

Equilibrium Selection via Adaptation: Using Genetic Programming to Model Learning in a Coordination Game*

Shu-Heng Chen
AI-ECON Research Center
Department of Economics
National Chengchi University
Taipei 11623, Taiwan
chchen@nccu.edu.tw

John Duffy
Department of Economics
University of Pittsburgh
Pittsburgh, PA 15260 USA
jduffy@pitt.edu

Chia-Hsuan Yeh
Department of Information Management
Yuan Ze University
Chungli, Taoyuan 320, Taiwan
imcyeh@saturn.yzu.edu.tw

Abstract

This paper models adaptive learning behavior in a simple coordination game that Van Huyck, Cook and Battalio (1994) have investigated in a controlled laboratory setting with human subjects. We consider how populations of artificially intelligent players behave when playing the same game. We use the genetic programming paradigm, as developed by Koza (1992, 1994), to model how a population of players might learn over time. In genetic programming one seeks to breed and evolve highly fit computer programs that are capable of solving a given problem. In our application, each computer program in the population can be viewed as an individual agent's forecast rule. The various forecast rules (programs) then repeatedly take part in the coordination game evolving and adapting over time according to principles of natural selection

*This project was initiated while Duffy was visiting National Chengchi University. A preliminary version of this paper, Chen, Duffy and Yeh (1996), was presented at the 1996 Evolutionary Programming Conference.

and population genetics. We argue that the genetic programming paradigm that we use has certain advantages over other models of adaptive learning behavior in the context of the coordination game that we consider. We find that the pattern of behavior generated by our population of artificially intelligent players is remarkably similar to that followed by the human subjects who played the same game. In particular, we find that a steady state that is theoretically unstable under a myopic, best-response learning dynamic turns out to be stable under our genetic-programming-based learning system, in accordance with Van Huyck et al.'s (1994) finding using human subjects. We conclude that genetic programming techniques may serve as a plausible mechanism for modeling human behavior, and may also serve as a useful selection criterion in environments with multiple equilibria.

AMS Subject Classifications. C63, D83.

1 Introduction

The empirical usefulness of static equilibrium analysis is compromised when economic models have multiple equilibria. Consequently, extensive efforts have been made to identify ways of reducing the set of equilibria that are focal in models with multiple equilibria. There seems to be some consensus emerging that a sensible selection criterion for choosing among multiple equilibria is to determine which of the candidate equilibria are stable with respect to some kind of disequilibrium, "learning" dynamic.¹ A number of such learning dynamics have been proposed and used to reduce or eliminate multiple equilibria as empirically relevant candidates. However, the notion that these learning dynamics accurately reflect the behavior of individual economic agents or groups of agents has only very recently begun to be examined through a number of controlled laboratory experiments with human subjects.²

This paper focuses on results obtained from one such experiment conducted by Van Huyck, Cook and Battalio (1994) that tested the predictions of a class of selection dynamics in a generic coordination game against the behavior of human subjects who played the same coordination game. Van Huyck et al. postulated that one of two candidate learning processes could describe the behavior of human subjects playing the coordination game. The first learning process is a Cournot-type, myopic best-response dynamic, and the second is an inertial learning algorithm that allows for slowing changing beliefs.³ Both learning models are special versions of a large class of relaxation algorithms that have frequently

¹For references, see, e.g. the surveys by Kreps (1990), Sargent (1993), and Marimon (1997).

²For a survey of some of these experiments, see, e.g., Kagel and Roth (1995).

³Van Huyck et al. (1994) refer to this inertial learning dynamic as the "L-map" which is a reference to Lucas' (1986) use of this type of learning dynamic.

appeared in the learning literature.⁴ Under certain parameterizations, these two learning processes yield different predictions for the stability of one of the game's two Nash equilibria. Van Huyck et al.'s (1994) experimental results suggest that in those parameterizations where the two learning algorithms yield different predictions, the inertial learning algorithm provides a better characterization of the behavior of human subjects in the coordination game than the myopic best response dynamic.

In this paper, we adopt a computational approach, using Koza's (1992, 1994) genetic programming techniques to model the behavior of artificial economic agents playing the same simple coordination game that was studied by Van Huyck et al. (1994). The computational approach that we take to model agent behavior allows for a considerably more flexible experimental design than is possible with experiments involving human subjects. Moreover, unlike most rule-based models of adaptive learning behavior, the artificial players in our genetic programming implementation of the coordination game are explicitly endowed with the ability to "think" nonlinearly, and are given all the "building blocks" necessary to construct a vast array of both linear and nonlinear forecasting rules including the myopic best response and the inertial learning algorithms considered by Van Huyck et al. (1994). Thus we know, at the outset, that our artificial players are capable of both choosing and coordinating upon linear or nonlinear forecasting rules that may result in stationary, periodic or aperiodic trajectories. We find that our more general computational approach to modeling learning behavior in the coordination game results in behavior that is qualitatively similar to that of the subjects in Van Huyck et al.'s (1994) coordination game experiment. Indeed, we think of our genetic programming implementation of learning in the coordination game as a kind of robustness check on the experimental results reported for the same game. Finally, we argue that the genetic programming techniques we illustrate in this application have certain advantages over other artificial intelligence techniques that have been applied to economic models, namely, genetic algorithms.

The coordination game found in Van Huyck et al. (1994) differs from previous coordination games that have been studied experimentally (e.g. Cooper, DeJong, Forsythe and Ross (1990) and Van Huyck, Battalio and Beil (1990, 1991)) in that 1) the set of agent actions is considerably larger (indeed, there can be a continuum of possible actions), and 2) the stability of one of the game's two equilibria cannot be ascertained *a priori*. The first difference makes it difficult to formulate and enumerate strategies that are based upon all possible actions as is often done in adaptive learning models. The second difference arises because Van Huyck et al. (1994) entertain the notion that agents might adopt nonlinear rules to choose actions. Because of these differences, learning models that have been used to

⁴The class of relaxation algorithms includes, for example, the past averaging algorithm of Bray (1982) and Lucas (1986), and the least squares learning algorithm of Marcet and Sargent (1989).

explain behavior in the early coordination game experiments, for example, the linear learning models of Crawford (1991, 1995) and the genetic algorithm approach of Arifovic (1997) are not as well suited to the coordination game environment studied by Van Huyck et al. (1994). By contrast, we argue that the genetic-programming approach that we take to modeling learning behavior is particularly well suited to the coordination game environment of Van Huyck et al. (1994). We now turn to a description of this coordination game.

2 The Coordination Game

Consider the generic coordination game $\Gamma(\omega)$, studied by Van Huyck et al. (1994). There are n players, each of whom chooses some action $e^i \in [0, 1]$, $i = 1, 2, \dots, n$. The individual player i 's payoff function in every round of play, t , is described by:

$$\pi_{i,t}(e^i, e^{-i}) = c_1 - c_2|e^i - \omega M_t(e)[1 - M_t(e)]|, \quad (1)$$

where c_1 and c_2 are constants, $M_t(e)$ denotes the median of all n players' actions in round t , $\omega \in (1, 4]$ is a given parameter and e^{-i} denotes the vector of actions taken by the other $n - 1$ players in the same round. Both the payoff function and the set of feasible actions are assumed to be common knowledge.

It is clear from the structure of the payoff function that the individual player in this game should seek to minimize the expression that lies between the absolute value signs. That is, for a given value of the median, M , the individual player's best response function $b(M)$, is:

$$b(M) = \omega M(1 - M).$$

This best response function gives rise to two Nash equilibria: a corner equilibrium, where $e^i = 0$ for all i , and an interior equilibrium where $e^i = 1 - (1/\omega)$ for all i . The best response function $b(M)$ is easily recognized to be a member of the family of quadratic maps, where the degree of curvature is determined by the tuning parameter ω .

3 Selection Dynamics

Van Huyck et al. (1994) suggested that a certain class of "relaxation algorithms" that are frequently encountered in the learning literature could be used to characterize the evolution of play of this coordination game over time. This class of relaxation algorithms is described by the simple dynamical system:

$$\begin{aligned} M_t &= b(\widehat{M}_t), \\ \widehat{M}_t &= \widehat{M}_{t-1} + \alpha_t(M_{t-1} - \widehat{M}_{t-1}), \end{aligned}$$

where \widehat{M}_t is the representative agent's *expected* value for the median at time $t > 1$, and $\alpha_t \in [0, 1]$ is a given forgetting factor. Van Huyck et al. (1994) consider two specific versions of this relaxation algorithm: 1) a "myopic" best response algorithm where $\alpha_t = 1$ for all $t > 1$, and 2) an "inertial" algorithm, where $\alpha_t = 1/t$ for all $t > 1$.

