From: AAAI Technical Report SS-94-03. Compilation copyright © 1994, AAAI (www.aaai.org). All rights reserved.

An Open Agent Architecture*

Philip R. Cohen
Adam Cheyer
SRI International
(pcohen@ai.sri.com)
Michelle Wang
Stanford University
Soon Cheol Baeg
ETRI

ABSTRACT

The goal of this ongoing project is to develop an open agent architecture and accompanying user interface for networked desktop and handheld machines. The system we are building should support distributed execution of a user's requests, interoperability of multiple application subsystems, addition of new agents, and incorporation of existing applications. It should also be transparent; users should not need to know where their requests are being executed, nor how. Finally, in order to facilitate the user's delegating tasks to agents, the architecture will be served by a multimodal interface, including pen, voice, and direct manipulation. Design considerations taken to support this functionality will be discussed below.

INTRODUCTION

Agents are all the rage. "Visioneering" videos, such as Apple Computer's Knowledge Navigator, have helped to popularize the notion that programs endowed with agency, if not intelligence, are just around the corner. Soon, users need not themselves wade into the vast swamp of data in search of information, but rather the desired, or better yet, needed information will be presented to the user by an intelligent agent in the most comprehensible form, at just the right time.

Although rosy scenarios are easy to come by, intelligent agents are considerably more difficult to obtain. Still, substantial progress is being made on a variety of aspects of the agent story. At least three general conceptions of agent-based software systems can be found in current thinking:

1. Agents are programs sent out over the network to be executed on a remote machine.
2. Agents are programs on a given machine that offer services to others.
3. Agents are programs that assist the user in performing a task.

Each of these models can be found to some extent in present-day software products, for example, in (1) General Magic's emerging TELESCRIPT interpreter, (2) Microsoft's OLE 2.0 and (3) Apple Computer's Newton and Hewlett Packard's New Wave desktop, respectively. Given this space of conceptualizations, we need to be specific about ours.

Definitions and Objectives

Listed below are characteristics of what we are terming agents followed by an example of those characteristics as found in our system:

- *Delegation* — e.g., the ability to receive a task to be performed without the user's having to state all the details
- *Data-directed Execution* — e.g., the ability to monitor local or remote events, such as database updates, OS, or network activities, determining for itself the appropriate time to execute.
- *Communication* — e.g., the ability to enlist other agents (including people) in order to accomplish a task.
- *Reasoning* — e.g., the ability to prove whether its invocation condition is true, and to determine what are its arguments.
- *Planning* — e.g., the ability to determine which agent capabilities can be combined in order to achieve a goal.

Our initial prototype includes agents that exhibit aspects of all the above capabilities, except planning (but see [7]). Our goal is to develop an open agent architecture for networked desktop and handheld machines. The system we are building should support distributed execution of a user's requests, interoperability of multiple application subsystems, addition of new agents, and incorporation of existing applications. Finally, it should be transparent; users should not need to know where their requests are being executed, nor how.

*This paper was supported by a contract from the Electronics and Telecommunications Research Institute (Korea). Our thanks are also extended to AT&T for use of their text-to-speech system.

DOCKET ALARM Find authenticated court documents without watermarks at docketalarm.com.

AGENT ARCHITECTURE

Based loosely on Schwartz's FLiPSiDE system [17], the Open Agent Architecture is a blackboard-based framework allowing individual software "client" agents to communicate by means of goals posted on a blackboard controlled by a "Server" process.

The Server is responsible both for storing data that is global to the agents, for identifying agents that can achieve various goals, and for scheduling and maintaining the flow of communication during distributed computation. All communication between client agents must pass through the blackboard. An extension of Prolog has been chosen as the interagent communication language (ICL) to take advantage of unification and backtracking when posting queries. The primary job of the Server is to decompose ICL expressions and route them to agents who have indicated a capability in resolving them. Thus, agents can communicate in an undirected fashion, with the blackboard acting as a broker. Communication can also take place also in a directed mode if the originating agent specifies the identity of a target agent.

An agent consists of a Prolog meta-layer above a knowledge layer written in Prolog, C or Lisp. The knowledge layer, in turn, may lie on top of existing standalone applications (e.g. mailers, calendar programs, databases). The knowledge layer can access the functionality of the underlying application through the manipulation of files (e.g., mail spool, calendar datafiles), through calls to an application's API interface (e.g. MAPI in Microsoft Windows), through a scripting language, or through interpretation of an operating system's message events (Apple Events or Microsoft Windows Messages).

