Novelty And Surprises In Complex Adaptive System (CAS) Dynamics: A Computational Theory of Actor Innovation

Sheri M. Markose
Economics Department and Institute of Studies in Finance (ISF)
University of Essex,
Wivenhoe Park, Colchester, C04 3SQ, UK.
scher@essex.ac.uk

INVITED TALK: Applications of Physics in Financial Analysis 4 (APFA4) Warsaw, November 13-15 2003.

ABSTRACT

The work of John von Neumann in the 1940's on self-reproducing machines as models for biological systems and self-organized complexity provides the computational legacy for **CAS**. Following this, the major hypothesis emanating from Wolfram (1984), Langton (1992, 1994), Kaufmann (1993) and Casti (1994) is that the *sine qua non* of complex adaptive systems is their capacity to produce novelty or 'surprises' and the so called **Type IV** innovation based structure changing dynamics of the Wolfram-Chomsky schema. The Wolfram-Chomsky schema postulates that on varying the computational capabilities of agents, different system wide dynamics can be generated: finite automata produce **Type I** dynamics with unique limit points or homogeneity; push down automata produce **Type II** dynamics with limit cycles; linear bounded automata generate **Type III** chaotic trajectories with strange attractors. The significance of this schema is that it postulates that only agents with the full powers of Turing Machines capable of simulating other Turing Machines, which Wolfram calls *computational universality* can produce **Type IV** irregular innovation based structure changing dynamics associated with the three main natural exponents of **CAS**, evolutionary biology, immunology and capitalist growth.

Langton (1990,1992) identifies the above complexity classes for dynamical systems with the halting problem of Turing machines and famously calls the phase transition or the domain on which novel objects emerge as 'life at the edge of chaos'. This paper develops the formal foundations for the emergence of novelty or innovation. Remarkably, following Binmore(1987) who first introduced to game theory the requisite dose of mechanism with players modelled as Turing Machines with the Gödel (1931) logic involving the Liar or the pure logic of opposition, we will see that only agents qua universal Turing Machines which can make self-referential calculation of hostile objectives can bring about adaptive novelty or strategic innovation. .

Novelty And Surprises In Complex Adaptive System (CAS) Dynamics: A Computational Theory of Actor Innovation

Sheri M. Markose
Economics Department and Institute of Studies in Finance (ISF)
University of Essex, UK.
scher@essex.ac.uk

INVITED TALK: Applications of Physics in Financial Analysis 4 (APFA4) Warsaw, November 13-15 2003.

I. Introduction

The work of John von Neumann in the 1940's on self-reproducing machines as models for biological systems and self- organized complexity provides a landmark transformation of dynamical systems theory based on motion, force and energy to the capabilities and constraints of information processors modelled as computing machines. [1] Following this von Neumann computational legacy on **CAS**, the major hypothesis emanating from Wolfram (1984), Langton (1992, 1994), Kaufmann (1993) and Casti (1994) is that the *sine qua non* of complex adaptive systems is their capacity to produce novelty or 'surprises' and the so called **Type IV** innovation based structure changing dynamics of the Wolfram-Chomsky schema. The Wolfram-Chomsky schema postulates that on varying the computational capabilities of agents, different system wide dynamics can be generated: finite automata produce **Type I** dynamics with unique limit points or homogeneity; push down automata produce **Type II** dynamics with limit cycles; linear bounded automata generate **Type III** chaotic output trajectories with strange attractors. The significance of this schema is that it postulates that only agents with the full powers of Turing Machines capable of simulating other Turing Machines, which Wolfram calls *computational universality* can produce **Type IV** irregular innovation based structure changing dynamics associated with the three main natural exponents of **CAS**, evolutionary biology, immunology and capitalist growth. Indeed, Goldberg (1995) claims that the mystery shrouding innovation can be dispelled .. "by a heavy dose of *mechanism*. Many of the difficulties in the social sciences comes from a lack of a *computational theory* of actor innovation …. . population oriented systems are dominated by what economists call the law of unintended consequences (which is itself largely the result of the innovative capability of the actors)" (*ibid.* p.28, italics added).

This paper builds on Binmore's (1987) seminal work that introduced to game theory this requisite dose of mechanism with players as Turing Machines and the Gödel (1931) logic involving the Liar or the rule breaker. Binmore(1987) indicated that that latter will provide the generic framework for the strategic necessity for the endogenous generation of indeterminism in the system by the actions of highly computationally intelligent agents. There is also a long standing tradition, albeit an informal one, in the macro economics policy literature introduced by Lucas(1972) on the strategic use of 'surprises'. However, in extant game theory whether eductive or evolutionary there is no notion of innovation being a Nash equilibrium strategy let

[1] Mirowski(2002) discusses how radical a shift this has been for the methodology of science and also from the perspective of von Neumann's earlier work with Oscar Morgenstern on the *Theory of Games and Economic Behaviour* .

alone one that is necessitated as a best response by a structure of opposition. As in traditional Darwinian evolution, in economic models innovation is either introduced at a random rate or as an *ad hoc* addition in the form of trend growth.

This paper develops a computational theory of actor innovation by combining a game theoretic framework with the Emil Post (1945) set theoretic proof of the epochal Gödel (1931) incompleteness result. The latter shows that the conditions of opposition between two Turing Machines *and* for these machines to recognize their mutual hostility are the *logically* necessary conditions for innovative outcomes that have an encoding beyond algorithmic enumeration. Gödel (1931) had seminally used the notion of the Liar, the agent who falsifies or controverts, to embody the pure logic of opposition. However, the Liar can falsify or contravene with certainty only from a computable fixed point. This is intuitively well understood in the Lucasian thesis on policy ineffectiveness that regulatees can contravene policy only if the policy outcomes can be rationally expected. When there is mutual recognition by the players of the structure of opposition, the so called fixed point with the Liar can be fully deduced to be a non-computable fixed point. [2] Any total computable function from this non-computable fixed point referred to as the *productive function* in Post's set theoretic proof of Gödel incompleteness result, shown to represent the best response function in a game theoretic framework, can only map into a set that cannot be recursively enumerated, viz. by a Turing Machine. This coincides with the notion of the strategic use of surprise as intuitively proposed by Lucas (1972). The corresponding equilibrium **Type IV** dynamics converges to the non-computable domain within the *productive set* which is disjoint from the so called *creative* set first defined by Post (1944) in the context of undecidable decision problems. It is in this context, that Casti (1994, pp. 143-149) makes the connection between complex undecidable dynamics and 'surprises'.

Langton (1990, 1992) identifies the analog between the Wolfram-Chomsky complexity classes, the halting problem and the phenomenon of phase transitions and colourfully refers to the phase transition associated with **Type IV** dynamics as "life at the edge of chaos". The latter epithet arises for the following reason that the creative set on which Turing Machines halt is associated with **Type I** and **Type II** dynamical systems with limit points and limit cycles, which Langton calls (computable) order. There is a set disjoint to the creative set on which Turing Machines[3] can be logically deduced to be incapable of halting and it represents systems with **Type III** chaotic dynamics. Both of these sets represent attractors for dynamical systems that cannot produce novelty. The domain for novelty producing **Type IV** dynamics lies outside both these sets. No finite meta model can ever computably identify the novelty based change in the structure of this system. Though not fully formally understood as this, again the predictive failure of econometric meta models is well known to economists as the Lucas Critique (1976) due to a lack of structural invariance that follow from strategically induced innovation. Finally, systems capable of endogenous novelty generation experience a critical slowing down at the phase transition between the two other domains of (computable) order and chaos as the system at this juncture is

---

[2] Gödel's analogue of the Liar proposition is the undecidable proposition , say A, which has the following structure : $A \leftrightarrow \sim P(A)$. That is, A says of itself that it is not provable ($\sim P$). However, there is no paradox here as it is indeed true that this is so. Any attempt to prove the proposition A results in a contradiction with both A and $\sim A$, its negation, being provable in the system.

[3] Technically, these disjoint sets are recursively enumerable but their complements are not. If this were not to be the case, the system will be complete in that no novel encoded objects (not already in these sets) can ever arise.

effectively involved in an irreducible process of calculation of an undecidable global ordering problem of Gödel's Diophantine degree of complexity.

The above indicates that Nash equilibria in which agents innovate as a best response to evade hostile objectives of other agents and produce novel objects not previously in action sets is currently outside the ambit of traditional game theory. The latter, without the scope of the mathematics of incompleteness, can only consider randomization and not innovation in zero sum and oppositional situations.

Remarkably, despite the deep mathematical foundations of **CAS** on the ubiquitous structure of opposition formalized in the Liar and the capacity for self-referential calculations by agents of hostile behaviour of other agents, systems capable of adaptive novelty are commonplace and by and large only involve the intuitively familiar notion of the need to evade hostile agents. Markose (2003) refers to this as the Red Queen dynamics that exists among coevolving species. In Ray's classic artificial life simulation called Tierra, Ray (1992), when some agents perceive that others are parasitic on them, they start hiding their whereabouts and also mutate to evade the parasite. Recent research on RNA virus (see, Solé *et. al.* 2001) has likewise identified a 'phenotype for mutation', viz. a behavioural or strategic response favouring novelty, which is a far cry from the notion of random mutation. Axelrod (1987) in his classic study on cooperative and non-cooperative behaviour in governing design principles behind evolution had raised this crucial question on the necessity of hostile agents :" we can begin asking about whether parasites are inherent to all complex systems, or merely the outcome of the way biological systems have happened to evolve" (*ibid.* p. 41). It is believed that with the computational theory of actor innovation developed in this paper, we have a formal solution of one of the long standing mysteries as to why agents with the highest level of computational intelligence are necessary to produce innovative outcomes in **Type IV** dynamics.