The myopic best response version of the dynamical system gives rise to a simple first-order difference equation that characterizes the evolution of the median over time:

$$M_t = \omega M_{t-1}(1 - M_{t-1})$$

It is easily shown that for $1 < \omega < 3$, the interior equilibrium, $1 - (1/\omega)$ is attracting (locally stable) while the corner equilibrium, 0, is repelling (locally unstable) under the myopic best response dynamics. However for $\omega > 3$, the dynamics of the myopic best response algorithm become increasingly more complicated, resulting in a dense set of periodic trajectories for the median that follows the Sarkovskii order. When $\omega = 3.839$, the myopic best response algorithm gives rise to a stable cycle of period 3, which according to the famous theorem of Li and Yorke implies that there are cycles of all periods and an uncountable set of non-periodic (chaotic) trajectories.⁵ Thus, for $\omega > 3$, the interior equilibrium is no longer stable under the myopic best response dynamics.

The inertial version of the relaxation algorithm gives rise to the dynamical system:

$$M_t = b(\widehat{M}_t)$$

$$\widehat{M}_t = \frac{t-1}{t} \widehat{M}_{t-1} + \frac{1}{t} M_{t-1}$$

Note that the inertial learning algorithm differs from the myopic best response algorithm in that the inertial algorithm gives most weight to the previous *expected* value of the median whereas the myopic best response algorithm gives all weight to the previous *realized* value of the median. It is easily shown that for all feasible values for ω ($\omega \in (1, 4]$), the interior equilibrium $1 - (1/\omega)$ is a global attractor under the dynamics of the inertial learning algorithm.

Thus if $1 < \omega \leq 3$, both the myopic best response and the inertial learning algorithms predict that the interior equilibrium $1 - (1/\omega)$ will be the equilibrium that agents eventually coordinate upon. However, for $3 < \omega \leq 4$, the myopic best response algorithm predicts that the interior equilibrium will be unstable, whereas the inertial learning algorithm predicts that the interior equilibrium will continue to be stable.

⁵For a detailed analysis of the first-order difference equation, $M_t = \omega M_{t-1}(1 - M_{t-1})$, that characterizes the myopic best response dynamic see, e.g. Devaney (1989).

4 Experimental Results and Experimental Design

Van Huyck et al. (1994) considered two experimental versions of the coordination game, $\Gamma(\omega)$, described above. In one version of the game, $\Gamma(2.4722)$, the interior equilibrium is predicted to be stable under both the myopic best response and inertial learning dynamics based on the choice of $\omega < 3$. In a second version of game, $\Gamma(3.86957)$, the interior equilibrium is predicted to be *unstable* under the myopic best response dynamics; starting from any initial condition, the myopic best response algorithm results in a chaotic trajectory for the median. By contrast, under the inertial learning dynamics, the interior equilibrium in the game, $\Gamma(3.86957)$, is predicted to remain stable since $\omega \leq 4$. Thus, the second game, $\Gamma(3.86957)$, is the more interesting one, as the stability predictions of the myopic best response and inertial learning dynamics differ for this particular game.

Van Huyck et al. (1994) report results from 2 experimental sessions of $\Gamma(2.4722)$ and 6 experimental sessions of $\Gamma(3.86957)$ using 5 subjects in each session. In all eight sessions they found that almost all subjects quickly coordinated on the interior equilibrium; that is, the interior equilibrium is judged to be stable in all treatments. The authors thus conclude that the inertial learning algorithm is a better selection device in the coordination game than the myopic best response algorithm, since the prediction of the inertial learning algorithm regarding the stability of the interior equilibrium is always consistent with the experimental findings.

Van Huyck et al.'s conclusion that the inertial learning algorithm serves as a reasonable learning model/selection criterion is subject to some criticism. First, while it is true that the inertial learning algorithm converges to the interior equilibrium in the game, $\Gamma(3.86957)$ (whereas myopic best response does not), the convergence trajectory taken by this algorithm is much too smooth when compared with the evolution of the median in the human subject experiments (see the experimental data reported in Appendix B of Van Huyck et al. (1994)). A second, related criticism is that it is apparent from the experimental data that the players in the coordination game do not all use the same learning scheme. If they all did use the same scheme, then for the same sequence of values for the median, we should expect to observe the same actions being taken. However, we observe players taking many different actions, especially in the early stages of the experiment, which suggests that they do not hold identical expectations. For this reason, it seems necessary to look beyond the predictions of representative agent learning models and to consider instead the performance of heterogeneous, multi-agent learning models. Our genetic-programming-based learning model is an example of this kind of multi-agent approach.

We note also that in implementing the coordination game experimentally, Van Huyck et al. (1994) made the simplifying assumption that the action set, e^i consists of only a *finite* set of *discrete* choices; players were asked to choose an action e^i from the set of integers $\{1, 2, \dots, 90\}$. Each subject's action was then mapped into the unit interval using the function $f(e^i) = (90 - e^i)/89$. The discreteness of the action set, however, leads to some rather dramatic changes in the analysis

of the myopic best response dynamics for the interesting case where $\omega > 3$. First, the discreteness of the action set rules out the possibility of chaotic trajectories which require the continuum of the unit interval. Indeed, the restriction that the median take on one of 90 values implies that the median must repeat itself at least once every 91 periods. Second, the discreteness of the action set also leads to the possibility that the interior equilibrium of the game, $\Gamma(3.86957)$, is locally *stable* under the myopic best response dynamics. In particular, the stability of the interior equilibrium of the discrete choice coordination game, $\Gamma(3.86957)$, now depends crucially on the initial condition, i.e. the first median value M_1 . For almost all feasible values for $M_1 \in \{1, 2, \dots, 90\}$, the myopic best response dynamics for the game, $\Gamma(3.86957)$, converge to a stable *seven* cycle, implying that the interior equilibrium is unstable. However, for some initial values the interior and the corner equilibria can also be locally *stable* under the discrete choice, myopic best response dynamics.⁶

In the genetic programming implementation of the coordination game that we explore in this paper, we do not have to discretize the action set. The computer programs that we evolve are all capable of choosing actions on the continuum of the unit interval. Therefore, unlike the experimental implementation of the coordination game we do not rule out the possibility of chaotic trajectories. Moreover, by allowing a continuum for the set of feasible actions, the coordination problem faced by our artificial agents is much more complicated than that faced by the experimental subjects whose actions were limited to a finite, discrete choice set. Finally, we consider a much larger size population of players than is practically feasible in an experiment with human subjects. This larger population size should, again, make the coordination problem more difficult. Thus, our genetic programming implementation of learning in the coordination game can be viewed as a check of whether the experimental results are robust to a continuous action set with a large number of players that would be difficult to implement in an experiment involving human subjects.

5 Genetic Programming

Before describing how we model agent behavior in the coordination game using genetic programming techniques, we first provide a brief overview of genetic programming.

⁶In particular, for $M_1 \in \{24, 67\}$, the interior equilibrium is locally stable under the discrete choice, myopic best response dynamics, and for $M_1 \in \{1, 90\}$, the corner equilibrium is locally stable under these same dynamics. It is interesting to note that in one of Van Huyck et al.'s 6 treatments of the game $\Gamma(3.86957)$ - session 7 - the initial median was 24. With this value for M_1 , the discrete choice, myopic best response dynamic would predict that the system would stay at 24, the interior equilibrium forever, and indeed, this is roughly what occurred. See figure 13 of Van Huyck et al. (1994). Thus, one cannot dismiss altogether the possibility that discrete choice, myopic best response dynamic might also characterize the behavior of the experimental subjects in the game, $\Gamma(3.86957)$.

5.1 An overview

Genetic programming (GP) represents a new field in the artificial intelligence literature that was developed only recently by Koza (1992, 1994) and others.⁷ GP belongs to a class of evolutionary computing techniques based on principles of population genetics. These techniques combine Darwin's notion of natural selection through survival of the fittest with naturally occurring genetic operations of recombination (crossover) and mutation. Genetic programming techniques have already been widely applied to engineering type optimization problems (both theoretical and commercial), but have seen comparatively little application to economic problems, which are often similar in nature. The few economic applications of GP thus far include Allen and Karjalainen (1999), Chen and Yeh (1997a,b), Dworman et al. (1996) and Neely et al. (1997).

While GP techniques are often viewed as an offshoot of Holland's (1975) genetic algorithm (GA), they are perhaps more accurately viewed as a *generalization* of the genetic algorithm. The standard genetic algorithm operates on a population of structures, usually strings of bits. Each of the members of this population, the individual *bitstrings*, represents different candidate solutions to a well-defined optimization problem. The genetic algorithm evaluates the fitness of these various candidate solutions using the given objective function of the optimization problem and retains solutions that have, on average, higher fitness values. Operations of crossover (recombination) and mutation are then applied to some of these more fit solutions as a means of creating a new "generation" of candidate solutions. The whole process is repeated over many generations, in order to evolve solutions that are as close to optimal as possible. In analyzing the evolution of solutions over time, it is typical to report the solution in each generation that has the highest fitness value – this solution is designated the "best-of-generation" solution. The algorithm is ended when this best-of-generation solution satisfies a certain criterion (e.g. some tolerance) or after some maximum number of iterations has been reached.

