[an error occurred while processing this directive]

Software Agents: An Overview

Hyacinth S. Nwana

Intelligent Systems Research
Advanced Applications & Technology Department
BT Laboratories, Martlesham Heath
Ipswich, Suffolk, IP5 7RE, U.K.
e-mail: hyacinth@info.bt.co.uk
Tel: (+44 1 473) 605457
fax: (+44 1 473) 642459

Knowledge Engineering Review, Vol. **11**, No 3, pp.1-40, Sept 1996. © Cambridge University Press, 1996

Abstract

Agent software is a rapidly developing area of research. However, the overuse of the word ëagentí has tended to mask the fact that, in reality, there is a truly heterogeneous body of research being carried out under this banner. This overview paper presents a typology of agents. Next, it places agents in context, defines them and then goes on, inter alia, to overview critically the rationales, hypotheses, goals, challenges and state-of-the-art demonstrators of the various agent types in our typology. Hence, it attempts to make explicit much of what is usually implicit in the agents literature. It also proceeds to overview some other general issues which pertain to all the types of agents in the typology. This paper largely reviews software agents, and it also contains some strong opinions that are not necessarily widely accepted by the agent community.

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

[an error occurred while processing this directive]

4 What is an agent?

We have as much chance of agreeing on a consensus definition for the word ëagentí as AI researchers have of arriving at one for ëartificial intelligenceí itself - nil! Recent postings to the software agents mailing list (agents@sunlabs.eng.Sun.COM) prove this. Indeed, in a couple of these postings, some propounded the introduction of a financial and/or legal aspect to the definition of agents, much to the derision of others. There are at least two reasons why it is so difficult to define precisely what agents are. Firstly, agent researchers do not ëowní this term in the same way as fuzzy logicians/AI researchers, for example, own the term ëfuzzy logicí - it is one that is used widely in everyday parlance as in travel agents, estate agents, etc. Secondly, even within the software fraternity, the word ëagentí is really an umbrella term for a heterogeneous body of research and development. The response of some agent researchers to this lack of definition has been to invent yet some more synonyms, and it is arguable if these solve anything or just further add to the confusion. So we now have synonyms including knowbots (i.e. knowledge-based robots), softbots (software robot), taskbots (task-based robots), userbots, robots, personal agents, autonomous agents and personal assistants. To be fair, there are some good reasons for having such synonyms. Firstly, agents come in many physical guises: for example, those that inhabit the physical world, some factory say, are called robots; those that inhabit vast computer networks are sometimes referred to as softbots; those that perform specific tasks are sometimes called taskbots; and autonomous agents refer typically to mobile agents or robots which operate in dynamic and uncertain environments. Secondly, agents can play many roles, hence personal assistants or knowbots, which have expert knowledge in some specific domain. Furthermore, due to the multiplicity of roles that agents can play, there is now a plethora of adjectives which precede the word ëagentí, as in the following drawn only from Kingís (1995) paper: search agents, report agents, presentation agents, navigation agents, role-playing agents, management agents, search and retrieval agents, domain-specific agents, development agents, analysis and design agents, testing agents, packaging agents and help agents. Kingís paper is futuristic and provides a *role-specific* classification of agents, and so such rampant metaphorical use of the word is fine. But there is also another view that it gives currency to others to refer to just about anything as an agent. For example, he considers "print monitors for open printing, fax redial, and others" (p. 18) as agents, albeit simple ones. As Wayner & Joch (1995) write, somewhat facetiously,

> "the metaphor has become so pervasive that weíre waiting for some enterprising company to advertise its computer switches as *empowerment agents*" (p. 95).

We tend to use the word slightly more carefully and selectively as we explain later.

When we really have to, we define an agent as referring to a component of software and/or hardware which is capable of acting exactingly in order to accomplish tasks on behalf of its user. Given a choice, we would rather say it is an umbrella term, meta-term or class, which covers a range of other more specific agent types, and then go on to list and define what these other agent types are. This way, we reduce the chances of getting into the usual prolonged philosophical and sterile arguments which usually proceed the former definition, when any old software is conceivably recastable as agent-based software.

