Avoiding Global Congestion Using Decentralized Adaptive Agents

A. M. Bell and W. A. Sethares

***Abstract*—Everyone wants to go to a bar called *El Farol* if it is not crowded but would rather stay home if it is. Unfortunately, the only way to know whether or not the bar is crowded is to go. While such a scenario appears far removed from the typical communications literature, it provides a simple paradigm for analyzing public goods like the Internet, which may simultaneously suffer from congestion and coordination problems, e.g., multiple users trying to connect to the same server or to use the same resource simultaneously. This paper reviews previous solutions to the *El Farol* Santa Fe bar problem, which typically involve complex learning algorithms. A simple adaptive strategy similar to many signal processing algorithms such as LMS and its signed variants is proposed. The strategy is investigated via simulation, and the algorithm is analyzed in a few simple cases. Unlike most signal processing applications, the objective of the adaptation is not fast and accurate parameter estimation but rather the achievment of a degree of global coordination among users.**

***Index Terms*—Coordination failure, El Farol, multiagent systems, network congestion.**

I. INTRODUCTION

EL Farol is a bar in Santa Fe. The bar is popular but becomes overcrowded when more than 60 people are there on any given evening. Everyone enjoys themselves when fewer than 60 people go, but no one has a good time when the bar is overcrowded. How can, or how do, people choose whether to go to the bar on any given evening?

This simple scenario provides useful and counterintuitive insights into the performance of large-scale information technology systems that have decentralized decision making and rapid endogenous changes in the operating environment. It is a simplified model of a class of congestion and coordination problems that arise in modern engineering and economic systems. For example, despite rapid technological advances and constantly expanding bandwidth, the Internet can become congested when a large number of people independently decide to visit the same web site or to download files at the same time. The level of congestion is endogenously determined by the actions of hundreds or thousands of users [8].

Standard models of congested public, or freely available, resources focus on the marginal costs that an individual user imposes on other potential users (e.g., the classic papers by Hardin [6] and Vickrey [18]). For example, each person who visits a popular web site increases the download time of other users. Explicitly charging users for these unobserved costs can help eliminate the socially inefficient congestion of a scarce resource [11]. However, these models typically analyze equilibrium solutions in which all agents are fully informed about the structure of the problem and the behavior of other agents, which limits their applicability to modern information systems such as the Internet.

In contrast, this paper focuses on the interaction between individual learning strategies and the environment that agents face. Congestion arises in the deterministic version of the *El Farol* or Santa Fe bar problem because agents constantly attempt to predict the aggregate behavior of other agents, which simultaneously depends on all agents' predictions. If agents could perfectly predict the behavior of all other potential bar goers, *El Farol* would never experience congestion. In other words, there need not be a "tragedy of the commons" [6] or a congestion externality at the bar.

The structure of the problem resembles the market entry game found in the economics literature. Agents choose between entering a market, where the payoff depends on the total number of entrants, and not entering a market, which has a fixed payoff. The market entry game has been studied theoretically by Selten and Guth [13] and Gary-Bobo [4] and experimentally by Sundali *et al.* [17], Rapoport *et al.* [12], and Er'ev and Rapoport [3].

Using the bar problem to model the large-scale Internet environment focuses attention on congestion that arises from coordination failure across agents as well as from absolute constraints on bandwidth. For instance, consider logging onto a local internet service provider. If only a few people are also online, then e-mails and downloads occur rapidly, and the user has a "good time," but if too many people are online, then service is slow, graphics files are tedious to download, and the user has a bad experience. The desire to utilize Internet services in the future may depend crucially on past experiences. Observe that this model is not appropriate for packet-level features of the network (such as TCP/IP, ALOHA, etc.).

Arthur's [1] original formulation of the bar problem proposed a deterministic model of "inductive learning" at the level of individual agents. Agents predict how many others will attend the bar each time using a simple kind of inductive reasoning and decide whether or not to attend themselves accordingly. Each agent maintains a number of "rules of thumb" such as simple averages, moving averages, and linear or nonlinear filters to formulate predictions and then acts on the rule with the best performance in the recent past. If they predict attendance will be less than 60, then they go to the bar; if they predict attendance will be greater than 60 then they

Manuscript received September 15, 2000; revised August 20, 2001. The associate editor coordinating the review of this paper and approving it for publication was Dr. Gonzalo R. Arce.

A. M. Bell is with NASA Ames Research Center, Moffett Field, CA 94035-1000 USA (e-mail: abell@mail.arc.nasa.gov).

W. A. Sethares is with the Department of Electrical and Computer Engineering, University of Wisconsin, Madison, WI 53706 USA (e-mail: sethares@ece.wisc.edu).