Theoretical analyses of genetic algorithms suggest that they are capable of quickly locating regions of large and complex search spaces that yield highly fit solutions to optimization problems. That is because the genetic operators of the GA work together to optimize on the trade-off between discovering new solutions and using solutions that have worked well in the past (Holland (1975)).⁸

Koza's idea in developing genetic programming techniques was to take the genetic algorithm a step further and ask whether the same genetic operators used in GAs could be applied to a population of *computer programs* so as to evolve highly fit computer programs. There are several advantages to using computer

⁷See also Kinnear (1994).

⁸For an introduction to the theory of genetic algorithms, see, e.g. Goldberg (1989) or Mitchell (1996). Economic applications of genetic algorithms can be found in the work of Arifovic (1994–97), Arthur et al. (1997), Bullard and Duffy (1998a,b), Dawid (1996), Miller (1996) and Tesfatsion (1997) among others and are also discussed in Sargent (1993) and Birchenhall (1995).

programs rather than bitstrings as the structures to be evolved. First, the computer programs of GP have an *explicit, dynamic structure* that can be easily represented in a *decision tree format*. By contrast, the bitstrings of GAs typically encode passive yes/no type decisions or parameter values for pre-specified, often static functional forms. The dynamic nature of the computer programs of GP makes them capable of much more sophisticated and nonlinear decision-making than in generally possible using the bitstrings of GAs. Second, the computer programs of GP are *immediately implementable structures*; as such, they can be readily interpreted as the forecast rules used by a heterogeneous population of agents. For example, in GP, a computer program used by player i in round t , $gpi_{i,t}$, might take the form:

$$gpi_{i,t} = 0.31 + M_{t-1}(M_{t-1} - M_{t-2}).$$

Here, M_{t-j} represents the value of the median j periods in the past. Given these lagged median values, this program can be immediately executed and delivers a forecast of the median in period t , equal to the value of $0.31 + M_{t-1}(M_{t-1} - M_{t-2})$. This forecast then becomes the action taken by player i in round t . Note that this program is readily interpreted as the agent's *forecast function*. By contrast, the bitstrings used in GAs are not immediately implementable and their interpretation is less clear; these bitstrings first have to be decoded and then the decoded values must be applied to some prespecified functional form before the solution the bitstrings represent can be implemented. Finally, while the length of the bitstrings used in GAs is fixed, the length of the computer programs used in GP is free to vary (up to some limit, of course) providing for a much richer range of structures.⁹

Koza (1992, 1994) chose to develop GP techniques using the Lisp programming language because the syntax of Lisp allows computer programs to be easily manipulated like bitstrings, so that the same genetic operations used on bitstrings in GAs can also be applied to the computer programs that serve as the evolutionary structures in GP. Moreover, the new computer programs that result from application of these genetic operations are immediately executable programs.

Lisp has a single syntactic form, the symbolic expression (S-expression), that consists of a number of *atoms*. These atoms are either members of a *terminal* set, that comprise the *inputs* (e.g. data) to be used in the computer programs, or they are members of a *function* set that consists of a number of pre-specified functions or operators that are capable of processing any data value from the terminal set and any data value that results from the application of any function or operator in the function set. Each Lisp S-expression has the property that it is immediately executable as a computer program, and can be readily depicted as a rooted, point-labeled tree. Moreover, the S-expressions are easily manipulated like data; cutting a tree at any point and recombining the cut portion with another tree (S-expression) results in a new S-expression that is immediately executable.

⁹While in principle it is possible to represent dynamic, variable length expressions using the bitstrings of genetic algorithms, this has not been the practice. See Angeline (1994) for a further discussion.

As Koza and others have noted, the use of Lisp is not necessary for genetic programming; what is important for genetic programming is the implementation of a *Lisp-like environment*, where individual expressions can be manipulated like data, and are immediately executable. For the results reported in this paper, we have chosen to implement the Lisp environment using Pascal 4.0.¹⁰

5.2 Using genetic programming to model learning in the coordination game

In this section, we explain how we use genetic programming to model population learning in the coordination game. The version of genetic programming used here is the *simple genetic programming* that is described in detail in Koza (1992).

Let GP_t , denote a population of trees (S-expressions), representing a collection of players' forecasting functions. A player $i, i = 1, \dots, n$, makes a decision about his action for time t using a *parse tree*, $gp_{i,t} \in GP_t$, written over the *function* and *terminal* sets that are given in Table 1.

Table 1: Tableau for the GP-Based Learning Algorithm

Population size	500
The number of initial trees generated by the full method	250
The number of initial trees generated by the grow method	250
The maximum depth of a tree	17
Function set	{+, -, ×, %, Exp, Rlog, Sin, Cos}
Terminal set	{ℝ, $M_{t-1}, M_{t-2}, M_{t-3}, M_{t-4}, M_{t-5}$ }
The maximum number in the domain of Exp	1,700
The number of trees created by reproduction	50
The number of trees created by crossover	350
The number of trees created by mutation	100
The probability of mutation	0.0033
The probability of leaf selection under crossover	0.5
The maximum number of generations	1,000
Fitness criterion	Payoff function: $\pi_{i,t}$

¹⁰Other programming languages, e.g. C, C++, and Mathematica have also been used to implement Lisp environments.

As Table 1 indicates, the *function set*, includes the standard mathematical operations of addition (+), subtraction (−), multiplication (×) and protected division (%), and also includes the exponential function (*Exp*), a protected natural logarithm function (*Rlog*) and the sin and cosine functions (*Sin* and *Cos*).¹¹ This set of operators and functions is the one that the artificial agents in our experiments are “endowed” with.

The *terminal set* includes the set of constants and variables that the artificial agents may use in combination with the operators and functions from the function set to build forecast rules. As indicated in Table 1, the terminal set includes the random floating-point constant \mathbb{R} which is restricted to range over the interval $[-9.99, 9.99]$, as well as the population *mean* choice of action lagged up to h periods, i.e., M_{t-1}, \dots, M_{t-h} . Note that in our version of the coordination game, M refers to the *mean* rather than the *median* choice of action as in Van Huyck et al. (1994).¹² The choice of the lag length, h , determines players’ ability to recall the past. We set h equal to 5, so that agents may consider as many as 5 past lagged values of the mean in their forecast functions.

The forecasting functions that players may construct and use are linear and nonlinear functions of $M_{t-1}, \dots, M_{t-h}, \mathfrak{R}$, and, as we shall see later, they may also be functions, in whole or in part, of past forecast rules $gp_{i,t}(M_{t-1}, \dots, M_{t-h})$. We note that the set of forecast functions that our artificial players may adopt includes the myopic best response, but not the inertial learning algorithm, as the latter requires knowledge of the previous period’s forecast of the mean value, \hat{M}_{t-1} .

Indeed, Chen and Yeh (1997a) have shown that GP techniques can be used to uncover a variety of nonlinear data generating functions. In one demonstration, they generated a time series for the nonlinear, chaotic dynamical system $x_{t+1} = 4x_t(1 - x_t)$, which is the same as our myopic best response law of motion with $\omega = 4$. They then used a GP-based search in an effort to recover this exogenously given system. Fitness was based on how close the forecast functions in the population

¹¹The protected division operator protects against division by zero by returning the value 1 if its denominator argument is 0; otherwise, it returns the value from dividing its first argument (the numerator) by its second argument (the denominator). Similarly, the protected natural logarithm function avoids non-positive arguments by returning the natural logarithm of the absolute value of its argument, and returning the value 0 if its argument is 0. The exponential function, which takes the argument x and returns the value e^x , allows a maximum argument value of 1,700 as indicated in Table 1. Such function modifications and restrictions are necessary to avoid ill-defined forecasts; these types of modifications are quite standard in the GP literature. See, e.g., Koza (1992).

¹²Van Huyck et al. (1994) used the median rather than the mean in the coordination game because in experiments involving small numbers of human subjects (they only had 5 subjects in each experimental session), the mean can be easily influenced by the behavior of a single subject. By contrast, the computational coordination game experiments that we perform involve hundreds of artificial agents, so that the use of the mean rather than the median is no longer a concern.

came to matching the given time series behavior, and the GP function and terminal sets were nearly identical to those used in this paper. Chen and Yeh (1997a) report that the GP-based search was able to uncover the data generating process in no more than 19 generations. In this paper, by contrast, the data generating process for the mean is *endogenously* determined by the actions chosen by all of the individual players. Nevertheless, it is nice to know that a GP-based search algorithm can deduce a nonlinear data generating function such as the myopic best response law of motion.

