To appear in: *Complexity*

Non-Explanatory Equilibria

An Extremely Simple Game
With (Mostly) Unattainable Fixed Points

Joshua M. Epstein[1] and Ross A. Hammond[2]

August, 2001

ABSTRACT

Equilibrium analysis pervades mathematical social science. This paper calls into question the explanatory significance of equilibrium by offering an extremely simple game, most of whose equilibria are unattainable in principle from any of its initial conditions. Moreover, the number of computation steps required to reach those (few) equilibria that are attainable is shown to grow exponentially with the number of players—making long-run equilibrium a poor predictor of the game's observed state. The paper also poses a number of combinatorially challenging problems raised by the model.

[1] *Economic Studies Program, The Brookings Institution and External Faculty, Santa Fe Institute*
[2] *Department of Political Science, University of Michigan*

ACKNOWLEDGMENT: For insightful comments, the authors thank Robert Axtell, Jim Crutchfield, William Dickens, Scott Page, Duncan Watts, and Peyton Young. This research was conducted at The Brookings Institution-Johns Hopkins University Center on Social and Economic Dynamics.

Much of game theory and mathematical economics is concerned with equilibria[1]. Nash equilibrium is an important example. Indeed, in many quarters, "explaining an observed social pattern" is understood to mean "demonstrating that it is the Nash equilibrium of some game." But, there is no explanatory significance to an equilibrium that is unattainable in principle. And there is debatable significance to equilibria that are attainable only on astronomical time scales. Yet, in a great many instances, the social pattern to be explained is simply shown to be an equilibrium. The questions, "Is the equilibrium attainable?" and "On what time scale is it attainable?" are not raised.

There is a literature on unattainability—or uncomputability—of equilibria, undecidability in games, and related topics. But it is quite technical[2]. The aim of the present paper is to offer an extremely simple game—easily played by school children—that drives home the core distinction between attainable and unattainable equilibria. Indeed, the overwhelming preponderance of this game's equilibria are unattainable from any initial configuration of the game.

We hope this arresting example stimulates skepticism about the explanatory significance of equilibrium[3]. As we will show, the game—despite its surface simplicity—also raises a number of combinatorially very challenging questions.

Description of the Game

The game's ingredients are few and simple:

[1] Events transpire on a linear array of sites, extending from an origin (the leftmost site) to the right.

[2] Agents are numbered consecutively from 1 to n. These numbers do not change in the course of the game.

---

[1] See Kreps (1990), p. 405

[2] See, for example, Foster and Young (2001), Saari and Simon (1978), Prasad (1991), Jordan (1993), and Nachbar (1997)

[3] For an insightful discussion of this issue in the context of chaos and evolutionary games, see Skyrms (1997)

[3] Initially, we require that agents be arrayed in a contiguous row, beginning at the origin, in some arbitrary order. Figure 1 gives one such admissible initial configuration for three agents. Each agent is represented as a number, and each empty site is represented as an asterisk.

<u>Figure 1. An Initial 3-Agent Configuration</u>

3 2 1 * * *

[4] The agents' only rule of behavior is as follows:

AGENT RULE: If there is a lower-numbered agent anywhere to your right, go to the head of the line (the site immediately to the right of the rightmost agent).

The rule is reminiscent of the Schelling segregation model Schelling (1971, 1978) and the variant of Young (1998)[4]. In each case, agents have some preference for immediate neighbors. In our case, agents hate living anywhere with a lower-numbered agent to their right. And (with bounded rationality) they move to the one site that is certain to remove the problem, at least in the immediate term— the front site.

[5] In any given round, agents are queried in order from highest numbered to lowest numbered. As we shall see below, not all agents may wish to move. The first agent who does wish to move does so, resulting in a new configuration. That ends the round. Play continues until *equilibrium* is reached, where:

[6] An *equilibrium* is a configuration from which no agent would move further under the rule. It is a fixed point. An equilibrium is termed *attainable* if there is some *initial* configuration (see under [3]) from which it can be attained. An *unattainable* equilibrium is an equilibrium for which no such *initial* configuration exists.

