COMPLEX TYPE 4 STRUCTURE CHANGING DYNAMICS OF DIGITAL AGENTS: NASH EQUILIBRIA OF A GAME WITH ARMS RACE IN INNOVATIONS

Sheri M. Markose*

Economics Department, University of Essex
Wivenhoe Park, Colchester, UK

(Communicated by Athanasios Yannacopoulos)

Abstract. The new digital economy has renewed interest in how digital agents can innovate. This follows the legacy of John von Neumann dynamical systems theory on complex biological systems as computation. The Gödel-Turing-Post (GTP) logic is shown to be necessary to generate innovation based structure changing Type 4 dynamics of the Wolfram-Chomsky schema. Two syntactic procedures of GTP logic permit digital agents to exit from listable sets of digital technologies to produce novelty and surprises. The first is meta-analyses or offline simulations. The second is a fixed point with a two place encoding of negation or opposition, referred to as the Gödel sentence. It is postulated that in phenomena ranging from the genome to human proteanism, the Gödel sentence is a ubiquitous syntactic construction without which escape from hostile agents qua the Liar is impossible and digital agents become entrained within fixed repertoires. The only recursive best response function of a 2-person adversarial game that can implement strategic innovation in lock-step formation of an arms race is the productive function of the Emil Post [58] set theoretic proof of the Gödel incompleteness result. This overturns the view of game theorists that surprise and innovation cannot be a Nash equilibrium of a game.

1. **Introduction.** Operations on discrete chunks of encoded information are the hallmark of digital and computational phenomena, which are known for their capacity to produce error free copies. Recently, economists Brynjolfsson and McAfee [14]

described the implications of the widespread use of digitization in processes and products of the so called Second Machine Age[1]. In addition to important insights in their work on how digital technologies with their network effects and winner take all prizes will lead to some "weird and wonderful economics", Brynjolfsson and McAfee [14] assert that in the digital age, the easy access to a growing library of encoded processes and products gives rise to 'recombinant innovations'. There is now a widely held view that operations on digital information, also referred to as recursive or computable functions[2], contribute to the key building blocks for innovation in wide ranging areas. These include evolution and operations of the genome (Holland [36], Ben-Jacob [11], Mirowski [49], Pellionisz [57], Lieberman-Aiden [40]); protean behaviours associated with the growth of the neocortex, complex neural dynamics in cognition (Markose [44], Corballis [20], Tsuda [70]); human language (Chomsky [19, 18], Hauser et. al [32]); and, human technological progress (Arthur [6], Beinhocker [10], Youn et. al [74]).

The focus of the paper is on *How Can Digital Agents Innovate*? Digital agents operate on encoded information and represent a wide array of phenomena covering the artificial and biological. Following the early provenance of John von Neumann [73] in the 1940s, models of dynamical systems associated with digital agents involve recursive function theory rather than differential equations. Von Neumann drew on his work on self-reproducing automata as models for complex biological systems to engineer a radical transformation of dynamical systems theory from one that is based on motion, force and energy to the capabilities and constraints of information processing agents modelled as computing machines which operate on information in an encoded form[3]. Hence, the paper will give an overview of some of the key mathematical and scientific developments on the role of digital agents and recursive function theory, especially in the context of novelty production, in genomic evolution, cognition, economic systems and game theory. The paper contributes to the nexus on dynamics, games and science, by formalizing an important class of dynamics associated with a strategic game between oppositional digital agents in which innovations occur.

The significance of the von Neumann computational legacy of complex adaptive systems (CAS) is that it covers all substrata, ranging from an intra-cellular biochemical medium to an artificial one of silicon chips, in which effective procedures or computation reside. It is the pioneering work in mathematical logic by Gödel [28], Turing [71], and Post [58] (GTP logic from here on) that laid the foundations of what digital agents can and cannot do. By the Church-Turing thesis (Cutland [21]), the intuitive notion of effective procedures or an algorithm is identified with the class

---

[1]Brynjolfsson and McAfee [14] discuss how the low cost replication with digital technology with its rapid ubiquity and perfect fidelity result in "weird and wonderful economics" that include Power Law outcomes for income distribution.

[2]See Soare [66] for an interesting discussion of the history behind the wider use of the term recursive function theory for computability theory. As two popular textbooks by Rogers [61] and Cutland [21], extensively referred to in my paper, respectively, use each of these terms as titles of their books and proceed to discuss the same subject matter relating to agents that conduct operations on encoded information, I will use the terms equivalently. Also reference to agents being digital or computational will be made equivalently in so far as both classes of agents operate on discrete alpha-numeric bases for encoding information.

[3]Mirowski [49] 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.

of general recursive functions[4] and represent finitely encodable programs executed in a number of equivalent ways referred to as automata or mechanism.

Turing machines (TMs) (Turing [71]) are the best known ideal class of the most powerful digital agents. Following the epochal work of Gödel [28] which laid the foundations of encoding all syntactic objects in terms of integers called Gödel numbers (g.ns) or indexes, Turing machine operations along with their inputs and outputs can be given a meta-representation by using further codes in what is called a formal system, Smullyan [65]. Post [58] developed the theory of sets of indexes of TMs, both in terms of the domain and range of TM operations that can be listed by TMs when they halt. These are called recursively enumerable sets. This led to an ingenious formulation, from well-known logical antecedents of the Cantor diagonal lemma and the embodiment of opposition given by the Liar[5], (Gödel [28]), of a fixed point logic gate involving negation, typically referred to as the Gödel sentence, that marks the exit route from recursively enumerable sets of syntactic objects, in the formal system. This is the celebrated Gödel incompleteness result in that TMs produce syntactic objects, called undecidable propositions that no TM can enumerate. Though often regarded as a limitative result on what computers can do, as Post [58] shows with his aptly named creative and productive sets in the set theoretic proof of the Gödel incompleteness result, a recursive mapping from a Gödel sentence entails a productive function that provides an exit strategy from known recursively listable sets of encoded technologies to produce new digital objects that can be arranged in a structure of a never ending arms race (see, Smullyan [65], Cutland [21]).

Hence, on considering the von Neumann nexus on dynamical systems as outcomes of computing agents, we have the so called Wolfram-Chomsky schema (see, Wolfram [75], Chomsky [18][6], Casti [17], Langton [39], Albin [2], Markose [45, 44]). The Wolfram-Chomsky schema shows that on varying the computational capabilities of agents, different system wide dynamics can be generated: finite automata produce **Type 1** dynamics with unique limit points or homogeneity; push down automata produce **Type 2** dynamics with limit cycles; linear-bounded automata generate **Type 3** chaotic trajectories with strange attractors. The significance of

---

[4]General recursive functions include all elementary arithmetic, logical operations and also functions obtained from substitution, iteration and recursion. In the latter, functions call on themselves and use inputs that are outputs from previous calculations. See, Cutland [21] and Rogers [61], which are well known text books on this.

[5]The Liar is the embodiment of a hostile agent which can negate, controvert or falsify a code that it can successfully enumerate or predict. Since antiquity, it has been known that self-refuting statements generate paradoxes as in the Cretan Liar proposition: *this is false*. Gödel's analogue of the Liar proposition is the undecidable proposition. The latter, denoted as **A** in some appropriate code, has the following structure : $\mathbf{A} \leftrightarrow \neg Prov(\mathbf{A})$. $Prov(\mathbf{A})$ is a recursive function and **A** is regarded to be the fixed point of $\neg Prov(\mathbf{A})$ (see, Moschovakis [50], Gaifman [24]). **A** effectively says of itself that it is not provable in the formal system. However, unlike the Cretan Liar there is no paradox in Gödel's undecidable proposition as it can be proved that this is so, viz. $\vdash \mathbf{A} \leftrightarrow \neg Prov(\mathbf{A})$, where $\vdash$ denotes the symbol for theoremhood in the formal system. Any attempt to prove the proposition **A** results in a contradiction with both **A** and $\neg\mathbf{A}$, its negation, being provable in the system. Simmons [64, p.29] has noted how with the Cantor diagonal lemma (which proves that the power set of a set has greater cardinality than a set) we begin to have so called "good" uses of self-refuting structures that result in theorems rather than paradoxes.

[6]In Chomsky [18], the schema for the grammar of languages, which is called the Chomsky hierarchy, reverses the sequence of numbering with **Type 0** being the most complex. **Type 0** represents languages such as English, the sentences of which require the generative powers of Turing Machines. More restrictive automata identical to the ones listed in the Wolfram schema, respectively, produce regular, context free and context sensitive grammars.

this schema is that it postulates that only agents with the full powers of Turing machines capable of simulating other Turing machines can produce **Type 4** irregular innovation-based structure changing undecidable dynamics. The epithet of complex adaptive system (CAS) is applied exclusively to **Type 4** dynamics by Casti [17], Langton [39] and Albin [2], who also refer to the novelty and surprises that are manifest in **Type 4** dynamics[7].

In the spirit of von Neumann [73], Goldberg [29] claims that the mystery shrouding innovation can be dispelled ... "by a heavy dose of mechanism. Many of the difficulties (in this area) ....come from a lack of a computational theory of actor innovation". In Binmore [13], the eminent game theorist proposed the requisite dose of mechanism and pointed to the fundamental strategic game theoretic aspect behind innovation in digital agents. He was seminal in proposing the need to inject Gödel logic, specifically in the form of the adversarial or contrarian agent of the Liar, who can negate what it can predict/compute, into the extant game theory framework. This would then engender radical indeterminism in the form of Gödel incompleteness. However, to date game theory has remained resolutely closed and complete in the logical sense with no scope for novelty and surprises in the form of new digital objects. Indeed, as will be reviewed, it is alleged that a central solution concept of game theory, viz. Nash equilibrium, cannot accommodate novelty or surprises.

To rigorously define novel digital objects, depending on the scientific domain, it is useful to consider these to include the set of all potential actions or behavioural phenotypes and technologies or artefacts, which are extended phenotypes, Dawkins [22][8]. These can be represented as indexes of TMs that always halt (also known as total computable functions). Such a number set of all potential technologies denoted by $\mathfrak{R}$, is infinite and also not recursively enumerable, viz. there is no systematic way of 'searching' or listing this set. Some finite subset of this set entails g.ns of known technologies and can represent a given listable action set **A** of traditional game theory. A novelty or a surprise is an encoded object in the set $(\mathfrak{R} - \mathbf{A})$, ie. outside of set **A** that contains known technologies. The question is how can a formal system of the game involving digital agents, embedded in each of the agents, produce a total computable strategy function, which is called a surprise strategy, to exit the given action set **A** in the Nash equilibrium of the game?

Following Binmore [13], the objective of the paper is to use the GTP logic to formalize a 2-person adversarial game involving the Liar to examine how the wherewithal of Gödel meta-mathematics to encode the Gödel sentence becomes the necessary condition for digital agents to exit from known listable technologies or phenotypes and implement **Type 4** dynamics. In the absence of this, digital agents will be entrained within a fixed repertoire of actions. The paper will base the proof of novelty production and an arms race in it on what many regard (Smullyan [65,

---

[7]With regard to this, Mirowski [49, p.141] has asserted that mathematicians finally have blazed the trail to a formalised logical theory of evolution. Given extensive discussions on what constitutes a novelty in evolution (see, for example Müller and Newman [51]), **Type 4** novelty production rules out all TM listable codes for phenotypes and singles out only those digital outputs that are undecidable.

[8]Dawkins [22] uses the notion of extended phenotype to include the myriad artefacts made by organisms that reside external to themselves. Examples of this range from ant hills, bees hives, beaver dams, on the one hand, to gun powder, skyscrapers and iPhone devices.

p.58], Soare [66][9]) to be the most elegant and intuitive approach to Gödel incompleteness using the Post [58] productive function. One of the main contributions of the paper is to show that the Post productive function implements the Nash equilibrium surprise strategy function of the adversarial game and therefore can provide the arms race like arrangements observed in the exponents of CAS which manifest **Type 4** dynamics.