The rest of the paper is organized as follows. Section 2.1 gives a brief overview of the von Neumann computational legacy of modern complex adaptive systems theory. In 2.2, the mathematical prerequisites are given for the formal modelling of computationally intelligent agents as "parametrized decision algorithms", Arthur (1991), that is, agents whose behaviour is brought about by finitely encodable algorithms. Some elements of Gödel *meta-mathematics* and the limitative results on computability are also introduced with a brief discussion of their relevance for game theory. Specific to this is the capacity of Turing Machines to make self-referential calculations. Further, computability constraints on strategic decisions and best response functions enable us to use some classic results from computability theory such as the Second Recursion Theorem for the specification of fixed points. This was first introduced to Economics in the context of the generic non-computability problem of rational expectations equilibria by Spear (1989). In section 3.1, I extend this framework for the characterization of dynamical system changes in a two person game under conditions of cooperation and opposition. In section 4, a formalization of Gödel's logic of pure opposition with the Liar or the contrarian/hostile agent is given. The mutual self-referential recognition of hostility that only agents with the highest powers of computational intelligence qua Turing Machines are capable of doing, will be shown to be a necessary condition for the strategic use of 'surprise'.

2. Computation and CAS Complex Dynamics

2.1 von Neumann Computational Legacy of CAS

The von Neumann models based on cellular automata[4] have laid the ground rules of modern complex systems theory regarding -(i) the use of large ensembles of micro level computational entities or automata following simple rules of local interaction and connectivity, (ii) the capacity of these computational entities to self-reproduce and also to produce automata of greater complexity than themselves and (iii) use of the principles of computing machines to explain diverse system wide or global dynamics.

The significance of the von Neumann computational legacy of **CAS** is that it covers all substrata, ranging from the bio-chemical to the artificial, in which effective procedures or computation reside. By the Church-Turing thesis (see, Cutland 1980) the intuitive notion of an effective procedure or an algorithm can be identified with the class of general recursive functions and represent finitely encodable programs implemented in a number of equivalent ways referred to as automata or mechanism. The best known among these idealizations of mechanism is the Turing Machine (TM, for short) and no mechanism can exceed the computational powers of Turing Machines. Such a definition of mechanism or formalistic calculation is necessary before complexity measures of the disjunction between the microscopic elements of the system and their macroscopic properties can be ascertained and also on what constitutes an innovation or surprise in the system.

In keeping with (i) above, as observed by Arthur (1991), the units of modern adaptive models are “parametrized decision algorithms” or units whose behaviour is brought about by finitely encodeable algorithms. Indeed, Langton (1992) notes that physical dynamical systems “are bound by the same *in principle* limitations as computing devices” (*ibid.*p82). These limitative results of computing devices are generically referred to as the halting problem. Church’s Theorem and in particular the Gödel (1931) First Incompleteness Theorem show how Turing machines themselves can produce encoded objects (viz. by mechanizing the exit route in Georg Cantor’s famous diagonal method) that cannot be enumerated by any machine. Such objects are innovations in the system and technically do not belong to recursively or algorithmically enumerable sets on which Turing machines halt. With regard to this Mirowski (2002) has correctly asserted that mathematicians “finally have blazed the trail to a formalized logical theory of evolution ”(*ibid.* p.141). In other words, dynamical system outcomes produced by algorithmic agents need not be computable and fail to be systematically identified by codifiable meta models. This is referred to as undecidable dynamics. Gödel’s Second Incompleteness Result shows that it is precisely when systems isomorphic to number theory are consistent that internal consistency, which is a strongly self-referential system wide property often regarded as the hallmark of rational order, cannot be established by an algorithmic decision procedure. Gödel (1931) axiomatically derived the undecidable proposition, the encoding of which represents the diophantine equation which has no algorithmic solution.[5] This class well known as Hilbert’s Tenth problem has the highest degree of algorithmic unsolvability.

---

[4] Cellular automata were developed by von Neumann and Stanislav Ulam to represent biological systems and for the purpose of modelling biological self-reproduction.

[5] Diophantine equations are polynomial equations with integer solutions. The irreducible nature of the computation here is that short of letting the system run its course there is no *a priori* systematic way to determine the solution to the problem.

Penrose(1988) was amongst the first to identify so called non-computable patterns or tiling problems, that nevertheless emerge from the execution of simple rules, with the Gödel incompleteness result. However, what continues to remain a matter of considerable mystery, to all scientists concerned with adaptive novelty , be they evolutionary biologists, immunologists, economists or physicists concerned with novelty producing self-organization systems , is why agents with the highest level of computational intelligence qua Turing Machines as postulated in the Wolfram-Chomsky schema are necessary to produce adaptive novelty.

Albin(1988, see Foley, *ibid* , pp. 42-45) consider the use of computationally intelligent agents in a game theoretic framework so as to resolve this matter. However, it was Binmore (1987) who made the seminal connection that the structure of the game in question that necessitates the use of surprise or innovation was precisely the one in Gödel (1931) that leads to the construction of the undecidable proposition.

2.2 *Gödel Meta- Mathematics And Prerequisites On Computability*

The main purpose of the formal analysis is to show the relevance of the Gödel paradigm and the mathematics of incompleteness for the characterization of systems capable of novelty based complex **Type IV** dynamics. Gödel (1931) pioneered the framework of analysis called *meta mathematics* pertinent to self-referential structures where he obtains epochal results on the sort of statements an internal observer can make as a meta-theorist if he is constrained to be very precise in what he can know and how he can make inferences. As highlighted by Binmore(1987), the theoretical significance of the analogue of the Gödel type incompleteness or indeterminacy result for formalized game theory stems precisely because this can be proven to arise not from incorrect or inconsistent reasoning or calculation but rather to avoid strategic irrationality and logical contradiction. To this end instrumentally rational players are accorded the full powers of an idealised computation machine in the calculation of Nash equilibrium strategies and all information has to be in a codifiable form. Following from the Church-Turing thesis, the computability constraint on the decision procedures implies that these are computable functions that can only entail finitely specified set of instructions in the computation. Again by a method introduced by Gödel (1931) called Gödel numbering, all objects of a formalisable system describable on the basis of a countable alphabet are put into 1-1 mapping with the set of natural numbers referred to as their Gödel numbers (g.ns, for short). Thus, computable functions can be indexed by the g.n of their finitely encoded program. Impossibility results on computation therefore become the only constraints on what rational/optimizing players cannot calculate given the same information on the encoded primitives on the game.

By the Church-Turing thesis computable functions are number theoretic functions, $f : \mathbf{N} \rightarrow \mathbf{N}$ where **N** is the set of all integers.[6] Each computable function is identified by the index or g.n of the program that computes it when operating on an input and producing an output if the function is defined or the calculation terminates at this point. Following a well known notational convention, we state this for a single valued computable function as follows

---

[6] The first limitative result on functions computable by T.Ms is that at most there can only be a countable number of these with the cardinality of $\aleph$ being denoted by $\aleph_0$, while from Cantor we know that the set of all number theoretic functions have cardinality of $2^{\aleph_0}$. Hence, not all number theoretic functions are computable (see,Cutland,1980).

$$f(x) \cong \phi_a(x) = q . \qquad (1.a)$$

That is, the value of a computable function f(x) when computed using the program/TM with index a is equal to an integer $\phi_a(x) = q$, if $\phi_a(x)$ is defined or halts (denoted as $\phi_a(x) \downarrow$) or the function f(x) is undefined (~) when $\phi_a(x)$ does not halt (denoted as $\phi_a(x) \uparrow$). The domain of the function f(x) denoted by Dom ϕ_a or W_a is such that,

$$\text{Dom } \phi_a = W_a = \{ x \mid \phi_a(x) \downarrow : TM_a(x) \text{ halts}\}. \qquad (1.b)$$

Definition 1: *Computable functions that are defined on the full domain of N are called **total computable functions**. **Partial computable functions** are those functions that are defined only on some subset of **N**.*
Related to (1.b) is the notion of sets whose members can be enumerated by an algorithm or a TM.
Definition 2: A *set which is the null set or the domain or the range of a recursive/computable function is a recursively enumerable set. Sets that cannot be enumerated by T.Ms are not r.e .*

The one feature of computation theory that is crucial to eductive game theory where players have to simulate the decision procedure of other players, is the notion of the Universal Turing Machine(UTM).
Definition 3: *The UTM is a partial computable function, defined as $\psi(a,x)$, which uses the index a of the TM whose behaviour it has to simulate. By what is called the Parameter or Iteration Theorem, there is a total computable function u(a) which determines the index of the UTM such that*

$$\psi(a,x) = \phi_{u(a)}(x) \cong \phi_a(x) . \qquad (2)$$

Equation (2) says that the UTM, on the left-hand side of (2) on input x will halt and output what the TM_a on the right-hand side does when the latter halts and otherwise both are undefined.