Individual agents can respond to requests for information, perform actions for the user or for another agent, and can install triggers to monitor whether a condition is satisfied. Triggers may make reference to blackboard messages (e.g. when a remote computation is completed), blackboard data, or agent-specific test conditions (e.g. "when mail arrives...").

The creation of new agents is facilitated by a client library furnishing common functionality to all agents. This library provides methods for defining an agent's capabilities (used by the blackboard to determine when this agent should participate in the solving of a subgoal), natural language vocabulary (used by the interface agent), and polling status. It also provides functionality allowing an agent to read and write information to the blackboard, to receive requests for information or action, and to post such requests to the blackboard, a specific agent, or an entire population of appropriate agents.

When attempting to solve a goal, an agent may find itself lacking certain necessary information. The agent can either post a request of a specific agent for the information, or it may post a general request on the blackboard. In the latter case, all agents who can contribute to the search will send solutions to the blackboard for routing to the originator of the request. The agent initiating the search may choose either to wait until all answers return before continuing processing, or may set a trigger indicating that when the remote computation is finished, a notification should interrupt local work in progress. An agent also has access to primitives permitting distributed AND and OR-parallel solving of a list of goals.

Distributed Blackboard Architecture

As discussed above, the Open Agent Architecture contains one blackboard "server" process, and many client agents; client agents are permitted to execute on different host machines. We are investigating an architecture in which a server may itself be a client in a hierarchy of servers; if none of its client agents can solve a particular goal, this goal may be passed further along in the hierarchy. Following Gelerntner's LINDA model [8], blackboard systems themselves can be structured in a hierarchy, which could be distributed over a network (see Figure 1).[1]

When a goal (G) is requested to be posted on a local blackboard (BB1), and the blackboard server agent at BB1 determines that none of its child agents has the requisite capabilities to achieve the goal, it propagates the goal to a more senior blackboard server agent (BB4) in the hierarchy. BB4 maintains a knowledge base of the predicates that its lower level blackboards can evaluate. When a senior server receives such a request, it in turn will propagate the request down to its subsidiary servers. These subsidiary servers either have immediate client agents who can evaluate the goal, or can themselves pass on the goal to another subsidiary server. In the case illustrated in Figure 1, BB4 determines that none of its subsidiary blackboards can handle the goal, and thus sends the goal to its superior agent (BB5). BB5 passes the goal to BB6, who in turn passes it to BB9. When such a referred goal is passed through the hierarchy of blackboards, it is accompanied by information about the originating blackboard (indicated by the BB1 subscript on G), including information identifying its input port, host machine, etc. This continuation information will enable a return communication (with answers or failure) to be routed to the originating blackboard. Also, the identity of the responding knowledge source BB9 can be sent back to the originator, so that future queries of the same type from BB1 may be addressed directly to BB9 without passing through the hierarchy of blackboards.

Operational Agents

A variety of agents have been integrated into the Open Agent Architecture:

[1]This is referred to as a "federation architecture" in [9].

DOCKET ALARM Find authenticated court documents without watermarks at docketalarm.com.