The rest of the paper is organized as follows. Section 2 will survey recent developments in science, which can help demarcate between a specific application of recursion in many important morphological arrangements in organic and non-organic forms in nature and the necessity for the full gamut of GTP logic underpinning the **Type 4** dynamics associated with three main exponents of CAS. These include evolutionary biology and immunology; the brain and cognition; and strategic interactions in socio-economic systems with capitalism, which has long been identified by Schumpeter [63] as manifesting disruptive novelty production. Two features stand out. The concept of protean behaviours, so called after the Greek God Proteus, has been mainstreamed by Driver and Humphries [23], Byrne and Whiten [15] and Miller [48], to give an evolutionary explanation for why living organisms resort to surprise and change as a strategy to escape from hostile agents. Secondly, with the Red Queen[10] having become emblematic of competitive co-evolution, there is growing evidence of the Red Queen type arms race in all these exponents of CAS. Section 2, therefore, provides some key evidence for the postulate that phenomena ranging from the quintessential digital agency of the genome, which Ben-Jacob [11] presciently referred to as the 'creative genome' from which emanates a plethora of junk DNA[11] , to the recent discovery of the mirror neuron system in human cognition and complex protean interactions, the innovations far from being the consequence of being random mutation, can be formalized as a manifestation of the GTP logic.

The GTP logic underpinning the 2-person game with TMs, which purports to produce an arms race with innovations in phenotype or technology when oppositional interests prevail, will be outlined and organized in the following sections.

(i) GTP logic requires that agents operate on encoded information; they can encode all operations and store codes. Section 3 starts with the GTP preliminaries on encoding in terms of integers famously called g.ns/indexes and the Post [58] set theory for Gödel incompleteness results based on the enumeration of indexes of

---

[9]Soare [66] states: "Post's papers brought excitement, intuitive appeal, and an informal style of proof, much closer to ordinary mathematical proofs, and represented the real birth of the subject of recursively enumerable sets". Smullyan [65, p.58] has noted that the pair of recursively enumerable disjoint sets entailed in the creative and productive sets, play a fundamental role in modern approaches to incompleteness and undecidability

[10]The Red Queen is the character in Lewis Carol's *Alice Through the Looking Glass*, who signifies the need 'to run faster and faster to stay in the square'. The Red Queen hypothesis first proposed by van Valen [72] has become emblematic of the outcome of competitive co-evolution for evolutionary biologists in that no competitor gains absolute ground, see Markose [44].

[11]See, the July 2016 feature in the *New Scientist* on the puzzle regarding the relative small size of the number of genes in the human genome relative to the more numerous accretions of junk DNA which account for the characteristics that make us human. This was first discussed in Ohno [55]. This paper gives precedence to the work of Ben-Jacob [11] in the context of GTP logic and the genome. However, it must be noted that the biologist Barbara McClintock [47] in her Nobel Prize lecture of 1984 made a bold conjecture that has led to the notion of a 'dynamic genome', which can initiate mutations in response to specific stresses being experienced, thereby suggesting that genomic mutations need not all be random.

TMs. For our purposes to show that the Nash equilibrium surprise strategy is a non-trivial Post productive function, **Lemma 3.8** in Section 3 is instrumental in showing how incompleteness can be formalized in a general setting using recursive reduction to the archetypical creative and productive sets.

(ii) The second ingredient of GTP logic is the capacity of digital agents to do Gödel meta-mathematics. In Section 4, this is shown to be crucial for TM agents to play an interactive game. Firstly, agents have the capacity to do *offline* simulations corresponding to a 1-1 recursive mapping with online machine execution, which precipitates an action or phenotype in the agent. This Gödel meta-mathematics, well known from Rogers [61], which also involves the two place Gödel substitution function, will be shown to have a close bearing on the recent discovery of the mirror neuron system (MNS) on how self and the other can be integrated in a meta-system. The so called MNS hypothesis postulates that understanding others involve self-referential meta-calculations arising from encoded neural imprints emanating from agent's own execution of procedures via the canonical neurons. Section 4 provides an original graphical illustration of the so called MNS hypothesis for action prediction in the other and the two place Gödel substitution function of Gödel meta-mathematics, Rogers [61]. The stock in trade of recursive function theory to do with diagonal arrangements and the Second Recursion Theorem is shown to characterize Nash equilibria in a two place diagonal array of the fixed point mappings of recursive strategy functions.

(iii) The third ingredient of the GTP logic is that agents record negation and can process the logical archetype of the Liar in a fixed point setting. In Section 5, **Theorem 5.2** will prove that the only exit route from relevant recursively enumerable strategy sets requires the negation operation of the Liar or contrarian strategy function, which will be denoted by $f^\neg$. This Liar or contrarian strategy function, $f^\neg$ operates on a specific code of the opponent, which is amenable to action prediction, and will negate or subvert it. The first significant point is that the Liar can win only *out of equilibrium* when the identity of the Liar is not recorded by the other. The second and more famous result is that when there is mutual or common encoding of the Liar, we are at Gödel's non-computable fixed point encapsulated in the Gödel sentence.

Thereafter, as shown in **Theorem 6.2** in Section 6, any non-trivial recursive strategy function based on the Gödel sentence is the so called Post productive function (Post [58]) of **Lemma 3.8** of Section 3 that determines the logical necessity for surprise mappings into the set $(\mathfrak{R} - \mathbf{A})$. As the productive function in logic maps outside recursively enumerable strategy sets to implement novelty and surprise, it will be denoted as $f^!$, and it will be shown to be the Nash equilibrium surprise strategy function.

2. Recursion and GTP logic in novelty production: A survey.
This section will provide a brief overview of the use of recursive function theory in nature and evolutionary frameworks and also in cognition and game theory.

2.1. Recursion in nature and evolution.
Many accounts of the power of recursion refer to the role of recombinations or concatenations of previously encoded algorithms which in turn can be reused as inputs for further outputs. Many economists (Arthur [6], Youn et. al [74], Beinhocker [10], Brynjolfsson and McAfee [14]) give similar descriptions for the scope and speed of innovations in the digital revolution

unfolding in advanced economies. In the case where recursive functions are able to call on themselves to produce self-similarity at multiple (iterated) hierarchical scales, we have the famous discoveries that include how these encoded algorithms produce the fractal geometry of DNA packaging[12] (see, Lieberman-Aiden et. al. [40], Pellionisz [57]), as well as the morphology of many biological organisms and organs. Chomsky [19, 18] had adduced similar generative powers of TMs as a condition for human language usage to produce embedded and hierarchical sentence construction, but, in addition, he has postulated an unbounded creativity of the human language, which is part of the recursive framework, but can surpass what can be recursively enumerated. Indeed, Chomsky and co-authors (Hauser et al. [32]) famously proposed grammatical recursion as being unique to human language. Corballis [20] has argued that the incorporation of recursion into other cognitive domains predates its incorporation into language. In the case of genetic algorithms, Holland [36] highlighted the role for random recombination and concatenation of 'genetic' codes from parents as a source of diversity in progeny which are then selected for their fitness.

It is argued that while the machine execution of a simple recursive fractal code produces what are recognized to be phenomenal hierarchical self-similar patterns and clearly illustrate the power of recursion, both this and the Holland [36] model of random recombinations of syntactic objects in his genetic algorithms fall short of the important aspect of the GTP model for innovations as new syntactic objects that are forced outside extant recursively enumerable sets, which arise from the response of digital agents to an encoding of negation or opposition.

Two important features have been observed in the **Type 4** dynamics of well known examples of CAS. The first is the strategic aspect of protean behaviours arising from adversarial situations and the second is the Red Queen type co-evolutionary dynamics that accompanies the innovation process.

As seen in Kashiwagi and Yomo [37] and in a related growing literature (see, Samson et. al. [62]), experimental molecular studies on organisms such as E-coli, RNA virus and bacteriophages are designed so that the changes in phenotypes and genotypes of coevolving parasite-host pairs can be monitored through the arms race. Kashiwagi and Yomo [37] find that the virus and its host can coexist in an evolutionary arms race, despite a difference in genome mutability (i.e., mutations per replication) of approximately one to three orders of magnitude. The evidence that mutual genomic and phenotype changes are triggered in a synchronized manner in an arms race between adversarial microbes that permit coexistence, directly controverts random mutation as the sole driver of evolution of new forms, Ben-Jacob [11] and McClintock [47].

Baumol [8, 9], who has extensively documented and discussed the role of the relentless Red Queen type strategic arms race in innovation by firms of products and processes in capitalism, claims that this is not addressed in mainstream economics[13].

---

[12]On the cover page of **Science** for the article by Lieberman-Aiden et. al. [40], Mr. Lander, one of the co-authors and also Science Adviser to the US President, says: "*Mr. President, the Genome is Fractal!*" The important feature of the power of a recursive function, viz. an algorithm like the one that produces fractal morphology in biological systems, is that it only requires a few lines of code and a bio-chemical equivalent of a linear bounded automata to execute it. The same fractal geometry can be seen if the code for this algorithm is run on a personal computer. Note, fractals belong to the class of chaotic dynamics, Mandelbrot [42].

[13]In the wake of the 2007 Great Financial Crisis, renewed efforts are being made to come to terms with arms races in novelty production, especially in the monetary and financial system (see, Haldane [31]).

Recently, in the wake of the 2007 Global Financial Crisis, there has been criticism that economists gave insufficient attention to a regulator-regulatee arms race. This is arguably no different from a parasite host dynamics in immunology, as in for example, Kashiwagi and Yomo [37], which requires monitoring and production of countervailing new measures by authorities (comparable to the production of antibodies) if system failure is to be avoided and competitive coexistence of diversity is to be preserved, Axelrod [7]. Thus, though fully formalizable in the Gödel logic, insufficient attention has been paid to role of the adversarial structure of negation inherent in a fixed point formulation of the Gödel sentence as the logical driver of incompleteness in systems involving digital agents. Likewise, there has been little discussion on the arms race structure of the Post [58] productive sets and its correspondence with many empirical observations on the arms races in phenotypes and extended phenotypes (Dawkins [22]) of interacting organisms.

What evidence is there in manifestations of CAS that involve human and microbial agents that they have the wherewithal to produce syntactic objects equivalent to the Gödel sentence, which are the only known logical exit points from a fixed repertoire of actions for digital agents?

At the level of cell biology with regard to threats by microbial agents to a code relating to the genome of a host, Eshel Ben-Jacob [11][14] is one of the first to explicitly discuss the relevance of Gödel logic both in the detection of threats and also to adduce creative genomic changes that alter phenotype that can counter threats[15]. It is highly significant that Ben-Jacob [11] used the Gödel framework to throw light on the growing evidence that mutual genomic changes arising from adversarial interaction at the level of microbes can drive evolution, over and above random mutations. Ben-Jacob presciently discussed the following capabilities that 'genome units' should possess (*Ibid.* p.59) to be able to engineer, possibly as emergent outcomes, novel artifacts and behaviours: (i) Capacity to recognize the difficulties in environment or hostile conditions that are a threat to a code relating to some functionality; (ii) Need for advanced language that includes self-referential mapping; Finally and significantly, (iii) Even with earliest microbial forms of life, there is evidence of evolved social capabilities of recognizing signals between self and the other which requires a rich (chemical) language.