That is the complete model specification

[4] Note, however, that the model is not a Cellular Automaton, because it involves a non-local operation (agents go to the head of the line, and are queried in sequence order). We thank Jim Crutchfield for this observation. On Cellular Automata, see: Wolfram (1986), and Toffoli and Margolus (1987).

Child's Play

One can imagine the model as a children's game, played on a linear sequence of hopscotch squares. Assume the kids differ by height. They form a line extending out from the school wall into the playground, one in front of the next, in some random order by height. Then they move, as specified under [5] above, each according to the simple rule:

If there's a shorter kid anywhere in front of you, jump to the very head of the line (the square immediately in front of the front kid).

The game ends when equilibrium is attained— when no kid would move further under the rule[5]. (This equilibrium notion is Nash-like: no agent has any incentive to unilaterally depart under the rule).

A Numerical Example

As a simple illustration of how the configurations progress, let us walk the game forward from the Figure 1 configuration.

Table 1. A Complete Game

Configuration Number	Configuration
1	3 2 1 * * *
2	* 2 1 3 * *
3	* * 1 3 2 *
4	* * 1 * 2 3

Starting in Configuration 1, agent "3" (the highest numbered) is queried first. Since there is a lower numbered agent to her right, she jumps to the rightmost position— leaving a space in her former position— yielding Configuration 2. That ends the round. So, we begin a new round. As before, we query agent "3" (the highest numbered) first. This time, she declines to change position. So, we query the next highest numbered agent: "2." Since there is a lower numbered agent to her right, she now jumps to the rightmost position— leaving a space in her former spot—

[5] In effect, the kids have invented a type of decentralized (albeit highly inefficient) sorting algorithm.

yielding Configuration 3. This, of course, upsets agent “3”, who moves when queried at the beginning of the next round, and so it goes. In Configuration 4, agent 3 does not wish to move, so agent 2 is queried. She declines, so agent 1 is (at last) queried, but declines as well (as the lowest numbered agent always does). Configuration 4 is therefore an equilibrium. It is obviously attainable. Notice that it requires 6 spaces in total.

Space and Time Requirements for Attainable Equilibria

For n agents, how many spaces are required to ensure enough space for all attainable equilibria? Perhaps surprisingly, the answer is:

[1] $$s_{\max}(n) = n + \sum_{i=0}^{n-2} 2^i = (n-1) + 2^{(n-1)}$$

This space requirement grows exponentially in n. Values of $s_{\max}(n)$, for various n values are given in Table 2.

Table 2. Maximum Space Requirements for Attainable Equilibria, Various n

n	Sites
3	6
4	11
5	20
20	524,307
25	16,777,240
30	5.37×10^{8}
50	5.63×10^{14}
100	6.34×10^{29}

Regarding time (i.e., number of computation steps), the equilibrium of Table 1 required 3 rounds to compute, from the initial configuration 321***. In general, equilibria occupying

$s_{\max}(n)$ (as in equation 1) spaces will be obtainable in $s_{\max}(n)-n$ rounds, which, quite notably, is also exponential in *n*. Daunting numerical examples are left to the reader.

As prosaic examples with kids, assume each hopscotch square is 2 feet deep, and that the games begin on a playground in Cambridge, Massachusetts. Then, for 20 kids (an average kindergarten class), there are initial line-ups such that, when (after 524,287 moves) equilibrium is attained, the tallest kid is standing in Central Park. For 25 kids, there are initial line-ups such that, when (after about 17 million moves) equilibrium is attained the tallest kid is standing in Tokyo. For 30 kids, there are initial line-ups such that, when (after more than 500 million moves) equilibrium is attained, the tallest kid has circumnavigated the earth ten times. For 50 players, there are attainable equilibria extending over roughly 563 trillion sites. And for games involving 100 agents— a standard population size in the literature of n-person games and agent-based models— even the set of *attainable* equilibria is uncomputable on all practical time scales. And, in fact, most equilibria are unattainable in principle.