The decoding of a parse tree $gp_{i,t}$ gives the forecasting function used by player i at time period t , i.e., $gp_{i,t}(\Omega_{t-1})$ where Ω_{t-1} is the information set containing past mean values through time $t - 1$. Evaluating $gp_{i,t}(\Omega_{t-1})$ at the realization of Ω_{t-1} gives the mean action predicted by player i in round t , i.e., $gp_{i,t}$. Without any further restriction, the range of $gp_{i,t}$ is $(-\infty, \infty)$. However, since the action space for each player is restricted to $[0, 1]$, we must restrict $gp_{i,t}$ so that it also lies in $[0, 1]$. We chose to implement this restriction in two different ways. Our first approach was to use the *symmetric sigmoidal activation function* to map $(-\infty, \infty)$ to $[0, 1]$ so as to obtain a valid mean forecast, $\widehat{M}_{i,t}$, for player i in round t , i.e.

$$\widehat{M}_{i,t} = \frac{1}{1 + e^{-gp_{i,t}}}.$$

A second approach that we also considered was a simple *truncated linear transformation* where player i 's round t forecast was determined as follows:

$$\widehat{M}_{i,t} = \begin{cases} gp_{i,t} & \text{if } 0 \leq gp_{i,t} \leq 1, \\ 1 & \text{if } gp_{i,t} > 1, \\ 0 & \text{if } gp_{i,t} < 0. \end{cases}$$

Using either of these two approaches ensures that player i 's mean choice of action lies in the feasible $[0, 1]$ interval.

Once we have all n players' mean action choices (equivalent to their mean forecasts), it is possible to determine the actual value of the mean in round t , $M_t = \frac{1}{n} \sum_{i=1}^n \widehat{M}_{i,t}$. Given this mean value, we can calculate each player's *fitness value* in round t . The *raw fitness* of a parse tree $gp_{i,t}$ is determined by the value of the player's payoffs earned in round t as determined by the payoff function $\pi_{i,t}$, given in equation (1). To avoid negative fitness values, each raw fitness value is adjusted to produce an *adjusted fitness* measure $\mu_{i,t}$ that is described as follows:

$$\mu_{i,t} = \begin{cases} \pi_{i,t} + 0.25 & \text{if } \pi_{i,t} \geq -0.25, \\ 0 & \text{if } \pi_{i,t} < -0.25. \end{cases}$$

In making this adjustment, we are effectively eliminating from the population forecast functions $gp_{i,t}$ that lose more than \$0.25, since these rules have comparatively lower adjusted fitness values (equal to 0) than rules that did not perform so poorly.

Our decision to make the above adjustment to the fitness measure was due to the following considerations. In the early rounds of a game, players have very limited experience with the environment so their expectations essentially amount to random guessing. As a consequence, many of the players will lose money. If we only considered players with forecast functions that earned positive payoffs, the selection process would quickly come to be dominated by those few players (forecast functions) that were lucky enough to earn positive payoffs in the initial stages of the game. However, we want to maintain some heterogeneity in the population and avoid the possibility of *premature convergence*, a problem that can occur in populations lacking sufficient heterogeneity. For this reason, we allow some players to earn negative payoffs, but we restrict such losses so that they do not exceed \$0.25. After a few generations, when most of the players have begun to earn positive payoffs, this protection no longer plays any effective role. We have experimented with adjustment values other than 0.25. While small adjustment values do not significantly alter our simulation results, very large adjustment values do affect the results because these large values effectively nullify the adjusted fitness measure as an indicator of the relative success of a forecast function. Later in the paper, we examine what happens when we replace the adjustment value of 0.25 in the adjustment scheme described above with the much larger value of 200.00. However, unless otherwise indicated, all of the simulation results we report below involve an adjustment value equal to 0.25.

Once all of the adjusted fitness values are determined, each adjusted fitness value $\mu_{i,t}$ is then normalized. The *normalized fitness* value $p_{i,t}$ is given by:

$$p_{i,t} = \frac{\mu_{i,t}}{\sum_{i=1}^n \mu_{i,t}}.$$

It is clear that the normalized fitness value is a *probability measure*. Moreover, $p_{i,t}$ varies directly with the performance of the parse tree $gp_{i,t}$; the better the parse tree performs (in terms of its payoff), the higher is its normalized fitness value. The normalized fitness values $p_{i,t}$ are used to determine the next generation of agents (parse trees) GP_{t+1} from the current generation GP_t through application of the three primary genetic operators, i.e., *reproduction*, *crossover*, and *mutation*. We now describe these three genetic operators.

a. **Reproduction:**

The reproduction operator makes copies of individual parse trees from generation GP_t and places them in the next generation GP_{t+1} . The criterion used for copying is the normalized fitness value $p_{i,t}$. If $gp_{i,t}$ is an individual in the population GP_t with normalized fitness value $p_{i,t}$, then each time the reproduction operator is called, $gp_{i,t}$ is copied into the next generation with probability $p_{i,t}$. The reproduction operator does not create anything new in the population and the “offspring” generated by reproduction constitute only part of the population of the next generation of trees, GP_{t+1} . As specified in Table 1, the reproduction operator is used to create only 10% (50 out of 500)

of the next generation. The rest of the offspring are generated by the other two operators, *crossover* and *mutation*.

b. Crossover:

The crossover operation for the genetic programming paradigm is a sexual operation that starts with two parental parse trees that have been randomly selected from the population GP_t based upon their normalized fitness values as described above. Crossover involves exchanging different parts of these “parents” to produce two new “offspring.” This exchange begins by randomly and independently selecting a single point on each parental parse tree using a uniform distribution described below. By the syntax of Lisp, each point (atom) of a parse tree could be either a *leaf* (terminal) or a *inner code* (function). Thus, the point (atom) selected for crossover could either be a terminal or a function. As specified in Table 1, the probability that the crossover point is a terminal or a function is the same, i.e., one-half. Given that a terminal or function is to be the point chosen for crossover, the probability that any terminal or function is chosen as the crossover point is uniformly distributed. For example, if the crossover point is to be a terminal, and there are three terminals in the parse tree, the probability that any one of the three terminals is chosen for the crossover point is one-third. Unlike reproduction, the crossover operation adds new individuals (new forecasts rules) to the population. As indicated in Table 1, crossover is responsible for creating 70% (350 out of 500) of the next generation of parse trees, GP_{t+1} .

c. Mutation:

The operation of mutation also allows for the creation of new individuals. The mutation operator begins by selecting a parse tree $gp_{i,t}$ from the population GP_t based once again upon normalized fitness values $p_{i,t}$. Each point (atom) of the selected parse tree is then subjected to mutation (alteration) with a small, fixed probability. As specified in Table 1, this fixed probability of mutation is 0.0033. To ensure that the resulting expression is a syntactically and semantically valid Lisp S-expression, terminals can only be altered to another member from the terminal set and functions can only be altered to another member from the function set possessing the same number of arguments. The altered individual forecast rule (parse tree) is then copied into the next generation of the population. As indicated in Table 1, mutation is responsible for creating 20% (100 out of 500) of the next generation of parse trees.

The three operators combined create the population GP_{t+1} by copying, recombining and mutating the parse trees that make up the population GP_t . Once the new population GP_{t+1} has been created, the decoding of each parse tree $gp_{i,t+1}$ is performed to obtain the new mean, M_{t+1} . Once the new mean is determined, the raw, adjusted and normalized fitness values for each parse tree can be determined using the payoff function (1), and the GP operators can then be applied to create the population GP_{t+2} . The algorithm continues with successive generations, up to

the maximum number of generations. We set the maximum number of generations equal to 1,000 as indicated in Table 1.

The initial S-expressions were randomly generated using both the methods suggested by Koza (1992) – the *full* method and the *grow* method.¹³ Together, these two initialization methods provide for a great diversity of initial programs. Table 1 indicates that each method was responsible for creating one-half (250) of the initial population of trees, GP_1 .

We consider the same two coordination games studied by Van Huyck et al. (1994), although as mentioned previously, we do not restrict the action set to a finite set of discrete choices. Furthermore, we have many more (artificial) players. We refer to the game, $\Gamma(2.4722)$, studied by Van Huyck et al. (1994) as Case 1 and the other game these authors considered, $\Gamma(3.86957)$, as Case 2. The exact parameterizations of these two cases are reported in Table 2.

Table 2: Parameter Values for the Coordination Game Used in the Genetic Programming Simulations

Parameter	Case 1	Case 2
ω	2.47222	3.86957
c_1	0.5	0.5
c_2	1.0	1.0
n	500	500
e_i^*	0	0
e_{ii}^*	0.59551	0.74157

e_i^* : The optimal action under the strict equilibrium $e^i = 0 \forall i$.
 e_{ii}^* : The optimal action under the strict equilibrium $e^i = 1 - (1/\omega) \forall i$.