Publisher Item Identifier S 1053-587X(01)09604-0.

stay home.[1] In Arthur's simulations, the mean attendance is very close to 60. However, attendance varies greatly, often exceeding 70 or dropping below 50. About half of the time, more than 60 people visit the bar, all of whom have a bad time.

According to Arthur, the dynamic behavior observed in simulations arises because successful predictions shared by many agents are self-defeating. Whenever attendance exceeds or falls below 60, a large number of agents must be wrong in their predictions, no matter how the predictions are made. Note that the number of agents who make wrong predictions depends on the variance, not the mean, of attendance. These incorrect predictions drive the continual churning of rules that Arthur observes in his simulations. Furthermore, Arthur notes that "any commonality of expectations gets broken up. If all believe *few* will go, *all* will go, but this would invalidate that belief. Similarly, if all believe *most* will go, *nobody* will go, invalidating that belief. Expectations will be forced to differ." Thus, Arthur argues that agents must utilize a heterogeneous set of predictive rules, undermining the "rational expectations" approach that prevails in many economic models.

In this paper, we show that agents need not utilize different rules, nor must they constantly switch between rules, when deciding whether to attend the bar. We propose a simple adaptive scheme (based on habit formation and adaptive signal processing techniques similar in spirit to [19]) that, if followed by all agents, leads to an outcome close to the socially optimal attendance of 60 agents per night. The next section motivates this adaptive strategy and details the algorithm. Section III examines the generic behavior of the adaptive strategy using simulations. The analysis of Section IV concretizes the simulation results by examining the behavior of the algorithm in certain simple cases. The final section places the adaptive solution in the broader context of congestion problems on the Internet and suggests several areas for further investigation.

II. Adaptive Agent-Based Solution

The agent's decision process is determined by a simple behavioral algorithm. Agents do not predict the aggregate behavior of the system; rather, they base their decision to go to the bar or to stay home on their recent experiences. Agents' attendance decisions are parameterized by a rate or period that determines the frequency of attendance where each agent initally attends once every c time periods.

Assume that people prefer to experience good times rather than bad, to repeat the enjoyable, and to minimize the unpleasant. In response to a pleasant experience at an uncrowded bar, the agent goes more often (decreases c). Similarly, in response to an unpleasant experience at a crowded bar, the agent goes less often (increases c). The agent's behavior is expressed in the form "go to the bar every c periods." Over time, the agent's experiences become encoded in the parameter c. An agent who has had many pleasant experiences at the bar goes frequently (has a low c), whereas one who has had a series of bad experiences goes rarely (has a high c). The adaptive scheme proposed here is analogous to habit formation (or reinforcement learning) on the part of agents. Agents only observe the congestion level online or attendance at the bar when they themselves log on or attend.

When exactly 60 agents attend *El Farol*, the bar is neither overcrowded nor undercrowded. All bar-going agents enjoy themselves, but no agent who chose to stay home would have been better off at the bar. Indeed, all bar-going agents would have been worse off if one more agent had chosen to attend. How might agents respond to this knife-edge scenario? Interpreting the adaptive strategy as a type of habit-formation suggests that agents would continue to increase their frequency of attendance because all attending agents had a good time. Interpreting the strategy as a type of reinforcement learning suggests that agents would neither increase nor decrease their frequency of attendance because an increase in attendance beyond 60 would result in a worse outcome for all attendees. In order to allow for both scenarios, the specification of the adaptive strategy includes a "dead zone" [14] (indicated by a parameter d) below the point at which the bar becomes crowded. Agents neither increase nor decrease their frequency of attendance c when attendance at *El Farol* falls in the "dead zone." Results presented below indicate that the behavior of the system can depend critically on the treatment of such borderline cases[2] and on the assumption that agents only learn attendance at the bar on nights that they themselves attend.

The adaptive strategy for the bar problem resembles the algorithm for the "number guessing game" described in the first chapter of Johnson [9]. When new information reveals that the current guess is too low (or too high), then the guess is increased (or decreased). Although the unknown quantity in the bar scenario varies over time and depends directly on other agents' guesses, the simplicity of the adaptive scheme and its structural similarity to other algorithms suggest strategies for analyzing the resulting behavior.

A. Algorithm Statement

To write this strategy in symbolic form, suppose there are M agents and N spaces at the bar before it becomes crowded. Let $d \geq 0$ represent the size of the dead zone. Let c_i be the period at which the ith agent attends, and let p_i, which is the "phase," be the number of timesteps until the ith agent attends next (thus, p_i is always less than or equal to c_i). The parameter c_i indicates how often the agent attends the bar, for example, if $c = 20$, the agent attends once every 20 timesteps. The parameter p_i indicates how many timesteps remain before the agent goes to the bar. For example, if $c_i = 20$ and the agent attended the bar five timesteps ago, then $p_i = 15$, and the agent will next attend 15 timesteps from now.