Unattainable Equilibria

A full treatment of the n=3 case will be instructive. There are 3! acceptable initial configurations, and 5 distinct attainable equilibria, as shown in Table 3.

Table 3. The 5 Attainable Equilibria for n=3

Initial Configuration	Resulting Equilibrium
123***	123***
132***	1*23**
231***	**1*23
213***	*1*23*
312***	*123**
321***	**1*23

Notice that the equilibrium **1*23 is attainable from the initial configurations: 231*** and 321***. In general, a given attainable equilibrium may be attainable from multiple initial configurations[6].

Unattainable Equilibria

While (as shown in Table 3) there are 5 distinct attainable equilibria for the n=3 case, there are 20 equilibria in total (see equation 2). *Ipso facto*, there are 15 unattainable equilibria! They are listed below:

Table 4. The 15 Unattainable Equilibria for n=3

1.	1*2**3
2.	*12**3
3.	1**2*3
4.	*1*2*3
5.	**12*3
6.	1***23
7.	*1**23
8.	***123
9.	12*3**
10.	12**3*
11.	1*2*3*
12.	*12*3*
13.	1**23*
14.	12***3
15.	**123*

In each of these configurations, every agent is happy with her immediate neighborhood, but none of these configurations are attainable from any initial configuration.

[6] In this connection, the reader might find it interesting to consider the following general problem: Give a formula, f(n), for the number of *distinct* equilibria attainable from the n! distinct initial configurations of the n-agent game.

For n=3, then, unattainability is the norm among equilibria. This pattern only gets more dramatic as n increases. Indeed, the ratio of attainable to unattainable equilibria approaches zero very quickly. For n=4, there are 330 equilibria, of which 12 are attainable, a mere 4%. For n=5, there are 15,504 equilibria, of which 41 are attainable, or 0.2%. For n>5, the attainable percentage is effectively zero.

The formula for the total number of equilibria, T(n) , even for the n=4 case, turns out to be quite complex. It is:

$$T(4) = (\beta_4 + 1) + \sum_{i=1}^{\beta_4} i + \sum_{i=1}^{\beta_4}\sum_{j=1}^{i} j + \sum_{i=1}^{\beta_4}\sum_{j=1}^{i}\sum_{k=1}^{j} k,$$

where $\beta_4 = s_{\max}(4) - 4 = \sum_{i=0}^{2} 2^i = 7$[7]. For n agents, the appropriate generalization is as follows. First, the index variables will run from v_1 to v_{n-1}.Then,

[2] $$T(n) = (\beta_n + 1) + \sum_{v_1=1}^{\beta_n} v_1 + \sum_{v_1=1}^{\beta_n}\sum_{v_2=1}^{v_1} v_2 + \ldots + \sum_{v_1=1}^{\beta_n}\sum_{v_2=1}^{v_1} \cdots \sum_{v_{n-1}=1}^{v_{n-2}} v_{n-1},$$

where $\beta_n = s_{\max}(n) - n = \sum_{i=0}^{n-2} 2^i$, as before.

Now, for n agents, the number of distinct initial configurations is n!, but the number of attainable equilibria is less than n! (as illustrated in Table 4). Hence, the fraction of attainable to total is bounded above by $\frac{n!}{T(n)}$. Since $\frac{n!}{T(n)} \to 0$ extremely fast, so does the fraction of attainables. Hence the generic equilibrium is, in fact, unattainable from any initial conditions[8].

[7] A closed form representation of the result would obscure the iterative nature of the solution. Hence, the iterated summations shown.

