different computing platforms and operating systems. Specifically, the PeopleFinder's functionality is carried out by a collection of coordinated software agents which can

1. assume user-delegated tasks (e.g. interpreting spoken commands, invoking email and voice recording tools, dialing out on a phone);

2. perform a number of application level tasks (e.g. updating the graphical user interface, applying heuristics to decide whether and/or how system users want to be contacted); and

3. carry out a number of lower level operating system level activities, the majority of which will be beyond the level of expertise of most casual system users (e.g. translating between different speech encoding formats, establishing which users are logged on to the various multi-platform computer networks throughout the building).

The agents used in the PeopleFinder application are based on the *CALVIN* open agent framework (Ferguson and Davlouros 1995) which in turn is an extension of the TuringMachine architecture (Ferguson 1992). [1] This framework provides application developers with a powerful set of agent programming tools

including libraries of intra- and inter-agent protocols (e.g. KQML [2]), sensory and effectory apparatus, internal behavior APIs, persistent storage management, and (currently under consideration) CORBA compliance.

Besides providing an added level of flexibility and robustness to the overall system behavior (Ferguson 1995), the existence of multiple autonomous – and, therefore, concurrently operating – agents also facilitates the application of multiple channels of interaction between human and system (Faure and Julia 1994). A number of similar benefits and issues have been identified in the application of the Open Agent Architecture (Cohen *et al.* 1994) to the management of email; in particular, the roles of such agent skills as *delegation* ("the ability to receive a task to be performed without the user's having to state all the details"), *data-directed execution* ("the ability to monitor local or remote events, such as database updates, OS, or network activities"), and *communication* ("the ability to enlist other agents ... in order to accomplish a task"). In addition, the PeopleFinder can be seen to offer a number of desirable features which are characteristic of technology for *telepresence*; that is, technology which provides, despite geographical or temporal distance, a sense of social proximity – a kind of *social prosthesis* for overcoming gaps and weaknesses with an organizational structure (Giachino 1993; Buxton 1994).

Briefly, some of the key design features of the PeopleFinder include:

- **Agent-oriented design** The system is based on a number of coordinated autonomous agents, each specialized in one or more system tasks, a number of which have been mentioned above.

- **Multimodal interaction** The system makes use of various input and output modalities for human-

---

[2] The Knowledge Query and Manipulation Language (KQML) is a protocol intended to support interoperability among intelligent agents in distributed applications (Finin *et al.* 1992; Werkman 1994)

[1] The *Communicating Agents Living Vicariously In among Networks (CALVIN)* architecture is an agent fram

computer interaction; in particular, the keyboard, mouse, and speech for input, and audio and screen-based feedback for output.

- **Multimedia presentation** The system makes use of a number of different media for human-computer interaction; in particular text, graphics, animation, and prerecorded video (the latter acting as a "cheap", but nevertheless very effective, substitute for in-office video cameras and/or active badges [3])

- **Communications-oriented capabilities.** The system enables transparent communication across different computer platforms (Macintosh, Unix) and facilitates the interconnection of system users via telephone, email, and voice messaging.

A number of features of the *CALVIN* architecture have proven useful for developing multimodal applications that integrate a number of distributed media resources. In particular, rapid responses to users' commands are facilitated through integration of appropriate reactive behaviors in the system's Interface and User agents (Ferguson and Davlouros 1995); in addition, blending of complementary input modalities is facilitated through the execution of multiple concurrent agents (which in turn are able to execute multiple concurrent, task-specific behaviors).

Current work already underway includes porting the graphical user interface portion of the PeopleFinder to both PC and Unix platforms (in the interest of extending the tool's audience and ensuring a more thorough testing and empirical evaluation phase of the project); integrating a number of other software applications such as teleconferencing, voice dictation, and video camera-based face recognition; extending agents' capabilities for autonomously resolving run-time conflicts resulting from shared access to the different presentation and communications resources used by the system (see Werkman's KBN negotiation-based conflict resolution work for related issues (Werkman 1994)); and formalizing the various rules used by the PeopleFinder to combine multiple media with multiple modalities for both human-computer interaction and computer supported human-human communication, much along the lines of the work of Arens *et al.* (Arens *et al.* 1993) on allocating multiple media.

The tool is implemented using a variety of different scripting languages (AppleScript, Quickeys, and C shell) and runs on a Macintosh Quadra 840 AV. The tool also makes use of the Macintosh's Apple Phone tool and Geoport Telecom Adapter for performing its various computer-telephony integration tasks.

References

Arens, Yigal; Hovy, Eduard; and Vossers, Mira 1993. On the Knowledge Underlying Multimedia Presenta-

tions. In Mark T. Maybury, editor, *Intelligent Multimedia Interfaces.* AAAI Press: Menlo Park, CA

Buxton, William 1994. The Three Mirrors of Interaction: A Holistic Approach to User Interfaces. In L. W McDonald and J. Vince, editors, *Interacting with Virtual Environments.* New York: Wiley.

Cohen, Philip R.; Cheyer, Adam; Wang, Michelle; and Baeg, Soon Cheol 1994. An Open Agent Architecture. In Working Notes of the AAAI-94 Spring Symposium on Software Agents, Palo Alto, CA, pp. 1–8.

Faure, Claudie and Julia, Luc 1994. An Agent-Based Architecture for a Multimodal Interface. In Working Notes of the AAAI-94 Spring Symposium on Intelligent Multi-Media Multi-Modal Systems, Palo Alto, CA, pp. 82–86.

Ferguson, Innes A. 1992. Touring Machines: An Architecture for Dynamic, Rational, Mobile Agents. Ph. D diss., Computer Laboratory, University of Cambridge, Cambridge UK

Ferguson, Innes A. 1995. Integrating Models and Behaviors in Autonomous Agents: Some Lessons Learned on Action Control. In Working Notes of the AAAI-95 Spring Symposium on Lessons Learned from Implemented Software Architectures for Physical Agents, Palo Alto, CA, March 27–29.

Ferguson, Innes A. and Davlouros, James D. 1995. PeopleFinder: A Multimodal Multimedia Communications Tool for Interconnecting Network Users. In Working Notes of the IJCAI-95 Workshop on Intelligent Multimedia Information Retrieval, Montreal, PQ August.

Finin, Tim; McKay, Don; and Fritzson, Rich 1992. An Overview of KQML: A Knowledge Query and Manipulation Language. Available through the Stanford University Computer Science Department, Palo Alto, CA, March.

Giachino, Luca 1993. Activity Sensing Through Portholes Images: A Bridge between Passive Awareness and Active Awareness, Technical Report, OTP-93-08, Ontario Telepresence Project, University of Toronto, Toronto, ON August.

Rheingold, Howard 1994. PARC is back! *Wired,* 2(2), pp. 90-95.

Werkman, Keith J. 1994. A DAI Architecture for Coordinating Multimedia Applications. In Working Notes of the AAAI-94 Spring Symposium on Intelligent Multi-Media Multi-Modal Systems, Palo Alto, CA, pp. 93-97.

Acknowledgments

[3] Such as those used in various in-house applications at Xerox's Palo Alto Research Center (Rheingold 1994).