Of particular significance are Turing Machines that use their own code/g.n as inputs in their calculation. We will refer to these as self-referential calculations.
Definition 4: The set denoted by **C** is the set of g.ns of all TMs that halt when operating on their own g.ns or alternatively **C** contains the g.ns of those recursively enumerable sets that contain their own codes (see, Cutland , 1980, p.123, Rogers, 1967, p.62).

$$\mathbf{C} = \{ x \mid \phi_x(x)) \downarrow ; TM_x(x) \text{ halts} ; \ x \in W_x \} \qquad (3.a)$$

The complement of **C**

$$\mathbf{C}^{\sim} = \{ x \mid \phi_x(x) \uparrow; TM_x(x) \text{ does not halt}; \ x \notin W_x\} \qquad (3.b).$$

Theorem 1: *The set $\mathbf{C}^{\sim}$ is not recursively enumerable.*

In the proof that $C^{\sim}$ is not recursively enumerable, viz there is no computable function that will enumerate it, Cantor's diagonalization method is used. [7]

As indicated in the introduction, what is remarkable is that the formal character of systems capable of the endogenous production of novelty based complex dynamics corresponds to the notion of creative and productive sets first defined by Emil Post (1944) in the set theoretic proof of the Gödel incompleteness result.

Definition 5**:* A ***creative set Q is a recursively enumerable set whose compliment, $Q^{\sim}$, is a ***productive*** set. The set $Q^{\sim}$ is productive if there exists a recursively enumerable set W_x disjoint from Q (viz. $W_x \subset Q^{\sim}$) and there is a total computable function f(x) which belongs to $Q^{\sim} - W_x$. $f(x) \in Q^{\sim} - W_x$ is referred to as the ***productive function*** and is a 'witness' to the fact that $Q^{\sim}$ is not recursively enumerable. Any effective enumeration of $Q^{\sim}$ will fail to list f(x), Cutland (1980, p. 134-136).

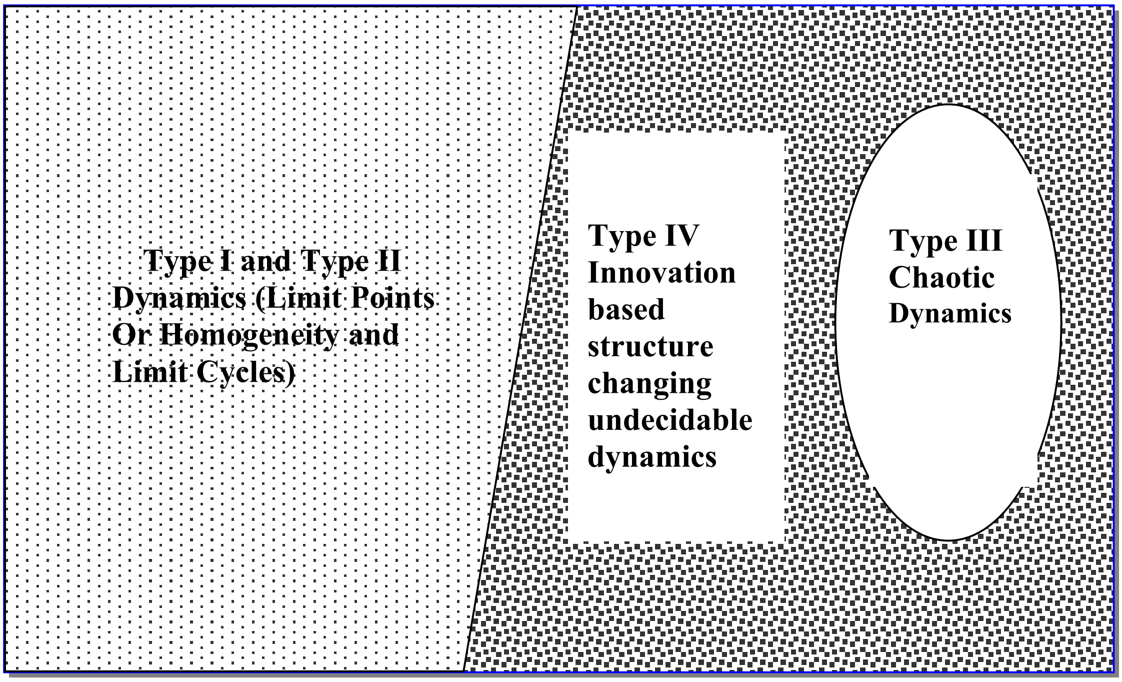
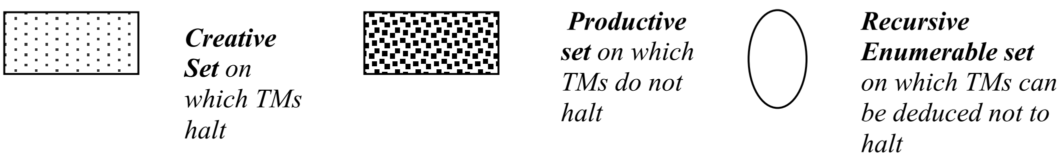


Figure 1 *Set Theoretic Representation of CAS Dynamics Being Outside Disjoint Recursively Enumerable Sets*



[7] Assume that there is a computable function $f = \phi_y$, whose domain $W_y = \mathbf{C}^{\sim}$. Now, if $y \in W_y$, then $y \in \mathbf{C}^{\sim}$ as we have assumed $\mathbf{C}^{\sim} = W_y$. But by the definition of $\mathbf{C}^{\sim}$ in (3.b) if $y \in W_y$, then $y \in \mathbf{C}$ and not to $\mathbf{C}^{\sim}$. Alternatively, if $y \notin W_y$, $y \notin \mathbf{C}^{\sim}$, given the assumption that $\mathbf{C}^{\sim} = W_y$. Then, again we have a contradiction, as since from (3.b) when $y \notin W_y$, $y \in \mathbf{C}^{\sim}$. Thus we have to reject the assumption that for some computable function $f = \phi_y$, its domain $W_y = \mathbf{C}^{\sim}$.

We propose to show that f(x) the productive function which is proof of the incompleteness of the formal system also corresponds to the best response surprise function in the Nash equilibrium of a game that produces the structure changing **Type IV** undecidable dynamics. The Langton (1990,1992) analog of the Wolfram-Chomsky complexity classes with the halting problem of Turing Machines may be given in terms of Post's disjoint creative and productive sets defined above.

The creative set on which Turing Machines halt is associated with **Type 1** and **Type II** dynamics which can be called (computable) order. The prototypical creative set is the set **C** in (3.a) which contains self-referential calculations that converge. They will be shown to correspond to computable fixed points. Negation or computable contrarian propositions of the latter, on account of consistency of the system belong to a set disjoint from **C** and hence though a subset of the complement of **C**, viz. $\mathbf{C}^{\sim}$ in (3.b)**,** its membership can be enumerated and also shown to be a set on which Turing Machines can be logically deduced to be incapable of halting. Thus, there is a recursively enumerable subset of $\mathbf{C}^{\sim}$ and it represents systems with **Type III** chaotic dynamics. The domain for novelty producing **Type IV** dynamics lies outside both these recursively enumerable disjoint sets.

Figure 1 gives the set theoretic representation of the Wolfram-Chomsky schema of complexity classes for dynamical systems which formally corresponds to Post's set theoretic proof of Gödel Incompleteness Result.

It is envisaged that with a minimum of details on computation theory and relying only on some familiarity with calculations involved in the determination of rational expectations equilibria which typically entail iteration and substitution, the main thrust of the paper can be followed. As discussed, all finitely codifiable information on the game can be assigned g.ns including the calculations used in determining optimal strategies. The advantage of this framework is that all calculations can be reduced to operations on g.ns of functions as in (2). Especially in the case of computability or non-computability of classes of fixed points or rational expectations equilibria we can obtain generic results, without the hard grind of producing any of the calculations or having to make *ad hoc* functional specifications.

3. The Game Under Computability Constraints

3.Implications of Computability Constraints on Decision Procedures

The primitives of the game, best interpreted as one in which both cooperation and opposition arise such as in a regulatory /policy game or a parasite host game, is codified as follows.

$$\mathbf{G}= \{(p,g), (A_p, A_g), s\in S\}.$$

Here,(p,g) denote the respective g.ns of the objective functions, to be specified, of players, p, the private sector/regulatee and g, government/regulator. The action sets denoted by A_i are finite and countable with $a_{il} \in A_i$, $i\in (g, p)$ being the g.n of an action rule of player i and l=0,1,2,.....,L. An element $s\in S$ denotes a finite vector of state variables and other archival information and S is a finite and countable set.

