

# Behaviosites: A Novel Paradigm for Affecting Distributed Behavior

Amit Shabtay  
ashabtay@gmail.com

Zinovi Rabinovich  
nomad@cs.huji.ac.il

Jeffrey S. Rosenschein  
jeff@cs.huji.ac.il

School of Engineering and Computer Science, Hebrew University, Jerusalem, Israel

## ABSTRACT

We present the Behaviosite paradigm, a new approach to affecting the behavior of distributed agents in a multiagent system, which is inspired by biological parasites with behavior manipulation properties. Behaviosites are special kinds of agents that “infect” a system composed of agents operating in that environment. The behaviosites facilitate behavioral changes in agents to achieve altered, potentially improved, performance of the overall system. Behaviosites need to be designed so that they are intimately familiar with the internal workings of the environment and of the agents operating within it, and behaviosites apply this knowledge for their manipulation, using various infection and manipulation strategies. To demonstrate and test this paradigm, we implemented a version of the El Farol problem, using behaviosites.

## Categories and Subject Descriptors

I.2.11 [Artificial Intelligence]: Distributed AI—*Intelligent Agents, Multiagent Systems*

## 1. INTRODUCTION

Parasites have been examined in biomimetics, but one of the most interesting abilities of biological parasites has received little attention in technological contexts — their ability to manipulate and alter their host’s behavior. In rare cases, biological parasites can actually benefit the host or the host’s society.

In this paper, we present a novel paradigm that employs a special kind of agent (called a *behaviosite*) that manipulates the behavior of other agents so as to achieve altered, possibly improved, performance of the entire system. The behaviosite (by definition) is not itself necessary for the normal conduct of the system; thus, it is termed a kind of “parasite”.

Within the field of MultiAgent Systems (MAS), the behaviosite is closely related to the idea of adjustable autonomy (AA) [6]. In the context of the Behaviosite paradigm, an agent usually transfers some of its autonomy to the behaviosite unwittingly, and in any case the agent is better described as *ceding autonomy*.

Introducing behaviosites to a system may be either planned as part of the overall system design, or added (after the fact) to an

already working system. When planned as part of the overall system, it is possible to choose between applying *internal behaviosites* (design hooks inside agents, so that behaviosites can attach to the agent and manipulate them) or *external behaviosites* (which manipulate only the input/output of the agent, vis-à-vis the environment). When applied to an already working system, external behaviosites are usually the only option.

Section 2 presents a formalization of the Behaviosite paradigm. Section 3 discusses the El Farol problem, and how behaviosites could be employed in its solution. Section 4 concludes with a discussion of our approach and of future work.

## 2. THE BEHAVIOSITE MODEL

### 2.1 Overall Structure

The Behaviosite paradigm is composed of three parts: the environment, agents (also referred to as hosts), and behaviosites. Environment, agents, and behaviosites will be referred to collectively as “the system” (as in ecosystem).

*Environment*: encapsulates all of the external factors, conditions, and influences that affect a community of agents and behaviosites.

*Agent society*: A society of agents is a system composed of multiple, interacting, possibly cooperating agents.

*Behaviosite*: Most basically, a specialized type of agent. A behaviosite is an additional property/information added to a system with a society of agents, and is not (and should not be) a property of the agent or the environment. The behaviosite is not required for the system, but should be beneficial in some sense; too many behaviosites can degrade some aspects of the system.

### 2.2 Required Traits

*Benefiting the system*: Behaviosites must prove beneficial to the system in some respect, since they come with costs. Behaviosites may influence a system in two desired ways: they may increase social utility, or they may create new features in a working system. Both are accomplished by altering the behavior of some of the agents.

*System knowledge*: A behaviosite must be designed with deep understanding of how the system works: agent-agent interactions, agent-environment interactions, and also internal workings of the agents in some cases. One type of system in which behaviosites can be beneficial is a system that works in a suboptimal equilibrium. The internet is one example of a system in suboptimal equilibrium, with regard to the congestion problem caused by packet routing. The El Farol problem presents another, similar example.