As will be shown, these correspond to the main ingredients of Gödel metaanalyses from which incompleteness follows. The recent remarkable molecular experiments on changes in phenotypes and genotypes of coevolving parasite-host pairs (see, Kashiwagi and Yomo [37] and Samson et. al. [62]) have recorded both the original codes in the respective genomes of the host and parasite and the novel codes that evolved in the course of the arms race. With the genome being a quintessential digital agent, it can be conjectured that from the very get-go of the genome itself, it has the recursive powers to generate syntactic objects called the Gödel sentence which identifies in a self-referential way that a code is under attack in order to exit from fixed predictable phenotypes.

2.2. **Gödel meta-mathematics and the mirror neuron system.** Just as there is some evidence for the relevance of Gödel logic for the genome, explicitly presented

---

[14]In conversation with the author, Ben-Jacob underscored that getting smart in evolutionary terms is war by other means.

[15]Despite his seminal insights, it is beyond the scope of this paper to discuss certain antimachine views of Ben-Jacob [11] which make claims for semantic knowledge that go beyond the sub-personal syntactic expression of the Gödel sentence.

as such by Ben-Jacob [11] for the necessity of explaining protean genomic arms races, the recent discovery of the mirror neuron system (MNS) by the Parma group (Gallese et. al. [26], Rizzolatti et. al. [60]), can be regarded as one of the most important scientific discoveries of the late 20th century to understand complex social behaviours and in particular the über proteanism of humans[16].

Gallese [25], Gallese and Sinigaglia [27] have characterized the MNS as a common neuronal platform for conducting *offline embodied simulations* for action prediction in the other based on a parallel set of neurons that fire during action execution by one-self[17,18]. The neurons that fire with actual action execution are called *canonical neurons* (Arbib and Fagg [5]). This discovery has revolutionized our understanding of social coordination. While Gallese and others make a descriptive analogy between MNS and offline simulation, they do not refer to the two place Gödel substitution function in Gödel meta-mathematics (Rogers [61]) which can provide the mathematical model of what is meant by offline simulation and also how it can implement action prediction regarding self and the other using recursive self-referential mapping[19]. Indeed, as noted, Gödel meta-mathematics is about the 1-1 syntactic offline recording of machine execution, with the latter in the MNS corresponding to the firing of canonical neurons during action execution.

In addition, in order for the mirror neuron system to fulfil the full scope of the GTP logic, the MNS must include operations that entail negation that underpins the Liar or opposition. Again, in a remarkable set of experiments by Scott Kelso and his group (see, Tognoli et. al. [69], Naeem et. al. [53]), the recording of neurophysiological markers of anti-coordination that arises from actions that need to be different/opposite from what is predicted of another, indicate that these are also part of the human mirror neuron system. This is an important piece of evidence for Gödel logic to be relevant as an analogue of cognitive incompleteness associated with the capacity to 'think outside the box' and produce protean behaviours[20].

---

[16]Byrne and Whiten [15] have presented extensive evidence on the development of the Machiavellian brain. Ramachandran [59] has famously underscored the importance of the discovery of the mirror neuron system as being crucial for explanations of complex social interactions especially in the context of man as the Machiavellian primate. The latter is conjectured to require mind reading and action prediction in others.

[17]The neurons that fire with actual action execution are called *canonical* neurons (Arbib and Fagg [5]) which will be shown to correspond to on-line machine executions in the GTP logic.

[18]Ramachandran [59] describes this as follows: "It's as if anytime you want to make a judgement about someone else's movements you have to run a VR (virtual reality) simulation of the corresponding movements in your own brain and without mirror neurons you cannot do this."

[19]Tsuda [70] identifies how neural systems which need to process a self-referential description use the mirror neuron system as in the mathematics of the Gödel's incompleteness theorem : "When neural systems process a self-referential description, they may first have to make a copy of the object of self-reference and then refer to this copy. This two-stage formulation can be realized mathematically in the proof of the Gödel's incompleteness theorem through the processes of projecting mathematical statements to natural numbers and of referring to meta-mathematical statements by providing mathematical statements about such numbers. The presence of mirror neurons in animal brains or mirror neuron systems in human brains may also be a realization of the above two-stage formulation in brains, because mirror neurons, or mirror-neuron systems, can be activated, not only by behavior in others similar to one's own behavior, but also by one's own behavior. This can be represented in a dynamical systems model." Despite this important insight on the mirror neuron system and Gödel logic in Tsuda [70], Tsuda does not refer to the MNS as the model of offline simulation given in Gödel meta-analysis, Rogers [61].

[20]F.A. Hayek is the first economist who discussed the implications for economics that arise from the problems of non-computability that he called the *limits of constructive reason* and on the possibility that the brain manifests Gödel incompleteness (Hayek [33, 35]). Due to the fortuitous

Only those well versed in GTP logic will see the logical necessity for negation to be part of the MNS for it to be the source of proteanism.

2.3. **Recursive function theory and game theory.** The final building block that is needed to fulfil the objective of this paper is the development of a two person oppositional game involving Turing machines in which **Type 4** dynamics of an arms race in technology or phenotype follows. As noted in Markose [45, 44], it is the game theorist Binmore [13], who first seminally raised the "spectre of Gödel" (*ibid*) in the context of a game between Turing machines. The question that Binmore asks is whether the scope of strategic behaviour can be restricted to a system that is logically closed and complete. In game theory, there are strategy mappings to a *fixed* action set and indeterminism extends only to randomizations between given actions. Binmore states that in outlawing strategic indeterminism or novelty production, the question that is pertinent here is the center piece of Gödel [28]: '*what of the Liar*?' By this is meant that when faced by a hostile agent who will falsify or negate one's actions if he could deduce what they are, can one rationally play such an action which can be formally deduced/predicted by the adversarial TM? Or does one innovate and 'surprise' the opposition? However, here we are confronted by the omission noted by Binmore [13] that relates to the assumption of the completeness paradigm in which extant game theory is couched[21].

As reiterated by many game theorists, it is held that it is not possible to produce novelty or surprises in a Nash equilibrium, let alone the structure of an arms race in strategic innovation. Bhatt and Camerer [12] succinctly state this: "in a Nash equilibrium nobody is surprised about what others actually do, or what others believe, because strategies and beliefs are synchronized, presumably due to introspection, communication or learning." What is missing in this statement is the category of mutual expectation of surprise and the characterization of a Nash equilibrium in which players logically expect that they will need to surprise and be surprised. As will be shown, there is nothing inherent to a Nash equilibrium in which the strategic necessity of a surprise cannot be formulated. Indeed, novelty and surprise become a logical necessity to avoid inconsistency.

The game theory papers such as Albin [1, 2], Anderlini [3], Anderlini and Sabourian [4], Canning [16], Nachbar and Zame [52], which use recursive function theory, tend to focus exclusively on defining the problem of indeterminacy associated with self-refuting decision structures. These game theory papers do not discuss the possibility for surprise strategies and innovations that are arranged in a structure of an arms race. The problems here are twofold. These papers do not utilize the major methodological triumph of Gödel [28] which is the meta-analysis that produces fully

---

Viennese connection between Hayek and Gödel, Hayek seminally redirected the discussion on the limits of deductive inference from Humean scepticism to the Gödel logic of incompleteness (Markose [46, 44]) and indeed brought this to my attention by instructing me to read his book on the *Sensory Order* (Hayek [33]). However, Hayek's own account of this did not go beyond the Cantor diagonal lemma (see footnote 5), which led him to a view on cognitive incompleteness in terms of much knowledge that cannot be formally enumerated. As with the many so called computational models of cognition, which do not go beyond descriptive analogy, Hayek's theory of cognitive incompleteness though seminal in linking it with the Gödel incompleteness result, falls far short of the full paraphernalia of GTP logic. Also, as Hayek's work predates the discovery of the mirror neuron system, there was scant evidence that the neurophysiology of the brain had any powers for *offline simulation* associated with Gödel meta-mathematics.

[21]In many respects, this can be seen as being analogous to the pre McClintock [47] era in gene research.

definable meta propositions, in an ever extendable sequence, in terms of the Post [58] productive functions that provide recursive mappings from self-refuting fixed points, viz. Gödel sentences, to the outside of given recursively enumerable sets in order to avoid logical inconsistency. Hence, the aphorism that sufficiently rich formal systems cannot be both consistent and complete. Secondly, the characterization of Nash equilibria as fixed points of recursive strategy functions seem not to be specified as such. Though Spear [67] does not explicitly depict a game, Spear [67] was the first economist to use the Second Recursion Theorem (Cutland [21]) to model fixed points of recursive functions. Arguably, without the proper formalism for characterizing fixed points of recursive functions, the Nash equilibria of a game which requires the identification of the meta-representations of mutual best response functions in a two place diagonal alignment, one could be forced into different "resolutions" of classic oppositional problems[22].

3. **Post [58] productive function approach to Wolfram-Chomsky type 4 dynamics.** Here we start to formalize a 2-person game for TM agents that operate on encoded information and also the Post [58] formalism for a model of novelty production.

3.1. **Mathematical preliminaries.** The method for encoding introduced by Gödel [28] called Gödel numbering, implies that encoded information is represented in terms of integers. All objects constructed from a countable alphabet can be put into $1-1$ mapping with the set of natural numbers referred to as their Gödel numbers (g.ns, for short). As computers involve operations on codes, they are number theoretic functions, $f : \mathcal{N} \to \mathcal{N}$, where $\mathcal{N}$ is the set of all integers[23].

Each computable function[24] 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 adopted in Cutland [21], we state this for a single valued computable function as follows:

$$f(x) \cong \phi_a(x) = q. \tag{1}$$

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 ($\sim$) when $\phi_a(x)$ does not

---

[22]Koppl and Rosser [38] attempt to characterize the Nash equilibrium of the zero sum game that depicts the machinations of the well-known oppositional game involving Holmes and Moriarty using recursive function theory. Moriarty who seeks the demise of Holmes has to be in proximity with him while Holmes needs to elude Moriaty. Koppl and Rosser [38] conclude as follows: "We can see that there are best-reply functions, $f(x)$, such that $f(x) \neq x$ for all x. That is, there are best-reply functions without a fixed point. A fixed point is defined by the condition that $f(x) = x$." As outlined in footnote 5, it will be shown that Gödel meta-mathematics has no problem in giving a syntactic expression of the fixed point of the best response function that seeks to negate or deceive as in the Holmes-Moriarty game. The important point here, therefore, is not that one or the other player has to find a best response function that does not have a fixed point, but that the fixed point of an important class of best response functions is a self-referencing computer code of a recursive function that will not halt and further this syntactic undecidable object falls outside recursively enumerable strategy sets.

[23]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 $\mathcal{N}$ being denoted by $\mathcal{N}_0$, while from Cantor we know that the set of all number theoretic functions have cardinality of $2^{\mathcal{N}_0}$. Hence, not all number theoretic functions are computable (see, Cutland [21]).

[24]Note, the terms computable functions and recursive functions will be used interchangeably.

halt (denoted as $\phi_a(x) \uparrow$). The domain of the function $f(x)$ denoted by Dom ϕ_a or $\mathbf{W}_a$ is such that,

$$\text{Dom } \phi_a = \mathbf{W}_a = \{\text{x} \mid \phi_a(x) \downarrow : TM_a(x) \text{ halts}\}. \tag{2}$$

Note, the range of computable functions $f(x) \cong \phi_a(x)$ will be denoted as $\mathbf{E}_a$.

Definition 3.1. Computable functions that are defined on the full domain of $\mathcal{N}$ are called total computable functions. Partial computable functions are those functions that are defined only on some subset of $\mathcal{N}$. Related to (2) is the notion of sets whose members can be enumerated by an algorithm or a TM.

Definition 3.2. 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 recursive enumerable.

The one feature of computation theory that is crucial to game theory where players have to simulate the decision procedure of other players, is the notion of the Universal Turing Machine (UTM).

