BENEVOLENT AGENTS

Michael N. Huhns • *University of South Carolina* • *huhns@sc.edu*
Abdulla Mohamed • *University of South Carolina* • *mohamed@engr.sc.edu*

People often behave altruistically, acting in ways that are not in their immediate self-interest—perhaps not even in their long-term self-interest. A mattress accidentally dropped on a road offers a classic example. Vehicles must slow down and detour around it, creating a traffic jam that inconveniences everyone equally. However, one driver might stop and move the mattress, thereby eliminating the jam. Everyone will benefit from this action except the person who incurred the inconvenience of removing the mattress. Clearly this is not in that person's self-interest, so why do it?

Definition and Motivation for Benevolence

Benevolence means voluntarily helping others without expecting an immediate reward or benefit for doing so. One motivation for acting benevolently is to help the overall society to which one belongs. Another is the belief that our actions may encourage others to act benevolently in the future, thereby providing compensation in the longer term. It is important to understand that a benevolent entity can exist only in an environment with other entities, never alone.

Benevolent Software Agents

Some of the many agents interacting and roaming about the Web these days are benevolent. How might a software agent in a computational environment demonstrate benevolence, since agents are unlikely to encounter a mattress on the road? The agent could clean up stalled or failed database transactions, close sockets left open by a process that terminated early, or remove locks set by failed or former processes. If it doesn't have the authority or ability to take action, a benevolent agent can simply notify agents or systems that do.

Benevolent Agents on the Web

One of the most common Web agents is a query agent. A query agent searches the Web to find an answer to a user's request, and in so doing it may visit many sites and databases. When asked, a benevolent query agent would freely share its query results with other agents on the Web, even though it may have consumed substantial resources to get this knowledge and might have to consume more to share it. In lieu of cybercash or some other mechanism for remuneration, the agent's only prospect for a return on its benevolence is either personal satisfaction (which is difficult for an entity lacking emotions[1,2]) or an improved Internet. The latter is tangible and can be significant. Through one agent's benevolence, other agents charged with similar queries would not have to explore all the sites or databases the first explored: they can simply use its results. Thus, benevolent agents can help reduce Internet traffic, leading to faster Web processing for all.

Survival Strategies

Parisi, Pedone, and Cecconi[3-5] discuss the ideas of individual survival strategies and social survival strategies. Social survival employs a collective store to which all individuals in a group contribute some of their resources. The collective store in turn redistributes the resources to group members by some allocation criteria or converts the resources into something new. Resources may include essential provisions, money, knowledge—or CPU time and data storage space.

Through simulations, the researchers concluded that a group using a collective store could survive severe environmental conditions, while individuals without a collective store would perish. In addition, the raw resources that individuals contributed could be transformed into a new resource that no single individual could produce.

Parisi[4] pointed out the major difference between cooperation and the collective store strategy. In cooperation, the individual has power and control over personal resources and may choose to stop cooperating, but in a collective store the individual relinquishes this power.

Figure 1 illustrates a collective store to which benevolent query agents contribute. It might be implemented as a large database of query results and information. When heavy Internet traffic degrades the search environment, the collective store database could help those agents seeking information on the Web. This is the basis for Internet search services such as Excite, Lycos, and AltaVista, except that users do not have to contribute anything in exchange for using these services. However, agents making greater contributions to a collective store might be given higher priorities in the subsequent use of the store. The collective store could refine the data submitted by different agents and derive new results through data mining techniques. Moreover, a collective store can gather data from agents that have better Web access capabilities and redistribute it to those with poorer capabilities, such as low-bandwidth PDAs.

Sen[6] investigated the circumstances under which one agent should help another perform a given task when that agent requests help. The decision criterion is that such assistance should enable the agent providing help to perform more effectively in the long run. For his experiment, Sen used the principle of reciprocity, which means agents help only those agents who helped them in the past or can help them in the future. Sen's analysis and experiments showed that compared with selfish behavior, reciprocal behavior improves an individual agent's performance, and thereby the group's performance, over the long run.