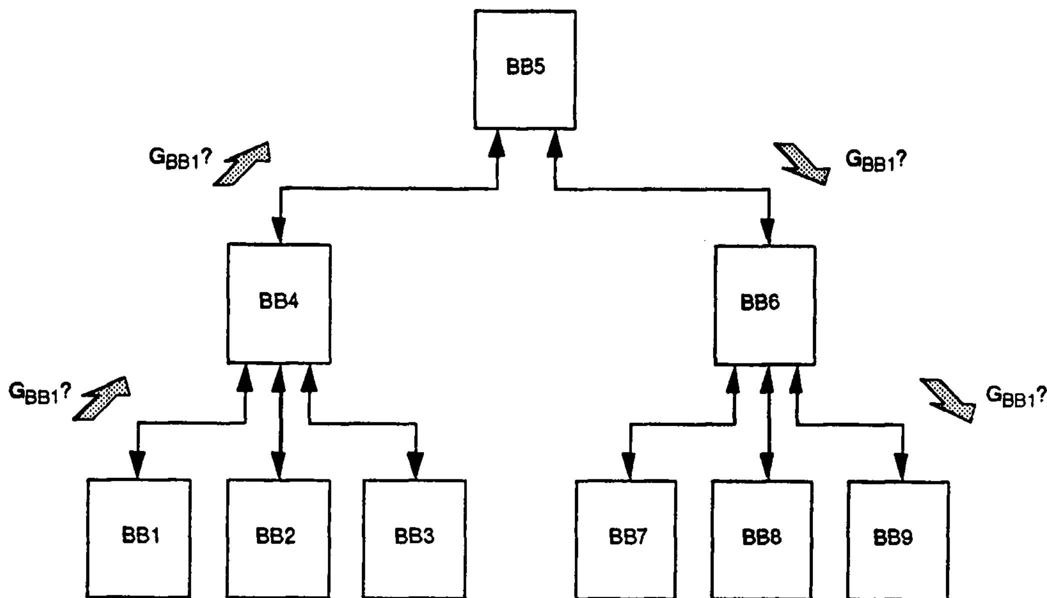

Figure 1: Hierarchy of Blackboard Servers

- a *User-interface* agent that accepts spoken or typed (and soon, handwritten) natural language queries from the user and presents responses to the queries.
- a *Database* agent, written in C, that interacts with a remote X.500 Directory System Agent database containing directory information.
- a *Calendar* agent, which can report upon where a person might be, or when they might be performing a particular action. This information is retrieved from data created by Sun Microsystem's CalenTool application.
- a *Mail* agent that can monitor incoming electronic messages, and forward or file them appropriately. The mail agent works with any Unix-compatible mail application (e.g. Sun's MailTool).
- a *News* agent that scans Internet newsgroups searching for specified topics or articles.
- a *Telephone* agent, that can dial a telephone using a ComputerPhone controller, and can communicate with users in English, using NewTTS, AT&T's text-to-speech system.

Communication Language

The key to a functioning agent architecture is the interagent communication language. We explain ours in terms of its form and content. Regarding the former, three speech act types are currently supported: `Solve` (i.e., a question), `Do` (a request) and `Post` (an assertion to the blackboard). For the time being, we have adopted little of the sophisticated semantics known to underlie such speech acts [5, 18, 19]. However, in attempting to protect an agent's internal state from being overwritten by uninvited information, we do not allow one agent to change another's internal state directly — only an agent that chooses to accept a speech act can do so. For example, a fact posted to the blackboard does not necessarily get placed in the database agent's files unless it so chooses, by placing a trigger on the blackboard asking to be notified of certain changes in certain predicates (analogous to Apple Computer's Publish and Subscribe protocol).

Although our interagent communication language is still evolving, we have adopted Horn clauses as the basic predicates that serve as arguments to the speech act types. However, for reasons discussed below, we have augmented the language beyond ordinary Prolog to include temporal information.

Because delegated tasks and rules will be executed at distant times and places, users may not be able simply to use direct manipulation techniques to select the items of interest, as those items may not yet exist, or their identities may be unknown. Rather, users will need to be able to *describe* arguments and invocation

DOCKET ALARM Find authenticated court documents without watermarks at docketalarm.com.

conditions, preferably in a natural language. Because these expressions will characterize events and their relationships, we expect natural language tense and aspect to be heavily employed [6]. Consequently, the meaning representation (or "logical form") produced by the multimodal interface will need to incorporate temporal information, which we do by extending a Horn clause representation with time-indexed predicates and temporal constraints. The blackboard server will need to decompose these expressions, distribute pieces to the various relevant agents, and engage in temporal reasoning to determine if the appropriate constraints are satisfied.

With regard to the content of the language, we need to specify the language of predicates that will be shared among the agents. For example, if one agent needs to know the location of the user, it will post an expression, such as `solve(location(user,U))`, that another agent knows how to evaluate. Here, agreement among agents would be needed that the predicate name is `location`, and its arguments are a person and a location. The language of nonlogical predicates need not be fixed in advance, it need only be common. Achieving such commonality across developers and applications is among the goals of the ARPA "Knowledge Sharing Initiative," [13] and a similar effort is underway by the "Object Management Group" (OMG) CORBA initiative to determine a common set of objects.

A difficult question is how the user interface can know about the English vocabulary of the various agents. When agents enter the system, they not only register their functional capabilities with the blackboard, they also post their natural language vocabulary to the the blackboard, where it can be read by the user interface. Although conceptually reasonable for local servers (and somewhat problematic for remote servers) the merging of vocabulary and knowledge is a difficult problem. In the last section, we comment on how we anticipate building agents to enforce communication and knowledge representation standards.

Example Scenario

The following is an example of an operational demonstration scenario that illustrates inter-agent communication (see Figure 2).