Let μ_i be a stepsize parameter that defines how much the ith agent changes c_i in response to new information. Note that the stepsize is small and varies across agents. The time (iteration) counter is denoted by k. Since the c_i and p_i change as time

[1]If agents' behavior has a random element, it is easy to achieve a mean attendence of 60 by setting each agent's probability of attending to 60%. Defining such a "mixed strategy," where agents' choices involve randomization, requires a reward structure that specifies the costs and benefits of staying at home and attending a crowded or uncrowded bar.

[2]Arthur does not specify how agents deal with borderline cases in his original paper.

evolves, $c_i(k)$ and $p_i(k)$ designate the instantaneous values of c_i and p_i at the time k. Thus, the pair $[c_i(k), p_i(k)]$ represents the state of the ith agent at time k, and the concatenation of all M pairs gives the state vector for the entire system.

Let

$$N(k) = \sum_{i=1}^{M} 1_{[p_i(k) \leq 1]} \tag{1}$$

count the number of agents attending at time k, where 1_A is the indicator function that is one if the event A is true and zero otherwise. The evolution of the $c_i(k)$ and $p_i(k)$ is defined by

$$\begin{aligned} c_i(k+1) &\\ &= \text{Max}\left(1, c_i(k) + \mu_i\, \text{Dsgn}(N(k) - N) 1_{[p_i(k) \leq 1]}\right) \\ p_i(k+1) &\\ &= (p_i(k) - 1) 1_{[p_i(k) > 1]} + c_i(k+1) 1_{[p_i(k) \leq 1]} \end{aligned} \tag{2}$$

where Dsgn(x) represents the signum function, with a dead zone as discussed above, that is positive when $x > 0$, negative when $x + d \leq 0$, and zero otherwise. The μ_i term is a stepsize that scales the agents response and may also depend on time k, although this dependence has been suppressed in (2).

For each agent, the initial values of $c_i(0)$ and $p_i(0)$ are chosen randomly. Each agent's "phase" term p_i counts down until it drops below one. Meanwhile, the counter c_i remains unchanged. Once p_i reaches one, the agent attends the bar. At this point, c_i is increased by μ_i if bar attendance exceeds N (the bar is crowded), decreased by μ_i if the bar attendance is lower than $N - d$ (the bar is uncrowded), and remains unchanged if attendance falls in the dead zone just below the cut-off point N. The phase p_i is then reset to the current (updated) value of c_i. The counter c_i always remains positive because of the Max$(\cdot)$ function.

B. Derivation of the Algorithm

The algorithm for individual agents detailed above is now derived as an approximation to an instantaneous gradient descent for minimization of the global cost or social welfare function

$$J(k) = |\text{avg}\{N(k)\} - N| \tag{3}$$

where

$$\text{avg}\{N(k)\} = \frac{1}{w} \sum_{j=k-w+1}^{k} N(j) \tag{4}$$

is the average of the number of attendees over a window of length w. In other words, when each individual agent uses the proposed adaptive strategy, the system as a whole tends to minimize deviations away from the moving average of attendance over time. Note that this global cost function weights deviations above and below the optimum of 60 equally. Substituting (1) into (4) gives

$$\begin{aligned} \text{avg}\{N(k)\} &= \frac{1}{w} \sum_{j=k-w+1}^{k} \sum_{i=1}^{M} 1_{[p_i(j) \leq 1]} \\ &= \frac{1}{w} \sum_{i=1}^{M} \sum_{j=k-w+1}^{k} 1_{[p_i(j) \leq 1]}. \end{aligned}$$

However, the event $p_i(j) \leq 1$ occurs once every c_i timesteps; therefore, for a large window length w and a small stepsize $\mu_i(k)$, the event $p_i(j) \leq 1$ occurs approximately $[w/c_i(k)]$ times, where $[x]$ is the largest integer contained in x. Hence

$$\text{avg}\{N(k)\} \approx \frac{1}{w} \sum_{i=1}^{M} \left[\frac{w}{c_i}\right] \approx \frac{1}{w} \sum_{i=1}^{M} \frac{w}{c_i} = \sum_{i=1}^{M} \frac{1}{c_i}$$

for large windows w. Hence

$$\frac{d\,\text{avg}\{N(k)\}}{dc_i} \approx -\frac{1}{c_i^2}. \tag{5}$$

The typical gradient strategy updates the state by

$$c_i(k+1) = c_i(k) - \mu_i(k) \frac{dJ(k)}{dc_i(k)} \tag{6}$$

although in the *El Farol* problem, this update occurs only when the agent attends the bar. With $J(k)$ as in (3) and ignoring the singularity at $N(k) = N$

$$\begin{aligned} \frac{dJ(k)}{dc_i(k)} &= \text{sgn}(\text{avg}\{N(k)\} - N) \frac{d\,\text{avg}\{N(k)\}}{dc_i(k)} \\ &\approx -\text{sgn}(N(k) - N) \frac{1}{c_i^2(k)}. \end{aligned}$$