Gödel meta analysis is analogous to that in chess or any other game where a unique correspondence can be established between moves played or those that can be potentially played with meta statements of these in some notation being stored in the meta system which is in the public domain.[8]

[8] See Albin(1982) for a more rigorous discussion that all strategic analysis is meta analysis of this kind.

A major implication of imposing computability constraints on all aspects of the decision procedures of the game is that all meta-information with regard to the outcomes of the game for any given set of state variables, $s \in S$, can be effectively organized by the so called prediction function $\phi_{\sigma(x,y)}(s)$ in an infinite matrix Ξ of the enumeration of all computable functions, given in Figure 1 (see, Cutland, 1980, p.208). The tuple (x,y) identifies the row and column of this matrix Ξ whose rows are denoted as Ξ_i, i= 0,1,2,...... .

Figure2 : Meta –Information on Outcomes of Games

$$
\begin{array}{llllllll}
\Xi_0 & \phi_{\sigma(0,0)} & \phi_{\sigma(0,1)} & \phi_{\sigma(0,2)} & \phi_{\sigma(0,3)} & \cdots & \phi_{\sigma(0,y)} & \cdots \\
\Xi_1 & \phi_{\sigma(1,0)} & \phi_{\sigma(1,1)} & \phi_{\sigma(1,2)} & \phi_{\sigma(1,3)} & \cdots & \phi_{\sigma(1,y)} & \cdots \\
\Xi_2 & \phi_{\sigma(2,0)} & \phi_{\sigma(2,1)} & \phi_{\sigma(2,2)} & \phi_{\sigma(1,3)} & \cdots & \phi_{\sigma(2,y)} & \cdots \\
. \\
\Xi_x & \phi_{\sigma(x,0)} & \phi_{\sigma(x,1)} & \phi_{\sigma(x,2)} & \phi_{\sigma(x,3)} & \cdots & \phi_{\sigma(x,x)} & \cdots \\
. \\
.
\end{array}
$$

The function $\phi_{\sigma(x,y)}(s)$ if defined at a given state s and $\sigma(x,y)$ yields

$$\phi_{\sigma(x,y)}(s) = q .$$

Here, q in some code, is the vector of state variables determining the outcome of the game. Note, $\sigma(x,y)$ is the index of the program for this function ϕ that produces the output of the game when one player plays strategy x and the other player plays a strategy *that is consistent with his belief that the first player has used strategy y*. Thus, (x,y) are the codes of the calculations involved for the determination of the strategies and the tuple also identifies a point on the matrix Ξ in Figure 2. The conditions under which the prediction function for each (x,y) point in the above matrix is defined is given in the following Theorem.

Theorem 2*: The representational system is a 1-1 mapping between meta information in matrix Ξ in Figure1 and internal calculations such that the conditions under which the prediction function which determines the output of the game for each (x,y) point is defined are as follows:*

$$\phi_{\sigma(x,y)}(s) \cong \phi_{\phi_x(y)}(s) = q, \quad \text{iff} \quad \phi_x(y) \downarrow . \quad (4)$$

Here the total computable function $\sigma(x,y)$ modelled along the lines of Gödel's substitution function[9] *(see, Rogers, 1967,p.202-204) has the feature that it names or 'signifies' in the meta system Ξ the points in the game that correspond to the different internal calculations on the right- hand side of (4) as we substitute different values for (x,y). The g.ns implemented by $\sigma(x,y)$ can always be obtained whether or not the partial recursive function $\phi_x(y)$ on the right-hand side of (4) which executes internal calculations halts or not.*

[9] This approach economizes on formalism and enables us to high light and exploit the Fixed Point Theorems of recursive function theory to determine Nash equilibrium outcomes more readily than has been the case in for instance in Anderlini(1991) , Canning(1992) and Albin(1982).

Proof : See Appendix.

By the necessary condition in (4) if the function $\phi_x(y)$ on the right-hand side executing the internal calculation is defined, we say the prediction function ϕ in the meta system on the left-hand side producing the output of the game is *computable* and the outcome q of the game at that point is *predictable.* Likewise, the 'only if ' condition in (4) implies that meta statements that are valid on the predictability of the outcomes of the game at any (x,y) must give the correct inference on whether internal calculations on the right-hand side terminate.

Only subsets of the sets defined by the rows of the above matrix, such that $E_{\sigma i} \subseteq \Xi_i$, i=0,1,2,..... , can be taken to contain the outputs or the range of the functions used internally by the players making the calculations.[10] What this means is that, though there is always some outcome existentially and reference to any point in Ξ can be made as a meta-proposition, there are points at which players in the system will fail to compute or predict these systematically or via an algorithm.

On account of the 'only if' condition in Theorem 1, many interesting aspects of the Nash equilibria of computable games can be established only with reference to the meta analysis and information in the matrix Ξ , with no explicit reference to internal calculations being made by the players. Thus, all Nash equilibria and other relevant fixed points of the game satisfying what has been referred to as consistent alignment of beliefs (CAB, for short, Osborne and Rubinstein,1994)have to be elements along the *diagonal array* of this matrix. A typical Nash equilibrium is at a point defined by $\sigma(x,x)$, viz. say player p plays x and then g correctly identifies this. Off diagonal elements along any row defined by strategy, say x, employed by the player p, cannot be Nash equilibria, as these off diagonal terms imply that g is choosing his strategy assuming the wrong meta representation of p's play.

3.2 Total Computable Best Response Functions and Optimal Strategy Functions

A major advantage of this framework is that the determination of Nash equilibrium strategies involve the use of total computable best response functions (f_p , f_g) which can be shown to operate directly on points such as $\sigma(x,x)$ to effect computable transformations of the system from one row to another of matrix Ξ with special reference to its diagonal array, see, Figure 2**.** Thus,

$$\phi_{f_i\,\sigma(x,x)}(s) \ , \qquad i \in (p,g). \qquad (5)$$

The specification in (5) of how the response functions apply has a number of important implications. As we will show, the Second Recursion Theorem can then be directly used for the determination of rational expectations equilibria where the constructive identification of the fixed point of the best response function enables both players to identify the same prediction function in the matrix Ξ as producing the outcome of the game.

Definition 5*: The best response functions f_i, $i \in (p,g)$ that are total computable functions can belong to one of the following classes –*

---

[10] Here σ_i is the code of the function $\sigma(i, j)$, j=0,1,2,3… that enumerates the ith row of matrix Ξ.

$$f_i = \begin{cases} 1\,(\text{Identity Function}) & - \quad \text{Rule} \quad \text{Abiding} \\ f_i^{+} & - \quad \text{Rule} \quad \text{Bending} \\ f_i^{\neg} & - \quad \text{Rule} \quad \text{Breaking} \\ f_i^{!} & - \quad \text{Surprise} \end{cases} \qquad (6.a)$$

such that the g.ns of f_i are contained in set $\Re$,

$$\Re = \{\, m \mid f_i = \phi_m \ ,\ \phi_m \ \text{ is total computable}\}. \qquad (6.b)$$

The set $\Re$ which is the set of all total computable functions is not recursively enumerable. The proof of this is standard, see, Cutland (1980).

The total computability of best response functions $f_i = \phi_m$, $m \in \Re$ in (6.a,b) yields the notion of constructible/effective action rules such that a finitely codifiable description of some (institutional) procedure which is defined for all mutually exclusive states of the world is obtained. As will be clear, (6.b) draws attention to issues on how innovative actions/institutions can be constructed from existing action sets. The remarkable nature of the set $\Re$ is that potentially there is an uncountable infinite number of ways in which 'new' institutions can be constructed from extant action sets A.

Definition 6 : *The objective functions of players are computable functions Π_i, $i \in$ (p,g) defined over the partial recursive payoff/outcome functions specified as in (4).*

$$\text{Arg}\max_{b_i \in B_i} \Pi_i \ (\phi_{\sigma(b_p,\, b_g{}^{\wedge})}(s)\ )\ , \qquad i \in (p,g)\ . \qquad (7)$$

The Nash equilibrium strategies (β_g^{E} , β_p^{E})with g.ns denoted by (b_p^{E}, b_g^{E}) entail two subroutines or iterations, to be specified later. In principle, the strategy functions (β_g , β_p) are Universal Turing Machines that simulate optimal strategies of the players that satisfy (7) and involve the total computable best response functions (f_p , f_g) which incorporate elements from the respective action sets A= (A_p ,A_g) *and* given mutual beliefs of one another's optimal strategy. In the two place notation given in (4), b_p is the g.n of p's strategy given that g has optimally chosen its strategy on the basis of g's metarepresentation/belief, $b_g^{\wedge}$, of p's strategy . Note that we will use g.ns z_i, $i \in$ (p,g) to represent encoding of the optimization calculus with respect to respective objective functions. The problem is that actions can in general be implemented by *any* total computable function response function, $f_i = \phi_m$, $m \in \Re$, $i \in$ (p,g).