6 Simulation Results

Our simulation experiments were organized as follows. For each of the two different transformation functions – the symmetric sigmoidal transformation function and the truncated linear transformation function – we conducted 10 simulations for a total of 20 simulations. Within each group of 10 simulations, 5 of the simulations were conducted under the Case 1 parameterization and 5 were conducted under the Case 2 parameterization.

We focus our attention first on the 10 simulations that we conducted using the *symmetric sigmoidal transformation function*. Means and standard deviations from these 10 simulations are reported in Table 3. In this table, simulation 1.1 refers to our first simulation of Case 1, while simulation 2.1 refers to our first simulation

¹³See Koza (1992), pp. 92–93.

of Case 2, and so on. Time series for the mean, M_t , from a single simulation of Case 1 and Case 2 are plotted in Figure 1. These time series plots are typical of the other simulations we conducted for the two cases. As these figures clearly indicate, the time series for M_t in both cases of the GP-based coordination game (Case 1 and Case 2) tend to converge to a neighborhood of the strict interior equilibrium $1 - (1/\omega)$, i.e., 0.59551 for Case 1 and 0.74157 for Case 2. In addition, the transition to $1 - (1/\omega)$ is remarkably brief; If one considers $(0.99 - (1/\omega), 1.01 - (1/\omega))$ as a neighborhood of $1 - (1/\omega)$ then, for all simulations, it takes no more than 50 generations to move into this neighborhood.

A second finding is that while M_t does not converge to $1 - (1/\omega)$ in a strict sense, due to the constant mutation rate, there appears to be a force that serves to stabilize the movement of M_t in a very small band around the interior equilibrium. In other words, GP-based coordination games have a self-stabilizing feature. These properties are also revealed by Table 3.

As Table 3 reveals, in almost all of our simulations, the average of the means, M_t , from generation 201 to 1,000, i.e. \bar{M}_b , does not deviate from the interior equilibrium value, $1 - (1/\omega)$, by more than 0.5%. Note also that if we compare

Table 3: GP Simulation Results Using the Symmetric Sigmoidal Transformation

Case		Simulation				
		1	2	3	4	5
1	\bar{M}_a	0.5917	0.5897	0.5910	0.5907	0.5901
	$\delta_{M,a}$	0.0118	0.0107	0.0091	0.0094	0.0122
	$\delta_{M^*,a}$	0.0124	0.0122	0.0102	0.0105	0.0133
1	\bar{M}_b	0.5958	0.5925	0.5936	0.5933	0.5946
	$\delta_{M,b}$	0.0037	0.0026	0.0033	0.0028	0.0021
	$\delta_{M^*,b}$	0.0038	0.0039	0.0038	0.0035	0.0023
2	\bar{M}_a	0.7406	0.7411	0.7415	0.7450	0.7447
	$\delta_{M,a}$	0.0061	0.0067	0.0069	0.0064	0.0064
	$\delta_{M^*,a}$	0.0061	0.0067	0.0069	0.0073	0.0071
2	\bar{M}_b	0.7394	0.7403	0.7399	0.7434	0.7431
	$\delta_{M,b}$	0.0033	0.0039	0.0040	0.0034	0.0024
	$\delta_{M^*,b}$	0.0039	0.0041	0.0043	0.0039	0.0029

\bar{M}_a = the average of M_t of a simulation from Generation 1 to 1,000.

\bar{M}_b = the average of M_t of a simulation from Generation 201 to 1,000.

$\delta_{M,a}$ = standard deviation about the M_a of a simulation from Generation 1 to 1,000.

$\delta_{M,b}$ = standard deviation about the M_b of a simulation from Generation 201 to 1,000.

$\delta_{M^*,a}$ = standard deviation about the *strict interior equilibrium* $1 - (1/\omega)$ from Generation 1 to 1,000.

$\delta_{M^*,b}$ = standard deviation about the *strict interior equilibrium* $1 - (1/\omega)$ from Generation 201 to 1,000.

the standard deviations, $\delta_{M,a}$ with $\delta_{M,b}$ or $\delta_{M^*,a}$ with $\delta_{M^*,b}$ for each simulation, we see that after the first 200 periods of learning, the stability of the mean in all of the GP-based coordination games improves.

A third result is that the chaotic trajectories for $\Gamma(3.86957)$ that are predicted by the myopic best response dynamic are *not apparent* in any of our simulations of Case 2. However, by comparing $\delta_{M,b}$ or $\delta_{M^*,b}$ across Case 1 and Case 2 in Table 3, we find that the standard deviations in Case 2 are generally somewhat larger than those in Case 1. Indeed, a rank order test reveals that $\delta_{M^*,b}$ is significantly larger in Case 2 as compared with Case 1 ($p \leq .10$).¹⁴ This difference between the two cases is also apparent from a visual comparison between Figure 1. Thus, while it appears that the aggregate outcome from the GP simulations is similar for both treatments, there appears to be some evidence that the coordination problem in Case 2 is more difficult for our artificial players than is the coordination problem in Case 1.

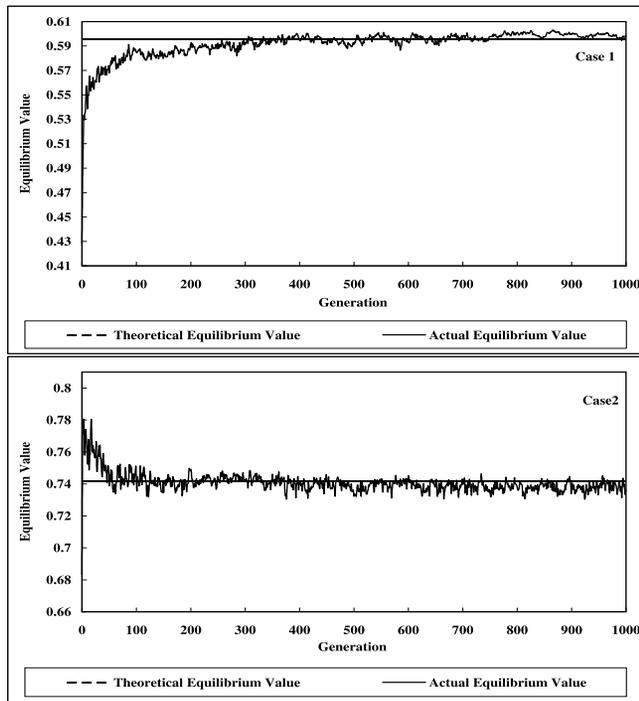


Figure 1: The Time Series of the Mean Choice of Action of the Coordination Game.

We have also considered how sensitive our results are to the use of payoff fitness as the main determinant of successive generations of forecast rules through

¹⁴No significant difference was found for $\delta_{M,b}$ between Cases 1 and 2. See Siegel and Castellan (1988) for an explanation of the nonparametric, robust rank order test used here.

application of the reproduction, crossover and mutation operations. Recall that we made an adjustment to the raw fitness values, so as to avoid excluding rules with negative payoffs. In all our simulations, we used an adjustment factor of 0.25. We also performed a simulation exercise where we considered what happens when we used a much larger adjustment factor of 200.00. That is, we adjusted raw fitness values, $\pi_{i,t}$ as follows:

$$\mu_{i,t} = \begin{cases} \pi_{i,t} + 200.00 & \text{if } \pi_{i,t} \geq -200.00, \\ 0 & \text{if } \pi_{i,t} < -200.00. \end{cases}$$

The effect of this adjustment is to nullify the usefulness of fitness as an indicator of the relative success of individual forecast functions. That is because the raw, unadjusted fitness values, $\pi_{i,t}$ can only take on values in the range $[-.50, .50]$. (See the payoff function (1) and the parameterizations of this function given in Table 2.) Adding 200.00 to these raw fitness values makes them essentially indistinguishable from one another, even after the adjusted fitness values have been converted into the normalized fitness values that are used to determine application of the reproduction, crossover and mutation operations.