The instantaneous approximation to the gradient replaces avg$\{N(k)\}$ with the instantaneous value $N(k)$. Combining this approximation with (5) and (6) gives

$$c_i(k+1) = c_i(k) + \frac{\mu_i}{c_i^2(k)}\, hboxsgn(N(k) - N) \tag{7}$$

which holds whenever $1_{[p_i(k) \leq 1]}$. Combining (7) with the phase terms $p_i(k)$ (so that the update to $c_i(k)$ occurs only when the ith agent is part of the sum $N(k)$) gives the complete algorithm (2), where $\mu_i(k)$ is identified with $\mu_i/(c_i^2(k))$. The "sgn" function can be readily generalized to incorporate the dead zone by suitable modification of the cost function (3), which is a standard procedure in the adaptive filtering literature; see, e.g., [14]. In this case, the global cost function treats deviations above the optimum and below the dead zone equivalently.

Alternatively, one could consider the squared cost function

$$J(k) = (\text{avg}\{N(k)\} - N)^2. \tag{8}$$

Repeating the logic of (5) to (7) leads to a kind of least mean square update [7]

$$c_i(k+1) = c_i(k) + \frac{\mu_i}{c_i^2(k)} (N(k) - N). \tag{9}$$

Here, the cost function weights large deviations away from the optimum more heavily but again treats overcrowding and undercrowding symmetrically. The full algorithm incorporating the phase is identical to (2) with the Dsgn$(\cdot)$ function removed.

C. Comparison with Standard Adaptive Algorithms

Iterative minimization of a cost function using approximate gradient descent methods is a standard signal processing strategy [7]. Perhaps the most common technique is the least mean square (LMS) algorithm [19] that has been modified in many ways to improve convergence or tracking (such as normalized LMS [9]), to decrease numerical complexity (such as the signed regressor LMS [16]), or to decrease its sensitivity to disturbances [5], [20]). This section details the relationship between standard signal processing algorithms and the similarly structured algorithms derived above for the *El Farol* problem.

All the algorithms have the basic gradient form that the new estimate is the old estimate plus a correction term defined as a scaled version of the (negative) gradient. The $\mu_i/(c_i^2(k))$ term is reminiscent of the stepsize update in the normalized LMS since it decreases as $c_i(k)$ grows. The $N(k)-N$ term is analogous to the "error" term in LMS, whereas the sgn$(N(k)-N)$ is the equivalent error term from the "signed-error" LMS. Although the $1/(c_i^2(k))$ may appear to be problematic, it is actually well behaved since $c_i(k)$ is never less than one. The presence of the "knife-edge" is paralleled by the sgn function, and sign functions with dead zones are commonplace in the adaptive signal processing literature [14].

Despite these similarities, the origin of the problem renders the *El Farol* algorithms quite different. First, the cost function $|\text{avg}\{N(k)\}-N|$ is not the same as avg$|\{\hat{y}(k)-y(k)\}|$ (as used to derive the signed-error LMS algorithm in [15]) since the direct analog would be avg$|\{N(k)-N\}|$. Similarly, the cost function for LMS has the average (or the expectation in a stochastic setting) outside the square rather than inside, as in (8).

Second, $N(k)$ is not a fixed (linear or nonlinear) target function with parameters to be identified. Since all agents are interchangeable, there are many possible equilibria that will minimize the cost. Hence, the error surface is not quadratic nor even unimodal in $c_i(k)$, and proving convergence is a more subtle affair. The presence of the phase terms $p_i(k)$ means that convergence is only possible to a periodic state and not to a fixed equilibrium. These issues are explored further in Section IV.

On a more fundamental level, unlike the majority of signal processing applications of adaptive algorithms, the objective in this case is not fast or accurate parameter estimation. Rather, the objective is to achieve a degree of global coordination among multiple independent agents.

III. Generic Behavior of the Adaptive Solution

A few example simulations illustrate the dynamics of a bar-going society in which each agent utilizes the strategy (2) defined above. Although details of the various simulations differ depending on the size of the population, the capacity of the bar, the size of the dead zone, and the initial conditions, Fig. 1 represents the typical behavior of attendance over time. In this case, $M=100, N=60$, and $d=5$. The counters and phases were initialized randomly.[3] In this case, too few people attend the bar initially. Other initializations that result in too many agents attending at the start exhibit similar periods of transience and similar long run behavior.

Perhaps the most striking aspect of these simulations is that the number of attendees converges within a few hundred iterations to the dead zone below the optimal value of attendance. After the initial transience resolves, spikes of attendance above 60 occur infrequently. All agents base their attendance decisions on local information, that is, on their own experiences, but nonetheless, attendance largely remains in the dead zone just below the optimal choice of a central planner. In comparison, both Arthur's inductive learning and Edmond's genetic programming approaches generate far greater excursions away from the optimal value.