Cesta, Miceli, and Rizzo[7] believe that social agents, those that provide and ask for help, are the most successful. To verify this, they investigated environments comprising social agents and exploiting agents—agents that ask for help but never give any. Interestingly, they discovered that the social system is endangered more by the presence of filters against exploitation than by the presence of exploiters. Moreover, the more robust the social strategy, the greater the usefulness of social agents to the entire system compared with that of exploiting agents.

Castelfranchi and Conte[8] considered the notion of joint activity and teamwork among agents. They believe that teamwork requires social dependence, including mutual and reciprocal dependence. Mutual dependence is when agents need each other's help to achieve their goals. Reciprocal dependence occurs when two agents realize each other's goals. One of the advantages of this notion is that it eliminates competition between agents, since they depend on each other.

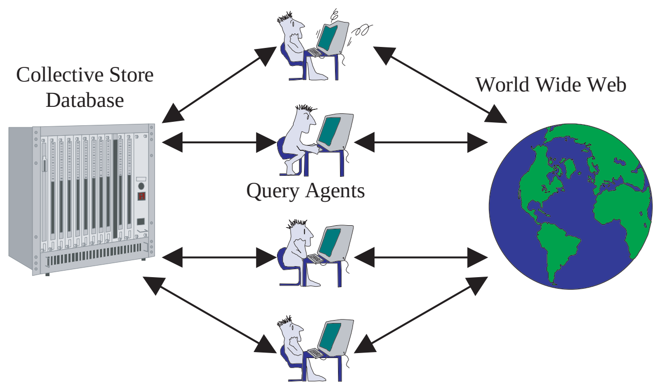

Figure 1. Benevolent query agents may contribute to a collective store database.

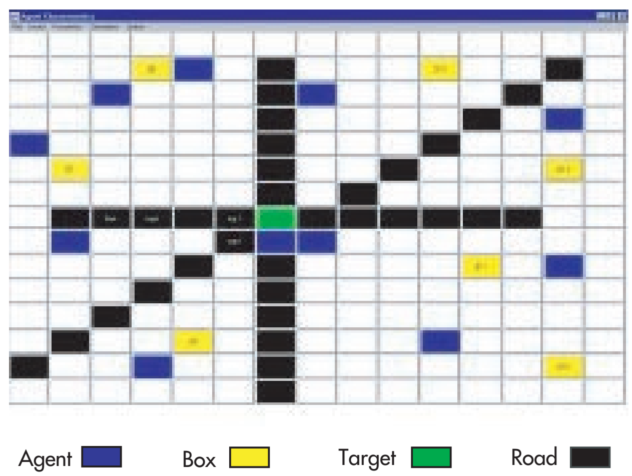

Figure 2. The Agent Behavior Testbed graphical environment.

Putting Benevolence to Work

We constructed the Agent Behavior Testbed (ABT) as a tool to simulate agent behaviors. ABT simulates a two-dimensional environment consisting of agents, boxes, obstacles, and targets (drop zones for boxes). Each agent's main goal is to pick up the boxes in its environment and deposit them at one of the targets. This is an example of a task-oriented domain,[9] in which each agent has all the resources it needs to achieve its goals. Think of it as a nuclear facility, where agents (robots) are used to pick up nuclear waste spilled at various locations and drop it at safe dumpsites. The agents generally try to move toward the targets nearest to them.

While an agent is carrying waste toward a dumpsite, it may encounter obstacles, as illustrated in Figure 2. The agent can choose a benevolent action by moving the obstacle out of the way. Moving rather than just avoiding the obstacle will delay the agent, but this benevolent action will clear a path for other agents navigating through the facility. Simulations show that such a behavior optimizes the efficiency of the agent group and helps to achieve the final goal: clearing the facility as fast as possible. Thus, benevolent action helps all of the agents and, in the long run, the benevolent agent itself.

A testbed user can instantiate as many agents from a selected agent society as desired. Each agent is implemented as a Java thread, so agents execute concurrently. Agents follow a set of rules based on their selected behaviors. Built-in behaviors encompass simple strategies for cooperation, movement, communication, and benevolence. The testbed user can modify the agents' behaviors on the basis of simulation results and repeat a scenario to obtain comparative results.