Thus, the experiment where the adjustment value is set at 200.00 rather than at 0.25 serves as a test of whether relative fitness values are the driving force behind the results reported above. Indeed, this experiment is a test of the explanatory power of GP techniques. Figure 2 presents the time series for the mean from the single experiment involving Case 1 where we set the adjustment factor equal to

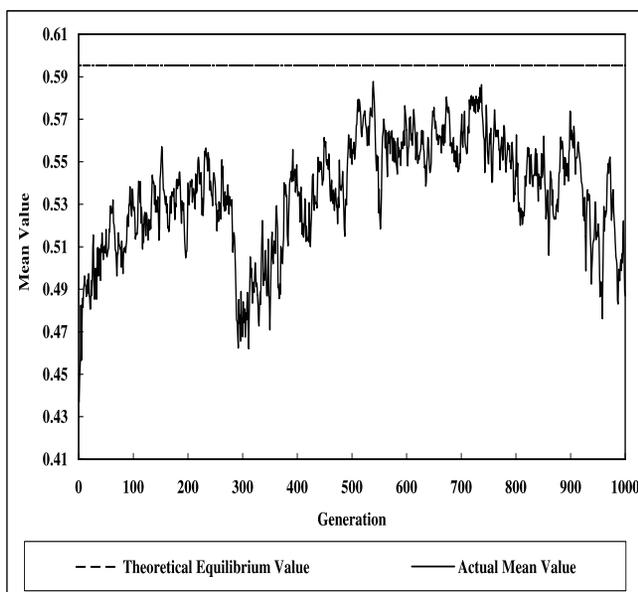


Figure 2: The Time Series of the Mean Choice of Action of their Coordination Game.

200.00 rather than 0.25. We see in this figure that the mean just wanders about randomly and has not settled down after 1,000 iterations. In particular, the mean does not approach either the corner equilibrium (0) or the interior equilibrium of Case 1 (0.59551). We may conclude from this exercise that the reliance of the genetic operators on relative fitness values is a driving force behind our simulation results, i.e. that fitness of forecast functions matters.

In addition to considering the dynamics of the mean choice of action, it is also interesting to examine the evolution of the population of forecast functions, i.e., GP_t . The *length* of the best-of-generation forecast function (Lisp S-expression) varies pretty widely. The length of a forecast function is measured by counting the number of elements (atoms) that are used in the program.¹⁵ Programs with longer lengths are more complex than those with shorter lengths, so the length of the program serves as a measure of the complexity of the forecast rule. Initially, the length of the best-of-generation program is rather small, but over time, the length increases substantially.

Consider, Simulation 2.5 for example. (Simulation 2.5 corresponds to the fifth simulation of Case 2). The length of the shortest best-of-generation program (S-expression) in this simulation is 15 and it appears in generations 16, 23, 27, 29, 32, 35, 37, 40, 41 and 50.¹⁶

$$\begin{aligned} gP_{\text{best},16} &= ((M_{t-5} + M_{t-4}) * M_{t-4}), \\ gP_{\text{best},23} &= ((M_{t-5} + M_{t-4}) * M_{t-4}), \\ gP_{\text{best},27} &= ((M_{t-5} + M_{t-3}) * M_{t-4}), \\ gP_{\text{best},29} &= (M_{t-3} * (M_{t-5} + M_{t-4})), \\ gP_{\text{best},32} &= ((M_{t-1} + M_{t-3}) * M_{t-3}), \\ gP_{\text{best},35} &= (M_{t-5} * (M_{t-5} + M_{t-4})), \\ gP_{\text{best},37} &= (M_{t-3} * (M_{t-4} + M_{t-5})), \\ gP_{\text{best},40} &= (M_{t-3} * (M_{t-4} + M_{t-5})), \\ gP_{\text{best},41} &= (M_{t-2} * (M_{t-5} + M_{t-5})), \\ gP_{\text{best},50} &= (M_{t-5} * (M_{t-4} + M_{t-5})), \end{aligned}$$

Programs with such a small program length continue to appear frequently after generation 50 but they are no longer selected as best-of-generation programs. Instead, increasingly complicated programs with lengths over 200 are more likely to be selected as the best-of-generation; the longest best-of-generation program with a length of 459 appears in generation 554:

¹⁵The length of a Lisp S-expression is distinct from the *depth* of a Lisp S-expression in tree form.

¹⁶All the GP programs below are represented as algebraic expressions (so that they can be more easily understood) rather than in the Lisp S-expression form in which they are encoded for GP.

$$\begin{aligned}
gp_{\text{best},554} = & (\text{SinCos}((\text{CosCosCos}(\text{Cos}M_{t-3} * M_{t-3}) * \text{CosSin}(\text{CosCos} \\
& (\text{Cos}(\text{Cos}M_{t-4}\%M_{t-4}) * \text{Cos}(M_{t-2} + \text{SinSin}M_{t-1})) + (M_{t-5} \\
& * \text{Cos}M_{t-2}))) * \text{CosCosCosSin}((M_{t-3} * M_{t-3}) - \text{Cos}M_{t-3})) \\
& \% \text{CosSinCos}(((\text{Cos}(\text{Cos}M_{t-1} * \text{Cos}(\text{ExpSinCos}M_{t-4}*M_{t-2})) \\
& \% \text{CosSinCosSin}(M_{t-4}\% \text{CosSinCos}M_{t-1})) * \text{CosSinRLog} \\
& ((\text{Sin}(M_{t-4} * \text{CosCos}M_{t-1}) * \text{Cos}(M_{t-4}\%((\text{Cos}M_{t-5} + M_{t-2})) \\
& * (\text{Cos}M_{t-3} * M_{t-4})))) * \text{Cos}((\text{Cos}(M_{t-3} * \text{Cos}M_{t-5}) * \text{Cos} \\
& \text{Cos}M_{t-4}) * M_{t-1})) * \text{Cos}(\text{Sin}(\text{SinCos}(\text{CosSinCos}M_{t-1} \\
& * (\text{CosSin}M_{t-2} * \text{CosCos}M_{t-4})) + \text{ExpSinCosCosSin} \\
& ((M_{t-2} * M_{t-3}) - \text{Cos}M_{t-3})) * \text{Cos}((\text{Sin}M_{t-3} + M_{t-5}) * \\
& ((M_{t-4} + (M_{t-3} + ((M_{t-2} * \text{Sin}M_{t-2}) * M_{t-1})))\% \text{Cos}M_{t-3}))))))
\end{aligned}$$

At generation 1,000 of simulation 2.5, the length of the individual programs was found to vary over the interval [3, 297]. This wide variation in program length implies that considerable heterogeneity remains in the population of forecast rules even after many generations.

Our findings that the best-of-generation programs become increasingly more complicated over time and that heterogeneity does not appear to diminish with time are perhaps attributable to our use of the symmetric sigmoidal activation function to map forecasts into the unit interval. The symmetric sigmoidal transformation function effectively “squashes” the output of forecast rules so that forecasts always lie within the unit interval. As a result, the forecasts of the various rules and their associated fitness values may not be all that distinct from one another even though the rules themselves may differ considerably. One consequence is that simple forecast rules, e.g. $gp_t = M_{t-i}$, $i = 1$ or 2 , may be unable to effectively compete with more complicated rules (programs with longer lengths), since these more complicated rules are better able to differentiate themselves from the simpler rules after being squashed, and therefore, these more complicated rules stand a better chance of being chosen for reproduction than the simpler rules.

As an alternative to the symmetric sigmoidal transformation function, we also considered the performance of our GP-based learning algorithm when the simple truncated linear transformation (discussed above) is used in place of the symmetric sigmoidal transformation. The truncated linear transformation is essentially a *linear* mapping into the unit interval whereas the symmetric sigmoidal transformation comprises a nonlinear mapping. Thus, with the truncated linear transformation there is less “squashing” of forecasts and associated fitness values. Indeed, squashing only occurs for forecasts that exceed the bounds of the unit interval; forecasts that lie within the unit interval are unaltered, and therefore remain more distinct (in terms of fitness) than under the symmetric sigmoidal transformation.

We conducted 10 simulations using the *truncated linear transformation function* in place of the symmetric sigmoidal transformation function – 5 simulations of Case 1 and 5 simulations of Case 2.¹⁷ Means and standard deviations from these 10 simulations are reported in Table 4. Here again, simulation 1.1 refers to our first simulation of Case 1, while simulation 2.1 refers to our first simulation of Case 2, and so on. Time series for the mean, M_t , from a single simulation of Case 1 and Case 2 are plotted in Figure 3. These time series plots are typical of the other simulations we conducted for the two cases using the truncated linear transformation.

Table 4: GP Simulation Results Using the Truncated Linear Transformation

Case		Simulation				
		1	2	3	4	5
1	\overline{M}_a	0.59281308	0.59279258	0.59328333	0.59295982	0.59249145
	$\delta_{M,a}$	0.02706819	0.02707828	0.02422021	0.02663368	0.02761726
	$\delta_{M^*,a}$	0.02720234	0.02721443	0.02432245	0.02675562	0.02778190
1	\overline{M}_b	0.59543389	0.59544433	0.59546691	0.59546368	0.59545463
	$\delta_{M,b}$	0.00033369	0.00033578	0.00025092	0.00028152	0.00029720
	$\delta_{M^*,b}$	0.00034227	0.00034215	0.00025460	0.00028531	0.00030232
2	\overline{M}_a	0.74079979	0.74081541	0.74069191	0.74081630	0.74080296
	$\delta_{M,a}$	0.01710630	0.01744866	0.01870916	0.01680324	0.01777976
	$\delta_{M^*,a}$	0.01712365	0.01746498	0.01872978	0.01682015	0.01779631
2	\overline{M}_b	0.74154575	0.74154649	0.74155158	0.74155041	0.74154058
	$\delta_{M,b}$	0.00043555	0.00045807	0.00044765	0.00048606	0.00050182
	$\delta_{M^*,b}$	0.00043622	0.00045867	0.00044803	0.00048645	0.00050268

\overline{M}_a = the average of M_t of a simulation from Generation 1 to 1,000.