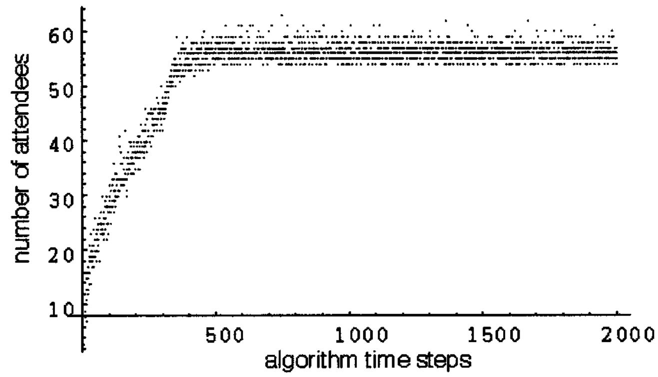

Fig. 1. When all agents use the adaptive solution, the number of attendees only rarely exceeds the critical $N=60$.

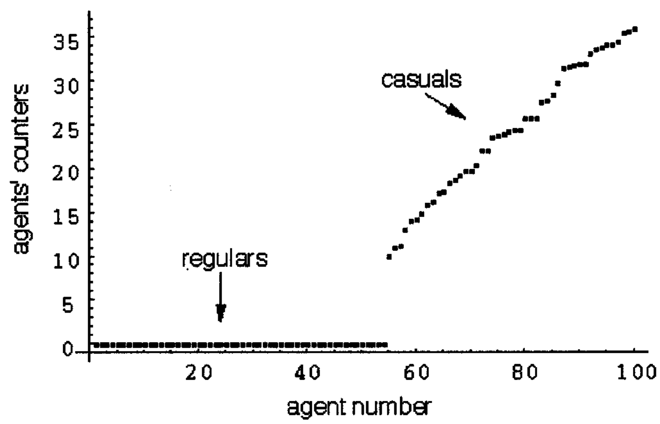

Fig. 2. Emergent property of the adaptive solution is that the population divides itself into "regulars" and "casuals."

The simulations also reveal an interesting pattern in the behavior of individual agents. Fig. 2 shows values of the counters $c_i(K)$, ranked from smallest to largest, at the final iteration $K=2000$ of a typical simulation run. Fifty-four agents have $c_i(K)$ equal to 1, indicating that they go to *El Farol* every time. The remaining agents with larger counters vie with each other for the remaining open slots. The occasional spikes in attendance in Fig. 1 occur when several of the remaining agents happen to go simultaneously. Thus, the casuals are collectively vying for the few remaining spaces left by the regulars. It is possible that the counters of the casuals will continue to increase without bound. We show one such case in the next section. It is also possible that the counters may settle into a periodic steady state where several of the casuals alternate with each other to fill one of the empty slots. (Such periodic states are easy to contrive by suitable choices of initial values, but they do not seem to occur generically in the simulations.) While initial conditions such as the stepsize μ_i and the phase $p_i(0)$ help determine which

[3]In this simulation, the $c_i(0)$ were uniformly chosen between 1 and 10, and the $p_i(0)$ were uniformly chosen between 1 and $c_i(0)$. The step sizes μ_i were chosen randomly between 0.1 and 1.0. The behavior of the algorithm appears insensitive to the particular values chosen.

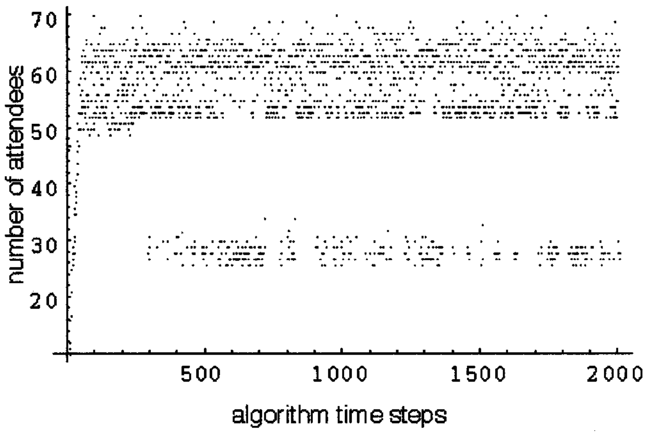

Fig. 3. In this simulation, each agent updates its counter at every time step. Using more information does not necessarily lead to better results.

agents attend every period and which agents attend infrequently, the same patterns of attendance develop.