Definition 3.3. 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 Parametrization Theorem, there is a total computable function $\tau(a)$ which determines the index or g.n of the UTM such that.

$$\psi(a, x) = \phi_{\tau(a)}(x) \cong \phi_a(x). \tag{3}$$

Equation (3) says that the UTM with index $\tau(a)$, on the left-hand side of (3) using 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. In what follows, the Parametrization Theorem will be used for effective constructions of indexes/g.ns for sets and functions, which in the case in (3), the total recursive numbering function τ operating on g.n of TM_a yields $\tau(a) = e = g.n\psi(a, x)$. The last equality states that the terms in the first two yield the index/g.n for the UTM[25].

Of particular significance are TMs that use their own code/g.n as inputs in their calculation. We will refer to these as self-referential calculations. The following set $\mathbf{C}$ and its complement, $\tilde{\mathbf{C}}$, that respectively, characterize TMs that halt or do not halt on their own codes play an important role in the proof of incompleteness and undecidability of formal systems.

Definition 3.4. The set denoted by $\mathbf{C}$ is the set of g.ns of all TMs that halt when operating on their own g.ns or alternatively $\mathbf{C}$ contains the g.ns of those recursively enumerable sets that contain their own codes (see, Cutland [21, p.123], Rogers [61, p.62]).

$$\mathbf{C} = \{\text{x} \mid \phi_x(x) \downarrow; TM_x(x) \text{ halts}; x \in \mathbf{W}_x\}, \tag{4}$$

$$\tilde{\mathbf{C}} = \{\text{x} \mid \phi_x(x) \uparrow; TM_x(x) \text{ does not halt}; x \notin \mathbf{W}_x\}. \tag{5}$$

Theorem 3.5. *The set $\tilde{\mathbf{C}}$ is not recursively enumerable.*

---

[25]It is customary to denote the indexes/g.ns as single alphabets, typically $\{a, e, x, n, i \ldots\}$ or as stand alone expressions such as, σ_n , while as in case of (3), the total computable encoding function of the Parameterization Theorem that outputs the indexes/g.ns will be depicted as a function of an input and denoted generically as $\tau(.)$.

In the proof that $\tilde{\mathbf{C}}$ is not recursively enumerable, viz there is no computable function that will enumerate it, the Cantor diagonalization method is used[26]. Hence, $\mathbf{C}$ is a recursively enumerable set that is not recursive, the latter would require both $\mathbf{C}$ and $\tilde{\mathbf{C}}$ to be recursively enumerable. This is not possible as $\tilde{\mathbf{C}}$ is not recursively enumerable by **Theorem 3.5**.

3.2. **Post** [58] **set theoretic representation of Gödel incompleteness.** As indicated in the introduction, we will now state the formal character of systems capable of the endogenous production of novelty or surprises in terms of creative and productive sets first defined by Emil Post [58].

Definition 3.6. A creative set $\mathbf{Q}$ is a recursively enumerable set whose compliment, $\tilde{\mathbf{Q}}$, is a productive set. The set $\tilde{\mathbf{Q}}$ is productive if there exists a recursively enumerable set $\mathbf{W}_x$ disjoint from $\mathbf{Q}$ (viz. $\mathbf{W}_x \subset \tilde{\mathbf{Q}}$) and there is a total computable function $f(x)$ which belongs to $\tilde{\mathbf{Q}} - \mathbf{W}_x$. Thus, $f(x) \in \tilde{\mathbf{Q}} - \mathbf{W}_x$ is referred to as the **productive function** and is a 'witness' to the fact that $\tilde{\mathbf{Q}}$ is not recursively enumerable. Any effective enumeration of $\tilde{\mathbf{Q}}$ will fail to list $f(x)$ even though by construction $f(x) \in \tilde{\mathbf{Q}} - \mathbf{W}_x$ (Cutland [21, p.134-136]).

Lemma 3.7. *Set $\mathbf{C}$ in (4) is an example of a creative set. Consider $\mathbf{W}_i$ to be a recursively enumerable set which is disjoint from $\mathbf{C}$. The productive function $f(i) = i$ is the identity function for the productive set $\tilde{\mathbf{C}}$, such that $f(i) \in \tilde{\mathbf{C}} - \mathbf{W}_i$.*

Proof is straightforward. By the definition of $\mathbf{C}$ in (4) if any number $i \in \mathbf{C} \leftrightarrow i \in \mathbf{W}_i$. Hence, for $f(i) = i$, if $f(i) \in \mathbf{C} \leftrightarrow i \in \mathbf{W}_i$ and $\mathbf{W}_i$ will not be disjoint from $\mathbf{C}$. As $\mathbf{W}_i$ is disjoint from $\mathbf{C}$, viz. $\mathbf{W}_i \cap \mathbf{C} = \emptyset$ and hence $f(i) \notin \mathbf{C} \cup \mathbf{W}_i$. Also, $f(i) = i \in \tilde{\mathbf{C}} - \mathbf{W}_i$. The Post productive function $f(i)$ provides the evidence that $\tilde{\mathbf{C}}$ has syntactic objects which cannot be exhaustively enumerated by any TM.

In formal systems, as proof is identical to computation, we can model a simple formal system given by the sets $\mathbf{C}$ and $\tilde{\mathbf{C}}$. The set $\mathbf{C}$ gives all theorems of the form $\phi_x(x) \downarrow$ with theoremhood defined by the halting of TMs. The negations of provable propositions in $\mathbf{C}$, denoted by g.n $x^\neg$ cannot be theorems in the formal system on pain of inconsistency. The refutable propositions (viz. propositions whose negations are provable) can be enumerated one by one as theorems are proved in $\mathbf{C}$ and we can collect them in a set $\mathbf{W}_q = \{x^\neg | \phi_x^\neg(x^\neg) \uparrow \text{ iff } \phi_x(x) \downarrow\}$. As the formal system is consistent, clearly the two recursively enumerable sets $\mathbf{C}$ and $\mathbf{W}_q$ are disjoint and $\mathbf{C} \cap \mathbf{W}_q = \emptyset$ and $\mathbf{W}_q \subset \tilde{\mathbf{C}}$. Also, $\mathbf{W}_q \neq \tilde{\mathbf{C}}$, for if $\mathbf{W}_q = \tilde{\mathbf{C}}$, then the formal system is complete, but inconsistent. The Post productive function $f(q) = q$ takes on the status of the undecidable proposition as it lies outside both the recursively enumerable sets $\mathbf{C}$ of provable propositions and $\mathbf{W}_q$ of refutable propositions. Cutland [21, p.148] calls this the Gödel incompleteness result in miniature.

As underscored by Smullyan [65, p.58] and is widely known (see, Cutland [21, Ch.3, Sec.3 and Ch.8]), the construction of two recursively inseparable disjoint subsets involving creative and productive sets, such as that of $\mathbf{C}$ and $\mathbf{W}_q \subset \tilde{\mathbf{C}}$ with

---

[26]Assume that there is a computable function $f = \phi_y$, whose domain $\mathbf{W}_y = \tilde{\mathbf{C}}$. If there is such a computable function ϕ_y, then the membership of $\tilde{\mathbf{C}}$ could be listed. Now, if $y \in \mathbf{W}_y$, then $y \in \tilde{\mathbf{C}}$ as we have assumed $\tilde{\mathbf{C}} = \mathbf{W}_y$. But, by the definition of $\tilde{\mathbf{C}}$ in (5) if $y \in \mathbf{W}_y$, then $y \in \mathbf{C}$ and not to $\tilde{\mathbf{C}}$. Alternatively, if $y \notin \mathbf{W}_y$ it implies that $y \notin \tilde{\mathbf{C}}$, given the assumption that $\tilde{\mathbf{C}} = \mathbf{W}_y$. Then, again we have a contradiction, as since from (5) when $y \notin \mathbf{W}_y$, $y \in \tilde{\mathbf{C}}$. Thus, we have to reject the assumption that for some computable function $f = \phi_y$, its domain $\mathbf{W}_y = \tilde{\mathbf{C}}$.

$\mathbf{W}_q$ disjoint from any subset of $\mathbf{C}$, plays a fundamental role in modern approaches to incompleteness and undecidability. In what follows in our proof that the Nash equilibrium surprise strategy functions are the Post productive functions, the g.ns of which lie outside extant recursively enumerable strategy sets, we will utilize what are called recursive reductions (Smullyan [65, Ch. V], Cutland [21, Ch.3, Ch.8]) of the archetypical sets $\mathbf{C}$ and $\tilde{\mathbf{C}}$.

The inseparability of two recursively enumerable disjoint number sets, $\mathbf{B}$ and $\mathbf{D}(\mathbf{B} \subset \mathbf{C}$ and $\mathbf{D} \subset \tilde{\mathbf{C}})$, arises from the property that a recursive reduction of these sets to two other sets, respectively, denoted as $\mathbf{B}'$ and $\mathbf{D}'$, will imply that the g.n for the productive function for $\tilde{\mathbf{B}}'$ lies outside of both $\mathbf{B}'$ and $\mathbf{D}'$ and will be a constructive 'witness' for incompleteness. The following **Lemma 3.8** will be used to engineer a recursive reduction between sets such as $\mathbf{B}$ and $\mathbf{D}$ and the respective sets $\mathbf{B}'$ and $\mathbf{D}'$ to prove that the Nash equilibrium surprise strategy is none other than the productive function, as in the case for the set $\tilde{\mathbf{B}}'$.

Lemma 3.8. *Consider two recursively enumerable disjoint number sets, $\mathbf{B}$ and $\mathbf{D}$ with $\mathbf{B} \subset \mathbf{C}$ and $\mathbf{D} \subset \tilde{\mathbf{C}}$, where $\mathbf{C}$ and $\tilde{\mathbf{C}}$ are respectively defined in (4) and (5), and let $\mathbf{D} = \mathbf{W}_{\sigma_n^{\neg}}$ be the recursively enumerable subset of $\tilde{\mathbf{C}}$ of **Lemma 3.7** with index $\sigma_n^{\neg}$.*

(i)As, $g(\sigma_n^{\neg}) =$ identity function is a productive function of $\tilde{\mathbf{C}}$ with $g.n(g(\sigma_n^{\neg})) = \sigma_n^{\neg} \notin \mathbf{C} \cup W_{\sigma_n^{\neg}}$, it also serves as the productive function for $\tilde{\mathbf{B}}$.

Let the total recursive function $h(i)$ define the following many-1 reduction of $\mathbf{B}$ to $\mathbf{B}'$: $i \in \mathbf{B}$ iff $h(i) \in \mathbf{B}'$. Hence, by this assumption,

(ii) $\mathbf{B} = h^{-1}\mathbf{B}'$ and

(iii) $\tilde{\mathbf{B}} = h^{-1}\tilde{\mathbf{B}}'$

Conditions (i)-(iii) implies that

(a) As $\mathbf{B}$ is creative so is $\mathbf{B}'$ and as $\tilde{\mathbf{B}}$ is productive, the recursive reduction function $h(i)$ implies that $\tilde{\mathbf{B}}'$ is productive with a recursively enumerable subset $\mathbf{D}' = \mathbf{W}_{\sigma_n^{!}}$ disjoint from $\mathbf{B}'$.

(b) Also, $h(g(\sigma_n^{\neg}))$[27] is the productive function for $\tilde{\mathbf{B}}'$ and with $\mathbf{D}' = \mathbf{W}_{\sigma_n^{!}}$, the $g.n(h(g(\sigma_n^{\neg})) \notin \mathbf{B}' \cup \mathbf{W}_{\sigma_n^{!}}$.