*Not a property of the agent nor of the environment*: Behaviosites exist in the middle ground between agents and the environment, and should not be a property of either of them.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

AAMAS’06 May 8–12 2006, Hakodate, Hokkaido, Japan.  
Copyright 2006 ACM 1-59593-303-4/06/0005 ...\$5.00.

## 2.3 Optional Traits

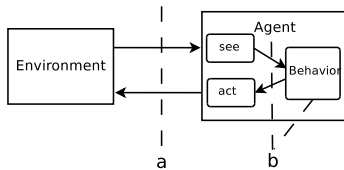
*Hidden or apparent infection:* We usually will not want agents to know who is infected by the behaviorsite, since such knowledge might help individual agents exploit the situation. However, there are some settings in which such knowledge can benefit the system (as also occurs in nature).

*Finding the host:* A behaviorsite designer may endow the environment with the responsibility of infecting agents in the system using some strategy (like in the parasitized El Farol problem), or it may leave the task of infection to the behaviorsites themselves, thus making the behaviorsites more autonomous.

*Behaviorsite communication:* Behaviorsites may communicate with one another within a host (the host may be infected with more than one behaviorsite), or across hosts. This may enable the behaviorsites to be a catalyst for the creation of norms or social laws.

## 2.4 Placing the Behaviorsite

To alter host behavior, the behaviorsite should be placed somewhere along the flow of the system, as can be seen in Figure 1.



**Figure 1:** (a) marks where external behaviorsite can alter host's behavior, by altering its input/output. (b) marks where internal behaviorsite can alter host's behavior, by altering internal data or by partial or full replacement of behavior modules.

Behaviorsites can be divided into two groups, external behaviorsites, and internal behaviorsites. External behaviorsites can alter the input or output of the agent vis-à-vis the environment (Figure 1a). In this way, the host designer(s) need not know of the possible existence of a behaviorsite now or in the future. Moreover, the behaviorsite designer may introduce them into an existing system.<sup>1</sup> On the other hand, internal behaviorsites usually require the system to be designed "plug and play" for the behaviorsite (Figure 1b).

Internal behaviorsites can have two main manipulation methods: changing the host's internal data (Figure 1b, interaction between see-behavior-act modules) or replacing some or all of its behavioral modules (Figure 1b, behavior module). The hook for the second manipulation method is very easy to program, by using the behavioral design patterns suggested by the "gang of four" [3].

## 2.5 Cost of Manipulation

In different scenarios, behaviorsites have different types of costs. The first cost that should be considered is the cost to the system designer, for there is the immediate issue of designing behaviorsites and testing them. Other costs of behaviorsites are system specific, such as balancing behaviorsite complexity and number with the benefit they give to the system, cost of possibly increased variation in the society, oscillations, run-time costs, and so forth.

## 3. PARASITIZED EL FAROL PROBLEM

### 3.1 The El Farol Problem

<sup>1</sup>Of course, these external behaviorsites affect inter-agent communication only if it, too, travels via the environment.

In the El Farol problem [1],  $N$  (e.g., 100) agents decide independently each week whether to go to the El Farol bar or not. Comfortable capacity is limited, and the evening is enjoyable only if the bar is not overcrowded (specifically, fewer than 60 agents out of the possible 100 attend). No communication or collusion is possible, and the only information available to the agents is the number of attendees in the past. An agent will go if it predicts less than 60 agents will go, and will stay home otherwise. Arthur's [1] utility function was:

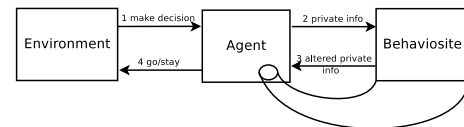
$$Util(ag[i]) = \begin{cases} x & \text{attended and undercrowded} \\ 0 & \text{did not attend} \\ -y & \text{attended and overcrowded} \end{cases}$$

where  $x = y = 0.5$ . A set of personal deterministic strategies combined with a simple learning algorithm facilitated system convergence to the capacity after some initial learning time. Moreover, membership kept changing, and some degree of fairness was maintained. However, the system fluctuated chaotically around the capacity, resulting in low personal and social utility.