Thus, the population has segregated itself into a group of "regulars" who go to the bar every time and a group of "casuals" who only attend occasionally. Asked about their experiences at *El Farol*, the regulars say "It's great, I go there all the time, and it's almost never crowded." Asked the same question the casuals say "I hardly ever go, and when I do, it often seems crowded." Both groups are right: Their different perspectives reflect different experiences and different data about the environment. This division of the population does not appear in the algorithm statement or the corresponding global cost function; rather, it is an emergent property of the adaptive solution to the *El Farol* problem. When agents follow this adaptive strategy, *El Farol* looks more like *Cheers*. As the following section demonstrates, the system defined by the algorithm (2) has a large number of equilibria; this type of habit formation or reinforcement learning tends to select those equilibria in which one group of agents attends regularly and another group attends sporadically.

In Arthur's simulations, agents know the history of attendance at the bar regardless of whether they attend or not. The final simulation, which is shown in Fig. 3, investigates the effect of allowing the agents to update their counters at every iteration, whether they have personally attended the bar or not. Observe that the variance about $N = 60$ is much larger than in Fig. 1, and there is a cloud of values around 30, indicating that seats in the bar often remain unfilled.[4] When all agents have access to the same information about attendance, they all respond simultaneously, and the individual decisions of the agents collectively overshoot or undershoot the optimum, creating inefficient variations in attendance. Somewhat paradoxically, the adaptive solution functions better when agents' have heterogeneous and limited information sets.

This insight suggests that universal access to information may reduce the global efficiency of some systems. For example, in financial markets in which current prices depend critically on individual expectations of future prices, the rapid simultaneous dissemination of information may lead to bubbles or overshooting relative to fundamental or underlying economic conditions. This issue may also arise in decentralized algorithms for managing traffic flows on the Internet; globally available information about the level of congestion may lead to an instantaneous over-response that merely shunts the congestion from point to point rather than distributing it evenly across possible routes. In contrast, more selective or slower distribution of information may induce a smaller response, giving the entire system time to adapt smoothly to changing conditions.

IV. Analysis of the Adaptive Solution

The analysis of the system focuses on defining and characterizing the equilibria of the system. In addition, under certain circumstances equilbria fail to exist: Some agents attend less and less frequently, their counters diverging toward infinity, albeit at an ever slower rate.

In a strict formal sense, the system defined by (2) with $M > N$ has no equilibria because the phase terms p_i continually decrease from c_i to zero, only to be reinitialized at c_i. However, if agents' counters (the vector of c_is) do not change, then neither does the pattern of attendance at the bar; the only externally observable quantity $N(k)$ is completely predictable. Consequently, an equilibrium in terms of attendance can be defined as follows.

Proposition IV.1: Consider (2) with $\mu_i(k) = \mu_i$ fixed for all k. Suppose that $c_i(k) = c_i^*$ for all $k \geq K$. Then, $N(k)$ is periodic for $k \geq K$.

In the following, an equilibrium of the system (2) refers to the c_i portion of the state and the corresponding periodic pattern of attendance, in accordance with the proposition.

The *El Farol* problem contains a knife-edge response to increased attendance; the transition from an uncrowded to a crowded bar depends on the behavior of any one agent.[5] Consequently, when agents use predictive behavioral rules like those utilized by Arthur, the definition and analysis of equilibria depends crucially on how the agent accounts for his or her own behavior. For example, suppose that attendance at *El Farol* has been 60 for the last ten periods and that all agents correctly predict attendance of 60 next period. How should an individual agent decide whether or not to attend in this case? Common sense suggests that agents who have attended the bar every period should continue to attend every period. On the other hand, agents who have not attended at all in the last ten periods should remain at home because the addition of another agent will result in attendance of 61. Agents who have attended irregularly should replicate their previous pattern of attendance.

While the adaptive algorithm proposed here does not rely directly on agents' prediction of attendance, similar issues arise in defining equilibria. When $d = 0$, the knife edge eliminates the possibility of equilibria. Unless the same N agents, all with $c_i = 1$, attend each period, at least one agents' counter changes.

[4]This behavior is observed in about half of the simulations, with the cloud of values that represents the clustering or herding of agents with similar c_i centered around different values.

[5]Although this is an extreme assumption, similiar nonlinearities arise in congestion models. No delays occur at a bottleneck when traffic is below capacity, but once the capacity is reached, long delays can appear rapidly. The performance of a computer network can degrade rapidly for all users with the addition of one more user [8]. Note that the introduction of a "dead zone" makes this transition less extreme.

This implies that the remaining $M - N$ agents never attend, which in turn implies that their counters have diverged. The following result demonstrates a simple case in which the c_i of certain agents must diverge toward infinity.

Proposition IV.2: Suppose M is given, $N = M - 1, d = 0$, and that $\mu_i \ll 1$ for all i. Suppose further that the system (2) has evolved (or is initialized) so that there are $M - 1$ "regulars" with $c_i < 1 - \mu_i$. The remaining agent (designated agent m) has $c_m \gg 1$. Then, $c_m(k) \to \infty$, as $k \to \infty$.