4.1 A Typology of Agents

This section attempts to place *existing* agents into different agent classes, i.e. its goal is to investigate a typology of agents. A typology refers to the study of types of entities. There are several dimensions to classify existing software agents.

Firstly, agents may be classified by their mobility, i.e. by their ability to move around some network. This yields the classes of *static* or *mobile* agents.

Secondly, they may be classed as either *deliberative* or *reactive*. Deliberative agents derive from the deliberative thinking paradigm: the agents possess an internal symbolic, reasoning model and they engage in planning and negotiation in order to achieve coordination with other agents. Work on reactive agents originate from research carried

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

out by Brooks (1986) and Agre & Chapman (1987). These agents on the contrary do not have any internal, symbolic models of their environment, and they act using a stimulus/response type of behaviour by responding to the present state of the environment in which they are embedded (Ferber, 1994). Indeed, Brooks has argued that intelligent behaviour can be realised without the sort of explicit, symbolic representations of traditional AI (Brooks, 1991b).

Thirdly, agents may be classified along several ideal and primary attributes which agents *should* exhibit. At BT Labs, we have identified a minimal list of three: autonomy, learning and cooperation. We appreciate that any such list is contentious, but it is no more or no less so than any other proposal. Hence, we are not claiming that this is a necessary or sufficient set. *Autonomy* refers to the principle that agents can operate on their own without the need for human guidance, even though this would sometimes be invaluable. Hence agents have individual internal states and goals, and they act in such a manner as to meet its goals on behalf of its user. A key element of their autonomy is their proactiveness, i.e. their ability to ëtake the initiativeí rather than acting simply in response to their environment (Wooldridge & Jennings, 1995a). *Cooperation* with other agents is paramount: it is the *raison díêtre* for having multiple agents in the first place in contrast to having just one. In order to cooperate, agents need to possess a social ability, i.e. the ability to interact with other agents and possibly humans via some communication language (Wooldridge & Jennings, 1995a). Having said this, it is possible for agents to coordinate their actions without cooperation (Nwana *et al.*, 1996). Lastly, for agent systems to be truly ësmartí, they would have to *learn* as they react and/or interact with their external environment. In our view, agents are (or should be) disembodied bits of ëintelligenceí. Though, we will not attempt to define what intelligence is, we maintain that a key attribute of any intelligent being is its ability to learn. The learning may also take the form of increased performance over time. We use these three minimal characteristics in Figure 1 to derive four types of agents to include in our typology: *collaborative agents, collaborative learning agents, interface agents* and truly *smart agents*.

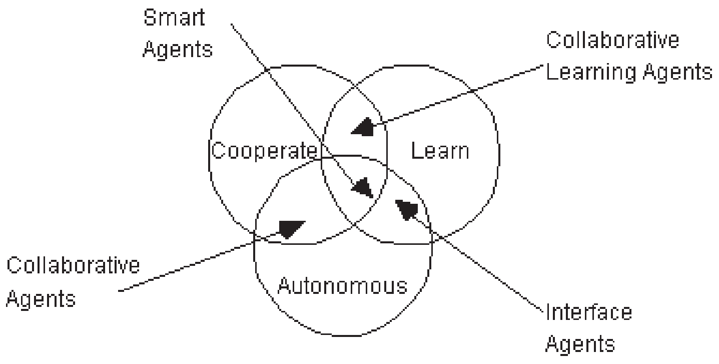

Figure 1 - A Part View of an Agent Typology

We emphasise that these distinctions are *not* definitive. For example, with collaborative agents, there is more emphasis on cooperation and autonomy than on learning; hence, we do not imply that collaborative agents never learn. Likewise, for interface agents, there is more emphasis on autonomy and learning than on cooperation. We do *not* consider anything else which lie outside the ëintersecting areasí to be agents. For example, most expert systems are largely ëautonomousí but, typically, they do not cooperate or learn. Ideally, in our view, agents should do all three equally well, but this is the *aspiration* rather than the reality. Truly smart agents do not yet exist: indeed, as Maes (1995a) notes "current commercially available agents barely justify the name", yet alone the adjective ëintelligentí. Foner (1993) is even more incandescent; though he wrote this in 1993, it still applies today:

> "... I find little justification for most of the commercial offerings that call themselves agents. Most of them tend to excessively anthromomorphize the software, and then conclude that it must be an agent because of that very anthropomorphization, while simultaneously failing to provide any sort of discourse or "social contract" between the user and the agent. Most are barely autonomous, unless a regularly-scheduled batch job counts. Many do not degrade gracefully, and therefore do not inspire enough trust to justify more than trivial delegation and its concomitant risks" (Foner, 1993, 39/40).