In standard rational choice models of game theory, the optimization calculus in the choice of best response requires choice to be restricted to given actions sets. Hence, strategy functions map from a relevant tuple that encodes meta information of the game into given action sets

$$\beta_i(\ f_i\sigma(x,x),\, z,\, s,\, A) \rightarrow A_i \ \text{ and } \ f_i = \phi_m\ ,\ m \in A,\ \ i \in (p,g)\ . \qquad (8.a)$$

Unless this is the case, as the set $\Re$ is not recursively enumerable there is in general no computable decision procedure that enables a player to determine the other player's response functions. However, in principle, a strategic decision procedure (β_g, β_p) for choice of best response, $f_i = \phi_m$, $m \in \Re$, $i \in (p,g)$, can map into $\Re$-A , implying that an innovative action not previously in given action sets is used.

$$\beta_i (f_i \sigma(x,x)), z, s, A) \rightarrow \Re\text{- } A \;\; and\, f_i = f_i^{\,!} = \phi_m \;,\; m \in \Re\text{-}A,\; i \in (p,g). \quad (8.b)$$

It has indeed been noted in passing by Anderlini and Sabourian (1995, p.1351), based on the work of Holland (1975), that heterogeneity in forms does not arise primarily by random mutation but by algorithmic recombinations that operate on existing patterns. However, a number of preconceptions from traditional game theory such as the 'givenness' of actions sets prevent Anderlini and Sabourian(1995) from positing that players who as in (8.b), equipped with the wherewithal for algorithmic recombinations of existing actions, do indeed innovate from strategic necessity rather than by random mutation. The innovation per se is emergent phenomena, but the strategic necessity for it is fully deducible. Indeed, it is the very function of the Gödel meta framework to ensure that no move in the game made by rational and calculating players can entail an unpredictable/surprise response function from set $\Re$- A unless players can mutually infer by strictly codifiable deductive means from σ(x.x) that (8.b) is a logical implication of the optimal strategy at the point in the game. In other words, the necessity of an innovative/surprise strategy as a best response and that an algorithmic decision procedure is impossible at this point are fully codifiable propositions in the meta analysis of the game. While it will be shown that the specific structure of opposition logically and strategically necessitates surprise strategies in the Nash equilibrium of the game, in keeping with the set theoretic formulation of novelty production in **Figure 1**, the corresponding creative and productive disjoint subsets of the strategy sets have also to be developed. These arise in the form of computable and non-computable Nash equilibria of the game.

3.3 Fixed Point/Second Recursion Theorem and Nash Equilibria

The symmetric structure of the strategy function implies that meta analysis in the determination of Nash equilibrium strategies (β_p^E , β_g^E) with g.ns (b_p^E, b_g^E) can proceed from the perspective of one player or the other. We will assume that the meta analysis proceeds form the perspective of g. Calculations start at so called basepoint

$$\phi_{\sigma\,(b_a\,,b_a)}(\,s\,) = q \quad . \qquad (9)$$

Here, the prediction function is computable and outcomes of a policy rule a is predictable and q is the desired outcome that g wants in state variables when applying this policy rule a . In our two place notation $\sigma(b_a , b_a)$, the first b_a is the code of the program, as adopted by p to simulate the impact of the policy rule a that p believes that g will follow and the second place b_a denotes that g believes and acts on the basis that the private sector has simulated policy rule a . It is convenient to assume that policy rule a is optimal for g *if* the private sector is rule abiding. By rule abiding is meant that p will leave the system unchanged in terms of the row b_a of matrix Ξ.

Recursive identification of Nash equilibria or fixed points of the game utilizes the Second Recursion Theorem that any systematic/computable transformation by a total computable function f of the diagonal array in matrix Ξ is itself a row with a g.n, say m, in Ξ such that the m+1th. element in the mth row, E_m, is identical to the same entry in the diagonal array (see, Cutland, 1980, p. 208). We will show how the players by substitution and iteration identify the fixed points of the game using the two following subroutines.

Step1: Player g utilizes the UTM in (8.a) by replacing $\sigma(x,x)$ by $\sigma(b_a, b_a)$ and simulating p by setting i=p . The fixed point of p's best response function f_p is identified in the following way. The g.n of p's optimal strategy is obtained by a total computable function b_1,

$$b_p^* = b_1 (z_p ; \sigma(b_a , b_a)) \qquad (10.a)$$

If $b_p^* \neq b_a$, that is p's optimal strategy is different from g's initial belief that p is rule abiding, then g updates its belief and identifies $\sigma(b_p^*, b_p^*)$ to be the fixed point of p's best response function f_p.

Theorem 3 *:(The Fixed Point Theorem[11]) In a rational expectations equilibrium for the game at each stage, agents identify the same partial recursive prediction function ϕ for the outcome of the game*

$$\phi_{f_p\, \sigma(b_p^*, b_p^*)}(s) = \phi_{\sigma(b_p^*, b_p^*)}(s). \qquad (11)$$

Here, $\sigma(b_p^*, b_p^*)$ is said to be the fixed point of p's best response function f_p such that f_p $\sigma(b_p^*, b_p^*)$ and $\sigma(b_p^*, b_p^*)$ are indexes for the same partial recursive prediction function ϕ. Thus, the index on the latter will predict the same outcome of the game on the right-hand side as does the index of the prediction function on the left-hand side of (11) if the function is computable at the fixed point.

Proof proceeds by using the subroutine in (10.a) of **Step 1**and using the Second Recursion Theorem outline above in Cutland (1980).

Using the second subroutine or iteration the g.ns (b_p^E, b_g^E) of the Nash equilibrium strategy functions are fully definable in the system.

Step 2: g applies the optimality algorithm using the meta information from the fixed point in $\sigma(b_p^*, b_p^*)$ in (11) and obtains

$$b_g^E = b_2 (z_g;\ \sigma(b_p^*, b_p^*)) . \qquad (12.a)$$

Likewise, player p uses information in the fixed point in (11) and infers that g will play b_g^E in (12.a) and hence the g.n of p's Nash equilibrium strategy is given by

$$b_p^E = b_1 (z_p ; b_2 (z_g;\ \sigma(b_p^*, b_p^*))) = b_2 (z_p; b_g^E). \qquad (12.b)$$

4. Applications of the Fixed Point Theorem

In this section we first use Theorems 2 and 3 to state the following Lemma on the

[11] Spear (1989) has used similar ideas from recursive function theory to model the identification and learning problem in rational expectations equilibria. The theorem involved here is the Second Recursion Theorem, see Cutland (1980) and in particular Rogers(1967) who gives this specific form.

computability or not of the fixed points of p's best response function from the basepoint.

Lemma 1: *Corresponding to player p's best response function, $f_p \in \{1, f_p^{+}, f_p^{\neg}\}$ to g's base point strategy b_a that utilizes some $a \in A_g$ for any given s, the fixed point in (11) can be computable or non-computable. If the prediction function ϕ indexed by $\sigma(b_p*, b_p*)$ in (11) is computable, then by Theorem 1 the UTM in duplicator form $\phi_{b_p*}(b_p*)$ on the right-hand side of (4) halts with*

$$\phi_{\sigma(b_p*,b_p*)}(s) = \phi_{\phi_{b_p*}(b_p*)}(s) \text{ and } \phi_{b_p*}(b_p*) \downarrow. \quad (13.a)$$

If the prediction function ϕ indexed by $\sigma(b_p*, b_p*)$ in (10) is not computable then by Theorem 1 the UTM in duplicator form $\phi_{b_p*}(b_p*)$ on the right-hand side of (3) will not halt with

$$\phi_{\sigma(b_p*,b_p*)}(s) \sim \phi_{\phi_{b_p*}(b_p*)}(s) \quad \text{and } \phi_{b_p*}(b_g*) \uparrow. \quad (13.b)$$

Proof: Use Theorems 2 and 3.

4.1 Computable Fixed Points

Note, the g.ns for p's optimal strategy in the case $f_p \in \{ f_p^{+}, f_p^{\neg}\}$ will be respectively denoted by b_a^{+}, $b_a^{\neg}$ to correspond to whether p is rule bending or rule breaking.

(i)The rule abiding case: If it is optimal for the p to be rule abiding viz. f_p=1, *vis-à-vis* the generic predictable policy rule a in (9), then it simply leaves the system unchanged at the b_a th. row of matrix Ξ in Figure2. $\sigma(b_a, b_a)$ is a trivial computable fixed point and $b_g^E = b_a$ is g's Nash equilibrium strategy of the game.

(ii) The rule bending case: The rule bender applies the best response function denoted by f_p^{+} on $\sigma(b_a, b_a)$ and will move the system to a new row of the matrix Ξ with code b_a^{+} of the rule bending strategy $f_p^{+}\sigma(b_a, b_a)$, such that the predicted outcomes in q (or some subset of them) specified as changes in state variables of policy are amplified but not controverted/subverted. Formally, the predictable outcome of the game at the point $\sigma(b_a^{+}, b_a)$ belongs to the set which is not disjoint with the set $E_{\sigma_{b_a}}$, viz. $E_{\sigma_{b_a+}} \cap E_{\sigma_{b_a}} \neq \varnothing$. The fixed point result necessary to determine the mutual Nash equilibrium strategies when the private sector has used the rule bending strategy $f_p^{+}\sigma(b_a, b_a)$ is given as

$$\phi_{f_p^{+}\sigma(b_a^{+},b_a^{+})}(s) = \phi_{\sigma(b_a^{+},b_a^{+})}(s). \quad (14)$$

In the fixed point result in (14) as player g by using the index $\sigma(b_a^{+}, b_a^{+})$, (viz.when p has used b_a^{+} and g has updated its belief that p is indeed the rule bender) has identified the same function on the right hand side of (14) as the other player p who plays f_p^{+} and uses index $f_p^{+}\sigma(b_a^{+}, b_a^{+})$ on the left-hand side of (14), we say by Theorem 3 that both players have rational expectations of the game at this point. If the base point for the a-rule in (9) is computable, then there are no problems of computability regarding the fixed point in (13) for when the private sector is rule bending.