In addition, benefits could be amplified when combining benevolence with other behaviors, such as cooperation. Agents might cooperatively report to other agents the positions of wastes they encounter while moving toward a target. Such a combination of behaviors will result in more efficient searches and fewer move operations, yielding a better multiagent system.

Toward a More Sociable Web

More and more people are using the Web to search for information or buy goods and services. This can be time-consuming but is nevertheless tolerable because it is still faster than traditional methods. However, as we begin spending more time on the Web, the demand will rise for agents that can perform our daily Web chores for us. Each agent will represent its owner, serving as the owner's surrogate for Web tasks and transactions. To be an effective surrogate, agents will have to be imbued with their owners' preferences and characteristics, such as cooperation, friendliness, sociability, and benevolence. Research prototypes of such benevolent agents are already operating and will soon be making their way onto the Web. So we can look forward to a Web that will be a friendlier and more productive environment for work, learning, and recreation. ■

Benevolent Agents on the Web

You can find many of the systems described in this column on the Web. **//ok?//**

ABT • www.engr.sc.edu/grad/mohamed/agent.htm
Castelfranchi • pscs2.irmkant.rm.cnr.it/users/cristiano/home.html
Conte • pscs2.irmkant.rm.cnr.it/users/rosaria/home.html
Parisi • gracco.irmkant.rm.cnr.it/domenico
Univ. of South Carolina Center for Information Technology • www.engr.sc.edu/research/CIT

System of the Bimonth

Several systems now exist that owe part of their functionality to benevolence. One of these is the USC Agent Behavior Testbed.[1] Check it out!

Reference

1. A.M. Mohamed and M.N. Huhns, "Benevolent Agents," USC-CIT Tech. Report 98-02, Univ. of South Carolina Center for Information Technology, Columbia, S.C., Oct. 1998

REFERENCES

1. M.N. Huhns and M.P. Singh, "Anthropoid Agents," *IEEE Internet Computing*, Vol. 2, No. 1, Jan./Feb. 1998, pp. 94-95.
2. M.N. Huhns and M.P. Singh, *Readings in Agents*, Morgan Kaufmann, San Francisco, Calif., 1998.
3. F. Ceccini and D. Parisi, "Individual Versus Social Survival Strategies," *J. Artificial Societies and Social Simulation*, Vol. 1, No. 2, The SimSoc Consortium, Dept. of Sociology, Univ. of Surrey, Guildford, UK, 1998; http://www.soc.surrey.ac.uk/JASSS/JASSS.html.
4. D. Parisi, "What to Do with a Surplus," in *Simulating Social Phenomena*, R. Conte, R. Hegselmann, and P. Terno, eds., Springer, Berlin, 1997.
5. R. Pedone and D. Parisi, "In What Kinds of Social Groups Can Altruistic Behavior Evolve?" in *Simulating Social Phenomena*, R. Conte, R. Hegselmann, and P. Terno, eds., Springer, Berlin, 1997.
6. S. Sen, "Reciprocity: A Foundational Principle for Promoting Cooperative Behavior Among Self-Interested Agents," *Proc. Int'l Conf. Multi-Agent Systems*, AAAI Press, Menlo Park, Calif., 1996, pp. 322-329.
7. A. Cesta, M. Miceli, and P. Rizzo, "Help Under Risky Conditions: Robustness of the Social Attitude and System Performance," *Proc. Int'l Conf. Multi-Agent Systems*, AAAI Press, Menlo Park, Calif., 1996, pp. 18-25.
8. C. Castelfranchi and R. Conte, "Distributed Artificial Intelligence and Social Science: Critical Issues," *Foundations of Distributed Artificial Intelligence*, John Wiley & Sons, Somerset, N.J., 1996.
9. J.S. Rosenschein and G. Zlotkin, *Rules of Encounter: Designing Conventions for Automated Negotiation among Computers*, MIT Press, Cambridge, Mass., 1994.

Michael Huhns is professor of electrical and computer engineering at the University of South Carolina.

Abdulla Mohamed is a PhD candidate in the Department of Electrical and Computer Engineering at the University of South Carolina, where he received his BS and MS degrees. His current research interest is multiagent systems.