Such agents increase their counters forever. Presumably, real humans in such a situation would eventually withdraw and go to some other bar.

The existence of a "dead zone" allows for a multitude of equilibria. To get an idea of the number of possible equilibria, consider a single "slot" of the M possible. This could be occupied by a customer who attends every time. It could also be occupied by two customers who alternate, i.e., who have different phases. In general, any number could alternate on successive evenings and still only occupy a single space at the bar. For $M > N$, let $\phi(M, N)$ represent the number of ways that the integer M can be partitioned into a sum of exactly N integers. Let $\psi(n)$ represent the number of ways that 1 can be written as a sum of exactly n fractions of the form $(1/i_1) + (1/i_2) + \cdots + (1/i_n)$ for integers i_k. $\psi(n)$ counts the number of ways that n agents can alternate to occupy a single bar stool. The total number of equilibria is a combination of ϕ and ψ.

Proposition IV.3: The algorithm (2) with $d = 1$ has at most $\sum_{j=1}^{\phi(M,N)} \prod_{k=1}^{N} \psi(i_k^j)$ equilibria.

V. Conclusion

Arthur believed that any solution to the bar problem would require heterogeneous agents, that is, agents who pursue different strategies. In contrast, we have presented a simple adaptive solution that can be followed by all agents and can readily solve the problem. Each agent in the adaptive solution is characterized by a parameter that determines how often the agent attends and a stepsize that determines how much to change the parameter in response to each visit to the bar.[6] Our agents, like Arthur's, are boundedly rational and use completely deterministic decision rules.

The adaptive solution to the *El Farol* problem differs from Arthur's strategy in several ways. We do not require agents to make explicit predictions of the state of the bar. We introduce the "don't care" or dead zones to handle the transition from uncrowded to crowded. Most importantly, we restrict agents' to using only the information that they have readily available, i.e., their own experiences. This differential access to information helps create the diversity of actions that is crucial to an optimal solution of the *El Farol* problem. Arthur focuses on the global behavior associated with a particular modeling strategy for boundedly rational decision making. In contrast, the approach detailed here demonstrates conditions under which an optimal solution to a global optimization problem can be achieved in the absence of coordination across decentralized agents. Arthur focuses on modeling the behavior of people attending the bar; we focus on strategies designed to optimize attendance.

The adaptive solution thus provides a simple mechanism whereby a large collection of decentralized decision makers, each acting in their own best interests and with only limited knowledge, can solve a complex social coordination problem. Convergence to the solution is relatively rapid (depending on the initial conditions) and robust. Under certain singular initial conditions, it is possible for some of the agents to "withdraw" from the society. Nonetheless, the population of the bar remains near its optimal value.

Do we believe that bar goers tick off the time till they can go again, increasing or decreasing a counter with each new visit? Of course not. However, the behavioral algorithm is in agreement with the commonsense idea that people base their expectations of the value of an action on past experiences. The algorithm can be generalized to to a probabilistic or mixed-strategy decision-making framework where agents increase or decrease their probability of attending based on previous experience. Moreover, the global behavior of the population is consistent with certain kinds of coordination phenomena. For instance, users of an Internet provider can spread demand over much of the day, even though everyone might prefer (all else being equal) to log on in the middle of the afternoon. By developing certain habits (for instance, always logging on at the same time) users send signals to others to avoid these times. In this way, demand is smoothed.

There are many ways to generalize the algorithm. For instance, different people have different tolerances for what constitutes a crowd or an unacceptable delay. Each agent could also have a parameter that represents their tolerance for congestion. Additionally, to more closely model the Internet situation, one might incorporate time-of-day or day-of-week as a parameter in the process of logging on. It would also be instructive to create a hybrid situation in which a number of Arthur-like agents and a number of adaptive agents compete for spaces at the bar. We hypothesize that the more goal directed adaptive agents will again become regulars at the bar, relegating Arthur's more haphazard agents to squabbling over the few remaining seats.

Acknowledgment

The authors would like to acknowledge helpful comments from two anonymous referees, R. Erickson, R. Gaylord, D. Lucking-Reiley, C. Mullin, and seminar participants at NASA Ames Research Center, Tulane University, and Vanderbilt University.

References

[1] W. B. Arthur, "Inductive reasoning and bounded rationality: The *El Farol* problem," in *Proc. Amer. Econ. Rev.: Amer. Econ. Assoc. Papers*, vol. 84, May 1994, pp. 406–411.
[2] J. D. Carillo, "Coordination and externalities," *J. Econ. Theory*, vol. 78, no. 1, pp. 103–129, Jan. 1998.
[3] I. Erev and A. Rapoport, "Coordination, "magic," and reinforcement learning in a market entry game," Hong Kong Univ. Sci. Technol., Dept. Marketing, MKTG 97.103, 1997.
[4] R. Gary-Bobo, "On the existence of equilibrium points in a class of asymmetric market entry games," *Games Econ. Beh.*, vol. 2, pp. 239–246, 1990.