Proof. [28] Lemma 3.8(a): As $\mathbf{B}$ and $\mathbf{D}$ are disjoint and $\mathbf{D} \subset \tilde{\mathbf{C}}$ with g, the productive function of **Lemma 3.7** being an identity function, $\sigma_n^{\neg}$ cannot belong to either $\mathbf{B}$ or $\mathbf{D}$ as this will imply that $\sigma_n^{\neg} \in \mathbf{W}_{\sigma_n^{\neg}}$. By **Lemma 3.7**, we have already proven that $\sigma_n^{\neg} \notin \mathbf{C}$ and hence it cannot belong to $\mathbf{B}$ either. For, if $\sigma_n^{\neg} \in \mathbf{B}$, then as $\mathbf{B} \subset \mathbf{C}$, $\sigma_n^{\neg} \in \mathbf{W}_{\sigma_n^{\neg}}$ by definition of $\mathbf{C}$ in (4) and $\mathbf{W}_{\sigma_n^{\neg}} \subset \mathbf{C}$, which entails a contradiction. Hence, $\sigma_n^{\neg} \notin \mathbf{B}$ and $\sigma_n^{\neg} \notin \mathbf{B} \cup \mathbf{W}_{\sigma_n^{\neg}}$. This also implies that $\mathbf{B}$ is creative and as $\tilde{\mathbf{B}}$ is not recursively enumerable, $\tilde{\mathbf{B}}$ is productive as stated in (i).

Lemma 3.8(b): The definition of recursive reduction function $h(i)$ will guarantee that as $\tilde{\mathbf{B}}$ is productive, so is $\tilde{\mathbf{B}}'$. Hence, suppose there is a recursively enumerable subset $\mathbf{D}' = \mathbf{W}_{\sigma_n^{!}} \subset \tilde{\mathbf{B}}'$. Use (iii) that $\mathbf{D} = h^{-1}(\mathbf{W}_{\sigma_n^{!}}) = \mathbf{W}_{\sigma_n^{\neg}} \subset \tilde{\mathbf{B}}$. For every index $\sigma_n^{!}$ of a recursively enumerable subset of $\mathbf{W}_{\sigma_n^{!}}$ of $\tilde{\mathbf{B}}'$ and for any recursive function h^{-1}, by the Parameterization Theorem there is a total recursive numbering function $\tau(\sigma_n^{!}) = \sigma_n^{\neg}$. As $g.n(g(\sigma_n^{\neg})) = \sigma_n^{\neg} \notin \mathbf{B} \cup \mathbf{W}_{\sigma_n^{\neg}}$ from (i), the

---

[27]Note, $h(g(\sigma_n^{\neg}))$ is the total recursive productive function, we denote its g.n as $g.n(h(g(\sigma_n^{\neg})))$. See, footnote 24.

[28]This is analogous to the proof in Smullyan [65, p.96] Chapter V, Proposition 2 (a).

recursive reduction function h maps the productive function for $\tilde{\mathbf{B}}$ in such a way that $g.n(h(g(\sigma_n^{\neg}))) \notin \mathbf{B}'$ and also $g.n(h(g(\sigma_n^{\neg}))) \notin \mathbf{W}_{\sigma_n^!}$. Hence, $\tilde{\mathbf{B}}'$ is productive set with productive function $h(g(\sigma_n^{\neg}))$ and the g.n of this productive function is witness for a syntactic object that is a novelty. $\qquad\square$

The non-trivial productive function $h(g(\sigma_n^{\neg}))$ for $\tilde{\mathbf{B}}'$ in **Lemma 3.8**, which is not an identity function as in **Lemma 3.7**, will also be shown to arise from the surprise best response function in the Nash equilibrium of the 2-person oppositional game.

4. 2-person game with Turing Machines and Gödel meta-mathematics.

4.1. The 2-person game for Turing Machines.
In order to model a game conducted by TMs in which both cooperation and opposition can arise, such a game, as discussed in the introduction, can be interpreted as a parasite-host or regulatee-regulator game. The primitives of the game are codified as follows.

$$\mathbf{G} = \{(p, g), (A_p, A_g), s \in \mathbf{S}\} \tag{6}$$

Here, (p, g) denote the respective g.ns of the objective functions, to be specified, of players, p, the parasite/regulatee and g, host/regulator. The action sets denoted by A_i are finite and countable with $a_{il} \in A_i$, $i \in (p, g)$ being the g.n of an action of player i and $l = 0, 1, 2, \ldots, L$. An element $s \in \mathbf{S}$ denotes a finite vector of states of the world and other archival information and $\mathbf{S}$ is a finite and countable set. The action set $\mathbf{A} = A_p \cup A_g$ represents the known technologies. In order to highlight the fundamental recursive nature of actions as technologies and the potential for new technologies, the class of best response strategy functions will be defined as a set of total computable functions.

Definition 4.1. The best response strategy functions f_i, $i \in (p, g)$ that are total computable functions can belong to one of the following classes

$$f_i = \begin{cases} f_i^1 = 1 & \text{-(Unit Function) -Rule Abiding.} \\ f_i^{\neg} & \text{-Rule Breaking/Liar.} \\ f_i^! & \text{-Surprise.} \end{cases} \tag{7}$$

such that the g.ns of f_i are contained in set

$$\mathfrak{R} = \{m \mid f_i = \phi_m, \phi_m \text{is total computable}\}. \tag{8}$$

The set $\mathfrak{R}$, which is the set of all total computable functions, is not recursively enumerable. The proof of this is standard, see, Cutland [21, p.127].

As will be clear, (8) involves issues on how innovative actions/institutions can be constructed from existing action sets. The remarkable nature of the set $\mathfrak{R}$ is that there are an uncountable infinite number of ways for new technologies to be formed. The task is to show the conditions under which it is mutually deducible that the best response function f_i, $i \in (p, g)$ satisfies Post's productive function and is a surprise strategy, $f_i = f_i^! = \phi_m$, such that $m \in \mathfrak{R} - \mathbf{A}$. Only such innovations will be accorded with the status of strategic innovations. The trigger for the surprise strategy involves the identification of the negation or Liar strategy $f_i^{\neg}$. These total computable strategy functions with the exception of the unit function in (7) will be shown to generate dynamics in the system.

4.2. **Gödel meta-mathematics for interactive digital agents.** This section sets out how the Gödel meta-mathematics helps to organize encoded information involving self and other in the digital game. A major implication of imposing computability constraints on all aspects of the game is that all meta-information with regard to the outcomes of the game for any given set of states, $s \in \mathbf{S}$, 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 partial computable functions. This is given below in **Figure 1** (see, Cutland [21], p.208]). The tuple (x,y) identifies the row and column of this matrix Ξ, the rows of which are denoted as Ξ_i, $i = 0, 1, 2, \ldots$.

$$
\begin{array}{ccccccccc}
\Xi_0 & \phi_{\sigma(0,0)} & \phi_{\sigma(0,1)} & \phi_{\sigma(0,2)} & \phi_{\sigma(0,3)} & \cdots & \phi_{\sigma(0,x)} & \cdots \\
\Xi_1 & \phi_{\sigma(1,0)} & \phi_{\sigma(1,1)} & \phi_{\sigma(1,2)} & \phi_{\sigma(1,3)} & \cdots & \phi_{\sigma(1,x)} & \cdots \\
\Xi_2 & \phi_{\sigma(2,0)} & \phi_{\sigma(2,1)} & \phi_{\sigma(2,2)} & \phi_{\sigma(2,3)} & \cdots & \phi_{\sigma(2,x)} & \cdots \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\
\Xi_x & \phi_{\sigma(x,0)} & \phi_{\sigma(x,1)} & \phi_{\sigma(x,2)} & \phi_{\sigma(x,3)} & \cdots & \phi_{\sigma(x,x)} & \cdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \cdots \\
\end{array}
$$

FIGURE 1. Prediction Function, Meta–Information on Outcomes and Dynamics of 2-person Turing Machine Game.

The prediction 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, determines the outcome of the decision problem of the game and $q \in E_{\sigma_x}$. Note, $\sigma(x,y)$ is the index of the program for this function ϕ that produces the output of the strategic decision problem of the 2-person game. The tuple also identifies a point in the matrix Ξ in **Figure 1**. The conditions under which the output of the prediction function for each (x,y) point in the above matrix is defined is given in the following **Theorem 4.2**. As will be discussed, the points (x,y) will refer to a single player's perspective of self and the other.

Theorem 4.2. *The Gödel representational system is a $1-1$ mapping between meta information in matrix Ξ in **Figure 1** and executable computations such that the conditions under which the prediction function $\phi_{\sigma(x,y)}(s)$ which determines the output of the game for each (x,y) point is defined are as follows:*

$$\phi_{\sigma(x,y)}(s) \cong \phi_{\phi_x(y)}(s) = q, \ iff \ \phi_x(y) \downarrow . \tag{9}$$

Here, the total computable function $\sigma(x,y)$ modelled along the lines of the Gödel two place substitution function[29] (see, Rogers [61], 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 executed calculations on the right-hand-side of (9) as we substitute different values for (x,y) for a given state s. The g.ns representing $\sigma(x,y)$ can always be obtained whether or not the partial recursive function $\phi_x(y)$ on the right-hand side of (9) which executes the programs halts.

Proof. See Rogers [61].   □

---

[29]This approach economizes on formalism and enables us to highlight 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 [3] and Canning [16].

By the necessary condition in (9) if the function $\phi_x(y)$ on the right-hand side (RHS) executing the machine calculation is defined, we say the prediction function $\phi_{\sigma(x,y)}$ in the meta system on the left-hand side (LHS) simulating 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 (9) 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 program executions on the right-hand side terminate.

In view of the discovery of the mirror neuron system, as indicated in Section 2 and discussed further in Markose [43], the set up in (9) formalizes the relationship between a mirror/meta system on the LHS of (9) which records all machine executions on the RHS of (9) permitting action prediction, viz q, of self and other. In **Figure 2**, the RHS of (9) is shown to relate to the canonical neuronal system (pink neuron, $\phi_x(y)$) involving action execution by self and as observer of the other (using the machine execution). The LHS of (9), in **Figure 2** is shown to relate to the **offline** simulations being conducted by the yellow mirror neuron. The synchrony implied in (9) can integrate self and the other as actor and observer. **Figure 2** provides the schema of how common hardwiring of recursive meta-analysis in con-specifics as in the LHS of (9) with a $1 - 1$ synchrony with own action execution on the RHS of (9), will result in the same offline simulation for both players when handling the same encoded information $\sigma(x, y)$ via the Gödel two place substitution function.

Definition 4.3. The two place notation of the meta-system $\sigma(x, y)$ can be used to define a second-order encoding of the following kind defined entirely from a single agent's perspective:

(a) When player p has to determine her own best response function, the first place entry x in $\sigma(x, y)$ refers to what the player p does (viz the g.n of best response function f_p) given player g plays a best response strategy that is consistent with player p attributing player g with the belief that player p has used strategy y (second place entry in $\sigma(.)$). Note this is the self-referential second order belief in Bhatt and Carmerer [12] relating to player p's choice of action.

(b) All Nash equilibria and other relevant fixed points of the game that satisfy what has been referred to as consistent alignment of beliefs (CAB), Osborne and Rubinstein [56] have to be elements, $\sigma(x, x)$, along the diagonal array of this matrix in **Figure 1**. These correspond to self-referential TMs $\phi_x(x)$. Indexes $\sigma(x, y)$ which are off diagonal entries in matrix Ξ violate the CAB condition and cannot be Nash equilibria.