\overline{M}_b = the average of M_t of a simulation from Generation 201 to 1,000.

$\delta_{M,a}$ = standard deviation about the M_a of a simulation from Generation 1 to 1,000.

$\delta_{M,b}$ = standard deviation about the M_b of a simulation from Generation 201 to 1,000.

$\delta_{M^*,a}$ = standard deviation about the *strict interior equilibrium* $1 - (1/\omega)$ from Generation 1 to 1,000.

$\delta_{M^*,b}$ = standard deviation about the *strict interior equilibrium* $1 - (1/\omega)$ from Generation 201 to 1,000.

From Table 4 and Figure 3, we see that our use of the truncated linear transformation in place of the symmetric sigmoidal transformation results in several significant differences. First, a comparison between Figures 3 and 1 and between the results in Tables 4 and 3 reveals that after 1,000 generations, the GP learning algorithm with the truncated linear transformation is generally closer to achieving the interior equilibrium than is the GP learning algorithm with the symmetric sigmoidal transformation. This quicker convergence to the interior equilibrium is

¹⁷Here we are using an adjustment factor of 0.25 once again to obtain adjusted fitness values.

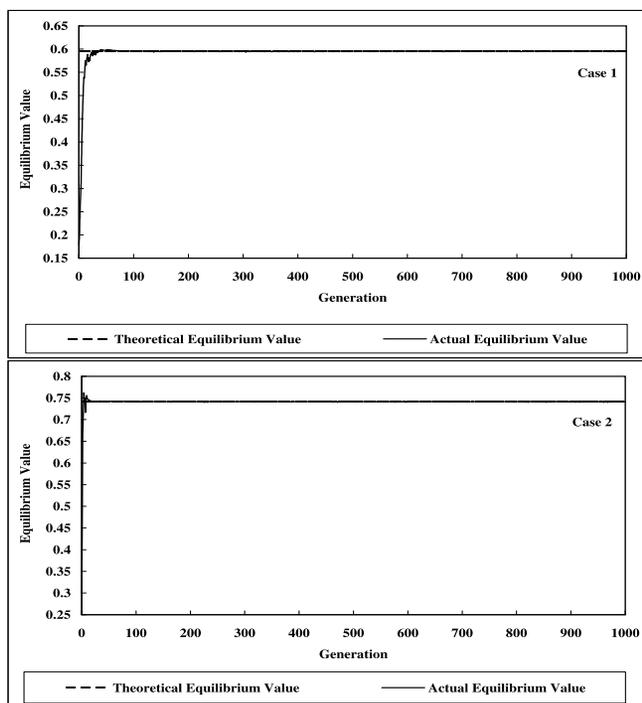


Figure 3: The Time Series of the Mean Choice of Action of the Coordination Game.

more consistent with the experimental findings of Van Huyck et al. (1994). Second, we observe that the deviation of the mean from the interior equilibrium, $1 - (1/\omega)$, in both Case 1 and Case 2 is *much smaller* when we use the truncated linear transformation in place of the symmetric sigmoidal transformation; in all 10 simulations, the average of the means, M_t , from generation 201 to 1,000, i.e. \bar{M}_b , does not deviate from the interior equilibrium value, $1 - (1/\omega)$ by more than 0.01%. This tighter distribution of forecasts around the interior equilibrium is again more consistent with the experimental findings of Van Huyck et al. (1994).¹⁸

As in the case of the symmetric sigmoidal transformation, we find that the chaotic trajectories for the game, $\Gamma(3.86957)$ that are predicted by the myopic best response dynamic are not apparent in any of our simulations of Case 2. We also find once again that a comparison of the standard deviations, $\delta_{M,b}$ or $\delta_{M^*,b}$ across Cases 1 and 2 in Table 4 reveals that these standard deviations are slightly larger in Case 2 than in Case 1. A rank order test confirms that both $\delta_{M,b}$ and $\delta_{M^*,b}$ are significantly larger in Case 2 than in Case 1 ($p \leq .01$ in both cases). This difference

¹⁸More direct comparisons between the experimental data and the simulated data from the GP-based learning algorithm are not really possible due to differences in the two experimental designs (e.g. the GP algorithm allows forecasts on the continuum of the unit interval, while the experimental subjects are limited to a finite set of discrete choices).

is also present, though difficult to see, in a visual comparison between Figure 3. We conclude that the coordination problem remains somewhat more difficult in Case 2 than in Case 1 regardless of whether we use truncated linear transformation or the symmetric sigmoidal transformation.

Finally, we note that under the truncated linear transformation, the length of the best-of-generation programs are considerably smaller than those discovered under the sigmoidal transformation. Simulation 2.2 (our second simulation of Case 2) is typical of the other simulation results we obtained using the truncated linear transformation. In this simulation, the longest best-of-generation program appeared in generation 2, and had a length of 29:

$$gp_{\text{best},2} = \text{Exp}(((M_{t-3} * M_{t-5}) - \text{Exp}M_{t-1}) * M_{t-1}) \\ - (M_{t-2} \% R\log(-2.66020 + M_{t-5})).$$

Following generation 100 of simulation 2.2, no best-of-generation program had a length that exceeded 3. In fact, the best-of-generation programs after generation 100 were always of the simplest form:

$$gp_{\text{best},t>100} = M_{t-i}, \quad i = 1, 2, 3, 4 \text{ or } 5.$$

Given such simple forecast rules, it is easy to understand why the distribution of forecasts becomes more tightly concentrated around the interior equilibrium when we use the truncated linear transformation.

7 Summary and Conclusions

We have considered a simple coordination game where the actions of the individual players are modeled and updated using GP techniques. Our GP-based coordination game allows for a considerably more flexible experimental design than is possible in experiments with human subjects. In particular, we do not have to restrict the choice set to a finite set of discrete actions, and we can have large numbers of players, e.g. $n = 500$. Moreover, players in our genetic programming implementation are explicitly endowed with the ability to formulate a vast number of both linear and nonlinear forecasting rules for the mean, including the myopic best response rule. This more flexible design allows for a possibly dense set of periodic and chaotic trajectories for the mean for values of $\omega > 3$. Despite this more flexible design, the evolution of play in our GP-based coordination game remains quite similar to that observed in the experiments that Van Huyck et al. (1994) conducted with human subjects. The mean choice of action eventually settles down to a small neighborhood of the interior equilibrium, even in Case 2, where the myopic best response dynamic predicts that this interior equilibrium should be unstable. There is evidence, however, that the coordination problem that our artificial agents face in Case 2 is somewhat more difficult than the coordination problem they face in

Case 1, as indicated by the different standard deviations about the mean/interior equilibrium for these two cases.

While these results cast some doubt on the plausibility of the myopic best response dynamic as a selection criterion (or any other learning schemes that would predict the interior equilibrium to be unstable), it is not yet clear that the myopic best response dynamic should be rejected on the basis of a “bad” prediction for a single game, namely $\Gamma(3.86957)$, or that the alternative, inertial learning algorithm should be accepted as a plausible selection dynamic on the same basis. While the inertial learning dynamic predicts that the interior equilibrium is always stable, the predicted trajectory for the mean/median is much too smooth when compared with the same trajectory from the experimental data. Moreover, the notion that a single, representative-agent-type learning algorithm can be used to characterize the evolution of the mean/median is at odds with the initial heterogeneity that is apparent in the experimental subjects’ actions. Finally, since our GP-based learning algorithm always “converges” to the interior equilibrium it is, by the criterion of Van Huyck et al. (1994), just as plausible a selection dynamic as the inertial learning algorithm. The initial heterogeneity of the forecasts that arise from our population-based GP algorithm makes it all the more plausible as a characterization of the experimental data.