[6] In any multiagent system, the agents may be homogeneous or heterogeneous in terms of structures or in terms of parameters. Our agents are completely homogeneous in terms of structure; the varying stepsizes and initial conditions makes them heterogeneous in terms of parameters. Arthur's are heterogenous in both parameters and structure.

[5] A. Gersho, "Adaptive filtering with binary reinforcement," *IEEE Trans. Inform. Theory*, vol. IT-30, Mar. 1984.

[6] G. Hardin, "The tragedy of the commons," *Science*, no. 162, pp. 1243–1248, 1968.

[7] S. Haykin, *Adaptive Filter Theory*. Englewood Cliifs, NJ: Prentice-Hall, 1991.

[8] B. Huberman and R. M. Lukose, "Traffic jams on the internet," *Science*, vol. 280, no. 5361, Apr. 1998.

[9] C. R. Johnson Jr., *Lectures on Adaptive Parameter Estimation*. Englewood Cliffs, NJ: Prentice-Hall, 1988.

[10] D. M. Kreps, *A Course in Microeconomic Theory*. Princeton, NJ: Princeton Univ. Press, 1990.

[11] J. K. MacKie-Mason and H. R. Varian, *Public Access to the Internet*, B. Kahin and J. Keller, Eds. Englewood Cliffs, NJ: Prentice-Hall, 1994.

[12] A. Rapoport, D. Seale, and E. Winter, "An experimental study of coordination and learning in iterated two-market entry games," Hong Kong Univ. Sci. Technol., Dept. Marketing, MKTG 98.116, 1998.

[13] R. Selten and W. Guth, "Equilibrium point selection in a class of market entry games," in *Games, Economic Dynamics, and Time Series Analysis*. Vienna, Austria: Physica-Verlag, 1982, pp. 101–116.

[14] W. A. Sethares, "Adaptive algorithms with nonlinear data and error functions," *IEEE Trans. Signal Processing*, vol. 40, pp. 2199–2206, Sept. 1992.

[15] W. A. Sethares and C. R. Johnson Jr., "A comparison of two quantized state adaptive algorithms," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 37, pp. 138–143, Jan. 1989.

[16] W. A. Sethares, I. M. Y. Mareels, B. D. O. Anderson, and C. R. Johnson Jr., "Excitation conditions for sign-regressor LMS," *IEEE Trans. Circuits Syst.*, vol. 35, pp. 613–624, June 1988.

[17] J. Sundali, A. Rapoport, and D. Seale, "Coordination in market entry games with symmetric players," *Organiz. Beh. Human Decision Process.*, vol. 64, no. 2, pp. 203–218, 1995.

[18] W. S. Vickrey, "Congestion theory and transport investment," *Amer. Econ. Rev.*, no. 59, pp. 251–260, 1969.

[19] B. Widrow, J. M. McCool, M. G. Larimore, and C. R. Johnson Jr., "Stationary and nonstationary learning characteristics of the LMS adaptive filter," *Proc. IEEE*, vol. 64, pp. 1151–1162, Aug. 1976.

[20] G. A. Williamson, P. Clarkson, and W. A. Sethares, "Performance characteristics of the adaptive median LMS filter," *IEEE Trans. Signal Processing*, vol. 41, pp. 667–680, Feb. 1993.

A. M. Bell received the B.A. degree from Cornell University, Ithaca, NY, and the M.A. and Ph.D. degrees from the University of Wisconsin, Madison, all in economics.

She is currently with Orbital Sciences doing research in Advanced Life Support at the NASA Ames Research Center, Moffet Field, CA, and for the non-profit organization Redefining Progress, where she analyzes the implications of new models of complex systems for public policy. She has also been an Assistant Professor of Economics at Vanderbilt University, Nashville, TN, and a Visiting Assistant Professor at the University of California, Santa Cruz.

W. A. Sethares received the B.A. degree in mathematics from Brandeis University, Waltham, MA and the M.S. and Ph.D. degrees in electrical engineering from Cornell University, Ithaca, NY.

He has worked at the Raytheon Company as a Systems Engineer and is currently Associate Professor in the Department of Electrical and Computer Engineering, University of Wisconsin, Madison. His research interests include adaptation and learning in signal processing, communications, and acoustics, and he is author of *Tuning, Timbre, Spectrum, Scale* (New York: Springer, 1998). His website at http://eceserv0.ece.wisc.edu/sethares/ contains further information about the *El Farol* problem and other fascinating topics.