Note, $\sigma(x, x)$ diagonal points in the meta system Ξ in **Figure 1** assume perfect mutual mirroring. Thus, GTP meta-analyses on the LHS of (9) are operations on Gödel numbers yielding simulated or virtual experience of the actual phenomena, and in principle bypass the online executions which involve canonical or motor activity. In other words, all permissible inferences are obtained in short hand by simulation from encoded information. Likewise, on account of the 'only if' condition in **Theorem 4.2**, many interesting aspects of the Nash equilibria of computable games can be established only with reference to the meta analyses, and information in the matrix Ξ with no explicit reference to physical executions of programs. An important out-of-equilibrium belief state will be defined here, which represents off diagonal terms in matrix Ξ.

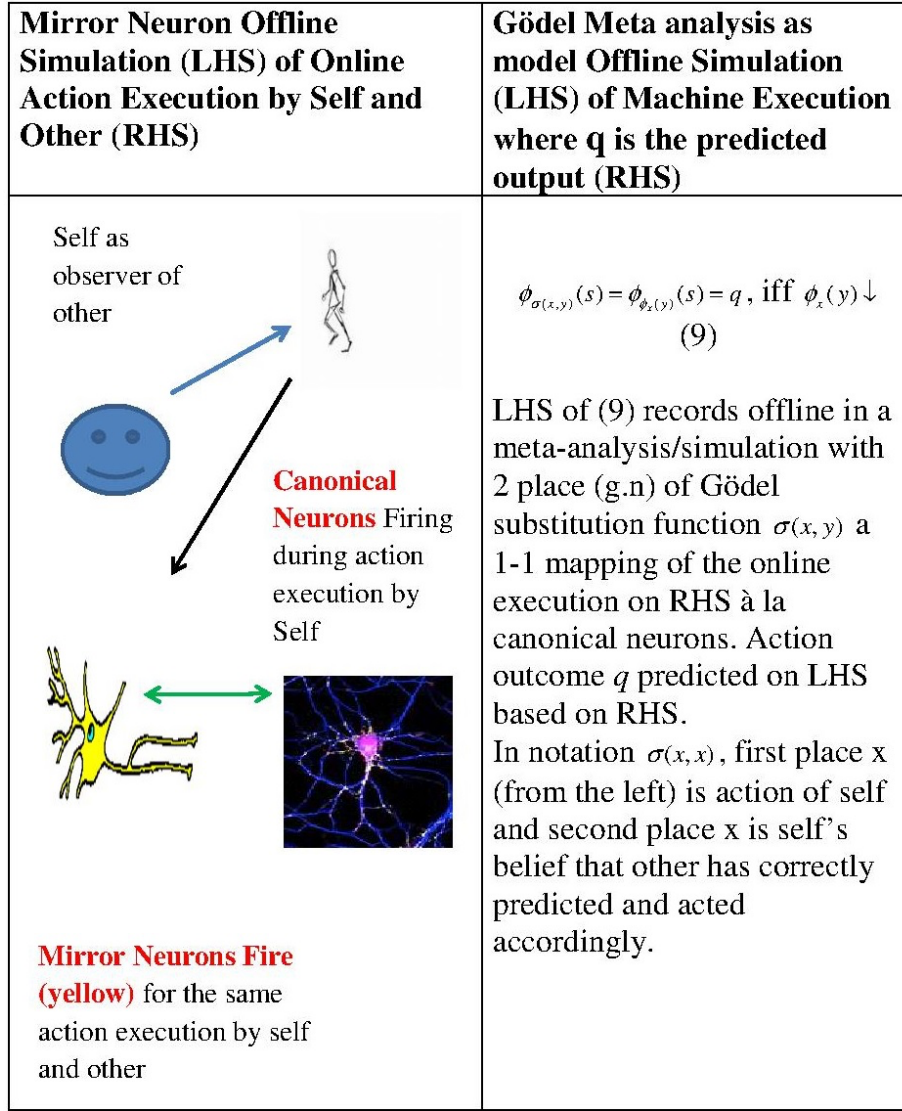| Mirror Neuron Offline Simulation (LHS) of Online Action Execution by Self and Other (RHS)	Gödel Meta analysis as model Offline Simulation (LHS) of Machine Execution where q is the predicted output (RHS)
Self as observer of other **Canonical Neurons** Firing during action execution by Self **Mirror Neurons Fire (yellow)** for the same action execution by self and other	$$\phi_{\sigma(x,y)}(s) = \phi_{\phi_x(y)}(s) = q \,, \text{ iff } \phi_x(y)\downarrow$$ (9) LHS of (9) records offline in a meta-analysis/simulation with 2 place (g.n) of Gödel substitution function $\sigma(x,y)$ a 1-1 mapping of the online execution on RHS à la canonical neurons. Action outcome q predicted on LHS based on RHS. In notation $\sigma(x,x)$, first place x (from the left) is action of self and second place x is self's belief that other has correctly predicted and acted accordingly.

FIGURE 2. Mirror Neurons As Offline Simulations for Mutual Predictions with Self and Other as Gödel 2-Place Substitution Function for Meta-Analysis (Rogers [61, p.202-204])

Definition 4.4. Deceit and False Belief: Denoting by $x^\neg$ the negation of x brought about by best response function, called the Liar strategy, $f_i^\neg$, $i \in (p,g)$ defined in (7), we have $\sigma(x^\neg, x)$ in the two place meta representation of the game by say p. In **Definition 4.3(a)**, this represents the case when player p records his negation of the action with g.n x and attributes to player g the false belief that

player p is playing x. This can also be viewed as p having to be deceitful in order that g maintains his false belief about p[30, 31].

It will be shown how total computable functions for the best response function f_i, $i \in (p,g)$ in a 2-person game when applied to the diagonal array of the matrix Ξ can dynamically move it to a specific row in matrix Ξ in Figure 1. The Nash equilibrium of the game requires the application of the Fixed Point or Second Recursion Theorem.

Theorem 4.5 (Fixed Point or Second Recursion Theorem (SRT) (Cutland [21, p.200])). *Let f be a total unary computable function, then there exists a number v, called the fixed point of $f(v)$, such that*

$$\phi_{f(v)} = \phi_v. \tag{10}$$

Note, $f(v) \neq v$ as they are codes for different programs, but they identify the same function ϕ and both sides of equation (10) will yield an identical output if the programs $f(v)$ and v represent TM calculations that halt, viz. $\phi_{f(v)} \downarrow$ and $\phi_v \downarrow$.

From a perspective of the dynamics implied by (10), the property that any computable function f has a recursively identifiable fixed point follows from the fact that this computable function f represents an encoded set of instructions when applied to the diagonal array of matrix Ξ in **Figure 1** can be found to yield a g.n m for the program for $\phi_m = f^o d$ where $d(x)$ denotes the g.n of the program for the partial recursive functions $\phi_x(x)$ along diagonal array of matrix Ξ in **Figure 1**. Let v be the code for $d(m)$ where $d(m)$ represents $\phi_m(m)$. The proof of (10) follows by noting

$$\phi_v = \phi_{d(m)} = \phi_{\emptyset_m(m)} = \phi_{f^o d(m)} = \phi_{f(v)}. \tag{11}$$

Note in the fixed point for the TM game, we will use the equivalence $d(m) = \sigma(m,m)$ and hence $d(m)$ is the fixed point of $f(d(m))$ in the proof of SRT above. As the application of the total computable strategy function $f_i(f_i \neq 1)$ in (7) to the diagonal array of Ξ in **Figure 1** represents a new code m, in the mth row of Ξ, the $m+1th$ element $\phi_{f_i\sigma(m,m)}$ will coincide with the $m+1th$ element in the diagonal array of Ξ, yielding

$$\phi_{f_i\sigma(m,m)} = \phi_{\sigma(m,m)}, i \in (p,g). \tag{12}$$

Note, these recursive transformations from row to row of Ξ in **Figure 1** can be viewed as digitally generated dynamics. **Theorem 4.5** is used in the determination of the fixed points for the total computable functions best response function f_i, $i \in (p,g)$. When one player applies his best response f_i, Nash equilibria require that other player identifies the same prediction function for producing the output of the game under conditions of consistent alignment of beliefs (that avoids false beliefs given in **Definition 4.4**) as a fixed point of f_i.

---

[30]Thagard [68] gives one of the earliest accounts of the role of second order meta analysis, especially regarding the cognitive processes involved in deception. Thagard is credited with the view that these meta models with recursion required "very little extra representation (space/memory) due to the necessary assumption that the opponent has (roughly) the same cognitive abilities as oneself", MacInnes [41].

[31]Gunnthorsdottir et. al. [30] have found evidence that high 'Mach' types who seem better equipped to conceive of deceit are also likely to deliver Nash equilibrium outcomes in well-known trust games.

5. Nash equilibria: When does one surprise the opposition?

5.1. Total computable best response functions and optimal strategy functions.

The optimization algorithms entailed in achieving best responses in the game arise from the objective functions of players.

Definition 5.1. The objective functions of Turing Machine players are computable functions Π_i, $i \in (p, g)$ defined over the partial recursive payoff/outcome functions specified in (13) and the strategy functions specified in (7):

$$\arg \max_{b_i \in \boldsymbol{B}_i} \Pi_i(\phi_{\sigma(b_i, b_{i|j})}(s)), i \in (p, g). \tag{13}$$

The choice set $\mathbf{B}_i$ contains the g.ns of strategy functions. The Nash equilibrium strategies (β_g^E, β_p^E) with g.ns denoted by (b_p^E, b_g^E) entail up to two subroutines or iterations, to be specified below. In principle, the strategy functions (β_g, β_p) are Universal Turing Machines that simulate optimal strategies of the players that satisfy (13) and involve the total computable best response functions (f_p, f_g) which incorporate elements from the respective action sets $\mathbf{A} = (A_p, A_g)$ and given 2nd order self-referential beliefs of one another's optimal strategy denoted in (13) as $b_{i|j}$ (see, Definitions 4.3a and 4.3b). In the Nash equilibrium best response calculus, the first subroutine denoted by g.n b^1 uses the optimization calculus to determine a player's own optimal action. The problem is that actions can in general be implemented by any total computable best response function, $f_i = \phi_m$, $m \in \mathfrak{R}$, $i \in (p, g)$ in (8).

In standard rational choice models of game theory, the optimization calculus (with g.n z) in the choice of best response restricts choice to *given* actions sets. Hence, starting from some point $\sigma(x, x)$, the strategy functions map from a relevant tuple that encodes meta information of the game into given action sets:

$$\beta_i(f_i\sigma(x, x), z, s, \mathbf{A}) \to A_i \text{ and } f_i = \phi_m, m \in \mathbf{A}, i \in (p, g). \tag{14}$$

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

$$\beta_i(f_i\sigma(x, x), z, s, \mathbf{A}) \to \mathfrak{R} - \mathbf{A} \text{ and } f_i = f_i^! = \phi_m, m \in \mathfrak{R} - \mathbf{A}, i \in (p, g). \tag{15}$$

The question is, which fixed point $\sigma(x, x)$, fully encodable in the meta-mathematics, will trigger such Nash equilibrium surprise strategies, $(\beta_g^{E!}, \beta_p^{E!})$, with g.ns denoted by $(b_p^{E!}, b_g^{E!})$?

It has been noted in passing by Anderlini and Sabourian [4, p.1351], based on the work of Holland [36], that heterogeneity in forms does not arise primarily by random mutation but by algorithmic recombinations that operate on extant 'technologies'. However, a number of preconceptions from traditional game theory such as the 'givenness' of actions sets prevent Anderlini and Sabourian [4] from positing that players who as in (15), equipped with the wherewithal for algorithmic recombinations of existing actions, do indeed innovate from strategic and logical necessity rather than by random mutation. 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 $\mathfrak{R} - \mathbf{A}$ unless players can mutually infer by strictly codifiable means from $\sigma(x, x)$ that (15) 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 which specific structure of opposition logically and strategically necessitates surprise strategies in the Nash equilibrium of the game, in keeping with the Emil Post set theoretic formulation of novelty production in **Lemmas 3.8** and **3.8**, the disjoint recursively enumerable strategy sets that correspond to the recursive reduction to the archetypical creative and productive sets in (4) and (5) have also to be developed.