In effect, like Foner, we assert that the arguments for most commercial offerings being agents suffer from the logical fallacy of *petitio principii* - they assume what they are trying to prove - or they are circular arguments. Indeed, this applies to other ëagentsí in the literature.

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

In principle, by combining the two constructs so far (i.e. static/mobile and reactive/deliberative) in conjunction with the agent types identified (i.e. collaborative agents, interface agents, etc.), we could have *static deliberative collaborative agents, mobile reactive collaborative agents, static deliberative interface agents, mobile reactive interface agents*, etc. But these categories, though quite a mouthful, may also be necessary to further classify existing agents. For example, Lashkari *et al.* (1994) presented a paper at AAAI on ëCollaborative interface agentsí which, in our classification, translates to *static collaborative interface agents*.

Fourthly, agents may sometimes be classified by their roles (preferably, if the roles are major ones), e.g. world wide web (WWW) information agents. This category of agents usually exploits internet search engines such as WebCrawlers, Lycos and Spiders. Essentially, they help manage the vast amount of information in wide area networks like the internet. We refer to these class of agents in this paper as *information* or *internet agents*. Again, information agents may be static, mobile or deliberative. Clearly, it is also pointless making classes of other minor roles as in report agents, presentation agents, analysis and design agents, testing agents, packaging agents and help agents - or else, the list of classes will be large.

Fifthly, we have also included the category of *hybrid* agents which combine of two or more agent philosophies in a single agent.

There are other attributes of agents which we consider *secondary* to those already mentioned. For example, is an agent versatile (i.e. does it have many goals or does it engage in a variety of tasks)? Is an agent benevolent or non-helpful, antagonistic or altruistic? Does an agent lie knowingly or is it always truthful (this attribute is termed veracity)? Can you trust the agent enough to (risk) delegate tasks to it? Is it temporally continuous? Does it degrade gracefully in contrast to failing drastically at the boundaries? Perhaps unbelievably, some researchers are also attributing emotional attitudes to agents - do they get ëfed upí being asked to do the same thing time and time again? What role does emotion have in constructing believable agents (Bates, 1994)? Some agents are also imbued with *mentalistic* attitudes or notions such as beliefs, desires and intentions - referred to typically as BDI agents (Rao & Georgeff, 1995). Such attributes as these provide for a stronger definition of agenthood.

In essence, *agents exist in a truly multi-dimensional space*, which is why we have not used a two or three-dimensional matrix to classify them - this would be incomplete and inaccurate. However, for the sake of clarity of understanding, we have ëcollapsedí this multi-dimensional space into a single list. In order to carry out such an audacious move, we have made use of our knowledge of the agents we know are currently ëout thereí and what we wish to aspire to. Therefore, the ensuing list is to some degree arbitrary, but we believe these types cover most of the agent types being investigated currently. We have left out collaborative learning agents, see Figure 1, on the grounds that we do not know of the existence ëout thereí of any such agents which collaborate and learn, but are not autonomous. Hence, we identify seven types of agents:

- ï Collaborative agents
- ï Interface agents
- ï Mobile agents
- ï Information/Internet agents
- ï Reactive agents
- ï Hybrid agents
- ï Smart Agents

There are some applications which combine agents from two or more of these categories, and we refer to these as *heterogeneous agent systems*. Such applications already exist even though they are relatively few. However, we also overview briefly such systems in the next section.

Another issue of note (for completeness sake) is that agents need not be benevolent to one another. It is quite possible that agents may be in competition with one another, or perhaps quite antagonistic towards each other. However, we view *competitive* agents as *potential* subclasses of all these types. That is, it is possible to have competitive collaborative-type agents, competitive interface agents, competitive information agents, etc.