We will now turn to the famous fixed point that fails to be computable such that the outcome of the game at this point is not predictable. This is brought about by the Liar/rule breaker strategy.

4.2 *The Liar/Rule Breaker Strategy :The Logic of Opposition*

For player p, for the given (a,s) it may be optimal for p to apply the Liar strategy, $f_p^{\neg}$ $\sigma(b_a, b_a)$, with code $b_a^{\neg}$. Formally, the Liar strategy has the following generic structure.

For any state s when the rule a applies,

$$\phi_{f_p^{\neg}\sigma(b_a,b_a)}(s) = q^{\sim}\ ,\ q^{\sim} \notin E_{\sigma_{b_a}} \leftrightarrow \phi_{\sigma(b_a,b_a)}(s) = q\,,\ q \in E_{\sigma_{b_a}} \quad (15.a)$$

For all s when policy rule a does not apply,

$$f_p^{\neg} = 0\ : \text{Do Nothing .} \quad (15.b)$$

The Liar can successfully subvert with certainty in (15.a) if and only if ($\leftrightarrow$) the policy rule a has predictable outcomes *and* $f_p^{\neg}$ itself is total computable. Also, $f_p^{\neg} = \phi_m$, $m \in A_p$, must include a codified description of an action rule if undertaken by the Liar can subvert the predictable outcomes of the policy rule a. Formally, if q is predicted then the application of $f_p^{\neg}$ to $\sigma(b_a, b_a)$ will bring about an outcome $q^{\sim}$ $\notin E_{\sigma_{\kappa_a}}$ which belongs to a set disjoint from the set that contains the desired output of rule a for all s for which rule a applies, viz. $E_{\sigma_{\kappa_a}} \cap E_{\sigma_{\kappa_a}\neg} = \varnothing$. The outcomes $(q^{\sim}, q)$ can be zero sum but in general we refer to property $q^{\sim} \notin E_{\sigma_{\kappa_a}}$ in (15.a) as being oppositional or subversive. This underpins the intuition behind Ray's Tierra (1992) simulation where agents who recognize that they are hosts to parasites adopt the strategy for secrecy. This is also well known from Lucas (1972) postulate on policy ineffectiveness in the case of fully anticipated policy and the wisdom behind the panacea that to forestall subversion, the policy rule must be undefined and fraught with ambiguity.

Thus, we come to the point as to why agents who precipitate the Wolfram-Chomsky **Type IV** dynamics with innovation have to have powers of self-referential calculation. Firstly, g acknowledges the identity of the Liar in (15.a) and understands that transparent rule a cannot be implemented rationally as the outcome now defined[12] by $\phi_{\sigma(b_a^{\neg}, b_a)}(s) = q^{\sim}$ is the opposite of what is optimal for g. The latter is out of equilibrium. Player g updates beliefs so that formally we have the fixed point involving the Liar which is $\sigma(b_a^{\neg}, b_a^{\neg})$ where $b_a^{\neg}$[13] is the code for the Liar strategy in (15.a). Now, the Liar, p, knows that g knows that p is the Liar.

---

[12] In out two place notation, the first $b_a^{\neg}$ is code for p's Liar strategy and the second b_a is code for g's mistaken belief of p's strategy.

[13] Formally, $b_a^{\neg}$ may be viewed as the g.n of a refutable proposition in a formal system. A refutable proposition is one whose negation (b_a here) is provable in the system. As theoremhood is a computable relationship, the g.n of the refutable proposition cannot belong to the domain of any computable function. However, as b_a is provable, the set of all such refutable functions is a recursively enumerable subset of the domain of calculations such as $\phi_x(x)$, for all x, do not terminate.

Theorem 4: *The prediction function indexed by the fixed point of the Liar/rule breaker best response function $f_p^{\neg}$ in (16) is not computable and corresponds to the famous Gödel non-computable fixed point.*

$$\phi_{f_p^{\neg}\sigma(b_a^{\neg},b_a^{\neg})}(s) = \phi_{\sigma(b_a^{\neg},b_a^{\neg})}(s). \qquad (16)$$

The proof is standard.[14]

4.3 Surprise Nash Equilibria

There is no paradox in stating that as both players can prove the non-computability of (16) they will have mutual knowledge that the only Nash equilibrium strategies for *both* players that is consistent with meta information in the fixed point in (16), is one that involves strategies that elude prediction from within the system. On substituting the fixed point $\sigma(b_a^{\neg}, b_a^{\neg})$ in (16) for $\sigma(x,x)$ in (8.b), g's Nash equilibrium strategy β_g^E with g.n b_g^E implemented by an appropriate total computable function such as (12.a) must be such that

$$\beta_g^E(f_g\sigma(b_a^{\neg}, b_a^{\neg}), z, s, A) \to \Re\text{- } A \text{ and } f_g = f_g^{E!} = \phi_m, \; m \in \Re\text{ -}A. \quad (17.a)$$

That is, $f_g!$ implements an innovation and $b_g^E!$ is the g.n of the surprise strategy function in (20a) hence is the fixed point of $f_g!$.
Likewise for player p, $f_p!$ implements an innovation in (17.b) and $b_p^E!$ is the g.n of the surprise strategy function viz. the fixed point of $f_p!$. Thus,

$$\beta_p^E(f_p\,\sigma(b_1(b_a^{\neg}), b_1(b_a^{\neg})), z, s, A) \to \Re\text{- } A \text{ and } f_p = f_p^{E!} = \phi_m, \; m \in \Re\text{ -}A. \; (17.b)$$

The intuition here is that from the non-computable fixed point with the Liar, the total computable best response function implementing the Nash equilibrium strategies can only map as above into domains of the action and strategy sets of the players that cannot be algorithmically enumerated in advance.

Using Theorem 4 and Lemma 1, we will now prove the incompleteness results for the strategy sets of the players from the Liar/rule breaking strategy. Analysis will be done for p's strategy set B_p as the productive subset of p's strategy set can be shown to correspond with that of g. The total computable function b_1 in (12.b) and Theorem (3.1,Cutland,1980, p.134) provides a 1-1 reduction between the former and the latter sets.

Corresponding to those (a_{gl}, s) tuples, $a_{gl} \in A_g$ of g's base point optimal strategy for which p's best response f_p is to be rule abiding or rule bending, viz. $f_p \in \{1, f_p^{+}\}$, the g.ns of these optimal strategies for p, $b_p^* \in B_p$ result in computable fixed

---

[14] Assume it is computable and the R.H.S of (18) produces the output $q^{\sim}$ and the L.H.S by the definition of the Liar strategy produces output q. However, if (18) is computable then we have $q=q^{\sim}$ which is a contradiction.

points . This set denoted by $ß_p^{+}$ can be generated by eductive/recursive methodology entailed in the proof of Theorem 3. Thus,

$$ß_p^{+} = \{\ b_p{*}\,|\phi_{b_p*}(b_p{*}) \downarrow \text{ for all } (a_{gl}\ , s),\ a_{gl} \in A_g\ ,\ f_p \in \{1, f_p^{+}\}\}. \quad (18.a)$$

Likewise by Theorem 4 and Lemma 1, we can recursively generate a set $ß_p^{\neg}$ that contains the g.ns of p's strategies for when it is optimal for p to use the Liar best response function $f_p^{\neg}$ to those $(a_{gl}\ , s)$ tuples, $a_{gl} \in A_g$ of g's base point optimal strategy. By Theorem 4, this is a set of p's strategies that can be proven to result in non-computable fixed points. Hence,

$$ß_p^{\neg} = \{b_p{*}\,|\phi_{b_p*}(b_p{*})\uparrow \text{ for all } (a_{gl}\ , s),\ a_{gl} \in A_g\ ,\ f_p = f_p^{\neg}\}. \quad (18.b)$$

For the same $(a_{gl}\ , s)$ tuple, $a_{gl} \in A_g$ constituting g's base point optimal strategy, p's optimal strategy b_p^{*} cannot belong to both $ß_p^{+}$ and $ß_p^{\neg}$. Hence, logical consistency of the meta analyis requires $ß_p^{+} \cap ß_p^{\neg} = \varnothing$ and these are disjoint sets.
Now, define the compliment set of $ß_p^{+}$ denoted by $ß_p^{+c}$ as