The learning algorithm used in the parasitized El Farol simulation was the additive updating learning algorithm, introduced in [5]. Basically, each agent has a pool of simple, personal deterministic strategies. Each such behavior has a weight, initially distributed uniformly, and updated each round.

### 3.2 Applying Behaviorsites

The El Farol bar problem presents a distributed system with sub-optimal social and personal utility. The main idea of the parasitized El Farol problem is to increase social utility with as few side effects as possible, using behaviorsites. The main problem was that agents learn and adapt themselves to new situations — behaviorsites' effects can vanish, or in each round behaviorsites will need to make a stronger impact to achieve the same effect.



**Figure 2:** The environment/agent/behaviorsite relationship

The system is composed of environment, agents (hosts) and behaviorsites (Figure 2). Each agent has an internal behaviorsite field that is always occupied, usually with the null behaviorsite (see [3]) who has no effect over his host's behavior. Each round, the environment infects some of the agents according to an infection strategy, and infection lasts only one round. Agents are asked to provide a decision and if the behaviorsite is not the null behaviorsite, it alters the relevant host's data (the behaviorsite is placed between "see" and "behavior" of Figure 1).

*Environment infection strategies:* The environment had three strategies of infection. In the first, all agents were candidates for infection (*infect all*). In the second, only the agents attending in the given round were candidates for infection (*infect attending*). In the third, all the agents attending in the given round were candidates for infection, but only when the bar was overcrowded (*infect overcrowded*). In each strategy, only a percentage of the candidates for infection actually got infected (responsibility of the behaviorsite design), depending on the infection rate (0%-100%).<sup>2</sup>

*Behaviorsite manipulation strategies:* The behaviorsite replaced the parasitized agents' belief regarding the current capacity with a

<sup>2</sup>Another possible design is that the entire infection strategy is a property of the behaviorsite, but this was not implemented.

lower one, common to all behaviorsites. Since capacity information was kept as the private history of the agent, this decrease was also considered in future rounds by the agent. This was intended to enforce a strict approach; agents were affected by the world according to their personal beliefs, not according to global truth.

### 3.3 Parasitized El Farol Simulation Results

The parasitized El Farol problem was simulated for 2000 rounds, with 100 agents in the system. 3 different capacities were tested: 50, 60, and 80, where the behaviorsite manipulation strategy was to decrease capacity for parasitized agents (50 decreased to 40, 60 also decreased to 40, and 80 decreased to 60). For each capacity, 3 different environment infection strategies were tested, as mentioned above: *infect all*, *infect attending*, and *infect overcrowded*, giving a total of 9 different simulations. For each of the 9 simulations, a different percentage of the agents who were candidates for infection were actually infected, with infection rates ranging from 0% to 100% with jumps of 10% (11 different simulations for each). Each of these 99 different simulations was repeated 50 times.

#### 3.3.1 Mean Attendance

In the parasitized El Farol problem, we would like to increase the social utility, without lowering by too much the mean attendance. As would be expected, the *infect all* strategy decreased the mean attendance severely, while the *infect overcrowded* strategy resulted in a relatively low effect on the mean attendance. The outcome of the *infect attending* strategy depended on the bar's capacity. For capacity of 80, the *infect all* and *infect attending* strategies yielded very similar results and for capacities of 50 and 60, *infect attending* strategy's effect resembles the *infect overcrowded* strategy.

#### 3.3.2 Mean Utility

The main objective in the parasitized El Farol problem simulation was to show that it is possible to increase social utility using behaviorsites. The *Infect all* strategy for capacities of 60 and 80 showed improvement of 7–9 times that which was achieved in a system without behaviorsites, and was relatively close to the optimal. For capacity of 50, improvement was 27 times the minimal, though still near zero (0.065 out of possible 0.25).

*Infect overcrowded* resulted in a moderate utility increase, with respect to other strategies. For capacities of 60 and 80, the maximal mean utility was about 2.5–3.5 times that which was achieved in a system without behaviorsites. For the capacity of 50, the maximal mean utility was 22 times better, but still near zero (0.034).

