



A Critique of Intelligent Agents

INTRODUCTION

One of the most discussed topics in computer literature today is agent technology, particularly intelligent agent technology. Much research and many articles have been devoted to this topic, and more and more new software products billed as having intelligent agent functionality are being introduced on the market every day. The articles and research that are being done on this topic, though, do not whole-heartedly endorse this trend. A growing number of computer information professionals recognize that there are certain problems and issues surrounding intelligent agent terminology and technology that must be resolved if agent technology is to continue to develop and mature. The current research into intelligent agent software technology can be broken up into two main areas: technological and social. The latter area is particularly important since, in the excitement of new and emergent technology, people often forget to examine what impact the new technology will have on people's lives. In fact, the social dimension of all technology is the driving force and most important consideration of technology itself. The technology is not created and produced for its own sake, but to improve people's lives. Technology and computers and software are not created simply to see what the human mind can achieve, but they are created for the sake of human beings. This paper will critique the current state of software intelligent agents by first examining the technological issues and then the social implications of intelligent agent software technology.

TECHNOLOGICAL ISSUES

The first and most fundamental technological aspect that must be considered is what constitutes a software intelligent agent: What is the definition of an intelligent agent? It is here that the first major problem for intelligent agent technology emerges. "In order for this term [intelligent agent] to have any effectiveness, there must first be a universal definition that can be agreed upon and used consistently." [i] Unfortunately, though, there is in fact no commonly agreed upon definition of an intelligent agent or even an (software) agent. Many proposals for a formal definition of "intelligent agent" have been made, but none has been widely accepted.[ii] The following are a few of the more promising definitions:

"An agent is anything that can be viewed as perceiving its environment

through sensors and acting upon that environment through effectors." [iii]

"Let us define an agent as a persistent software entity dedicated to a specific purpose. 'Persistent' distinguishes agents from subroutines; agents have their own ideas about how to accomplish tasks, their own agendas. 'Special purpose' distinguishes them from other entire multifunction applications; agents are typically much smaller." [iv]

"Intelligent agents are software entities that carry out some set of operations on behalf of a user or another program with some degree of independence or autonomy, and in so doing, employ some knowledge or representation of the user's goals or desires." [v]

"An **autonomous agent** is a system situated within and a part of an environment that senses that environment and acts on it, over time, in pursuit of its own agenda and so as to effect what it senses in the future." [vi]

Some of the key terms found in the preceding definitions are: sensing, environment, persistent, 'own agendas', autonomy, goals, and knowledge. Each of these terms seems to appropriately describe characteristics of what an intelligent agent is, yet none of them has gained wide recognition as *the* definition of a software (intelligent) agent. Woolridge and Jennings in their excellent article, "Pitfalls of Agent-Oriented Development" give a compelling reason why a definition consensus has not yet been reached. They point out that agent technology is so popular partly because the idea of an agent is extremely intuitive.[vii] The intuitive aspect of the term "intelligent agent" leads to many different people having different ideas of what an agent is. As Franklin and Graesser point out, most of the definitions

proposed thus far seem to have originated from particular examples that the people who have proposed of the definitions already had in mind.^[viii] It is important to note here that the same intuitive aspect of the term “intelligent agent”, while making it difficult to establish a broadly accepted formal definition, actually makes marketing a product billed as incorporating intelligent agent software technology much easier.

Another reason that a consensus has not been reached is that much of the agent research is proprietary. Companies sponsor the research and do not want to give away their work for free, since they have made significant monetary contributions to this research. The problem is analogous to that experienced by object technology. Standardization of new technology is difficult, because no one wants to give away their work, especially to would-be competitors. Intelligent agent technology will continue to suffer from this difficulty either until the companies and individuals with the proprietary information recognize that sharing it will benefit everyone, including themselves, or until the companies and individuals with the proprietary information recoup enough of their expenditures to feel justified in making available their research.

A third reason for the difficulty in reaching a generally approved definition of what comprises an intelligent software agent, which is probably the most important reason of the three outlined in this paper, is that so-called intelligent agent software does not seem to be qualitatively different from any other software. “Is it an agent, or just a program?”^[ix] In their article of the same title as the previous question, Franklin and Graesser note, correctly, that all software agents are programs. They go on to state that not all programs are agents, the implication being that some programs, then, are in fact agents.^[x]

This third reason for the difficulty in reaching a generally approved definition of “intelligent agent software” revolves around a key term, “quality”. The term “intelligent agent” does not simply mean a more complicated program. If it did, not much controversy would ever have been generated about what actually is an intelligent software agent. If “intelligent agent” only signified a more complicated program, the term “intelligent agent” would mean that a so-called intelligent agent software program was simply more complex and possibly more useful than other typical computer programs. This sense of “intelligent agent” is a *quantitative* sense.