5.2. **Fixed point/second recursion theorem: The base-point.** The meta-analysis in the determination of Nash equilibrium strategies (β_p^E, β_g^E) with g.ns (b_p^E, b_g^E) will be undertaken here. Calculations start at such a so called base-point which is the fixed point of f_g which has to be arrived at by player p in (16)

$$\phi_{f_g \sigma(b_a, b_a)}(s) = \phi_{\sigma(b_a, b_a)}(s) = q. \tag{16}$$

Here, b_a is the g.n of the strategy f_g that selects the optimal action a from set $\mathbf{A}$ in (14) when g is put in for the index i. In the two place notation in $\sigma(b_a, b_a)$ in (16), the first b_a from the left is the code of the program from (14) as adopted by p to simulate the optimal rule a for the host, g, and the second place b_a denotes that p believes that g believes and acts on the basis that p is rule abiding and has left the host's rule a unchanged. The fixed point for f_g in (16) is computable and the outcome of the host's rule a is predictable and q is the host g's desired outcome. It is convenient to assume that for specific states of the world, s, so called base-point rules are optimal for the host, g , if the parasite 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 in **Figure 1** for the case of rule a. However, the predictable outcome q may involve profitable 'arbitrage' opportunities for p in the form of the Liar strategy. This furnishes the conditions under which a transparent/predictable rule will fail to be a Nash equilibrium strategy.

5.3. **The Liar/rule breaker strategy: The logic of opposition.** For player p, for the given (a, s) it may be optimal 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 host's rule a applies,

$$\phi_{f_p^\neg \sigma(b_a, b_a)}(s) = q^\neg, q^\neg \notin E_{\sigma_{b_a}} \leftrightarrow \phi_{\sigma(b_a, b_a)} = q, q \in E_{\sigma_{b_a}}. \tag{17}$$

For all s when policy rule a does not apply,

$$f_p^\neg = 0 \quad : \text{Do Nothing.}$$

The Liar can successfully subvert with certainty (LHS of 17) if and only if ($\leftrightarrow$) the rule a has predictable outcomes (RHS of (17)) and $f_p^\neg$ itself is total computable. Thus, $f_p^\neg = \phi_m$, $m \in A_p$, must include a codified description of an action rule if undertaken by the Liar can subvert the predictable outcomes of the host's rule a. Formally, if q is predicted on the RHS of (17) then the application of $f_p^\neg$ to $\sigma(b_a, b_a)$ is equivalent to the condition of deliberate deceit in **Definition 4.4** and the g.n of the Liar strategy is $b_a^\neg$ in the first place of $\sigma(b_a^\neg, b_a)$, while the second place entry shows that g harbours a false belief about p, that p is rule abiding with b_a which is optimal for g. This out-of-equilibrium $\sigma(b_a^\neg, b_a)$ point in the game is off diagonal in

terms of the matrix Ξ in **Figure 1** and this will bring about an outcome $q^\neg \notin E_{\sigma_{b_a}}$ which belongs to a set disjoint from the set that contains the desired outcome for player g from rule a for all s for which rule a applies, viz. $E_{\sigma_{b_a}} \cap E_{\sigma_{b_a^\neg}} = \emptyset$. The outcomes $(q^\neg, q)$ can be zero sum, but in general we refer to property $q^\neg \notin E_{\sigma_{b_a}}$ in (17) as being oppositional or subversive.

As discussed, in a Nash equilibrium, p has to remove any attribution of false belief to the other. Both players can identify the fixed point of $f_p^\neg$ using the Second Recursion **Theorem 4.5** to get $\sigma(b_a^\neg, b_a^\neg)$ on the RHS of (18) where $b_a^\neg$ is the code for the Liar strategy in (17).

Theorem 5.2 (Gödel sentence for the oppositional game as a fixed point result)**.** *The prediction function $\phi_{\sigma((b_a^\neg, b_a^\neg)}$ indexed by the fixed point of the Liar/rule breaker best response function $f_p^\neg$ in (18) is not computable and the following corresponds to the famous Gödel sentence for the game.*

$$\phi_{f_g^\neg \sigma(b_a^\neg, b_a^\neg)}(s) = \phi_{\sigma(b_a^\neg, b_a^\neg)}(s). \tag{18}$$

Proof. The proof that the program representing the fixed point of $f_p^\neg$ will fail to compute the outcome of the game in (18) proceeds as follows. Recall from the Second Recursion **Theorem 4.5** and (10), that the indexes $\sigma(b_a^\neg, b_a^\neg)$ and $f_p^\neg \sigma(b_a^\neg, b_a^\neg)$ represents programs to compute the same prediction function ϕ[32]. However, by construction as $f_p^\neg$ is the Liar Strategy in (17), this implies that if on the RHS of (18) $\phi_{\sigma(b_a^\neg, b_a^\neg)}(s)$ halts with the output q, the LHS of (18) produces output $q^\neg$. Hence, if (18) is computable, we have $q = q^\neg$, which is a contradiction of the Second Recursion **Theorem 4.5**. Hence, the computation on both sides of (18) will not halt and the outcome of the game is undecidable. The Gödel sentence for the game in (18) therefore represents a fully encoded statement of this. $\square$

Note the condition of the out-of-equilibrium success of the Liar spelt out in (17) and the outcome of the game when player g, the host, is deceived is computable. In many fast moving competitive co-evolutionary systems, while the full digital molecular record of such adversarial encounters occurs[33], overt phenotypes from the 'original' predictable strategies such as b_a or $b_a^\neg$ may not be observed. Instead only the arms race in novelty given in the next section is what persists such that both players co-exist.

---

[32]This can be verified from the proof of **Second Recursion Theorem (SRT) 4.5**. Starting with $d(x) = \sigma(b_a, b_a)$, denote $\phi_{b_a^\neg} = f_p^\neg \circ d(b_a)$ where $b_a^\neg$ is the g.n of the Liar Strategy defined in (17). Following the proof of **SRT**, set v to be the g.n of $d(b_a^\neg) = \sigma(b_a^\neg, b_a^\neg) = \phi_{b_a^\neg}(b_a^\neg)$. Hence, $\phi_v = \phi_{d(b_a^\neg)} = \phi_{\sigma(b_a^\neg, b_a^\neg)} = \phi_{\phi_{b_a^\neg}(b_a^\neg)} = \phi_{f_p^\neg \circ d(b_a^\neg)} = \phi_{f_p^\neg \sigma(b_a^\neg, b_a^\neg)} = \phi_{f_p^\neg(v)}$. Thus, as required by **SRT**, $\sigma(b_a^\neg, b_a^\neg)$ and $f_p \sigma(b_a^\neg, b_a^\neg)$ are programs for the same prediction function on the RHS and LHS in (18).

[33]Indeed, remarkably, in the recent experimental studies of molecular co-evolution in arms race between parasite and host as in Kashiwagi and Yomo [37], both the original codes in the respective genomes of the host and parasite and the novel codes that evolved in the course of the arms race could be observed. In Markose [43] it is suggested that using environments suitable for neuro-physiological experiments of such a game with human subjects, it will be interesting to identify the brain waves arising at the juncture at which both players mutually know that say, player p can achieve his best payoffs from an out-of-equilibrium configuration wherein the other player g has to kept in a state of false belief $\sigma(b_a^\neg, b_a)$ given in **Definition 4.4**.

6. **Surprise Nash equilibria.** There is no paradox in stating that as both players can encode the non-computability of (18) they will be able to mutually deduce that the only Nash equilibrium strategies for both players that is consistent with meta information in the fixed point in (18), is one that involves strategies that elude prediction. On substituting the fixed point $\sigma(b_a^\neg, b_a^\neg)$ in (18) for $\sigma(x, x)$ in (15), g's Nash equilibrium strategy β_g^E with g.n b_g^E implemented by an appropriate total computable function such as in (15) must satisfy :

$$\beta_g^E(f_g \sigma(b_a^\neg, b_a^\neg), s, \mathbf{A}) \to \mathfrak{R} - \mathbf{A} \text{ and } f_g = f_g^{E!} = \phi_m, m \in \mathfrak{R} - \mathbf{A}. \qquad (19)$$

That is, $f_g^{E!}$ implements an innovation and $b_g^{E!}$ is the g.n of the surprise strategy function in (19). Hence $\sigma(b_g^{E!}, b_g^{E!})$ is the fixed point of $f_g^{E!}$.

Likewise, for player p, $f_p^{E!}$ implements an innovation in (20) and $b_p^{E!}$ is the g.n of the surprise strategy function and hence $\sigma(b_p^{E!}, b_p^{E!})$ is the fixed point of $f_p^{E!}$. Thus,

$$\beta_p^E(f_p \sigma(b_a^\neg, b_a^\neg), s, \mathbf{A}) \to \mathfrak{R} - \mathbf{A} \text{ and } f_p = f_p^{E!} = \phi_m, m \in \mathfrak{R} - \mathbf{A}. \qquad (20)$$

The intuition here is that from the non-computable fixed point with the Liar in (18), the total computable best response function implementing the Nash equilibrium surprise strategies can only map as above into the domains of the action and strategy sets of the players that cannot be algorithmically enumerated in advance.

6.1. **Recursively enumerable disjoint sets, productive functions, surprises and arms race.** Using **Theorem 5.2**, **Definition 3.6** and **Lemma 3.8**, we will now prove the incompleteness results for the strategy sets of the players arising from the Liar/rule breaking strategy. Analysis will be done for p's strategy set $\mathbf{B}_p$ as the strategy functions β_p and β_g, respectively, can be shown to implement a recursive reduction, as in **Lemma 3.8**, of the archetypal creative set $\mathbf{C}$ in (4).

Corresponding to those (a_{gl}, s) tuples, $a_{gl} \in A_g$ of g's base point optimal strategy, viz. those that are optimal for g, if p is assumed to be rule abiding and it is also the case that p's best response f_p is to be rule abiding viz. $f_p = 1$, the g.ns of these optimal strategies for p, $b_p^1 \in \mathbf{B}_p$ result in computable fixed points. Here, in the case when p is rule abiding, b^1 indicates the subroutine 1, which is sufficient for the determination of the Nash equilibrium strategy. This set denoted by $\boldsymbol{\beta}_p^1$ which contains g.ns of all of g's actions for which p is rule abiding can be generated by recursive methodology. Thus,

$$\boldsymbol{\beta}_p^1 = \left\{ b_p^1 | \phi_{\phi_{b_p^1}(b_p^1)}(s) \downarrow \text{ for all } (a_{gl}, s), a_{gl} \in A_g, f_p = 1 \right\}. \qquad (21)$$

Using logic in (17), a set $\boldsymbol{\beta}_p^\neg$ can be recursively generated so that it 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 base point (a_{gl}, s) tuples, $(a_{gl} \in A_g)$ of g's action set. By **Theorem 5.2**, this is a set of p's strategies that result in non-computable fixed points. Hence,

$$\boldsymbol{\beta}_p^\neg = \left\{ b_p^{1\neg} | \phi_{\phi_{b_p^{1\neg}}(b_p^{1\neg})}(s) \uparrow \text{ for all } (a_{gl}, s), a_{gl} \in A_g, f_p = f_p^\neg \right\}. \qquad (22)$$

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^1 cannot belong to both $\boldsymbol{\beta}_p^1$ and $\boldsymbol{\beta}_p^\neg$. Thus, logical consistency of the meta-analysis requires $\boldsymbol{\beta}_p^1 \cap \boldsymbol{\beta}_p^\neg = \emptyset$ and these are the recursively enumerable disjoint sets as required by **Lemma 3.8**. It is convenient, as shown in the following **Lemma 6.1**, to use indexes for the sets $\mathbf{W}_{\sigma_n}$ and $\mathbf{W}_{\sigma_n^\neg}$ that respectively enumerate $\boldsymbol{\beta}_p^1$ and $\boldsymbol{\beta}_p^\neg$. Here, n denotes the nth element of a sequence of elements in the

respective sets. In set $\mathbf{W}_{\sigma_n^{\neg}}$, $\sigma_n^{\neg}$ refers to $\sigma(b_a^{\neg}, b_a^{\neg})$ in (18), arguably the most recent encoding of the Gödel sentence denoting the fixed point of the adversarial attack on the opponent's code . The formalism behind the indexing of sets corresponding to $\mathbf{W}_{\sigma_n^{\neg}}$ is given in **Theorem 6.2**.