We also note that the predictions of our GP-based learning model, especially those involving the truncated linear transformation, compare quite favorably with some new coordination game experiments that Van Huyck, Battalio and Rankin (1996) have recently conducted with human subjects. These new experiments differ from the previous experiments conducted by Van Huyck et al. (1994) in that subjects are not informed of the game’s payoff function π ; the only information available to subjects is their own past action/payoff history and the discrete action set that they may choose from. The purpose of this new experimental treatment is to place the human subjects in an environment that is as close as possible to that of artificial learning algorithms such as GP. In this new treatment, the human subjects learn to coordinate on the interior equilibrium *even more quickly* than in the previous treatment where subjects are informed of the payoff function π , of the game. Van Huyck et al. (1996) compare the experimental behavior in the new treatment with the behavior of a representative-agent-type, linear, stochastic reinforcement algorithm. While this algorithm eventually achieves a neighborhood of the interior equilibrium, it takes much longer to achieve this equilibrium (750 iterations) than it takes the experimental subjects. In contrast, our multi-agent GP-based learning algorithm converges much more quickly to a neighborhood of the interior equilibrium (usually within 50 iterations) so that it comes closer to mimicking the behavior of the experimental subjects.

Finally, we note that our findings for the coordination game are consistent with some other coordination experiments that have involved overlapping generations economies. Marimon, Spear and Sunder (1993) for example, report that experimental subjects are unable to coordinate on two-state sunspot equilibria,

choosing instead to settle upon the steady state of an overlapping generations economy. Similarly, Bullard and Duffy (1998b) simulate behavior using a genetic algorithm-based learning model in an overlapping generations economy and find that their population of artificial agents is able to eventually coordinate on steady state and low-order cycles for inflation rates but not on the higher order periodic equilibria of their model. This paper extends these earlier findings by suggesting that it may not be possible for agents to coordinate on aperiodic, *chaotic* trajectories.

REFERENCES

- [1] Allen F. and Karjalainen R., 1999. Using Genetic Algorithms to Find Technical Trading Rules. *Journal of Financial Economics*, 51(2), 245–71.
- [2] Angeline P., (1994). Genetic Programming and Emergent Intelligence. Chapter 4 of Kinnear (1994).
- [3] Arifovic J., (1994). Genetic Algorithm Learning and the Cobweb Model. *Journal of Economic Dynamics and Control*, 18, 3–28.
- [4] Arifovic J., (1995). Genetic Algorithms and Inflationary Economies. *Journal of Monetary Economics*, 36, 219–243.
- [5] Arifovic J., (1996). The Behavior of the Exchange Rate in the Genetic Algorithm and Experimental Economies. *Journal of Political Economy*, 104, 510–541.
- [6] Arifovic J., (1997). Strategic Uncertainty and the Genetic Algorithm Adaptation. In H. Amman et al. (eds.), *Computational Approaches to Economic Problems*. Boston: Kluwer Academic Press, 225–36.
- [7] Arthur W. B., Holland J. H., LeBaron B., Palmer R. and Tayler P., (1997). Asset Pricing Under Endogenous Expectations in an Artificial Stock Market. In W.B. Arthur et al. (eds.), *The Economy as a Evolving Complex System II*. Reading, MA: Addison-Wesley, 15–44.
- [8] Birchenhall C. R., (1995). Genetic Algorithms, Classifier Systems and Genetic Programming and Their Use in Models of Adaptive Behavior and Learning. *Economic Journal*, 105, 788–795.
- [9] Bray M., (1982). Learning, Estimation, and the Stability of Rational Expectations. *Journal of Economic Theory*, 26, 318–339.
- [10] Bullard J. and Duffy J., (1998a). A Model of Learning and Emulation with Artificial Adaptive Agents. *Journal of Economic Dynamics and Control*, 22, 179–207.

- [11] Bullard J. and Duffy J., (1998). On Learning and the Stability of Cycles. *Macroeconomic Dynamics*, 2, 22–48.
- [12] Chen S. and Yeh C., (1997a). Toward a Computable Approach to the Efficient Market Hypothesis: An Application of Genetic Programming. *Journal of Economic Dynamics and Control*, 21, 1043–1063.
- [13] Chen S. and Yeh C., (1997b). On the Coordination and Adaptability of the Large Economy: An Application of Genetic Programming to the Cobweb Model. In P. Angeline and K.E. Kinnear, Jr., (eds.), *Advances in Genetic Programming II*, Chapter 22. Cambridge, MA: MIT Press.
- [14] Chen S., Duffy J. and Yeh C., (1996). Genetic Programming in the Coordination Game with a Chaotic Best-Response Function. In: *Proceedings of the 1996 Evolutionary Programming Conference*, San Diego, CA.
- [15] Cooper R., DeJong D., Forsythe R. and Ross T., (1990). Selection Criterion in Coordination Games: Some Experimental Results. *American Economic Review*, 80, 218–233.
- [16] Crawford V. P., (1991). An Evolutionary Interpretation of Van Huyck, Battalio and Beil's Experimental Results on Coordination. *Games and Economic Behavior*, 3, 25–59.
- [17] Crawford V. P., (1995). Adaptive Dynamics in Coordination Games. *Econometrica*, 63, 103–143.
- [18] Dawid H., (1996), *Adaptive Learning by Genetic Algorithms*, Lecture Notes in Economics and Mathematical Systems No. 441. New York: Springer.
- [19] Devaney R. L., (1989). *An Introduction to Chaotic Dynamical Systems*, 2nd Ed.. Reading, MA: Addison–Wesley.
- [20] Dworman G., Kimbrough S. O. and Laing J. D., (1996). On Automated Discovery of Models Using Genetic Programming: Bargaining in a Three–Agent Coalitions Game. *Journal of Management Information Systems*, 12, 97–125.
- [21] Goldberg D. E., (1989). *Genetic Algorithms in Search, Optimization and Machine Learning*. Reading, MA: Addison–Wesley.
- [22] Holland J. H., (1975). *Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence*. Ann Arbor: University of Michigan Press.
- [23] Kagel J. H. and Roth A. E., (1995), eds., *Handbook of Experimental Economics*. Princeton, NJ: Princeton University Press.
- [24] Kinnear K. E., Jr., (1994), (ed.) *Advances in Genetic Programming*. Cambridge, MA: MIT Press.

- [25] Koza J. R., (1992). *Genetic Programming*. Cambridge, MA: MIT Press.
- [26] Koza J. R., (1994). *Genetic Programming II*. Cambridge, MA: MIT Press.
- [27] Kreps D. M., (1990). *Game Theory and Economic Modelling*. New York: Oxford University Press.
- [28] Lucas R. E., Jr., (1986). Adaptive Behavior and Economic Theory. *Journal of Business*, 59, S401–S426.
- [29] Marcet A. and Sargent T. J., (1989). Convergence of Least Squares Learning Mechanisms in Self Referential Linear Stochastic Models. *Journal of Economic Theory*, 48, 337–368.
- [30] Marimon R., (1997). Learning From Learning in Economics. In D.M. Kreps and K.F. Wallis, eds., *Advances in Economics and Econometrics: Theory and Applications*, Vol. 1, Seventh World Congress, Econometric Society Monographs, No. 26, Cambridge: Cambridge University Press.
- [31] Marimon R., Spear S. E. and Sunder S., (1993). Expectationally Driven Market Volatility: An Experimental Study. *Journal of Economic Theory*, 61, 74–103.
- [32] Miller J. H., (1996). The Coevolution of Automata in the Repeated Prisoner's Dilemma. *Journal of Economic Behavior and Organization*, 29, 87–112.
- [33] Mitchell M., (1996). *An Introduction to Genetic Algorithms*. Cambridge, MA: MIT Press.
- [34] Neely C. J., Weller P. and Dittmar R., (1997) Is Technical Analysis in the Foreign Exchange Market Profitable?: A Genetic Programming Approach. *Journal of Financial and Quantitative Analysis*, 32, 405–26.
- [35] Sargent T. J., (1993), *Bounded Rationality in Macroeconomics*. New York: Oxford University Press.
- [36] Siegel S. and Castellan N. J., Jr., (1988). *Nonparametric Statistics for the Behavioral Sciences*, 2nd Ed. New York: McGraw Hill.
- [37] Tesfatsion L., (1997). A Trade Network Game with Endogenous Partner Selection. In H. Amman et al. (eds.), *Computational Approaches to Economic Problems*. Boston: Kluwer, 249-269.
- [38] Van Huyck J. B., Battalio R. C. and Beil R., (1990). Tacit Coordination Games, Strategic Uncertainty and Coordination Failure. *American Economic Review*, 80, 234-248.
- [39] Van Huyck J. B., Battalio R. C. and Beil R., (1991). Strategic Uncertainty, Equilibrium Selection Principles and Coordination Failure in Average Opinion Games. *Quarterly Journal of Economics*, 106, 885–910.

- [40] Van Huyck J. B., Cook, J. P. and Battalio R. C., (1994). Selection Dynamics, Asymptotic Stability, and Adaptive Behavior. *Journal of Political Economy*, 102, 975–1005.
- [41] Van Huyck J. B., Battalio, R. C. and Rankin F. W., (1996). Selection Dynamics and Adaptive Behavior Without Much Information. Working paper, Texas A&M University.