However, those who are doing research into so-called intelligent agent software technology do not mean that an intelligent agent is only more complicated than other computer programs. A so-called intelligent agent is two *qualitative* steps removed from a mere computer program. First it is an agent. An agent is, broadly speaking, someone or something that acts. In order to act, the thing that acts must have a purpose or a goal. This is included in the third and fourth proposed definitions set forth earlier. Do any computer programs have their own goals or purposes? Not really. All a computer program does is perform a set of instructions that were programmed into it. An intelligent software agent is no different from any other computer program in this respect. It simply has more possibilities than less complicated computer programs.

Even if we do grant that a computer program may act, it certainly does not act autonomously as the fourth definition asserts. For something to act autonomously it must have independence and freedom. Philosophically, for something to be autonomous, it must have knowledge of what it is doing and it must *will* to do what it is doing. Computer programs do not *will* to do anything. Whatever they do is programmed into them. Again, we reiterate that the program may be quite complex and be able to react to many different events, but the key is that the computer simply reacts, it does not act on its own.

The word, “react”, further clarifies the inherent limitations of computer programs, and why they cannot truly be called intelligent or autonomous agents. An agent, in the true sense of the word, initiates action. True agents are proactive as well as reactive. They have beliefs, intentions, and desires. It is absurd to speak of computer programs of any sort as having desires. This, then, sums up the problems with calling computer programs “intelligent agents”. Because of the various difficulties that have been raised concerning calling specific computer programs intelligent agents, I find it unlikely that a formal definition of an intelligent software agent will ever be satisfactorily defined.

This leaves us with a question. What are we to make of the all of the software currently on the market or in production that is billed as having intelligent agent functionality? Certainly we do not wish to demean all of the research that has gone into these products. Products such as email filters, help engines (ex: the Microsoft Paperclip), data warehousing tools, news filters, etc. all have the potential to be highly useful to human beings. But look at how they work. They are all based on the detection of patterns in conjunction with explicit user commands and preferences. At their core, all of these computer programs are based on pure and simple mathematics and logic. The email and filters reject messages which do not comply to the user’s defined preferences. The news filters may be more complex than the email filters in that they search on the internet for news in user-defined subject areas. The help engines and data warehousing tools search for built-in patterns, but the programs do not generate the patterns on their own. They have the patterns pre-built into them.

The news and searching tools (often marketed as intelligent agents), while having great potential given the explosion of information accessible on the internet, pose an interesting problem. The problem is this: if many users have news searching “intelligent agent” tools constantly searching for information on the internet, isn’t it likely that the internet may be clogged up by too many of these searching tools? It is likely that each person would have quite a number of these programs running in order to get a wide variety of information. To further complicate this picture, there is also the possibility that these “intelligent agents” will be programmed with the capability to spawn other agents. Imagine if one of the these “intelligent agents” had an error (bug) built in which caused the program to continuously spawn agents to search the internet. What if each one of these spawned programs also spawned other programs to search the internet?

Furthermore, the it is conceivable that a certain number of the “intelligent” agents searching the internet for information would get lost, that is, they would not return with the requested information to the entity which spawned them. Thus, one can begin see the technical dangers in having such “intelligent agents”. They might create severe bottlenecks on the already crowded internet.

SOCIAL IMPLICATIONS

The social implications, as might be expected with a relatively new and partially developed technology, include both positive and negative issues. The author recognizes that the current discussion and research into intelligent agent software technology deals quite sparingly on the topic of the social and ethical implications of this new technology. This lack of serious consideration of the impact intelligent agent software technology on people’s lives is a problem that this paper hopes to begin to address.

The first area that must be addressed is the philosophical nature of technology itself. Technology is created by human beings for human beings. Technology has always been pursued by mankind in order to improve his quality of life. Sometimes men engage in technological projects to help themselves only, but more often men pursue technological improvements for the good of mankind in general. All new technology must be tested in order to see if it meets the requirements of the person(s) who have invented them, and one of the tests must be: does the new technology provide significantly more benefit to mankind or a portion of mankind, than it causes harm? This is the most general test of any new technology, and it is also the most important test, since if the new technology fails this test, it should not be implemented at all. So let us apply this test to the new intelligent agent software technology.

To begin with, let us consider the positive aspects of intelligent agents. One of the biggest benefits to intelligent agents is that they have the potential to free humans from the tedious work of searching for information on the internet and in databases. The amount of information and data both on the internet and in corporate databases is already enormous, and it is continuing to grow exponentially. For any given search by a human being, however, much of the information available on the internet and in databases is of little value. The intelligent agent is supposed to aid in the searching by filtering out the information and data which is of little or no value with little human intervention. If the intelligent agents are successful in this task they certainly can provide a significant benefit to human beings.