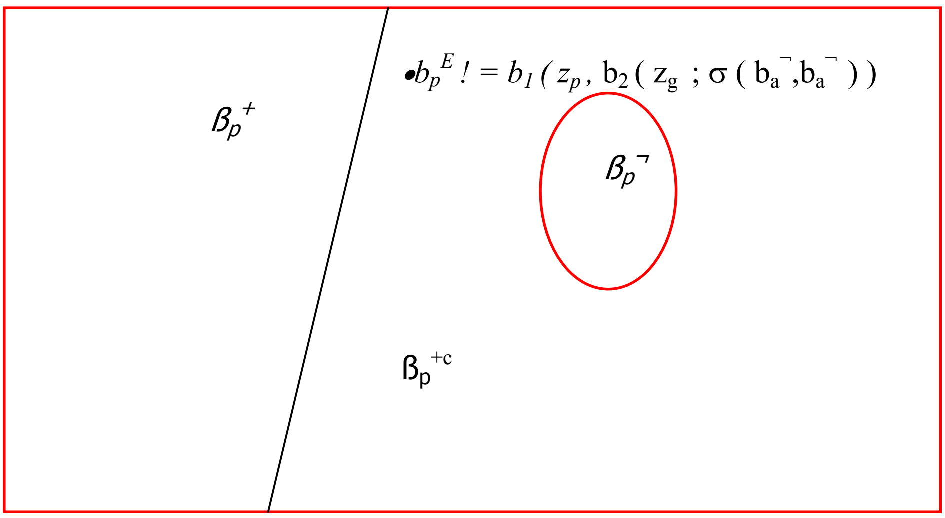
$$ß_p^{+c} = \{\ x \mid \phi_x(x) \uparrow,\ x \in B_p\ \}. \qquad (19)$$

As $ß_p^{+} \cap ß_p^{\neg} = \varnothing$, the two sets are recursively enumerable disjoint sets with $ß_p^{\neg} \subseteq ß_p^{+c}$ by definition in (18.b). Hence, the incompleteness of p's strategy set B_p that arises from the agency of the Liar strategy requires the proof that $ß_p^{+c}$ is productive as in Definition 5 with the g.n of the surprise strategy $b_p^{E}\,!$, $b_p^{E}\,! \in ß_p^{+c} - ß_p^{\neg}$.

FIGURE 3

THE INCOMPLETENESS OF p's STRATEGY SET B_p

$b_p^{E}\,! = b_1(z_p, b_2(z_g, b_a^{\neg}))$:SURPRISE STRATEGY



***Theorem* 5**: *The g.n of player p's Nash equilibrium surprise strategy is defined as* $b_p^E\,! = b_1(\,z_{p\,;}\, b_2(\,z_g\;;\,\sigma(\,b_a^{\neg}\,,\,b_a^{\neg}\,)\,)$ *from (12.b) having substituted in* $\sigma(\,b_a^{\neg}\,,\,b_a^{\neg}\,)$ *from the non-computable fixed point in (16). Then, by construction* b_p^E *is a 'witness' for the productivity of the set* β_p^{+c} *such that* $b_p^E\,! \in \beta_p^{+c} - \beta_p^{\neg}$ *and p's optimal strategy set* B_p *is incomplete. As* $b_p^E\,!$ *is the g.n of the total computable best response function* $f_p!$ *implementing the surprise or the innovation in the system as defined in (17.b),* $f_p!$ *is the productive function for the set* β_p^{+c}.

Proof: See Appendix.

The significance of Theorem 5 is that the surprise strategy is fully definable as a meta–proposition and is paradox free as the surprise strategy is indeed a pure innovation in the strategy set B_p and outside of sets $\beta_p^+ \cup \beta_p^{\neg}$ that can be enumerated by eductive calculation and information in **G**, see Figure 3. It is precisely the absence of logical inconsistency and strategic irrationality in the meta proposition on the surprise strategy that sustains the consistent alignment of beliefs condition of a Nash equilibrium with surprises. Thus, as already observed, for human players utilizing ideal reasoning provided by Gödel meta analysis, the set $\Re$ of best response functions in (6.b) should provide an inexhaustible source of surprise or innovative strategies. However, by the same token, by Theorem 5, there is no algorithmic way by which the prediction function with the index $\sigma(\,b_p^E!,\,b_p^E!)$ at the surprise equilibrium can be mutually identified or learnt ex ante by the players in the system. Indeed, $\sigma(\,b_p^E!,\,b_p^E!)$ says that this is so self-referentially.

Theorem 5 and Figure 3 on the surprise strategy in a Nash equilibrium of a game formally corresponds to the set theoretic proof of Gödel's undecidable proposition in miniature, Cutland (1980). We have succeeded in showing the formal equivalence between the Nash equilibrium with surprise or novelty in Figure 3 and the phase transition in dynamical systems theory that characterizes the endogenous production of novelty as in Figure 1.

Concluding Remarks

This paper sought to give the formal foundations for the phase transition that physicists (Langton,1990,1992) famously call "life at the edge of chaos", viz. the domain in which novel objects emerge. Following the seminal insight from Binmore(1987), we have shown the crucial significance of the Gödel (1931) incompleteness result for the formalization of a system with the endogenous capacity for novelty production. Indeed, as noted by Goldberg (1995) a computational theory of actor innovation is needed for this. In addition to the formal Gödel structure of the Liar on the pure logic of opposition as being a necessary condition for evasive and innovative behaviour, Wolfram (1984) had conjectured that the highest level of computational intelligence, the capacity for self-referential calculation of hostile behaviour was also necessary. Clearly, that agents with the highest level of computational intelligence qua Turing Machines are needed to produce adaptive novelty casts doubt on the Darwinian tradition that random mutation is the only source of variety.

In Markose (2003) it is argued that for systems to stay at the phase transition associated with novelty production requires the Red Queen dynamic of rivalrous coevolving species. In the Ray's Tierra(1992) and Hillis (1992)artificial life simulation models, once computational agents have enough capabilities to detect rivalrous behaviour that is inimical to them, they learn to use secrecy and surprises. In a two person game with computational agents, this can be fully formalized using the Gödelian result on incompleteness. To show how with parallel computing agents, we have cooperation and competition not simply as in Prisoners Dilemma, but with the use of periodic adoption of new institutions outside of extant action sets, we need the new technology of virtual models of emergent phenomena.

Finally, a matter that is beyond this paper, but is of crucial mathematical importance is that objects of adaptive novelty as in the Gödel (1931) result has the highest diophantine degree of algorithmic unsolvability of the Hilbert Tenth problem. This model of indeterminism is a far cry from extant modelling of adaptive innovation or strategic 'surprise' as white noise which in the framework of entropy represents perfect disorder, the antithesis of self-organized complexity. It can be conjectured that a lack of progress in our understanding of market incompleteness and arbitrage free institutions is related to these issues on indeterminism.

References

Albin, P. (1982), "The Metalogic of Economic Predictions, Calculations and Propositions", Mathematical Social Sciences, 3, 329 - 358.

Albin, P. (1998). *Barriers and Bounds to Rationality, Essays on Economic Complexity and Dynamics in Interactive Systems*, Edited and with an Introduction by Duncan Foley, Princeton University Press.

Anderlini, L. and Sabourian, H. (1995). "Cooperation and Effective Computability", *Econometrica* , 63, 1337-1369.

Arthur, W.B. (1991). "Designing Economic Agents that Act Like Human Agents: A Behavioral Approach To Bounded Rationality", *American Economic Review*, 81, 353- 359.

Arthur, W.B., (1993a). "On Learning and Adaptation in the Economy", Working Paper 92-07-038, Santa Fe Institute, Santa Fe,NM.

Arthur, W.B.(1993b). "On The Evolution of Complexity", Working Paper, 93-11-070, Santa Fe Institute, Santa Fe, NM.

Arthur, W.B., (1994). "Inductive Behaviour and Bounded Rationality", *American Economic Review*, 84, pp.406-411.

Axelrod, R. (1987). 'The Evolution of Strategies in the Iterated Prisoner's Dilemma', in Davis, L. (ed.) Genetic Algorithms and Simulated Annealing, Research Notes in AI, Pitman/Morgan Kaufmann, pp. 32-41.

Binmore, K., 1987, "Modelling Rational Players: Part 1", Journal of Economics and Philosophy, 3, 179-214.

Canning, D.1(992), "Rationality, Computability and Nash equilibrium", Econometrica, 60, 877- 895.

Casti, J., (1994). *Complexification : Explaining A Paradoxical World Through the Science of Surprises* , London Harper Collins.

Cutland, N.J., 1980, *Computability: An Introduction to Recursive Function Theory*, Cambridge University Press.

Gödel, K. (1931). 'On Formally Undecidable Propositions of *Principia Mathematica* and Related Systems '(Translation in English in *Gödel's Theorem in Focus,* ed. .S.G Shanker, 1988, Croom Helm).

Goldberg, D.E, (1995). "The Existential Pleasures of Genetic Algorithms , in G. Winter, J. Perioux, M. Galan and P. Cuesta (Eds), *Genetic Algorithms in Engineering and Computer Science,* pp. 23-31, Wiley and Sons.