The *Infect attending* outcome depended heavily on the level of the used capacity. For high capacities (specifically 80), the strategy's behavior was very much like *infect all*, and for low capacities (specifically 50), the strategy's outcome was very much like *infect overcrowded*. In between (specifically 60), the behavior of this strategy combined features of both.

#### 3.3.3 Forced Justice

Two previous attempts to increase social utility in the El Farol problem resulted in two proposed solutions, both with undesirable features in certain settings. One feature was the need to charge utility from attending agents [4], which basically kept the social utility within the system. Another led to the creation of bullies [2] — regulars and casuals. In the original El Farol problem, the membership of the agents kept changing, thus keeping some level of social fairness. Social fairness was calculated by the following formula:

$$1 - \frac{1}{\#trials} \sum_{t \in trials} \frac{Personal\ Attendance\ SD[t]}{Mean\ Attendance[t]}$$

For capacities of 50 and 60, *infect attending* and *infect over-*

*crowded* strategies increase social fairness as a function of infection rate. For capacity of 80, they improved social fairness only a small amount, since the vast majority of agents already go to the bar. However, the *infect all* strategy increased social fairness as a function of infection rate until a certain maximum was reached, and then decreased again. For capacity of 80, it reached exactly the level of social fairness of capacity 60, as would be expected.

## 4. DISCUSSION AND FUTURE WORK

We presented the Behaviorsite paradigm in the spirit of biomimetics — parasites manipulating their hosts' behavior. We specified the behaviorsite concept, which in essence is a special type of agent, which infects and manipulates other agents to achieve altered performance of the system. We described the parasitized El Farol problem, and showed that social and personal utilities and social fairness can be increased using behaviorsites; in some cases, mean attendance deviates from the capacity by only a small amount.

The parasitized El Farol problem simulated self-interested agents with external behaviorsites. However, behaviorsites could also be integrated into a cooperative society to form new norms or social laws, or to eliminate ill-functioning agents.

Behaviorsites are not just a way of propagating false information within a system. The intimacy of the agent-behaviorsite relationship induces far more powerful effects. Lies can be disregarded or overcome by agents. However, when a host is parasitized, the behaviorsite is considered an almost integral part of the agent. Agents can doubt external information, but do rely on their internal beliefs.

In the future, we would like to further strengthen the Behaviorsite paradigm by showing it is applicable and desirable in other scenarios. An appealing application is the automatic generation of stories; it is a rapidly growing field, with high potential in the gamers community. Behaviorsites are an excellent solution for altering stories by changing some of the characters' behavior and adding new, unpredictable system behavior in a distributed manner.

## 5. ACKNOWLEDGMENT

This work was partially supported by grant #039-7582 from the Israel Science Foundation.

## 6. REFERENCES

- [1] W. B. Arthur. Inductive reasoning and bounded rationality (the El Farol problem). *Amer. Econ. Review*, 84(406), 1994.
- [2] A. M. Bell and W. A. Sethares. The El Farol problem and the internet: Congestion and coordination failure. *Computing in economics and finance*, Society for Computational Economics, March 1999.
- [3] E. Gamma, R. Helm, R. Johnson, and J. Vlissides. *Design Patterns: Elements of Reusable Object-Oriented Software*. Addison-Wesley Professional Computing Series, 1994.
- [4] A. Greenwald, J. Farago, and K. Hall. Fair and efficient solutions to the Santa Fe bar problem. *Grace Hopper Celebration of Women in Computing*, October 2002.
- [5] A. Greenwald, B. Mishra, and R. Parikh. The Santa Fe bar problem revisited: Theoretical and practical implications. *Festival on Game Theory: Interactive Dynamics and Learning*, SUNY Stony Brook, July 1998.
- [6] M. Nickles, M. Rovatsos, and G. Weiss, editors. *Post-proceedings of the First International Workshop on Computational Autonomy — Potential, Risks, Solutions*, volume 2969 of *Lecture Notes in Computer Sciences*. Springer-Verlag, 2003.