[8] Whether or not an equilibrium can be easily *diagnosed as* unattainable is beside the point we are making here. But, to discuss this briefly, some cases are clear on inspection. For example, the equilibrium ***123 is unattainable, since the digit “1” never moves (as noted earlier) and appears too far to the right to be permissible initially. Similarly, the equilibrium *1*2*3 can be easily identified as a Garden of Eden configuration (see below), and is therefore not attainable. However, some cases are not so obvious: **123* is unattainable. Now, in *principle*, one can classify equilibria as unattainable by brute force. For each of

Clearly, restricting the space of permissible initial configurations is important to this result. While, at first glance, such restrictions may seem artificial, they are the norm in games and contests generally. Chess, checkers, and many other board games possess required initial set-ups. Straight pool, 9-Ball, and 8-Ball (stripes and solids) each begin with the billiard balls "racked" in a specified way. In racquet sports, such as tennis, squash, and ping-pong, players are not permitted to serve (i.e., begin a point) from "just anywhere." Football prohibits certain line-ups and allows others. Jousts and pistol duels had highly stylized initial positions, as do fencing matches. Further examples will come readily to mind. Indeed, on reflection, some restriction on initial configurations would seem to be the rule across formalized contests, rather than the exception. In this light, our restriction seems natural enough.

Equilibrium and Explanation

Here, then, is an extremely simple playground game that admits a huge number of equilibria, virtually all of which are not attainable from any initial configuration, once there are 5 or more players. So, returning to the central issue of explanatory significance, imagine being a theoretical playgroundologist. Your colleagues, the empirical playgroundologists, have documented a powerful regularity: They observe kids all over the world lined up from shortest to tallest on playgrounds; they are spaced in all sorts of bizarre ways, but they're lined up in order by height. What is the *explanation?* This is the central empirical puzzle of playgroundology.

Now, given an analogous empirical regularity, the standard and ostensibly *explanatory* practice in the formal social sciences is as

---

the n! initial conditions, one simply grinds out the attainable equilibria. Then, for any candidate equilibrium, one "simply" checks—by bitwise comparison—whether it is in the list of attainables or not. However, the number of required comparisons grows exponentially in n. Mechanical "space counting" tests for unattainability, while more direct, nonetheless require inspection of β_n sites, and will be computationally prohibitive in practice for agent populations of any significance.

follows: *Provide a game for which the observed regularity is an equilibrium.*

But, this is easily done for playgroundology— the game we've just been exploring fits the bill. Any line-up from shortest to tallest observed by our empiricists in the field will, indeed, be an equilibrium of this game. As we have shown, however, it will almost certainly *not* be attainable: kids could not have arrived there from any initial line-up. Clearly, then, the rules of this particular game are supremely unlikely to be those followed on real playgrounds.

Nonetheless, under the standard practice above, these rules would be regarded as explanatory! This seems unsatisfactory for playgroundology because the generic equilibrium of the game is not attainable even in principle, much less on time scales of any plausibility. So why, absent demonstrations of attainability, should the same practice be accepted as explanatory in social science? We believe it should not be.

An acceptable notion of "explanation" should include attainability. A candidate is the generative notion advanced in Epstein (1999), in which a set of individual rules, a microspecification, is regarded as explanatory only if it suffices to generate the observed regularity—incorporating the requirement of attainability.

Beyond its explanatory shortcomings, equilibrium may be a bad *predictor* of observed configurations[9]. Obviously, *unattainable* equilibria (since they will *never* be observed) are not predictive of the game's state on any time scale. But even *attainable* equilibria (given the exponential time complexity of the process) are, in almost all cases of this game, poor predictors on time scales of any interest to humans.

Conclusion

For the social sciences more broadly, there would appear to be two lessons of this simple exercise. First, implicit claims that

[9] Explanation and prediction are different matters: plate tectonics *explains* earthquakes, but does not predict when they'll occur. Similarly, electrostatics *explains* lightning, but doesn't predict where it will strike.

equilibrium analysis is explanatory or predictive should be challenged, and require the most careful defense. Second, a successful defense of any such claims must include a demonstration of attainability, on time scales of interest, by agents employing plausible rules[10].

[10] By plausible rules, we have in mind those involving bounded information and bounded individual computing capacity. See Simon (1982).

APPENDIX: FURTHER COMBINATORIAL QUESTIONS

While the playground game was contrived as a stark illustration of these points, it happens to raise a number of interesting combinatorial questions.