Hillis, W.D., (1992). “Co-Evolving Parasites Improve Simulated Evolution as an Optimization Procedure”, In G.C. Langton, C. Taylor, J. Doyne Farmer, and S, Rasmussen (Eds.) *Artificial Life II* (Sante Fe Institute Studies in the Sciences of Complexity, Vol 10) Reading MA : Addison-Wesley.
Holland, J.H., (1975). *Adaptation in Natural and Artificial Systems*, The University of Michigan Press, Ann Arbor, MI.
Kauffman, S.A, (1993) *The Origins of Order: Self Organization and Selection in Evolution,* Oxford University Press.
Langton, C.(1990). ‘Computation at the Edge of Chaos: Phase Transitions and Emergent Computation’, in *Emergent Computation*, Ed. By S. Forrest, North Holland.
Langton, C.(1992), “Life at the Edge of Chaos”, In G.C. Langton, C. Taylor, J. Doyne Farmer, and S, Rasmussen (Eds.) *Artificial Life II* (Sante Fe Institute Studies in the Sciences of Complexity, Vol 10) Reading MA : Addison-Wesley.
Lucas, R.(1972). “ Expectations and the Neutrality of Money”, Journal of Economic Theory , 4, pp.103-24.
Lucas, R.(1976), “Econometric Policy Evaluation: A Critique”, Carnegie-Rochester Conference Series on Public Policy, vol. 1, pp.19-46.
Ray, T.S. (1992). “An Approach to The Synthesis of Life”, In G.C. Langton, C. Taylor, J. Doyne Farmer, and S, Rasmussen (Eds.) *Artificial Life II* (Sante Fe Institute Studies in the Sciences of Complexity, Vol 10) Reading MA : Addison-Wesley.
Penrose, R. (1988). “On Physics and Mathematics of Thought”, In *The Universal Turing Machine: A Half Century Survey*, Edited by Rolf Herken, Oxford University Press.
Post, E.(1944). ‘Recursively Enumerable Sets of Positive Integers and Their Decision Problems’, *Bulletin of American Mathematical Society,* vol.50, pp.284-316.
Simon, H.A.,(1981), *The Sciences of the Artificial*, MIT Press, Cambridge, U.S.A.
Markose, S.(2001a) , Book Review of *Computable Economics* by K. Velupillai, In *The Economic Journal*, June.
Markose, S. (2001b), “The Liar and Surprises: So What *Is* The Lucas Critique- A New Perspective From Computability Theory”, University of Essex, Economics Department, Mimeo, http//www.essex.ac.uk/economics/staff/scher.html.
Markose, S. (2002), “ The New Evolutionary Computational Paradigm of Complex Adaptive Systems: Challenges and Prospects For Economics and Finance”, In *Genetic Algorithms and Genetic Programming in Computational Finance*, Edited by Shu-Heng Chen, Kluwer Academic Publishers. Also Essex University Discussion Paper no.552.
Markose, S. (2003), “Computability and Evolutionary Complexity: Markets as Complex Adpative Systems (CAS)”, *Forthcoming,* The Economic Journal Feature, 2004.
Mirowski, P. (2000). *Machine Dreams: How Economics Became a Cyborg Science*. Cambridge University Press, New York.
Rogers, H.(1967). *Theory of Recursive Functions and Effective Computability,* Mc Graw-Hill.
Solé, R., Ferrer, R., Gozalez-Garcia, I., Quer, J., Domingo, E., (1997) “The Red Queen Dynamics, Competition and Critical Points in a Model of RNA Virus Quasispecies”, Santa Fe Institute Discussion Paper No.97-12-086.
Spear, S.(1989). “Learning Rational Expectations Under Computability Constraints”, *Econometrica*, 57, 889-910.

Turing, A., (1936). "On Computable Numbers With an Application To the Entscheidungsproblem " In M. Davis, (ed.) *The Undecidable* , Raven, New York.
von Neumann, J. (1970). *Theory of Self-Reproducing Automata,* ed. Burks, A.W, Urbana: University of Ilinois Press.
Wolfram, S. (1984). "Cellular Automata As Models of Complexity", *Physica*, 10D,1-35.

APPENDIX

PROOF OF THEOREM 1: $\phi_{\sigma(x,y)}(s)$ on the left-hand side of (4) is a UTM, denoted here as $UTM_\sigma = \psi((x,y),s)$. As in (2) by the Iteration or the Parameter Theorem the g.n of UTM_σ is obtained uniformly by the total computable function σ from the indexes (x,y) of the function that it has to simulate on the right-hand side of (4). Thus, the so called substitution function σ (x,y) which implements the Gödel numbering of the prediction functions in matrix Ξ in Figure 2 enumerates the same sequence of computable functions as does the partial recursive function $\phi_X(y)$on the right-hand side of (4) which executes the internal calculations by agents in the system.

PROOF OF THEOREM 5: The proof entails showing that the best response function f_p in (17.b) is the productive function denoted as $f_p!$ with the '!' intended to focus on the feature that an innovation outside given action sets is involved, viz. $f_p! = \phi_m$, $m \in \Re - A$. We will use the two following Lemmas in the proof as well as the property of the set $ß_p^{+c}$ given in (19).

Lemma A.1: Since $ß_p^+$ contains those strategy functions with g.ns x such that $\phi_x(x)\downarrow$, any total computable response function $f_p(x)$ satisfies the condition that $\phi_{f_p(x)}(y)\downarrow \leftrightarrow \phi_x(y)\downarrow$. Note also that the g.n of the strategy function $\beta(f_p(x))$ is given by $b_1(.)$ in (10.a, 12.b). Therefore,

$$f_p(x) \in ß_p^+ \leftrightarrow \phi_{f_p(x)}(f_p(x))\downarrow \leftrightarrow \phi_x(f_p(x))\downarrow \text{ or } f_p(x) \in W_x = \text{Dom}\,\phi_x . \quad (A.1)$$

However, if W_x is disjoint from $ß_p^+$ such that $W_x \subset ß_p^{+c}$ as $\phi_x(x)\uparrow$, then we must have $f_p(x) \in ß_p^{+c} - W_x$. Assume the opposite viz. $f_p(x) \in W_x = \text{Dom}\,\phi_x$, then by the left hand side of (A.1), $f_p(x) \in ß_p^+$ and also that $x \in ß_p^+$ and that $\phi_x(x)\downarrow$. The latter contradicts our assumption that $W_x \subset ß_p^{+c}$ as $\phi_x(x)\uparrow$. Such a $f_p(x)$ the g.n of which cannot be in the recursively enumerable sets $ß_p^+$ and W_x is the *productive function* of the set $ß_p^{+c}$.

Lemma A.2 : Let W_x ,$W_x \subset ß_p^{+c}$ of Lemma A.1 be constructed as $W_{\sigma_{n+1}^{\neg}}$ to yield a non-repeating , recursive enumeration of $ß_p^{\neg}$ and surprise strategies thereof with $\sigma^{\neg}$ (.) denoting this total computable enumerating function such that

$$W_{\sigma_n^{\neg}} = ß_p^{\neg} \quad , \quad ß_p^{\neg} \subset W_{\sigma_{n+1}^{\neg}} , \qquad (A.2)$$

$$\text{and, } W_{\sigma_{n+1}^{\neg}} = W_{\sigma_n^{\neg}} \cup \{f_p(\sigma_n^{\neg})\} . \qquad (A.3)$$

Here with no loss of generality, let

$$\sigma_n^{\neg} \equiv \sigma^{\neg}(b_a^{\neg}) \qquad (A.4)$$

signify the non-computable fixed point from p's Liar/rule breaking strategy in (15.a) and Theorem 4 and

$$\sigma_{n+1}^{\neg} = \sigma^{\neg}(b_p^E !) . \quad (A.5)$$

In other words $W_{\sigma_n^{\neg}}$ contains a recursive listing of the members of $ß_p^{\neg}$ defined in (18.b) and surprise strategies thereof as $\{b_o^{\neg}, b_1^{\neg}, \ldots\ldots, b_n^{\neg}\}$. However, note in the construction of $W_{\sigma_n^{\neg}}$ at no time can the g.n of $f_p(\sigma_n^{\neg}) \in W_{\sigma_n^{\neg}}$. If the g.n of $f_p(\sigma_n^{\neg})$ $\in W_{\sigma_n^{\neg}}$, then following Lemma A.1 (left hand side), as f_p is total, $f_p(\sigma_n^{\neg}) \in ß_p^{+}$ and on using (A.4) we have $\phi_{b_a^{\neg}}(b_a^{\neg})\downarrow$. However, this leads to a contradiction as from (15.a) $\phi_{b_a^{\neg}}(b_a^{\neg})\uparrow$ and hence $f_p(\sigma_n^{\neg}) \notin W_{\sigma_n^{\neg}}$ and cannot be recursively enumerated . Hence, $f_p(\sigma_n^{\neg})$ is a productive function. Note, that $b_p^E ! = b_1 (z_p ; b_2 (z_g ;\sigma (b_a^{\neg}, b_a^{\neg})$ is the g.n of the Nash equilibrium surprise strategy with response function $f_p^E!$, and hence $f_p(\sigma_n^{\neg})$ must be as required in (17.b), viz. $f_p^E! = \phi_m$, $m \in \Re$ -A.