The user tells the interface agent (in spoken language) that "When mail arrives for me about a security break, get it to me". The interface agent translates this statement into a logical expression, and posts the expression to the blackboard. The blackboard server determines that a trigger should be installed on the mail agent, causing it to poll the user's mail database. Once the mail agent has determined that a message matching the requested topic has arrived for the user, it posts a query to find out the user's current location. The calendar agent responds, noting that the user is supposed to be in a meeting which is being held in a particular room; the database agent is then queried for the phone number of the room. Finally, the telephone agent is instructed to call the number, ask for the user (using voice synthesis), perform an identification verification by requesting a touchtone password, and then read the message to the user. We intend to add agents that would increase the number of ways in which a user might be contacted: agents to control fax machines, automatic pagers, and a notify agent that uses planning to determine which communication method is most appropriate in a given situation.

Comparison with Other Agent Architectures

The most similar agent architectures are FLiPSiDE [17] and that of Genesereth and Singh [9]. Like FLiPSiDE (Framework for Logic Programming Systems with Distributed Execution), our Open Agent Architecture uses Prolog as the interagent communication language, and introduces a uniform meta-layer between the blackboard Server and the individual agents. Some aspects of FLiPSiDE's blackboard architecture are more complex than in our system. It uses a multi-level locking scheme to try to reduce deadlock and minimize conflicts in blackboard access during moments of high concurrency. The system also uses separate knowledge sources for controlling triggers, ranking priorities and scheduling the executing of knowledge sources, whereas we incorporate these sorts of actions directly into the blackboard server. Some features important to our system that are not addressed by FLiPSiDE are the ability to handle temporal contraints over variables, and the possibility for an agent to explicitly request AND and OR-parallel solvingof a list of distributed goals.

Genesereth and Singh's architecture is more ambitious than ours in its employing a full first-order logic as the interagent communication language. As yet, we have not needed to expand our language beyond Horn clauses with temporal constraints, but this step may well be necessary. Genesereth and Singh use KIF (Knowledge Interchange Format) [13] as their basic language of predicates and as a knowledge integration strategy. Because of our user interface considerations, which in turn are heavily influenced by the form-factor constraints of future handheld devices, we will need to be able to merge contributions by different agents of their natural language vocabulary, related pronunciations, and semantic mappings of those vocabulary items to underlying predicates.

MAIL MANAGEMENT

In our earlier scenario, the mail agent was rather limited. To test our user interface and agent architecture more fully, we are creating a more substantive mail management agent, MAILTALK.

It has become common to develop mail managers that manipulate messages as they arrive according to a set

DOCKET ALARM
Find authenticated court documents without watermarks at docketalarm.com.