Garden of Eden Configurations

First, by way of definition, if there exists *no previous configuration* from which a given configuration can be attained, then the latter is termed a *Garden of Eden* (GE) configuration[11]. For example, the following configuration is GE:

* 1 * 2* 3

If 3 had been located anywhere to the left of 2 (or 1), it would have jumped to the site immediately to 2's right, *not* to the position shown. This is both an equilibrium and a GE state.

We know that there are unattainable equilibria (i.e., unattainable from any admissible *initial* configurations). Now, for many of these, there *are prior* configurations. So, beginning with such an unattainable equilibrium, if we back-calculate, we must stop short of the origin (i.e., the set of permissible initial configurations) since otherwise the equilibrium would have been attainable. Where we stop must therefore be a Garden of Eden configuration! So,

Proposition: For every non-GE unattainable equilibrium, there exists (at least one) GE non-equilibrium preceding configuration.

For example, consider the string: 1**23*. It is an equilibrium. But it is not attainable from any permitted initial condition. The non-equilibrium configurations from which it is derivable, however, are: 1*32** and 13*2**, both of which are GE, since neither one has a predecessor that could occur initially.

[11] According to E.F. Moore (1962), this term was first suggested by John Tukey.

The set of GE configurations from which a given configuration is attainable shall be referred to as its *basin of attraction*. Naturally, this suggests the following (evidently hard) question: For any equilibrium configuration *not* attainable from an initial configuration, determine its basin of attraction. Or, since all initial conditions are themselves GE, the general problem is simply:

Problem 1. For any equilibrium (attainable or not), determine its basin of attraction.

In pondering the computational complexity of this general problem, bear in mind that even for n=50 players there are many unattainable equilibria consuming 563 trillion sites— in general, $s_{\max}(n)$ sites.

For the sake of completeness, it would be of further interest to solve the following:

Problem 2. From each "point" of a given equilibrium's basin, how many computation steps are required to attain the equilibrium?

References

Epstein, Joshua M. 1999. "Agent-Based Computational Models and Generative Social Science." *Complexity*. 4: 5, 41-60.

Foster, Dean P. and H. Peyton Young. 2001. "On the Impossibility of Predicting the Behavior of Rational Agents." *Santa Fe Institute Working Paper* 01-08-039.

Jordan, James S. 1993. "Three Problems in Learning Mixed-Strategy Equilibria" *Games and Economic Behavior*. 5:368-386.

Kreps, David M. 1990. *A Course in Microeconomic Theory*. Princeton, NJ: Princeton University Press.

Moore, Edward F. 1962. "Machine Models of Self-Reproduction." In *Mathematical Problems in the Biological Sciences—Proceedings of Symposia in Applied Mathematics, Vol. XIV*. Providence: American Mathematical Society.

Nachbar, John H. 1997. "Prediction, Optimization, and Learning in Games." *Econometrica*. 65:275-309.

Prasad, K. 1997. "On the Computability of Nash Equilibria.' *Journal of Economic Dynamics and Control*. 21:943-953.

Saari, Donald G. and Carl P. Simon. 1978. "Effective Price Mechanisms". *Econometrica* 50:1097-1125

Schelling, Thomas C. 1971. "Dynamic Models of Segregation." *Journal of Mathematical Sociology* 1:143-86.

______. 1978. *Micromotives and Macrobehavior*. New York: Norton.

Simon, Herbert. 1982. *Models of Bounded Rationality*. Cambridge, MA: MIT Press.

Skyrms, Brian. 1997. "Chaos and the explanatory significance of equilibrium: Strange attractors in evolutionary game dynamics." In Bicchieri, C., R. Jeffrey and B. Skyrms (eds) 1997. *The Dynamics of Norms* NewYork: Cambridge University Press.

Toffoli, T. and N. Margolus. 1987. *Cellular Automata Machines: A New Environment for Modeling*. Cambridge: MIT Press.

Wolfram, Stephen. 1986. *Theory and Applications of Cellular Automata*. Singapore: World Scientific.

Young, H. Peyton. 1998. *Individual Strategy and Social Structure – An Evolutionary Theory of Institutions*. Princeton: Princeton University Press.