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

4.2 A Critique of Our Typology

As with our definition of agenthood, our typology of agents is bound to be contentious. Two official reviewers of this paper all took issue with it, but their suggestions are, in our opinion, either more debatable or unclear. One reviewer, reviewer 1, claimed that we have confused agents that are defined by what they *do* (information agents, interface agents and collaborative agents), and other types for the sort of *technology* that underpins these agents (mobile agents, reactive agents, hybrid agents). Thus, he/she would have preferred a 2-dimensional classification. The second reviewer mentioned a similar point but alluded to a different classification. To a large degree, we disagree with this criticism, though not fully. We believe we had already attempted, perhaps unsuccessfully, to pre-empt this criticism. Firstly, we would not group information agents, interface agents and collaborative agents in one large group: in our view, as we explained earlier, collaborative agents and interface agents are defined by what they *are*, while information agents are defined by what they *do*. Secondly, we do not agree fully with the assertion that mobile agents, reactive agents and hybrid agents are all underlying technologies for implementing the former classes. To this reviewer, interface agents are collaborative agents implemented using reactive technology! We simply disagree with this viewpoint. As we explain later in the paper, reactive agents for example, have a distinct philosophy, hypothesis, etc. which make it stand out from the rest. We have surveyed the area of technologies for building software agent systems in another paper, Nwana & Wooldridge (1996). However, we agree with the general thrust of the argument to some degree; for example, we fully accept the reviewersí viewpoint that mobility is not a necessary condition for agenthood - a point which is implicit in Section 4.1, and which we explain later. Thirdly, we address such issues when we discuss the individual types more fully in the rest of the paper. Fourthly, we point out, explicitly, in Section 4.1 that agents exist in a truly multi-dimensional space, and that for the sake of *clarity of understanding*, we have collapsed this multi-dimensional space into a single list. To produce this list, we used a set of criteria which included inate properties of agents which we would prefer to see (autonomy, cooperation, learning), other constructs (static/mobile, deliberative/reactive), major roles (as in information agents) and whether they are hybrid or heterogeneous. In a previous version of this paper where we had a more hierarchical breakdown, it turned out to be less clear. Fifthly, other typologies in the literature are equally as contentious. For example, Wooldridge & Jennings (1995a) broadly classify agents into the following: gopher agents, service performing agents and proactive agents. We believe this is too general and simplistic a classification. It is for these reasons that we opted for such a ëflatí breakdown. To be fair, apart from the typology, these two reviewers were *very* complementary about the paper.

In conclusion, our typology is not without its critics (but so are all others), but as reviewer 1 pointed out "while I agree that most agents in the literature can be categorised into these types, I think the types are themselves faulty". In this paper, we have deliberately traded in accuracy for clarity. Our typology highlights the key *contexts* in which the word ëagentí is used in the software literature.

4.3 What Agents are Not

In general, we have already noted that a software component which does not fall in one of the intersecting areas of Figure 1 does not count as an agent. In any case, before the word ëagentí came into vogue within the computing/AI fraternity, Minsky, in his *Society of Mind* (1985), had already used it to formulate his theory of human intelligence. However, Minsky used it to refer to much more *basic* entities:

> ". .to explain the mind, we have to show how minds are built from mindless stuff, from parts that are much smaller and simpler than anything weíd consider smart... But what could those simpler particles be - the "agents" that compose our minds? This is the subject of our book..." (Minsky, 1985, 18).

Clearly, Minskyís use of the word ëagentí is quite distinct from its use in this paper.

Furthermore, as noted earlier, expert systems do not meet the preconditions of agenthood, as do most knowledge-based system applications. Modules in distributed computing applications do not constitute agents either as Huhns & Singh (1994) explain. First, such modules are rarely ësmartí, and hence much less robust than agents are (or should be); they also do not degrade gracefully. Second, in agent-based systems generally, the communication involves involves high-level messages in contrast to the low-level messaging in distributed computing. The use of high-level messaging leads to lower communication costs, easy re-implementability and concurrency. Lastly, and perhaps most importantly, agent-

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

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**