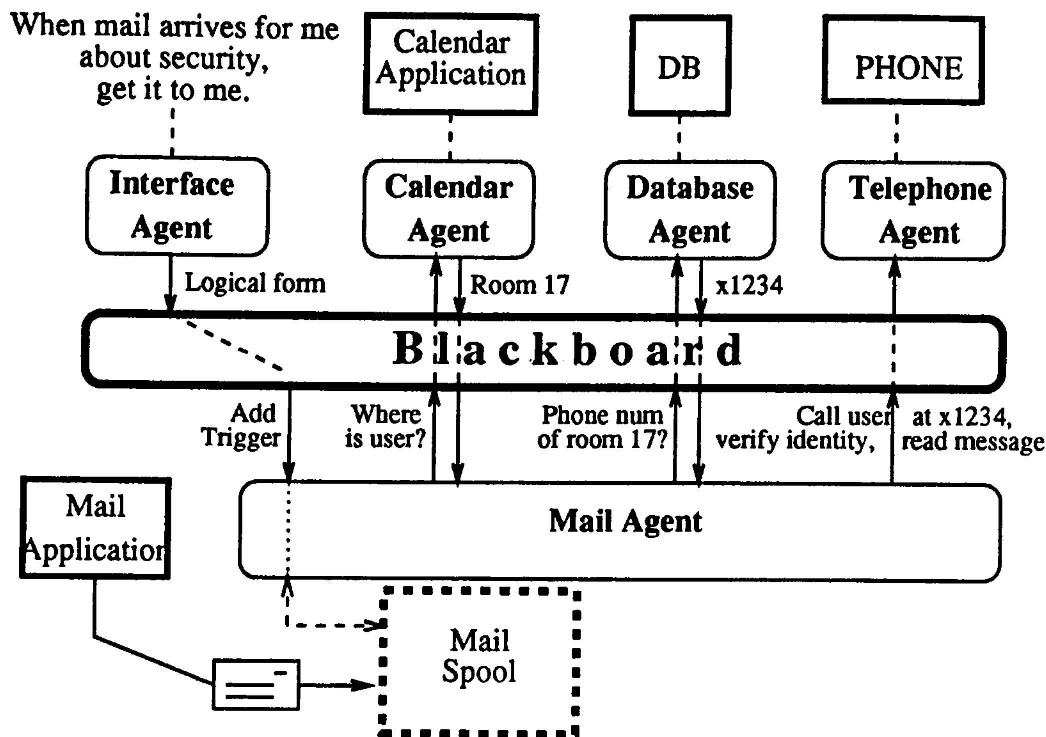

Figure 2: Example of agent interaction

of user-specified rules. The virtue of such systems is that users can make mail management decisions once, rather than consider each message in turn. However, a number of problems exist for such systems, as well as for all agent systems that we know of, especially when considered as tools for the general population.

- End users cannot easily specify the rules. In a number of current systems, a scripting language needs to be employed [1, 20], and in one system, users were required to write rules in a temporal query language [10]. We believe such methods for rule creation effectively eliminates the class of nontechnical users. Other systems employ templates that the user fills out [12]. Although this technique may work in many cases, it limits the power of the rules that users can create because they must search for an icon at which to point in order to specify the contents of a slot. Otherwise, they need to know or select the special syntax or concept name required. However, the selection of items from long menus is infeasible for hand-held devices with little screen territory.

- End users cannot determine in advance how the *collection* of rules will behave once a new rule is added. This lack of predictability and the lack of debugging tools will undermine the utility of agent-based systems, especially in a networked environment.

- End users cannot easily determine what happened. Generally, little or no history of the database of events and rule firings is kept, and few tools are provided for reviewing that history.[2]

- The mail manager is a special purpose system, interacting loosely, if at all, with other components. Without tighter integration, the architecture and user interface for dealing with mail rules may diverge from what is offered for other agents.

Our prototype MAILTALK was built to address these concerns.

Rule specification. Based on technology developed for the SHOPTALK factory simulation system [2, 3, 4], MAILTALK permits users to specify rules by describing complex invocation conditions, and arguments with a multimodal interface featuring typed and spoken natural language, combined with direct manipulation. For example, the user can delegate to the mail agent as follows: "When Jones replies to my message about 'acl tutorials', send his reply to the members of my group." Here, Jones's reply cannot be selected or pointed at since it does not yet exist. The English parser produces expressions in the temporal logic, which are evaluated against various

[2] An exception to this is the use of "Mission Status Reports" in the Envoy agent framework [15].

DOCKET ALARM
Find authenticated court documents without watermarks at docketalarm.com.

DOCKET ALARM

Explore Litigation Insights

Docket Alarm provides insights to develop a more informed litigation strategy and the peace of mind of knowing you're on top of things.

Real-Time Litigation Alerts

Keep your litigation team up-to-date with **real-time alerts** and advanced team management tools built for the enterprise, all while greatly reducing PACER spend.

Our comprehensive service means we can handle Federal, State, and Administrative courts across the country.

Advanced Docket Research

With over 230 million records, Docket Alarm's cloud-native docket research platform finds what other services can't. Coverage includes Federal, State, plus PTAB, TTAB, ITC and NLRB decisions, all in one place.

Identify arguments that have been successful in the past with full text, pinpoint searching. Link to case law cited within any court document via Fastcase.

Analytics At Your Fingertips

Learn what happened the last time a particular judge, opposing counsel or company faced cases similar to yours.

Advanced out-of-the-box PTAB and TTAB analytics are always at your fingertips.

API

Docket Alarm offers a powerful API (application programming interface) to developers that want to integrate case filings into their apps.

LAW FIRMS

Build custom dashboards for your attorneys and clients with live data direct from the court.

Automate many repetitive legal tasks like conflict checks, document management, and marketing.

FINANCIAL INSTITUTIONS

Litigation and bankruptcy checks for companies and debtors.

E-DISCOVERY AND LEGAL VENDORS

Sync your system to PACER to automate legal marketing.

WHAT WILL YOU BUILD? | **sales@docketalarm.com** | **1-866-77-FASTCASE**

fastcase®
Smarter legal research.