Unfortunately, intelligent agents also have the potential to harm human beings. First, if human beings rely too much on intelligent agents, they (human beings) may possibly lose too much freedom. This is a problem with technology and computers in general. A very good article by Jaron Lanier describes in detail the potentially harmful effects of technology on humans.^[xi] Lanier objects to the use of the words “intelligent agent” to describe any type of computer program. His argument centers around the concept that computers contribute substantially to the dulling of the human mind and human creativity. Confining oneself to an artificial world created by some human programmer(s) does limit human potential. This argument, though, lends itself more to the development of children’s minds rather than human beings in general.

Another objection Lanier raises is that human beings end up degrading and lowering themselves when they accept computer programs as “intelligent agents”.^[xii] This argument is more applicable than the first to human beings in general. While information technology professionals and those more aware of what intelligent agents actually are and how they function realize the limitations and scope of so-called intelligent agents, the general public who are the intended audience and users of intelligent agents are less aware of the inner-workings and limitations of intelligent agents. These are the people who Lanier suggests will be psychologically harmed by “intelligent agent” terminology. When individuals begin to think of computers as actually possessing intelligence and autonomy, they will begin to treat the computers like people rather than the (helpful) tools which they are intended to be. The result of treating computers like people will actually be that people begin to view themselves and others around them as computers. “As a consequence of unavoidable psychological algebra, the person starts to think of himself as being like the computer.”^[xiii] This is a serious problem that must be avoided at all

costs.

Another more technical problem that Lanier raises is this:

“If info-consumers see the world through agent’s eyes, then advertising will transform into the art of controlling agents, through bribing, hacking, whatever. You can imagine an “arms race” between armor-plated agents and hacker-laden ad agencies.”^[xiv]

The point here is that if intelligent agents are used to find useful information what will end up happening is that the agents themselves will be manipulated by producers of goods and services. Imagine an agent searching for information on airline flights being manipulated by the various advertising agents sent out by the more clever airline companies. This problem is not insurmountable for those creating “intelligent agents”, but it is a significant problem that needs to be addressed.

In conclusion, while intelligent agent technology has potential to be useful to mankind, many fundamental problems remain to be solved. These problems are both technical and social or ethical in origin, and require careful thought and consideration by those who are developing intelligent agent technology. This paper has been critical of the current state of intelligent agent software technology in the hope of making these developers aware of problems they do not seem to have taken into account. The issues raised by such authors as Lanier are often thrust aside as extremist, but, in fact, they are fundamental to the well-being of those for whom intelligent agent technology is ultimately intended and need to be carefully considered. Intelligent agent software technology has made some progress but has much, much further to go before it can and should be accepted as a tool to improve the quality of people’s lives.

[i] Roberto Vinaja and Sumit Sircar (1999), “Agents Delivering Business Intelligence,” 478.

[ii] Ibid: Stan Franklin and Art Graesser, “Is It an Agent, or Just a Program?: A Taxonomy for Autonomous Agents,” Proceedings of the Third International Workshop on Agent Theories, Architectures, and Languages, Springer-Verlag, 1996, 3.

[iii] Stuart Russell and Peter Norvig (1995), Artificial Intelligence: A Modern Approach, Engelwood Cliffs, NJ: Prentice Hall, 33.

[iv] D. C. Smith, A Cypher, and J. Spohrer (1994), “KidSim: Programming Agents Without a Programming Language,” Communications of the ACM, 37, 7, 58.

[v] “The IBM Agent,” <http://activist.gpl.ibm.com:81/WhitePaper/ptc2.htm>.

[vi] Franklin and Graesser, “Is It an Agent or Just a Program?” 3.

[vii] Michael Woolridge and Nicholas Jennings, “Pitfalls of Agent-Oriented Development,” Section 4.2.

[viii] Ibid.

[ix] Franklin and Graesser, “Is It an Agent or Just a Program?”, Title.

[x] Ibid, 4.

[xi] Jaron Lanier, “Agents of Alienation,” <http://www.well.com/user/haron/agentalien.html>.

[xii] Ibid.

[xiii] Ibid.

[xiv] Ibid.

WORKS CITED

Franklin, Stan and Art Graesser, “Is It an Agent, or Just a Program?: A Taxonomy for Autonomous Agents,” Proceedings of the Third International Workshop on Agent Theories, Architectures, and Languages, Springer-Verlag, 1996.

Lanier, Jason, “Agents of Alienation,” <http://www.well.com/user/jaron/agentalien.html>.

Russell, Stuart J. and Peter Norvig (1995), Artificial Intelligence: A Modern Approach,

Englewood Cliffs, NJ: Prentice Hall.

Smith, D. C., A. Cypher, and J. Spohrer (1994), "KidSim: Programming Agents Without a Programming Language," Communications of the ACM, 37, 7, 55-67.

"The IBM Agent", <http://activist.gpl.ibm.com:81/WhitePaper/ptc2.htm>.

Vinaja, Roberto and Sumit Sircar (1999), "Agents Delivering Business Intelligence," 477-490.

Woolridge, Michael and Nicholas R. Jennings, "Pitfalls of Agent-Oriented Development".

© 2003 **The School Of Management** The University of Texas at Dallas