Lemma 6.1. *The set* $\mathbf{W}_{\sigma_n}$ *which represents the enumerable set* $\boldsymbol{\beta}_p^1$ *given in (21) corresponds to set* **B** *in **Lemma 3.8** and hence* $\mathbf{W}_{\sigma_n} \subset \mathbf{C}$ *and set* $\mathbf{W}_{\sigma_n^{\neg}}$ *which represents the enumerable* $\boldsymbol{\beta}_p^{\neg}$ *in (22) corresponds to set* **D** *in **Lemma 3.8** and is the recursively enumerable disjoint set from* $\mathbf{W}_{\sigma_n}$ *with* $\mathbf{W}_{\sigma_n^{\neg}} \subset \tilde{\mathbf{C}}$. *Hence,*

a) $\sigma_n^{\neg} \notin \mathbf{W}_{\sigma_n} \subset \mathbf{C}$.

b) $\sigma_n^{\neg} \notin \mathbf{W}_{\sigma_n^{\neg}}$ *and* $\sigma_n^{\neg} \notin \mathbf{W}_{\sigma_n} \cup \mathbf{W}_{\sigma_n^{\neg}}$. *This implies that* $\tilde{\boldsymbol{\beta}}_p^1$ *(the complement set of* $\boldsymbol{\beta}_p^1$*) is productive corresponding to the set* $\mathbf{W}_{\sigma_n^{\neg}}$ *and the productive function of which is* $g(\sigma_n^{\neg}) = $ *identity function.*

Proof of Lemma 6.1(a). By the Parameterization Theorem, as $\boldsymbol{\beta}_p^1$ is recursively enumerable, the index function $\tau(b_a) = \sigma_n$ for some $b_p^1 = b_a$ in (21) is such that :

$$b_a \in \boldsymbol{\beta}_p^1 \leftrightarrow \phi_{\sigma_n}(\sigma_n) \downarrow \leftrightarrow \sigma_n \in \mathbf{C}. \tag{23}$$

Set $\mathbf{C}$ is the archetypical creative set in (4). By construction for some $b_a^{\neg} = b_p^{1\neg}$, $b_a^{\neg} \notin \boldsymbol{\beta}_p^1$ from (21). The Parameterization Theorem yields $\tau(b_a^{\neg}) = \sigma_n^{\neg}$ for the recursively enumerable set $\mathbf{W}_{\sigma_n^{\neg}}$ which is disjoint from $\mathbf{W}_{\sigma_n}$ and $\mathbf{W}_{\sigma_n^{\neg}} \subset \tilde{\mathbf{C}}$. **Lemma 6.1**(b) follows immediately as $\sigma_n^{\neg} \notin \mathbf{C}$. **Lemma 3.8**(i) yields $\sigma_n^{\neg} \notin \mathbf{W}_{\sigma_n} \cup \mathbf{W}_{\sigma_n^{\neg}}$ and the productive function $g(\sigma_n^{\neg}) = $ identity function for the productive set $\tilde{\boldsymbol{\beta}}_p^1$. $\qquad\square$

The feature that the productive function $g(\sigma_n^{\neg})$ is an identity function at the fixed point relating to the Gödel sentence is interesting. This implies that the digital agent passively enumerates, in set $\mathbf{W}_{\sigma_n^{\neg}}$, all codes in $\boldsymbol{\beta}_p^{\neg}$ that suffer adversarial attack corresponding to the self-referential mutual representations of hostility in non-computable fixed point of **Theorem 5.2**. As will be shown in **Figure 4**, sets like $\mathbf{W}_{\sigma_n^{\neg}}$ which are subsets of productive sets have a specific arms race like structure in their recursive enumeration.

Finally, as the set $\boldsymbol{\beta}_p^{\neg}$ contains negations of known strategies b_p^1 that imply the negation of g's objectives, denoted generically as q in (17), for p and g to achieve their, respective, desired objectives of $q^{\neg}$ and q, in a Nash equilibrium, these can only be done by implementing surprise strategies that map outside of action set $\mathbf{A}$. The second iteration or subroutine denoted as b^2 implementing $f_p\sigma(b_a^{\neg}, b_a^{\neg})$ for the Nash equilibrium surprise strategy $f_p^{E!}$ in (21) is a consistent many-one recursive reduction (analogous to h in **Lemma 3.8**) of the recursively enumerable subset $\mathbf{W}_{\sigma_n^{\neg}}$ of the productive set $\tilde{\mathbf{C}}$ in **Lemma 6.1**. The productive set corresponding to $\tilde{\mathbf{C}}$ in the many one recursive reduction will be denoted as $\tilde{\boldsymbol{\beta}_p}^+$.

Theorem 6.2. *(a) By construction of the recursive reduction, as shown in **Figure 3**, using the second iteration of the Nash equilibrium strategy function b^2 on the set $\mathbf{W}_{\sigma_n^{\neg}}, \mathbf{W}_{\sigma_n^{\neg}} \in \tilde{\mathbf{C}}$ in **Lemmas 6.1** and **3.8** the set $\mathbf{W}_{\sigma_n^!}$ which is a recursively enumerable subset of $\tilde{\boldsymbol{\beta}}_p^+$ is generated and $b_p^E! = b_n^! = g.n(b^2(f_p(\sigma_n^{\neg})))$ is g.n of the Nash equilibrium total computable best response surprise strategy function, $f_p^{E!}$, which implements an innovation as in (20). Hence, $f_p^{E!}$ is the productive function of the set $\tilde{\boldsymbol{\beta}}_p^+$ such that $b_n^! \notin \mathbf{W}_{\sigma_n^!}$ and p's Nash equilibrium strategy set $\mathbf{B}_p$ is incomplete and non-trivially productive.*
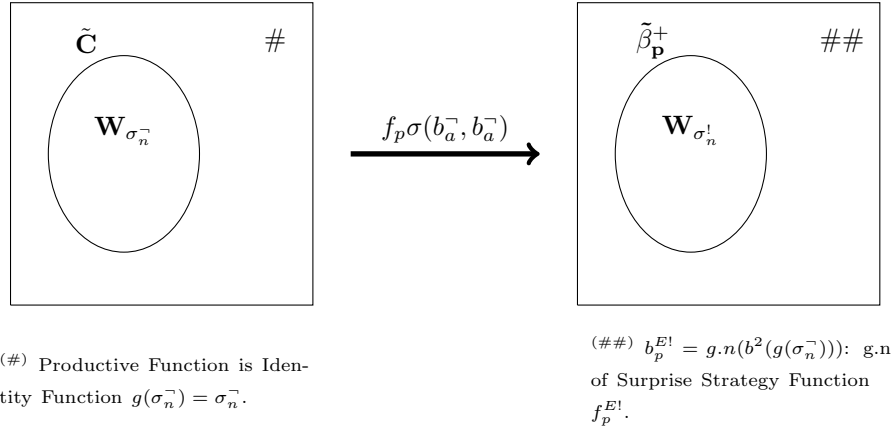
$^{(\#)}$ Productive Function is Identity Function $g(\sigma_n^-) = \sigma_n^-$.

$^{(\#\#)}$ $b_p^{E!} = g.n(b^2(g(\sigma_n^-)))$: g.n of Surprise Strategy Function $f_p^{E!}$.

FIGURE 3. The Incompleteness of p's Nash Equilibrium Strategy Set $\mathbf{B}_p$. Note that the arrow denotes the many-one recursive reduction of **Lemma 3.8** using the second subroutine $f_p\sigma(b_a^-, b_a^-) = b^2$ for the surprise strategy function in (20) from the recursively enumerable subset $\mathbf{W}_{\sigma_n^-}$ of the archetypical productive set $\tilde{\mathbf{C}}$ in **Lemma 6.1** to the Surprise Strategy set $\mathbf{W}_{\sigma_n^!}$, of **Theorem 6.2** yielding the productive surprise strategy function $f_p^{E!}$ with $g.n(b^2(g(\sigma_n^-)))$

(b) Once the surprise Nash equilibrium strategy has been implemented by p with g.n $b_n^!$, the growth of the strategy set can be proven to take the following non-anticipating form and is shown in **Figure 4**[34]:

$$\mathbf{W}_{\sigma_{n+1}^!} = \mathbf{W}_{\sigma_n^!} \cup \left\{ b_n^! = g.n(b^2(f_p(\sigma_n^-))) \right\}. \tag{24}$$

Proof of **Theorem 6.2***(a).* Proof follows from condition (ii) in **Lemma 3.8**. We use the preimage mapping provided by the recursive reduction function b^2 implementing the second subroutine of the surprise strategy function $f_p^{E!}$ in (20) with $\mathbf{W}_{\sigma_n^-} = (b^2)^{-1}(\mathbf{W}_{\sigma_n^!})$, implying that the index $b_n^! = g.n(b^2(g(\sigma_n^-))$ is the g.n of the productive Nash equilibrium surprise function $f_p\sigma(b_a^-, b_a^-)$ as $g(\sigma_n^-) = \sigma_n^-$ is the identity productive function for $\mathbf{W}_{\sigma_n^-}$ from **Lemma 3.8** and **Lemma 6.1**. The Parameterization Theorem provides a total recursive indexing function $\tau(\sigma_n^-, b^2)$ such that $\tau(\sigma_n^-, b^2) = \sigma_n^!$. $\qquad\square$

Proof of **Theorem 6.2***(b).* Proof requires showing that surprise strategy functions have g.ns $b_n^!$ that can only be added on to the extant set $\mathbf{W}_{\sigma_n^!}$ and cannot belong to $\mathbf{W}_{\sigma_n^!}$ itself. As in **Figure 4**, let $\mathbf{W}_{\sigma_n^!} = \left\{ b_0^!, b_1^!, \ldots, b_{n-1}^! \right\}$. Then to get (24), use the Parameterization Theorem on indexes in $\mathbf{W}_{\sigma_n^!} \cup \left\{ b_n^! = g.n(b^2(f_p(\sigma_n^-))) \right\}$ such that $\tau(\sigma_n!, b_n^!) = \sigma_{n+1}^!$. Thus,

$$\mathbf{W}_{\sigma_{n+1}^!} = \mathbf{W}_{\sigma_n^!} \cup \left\{ b_n^! = g.n(b^2(f_p(\sigma_n^-))) \right\} = \mathbf{W}_{\tau(\sigma_n!, b_n^!)}. \tag{25}$$

Hence, for any n, if $b_n^! \in \mathbf{W}_{\sigma_n^!}$ will imply that $\sigma_n^- \in \mathbf{W}_{\sigma_n^-}$, which was shown not to be possible in **Lemma 6.1** as it will lead to a conclusion that the fixed

---

[34]**Figure 4** corresponds closely with Figure 7c of Theorem 3.11 in Cutland [21, p.138].
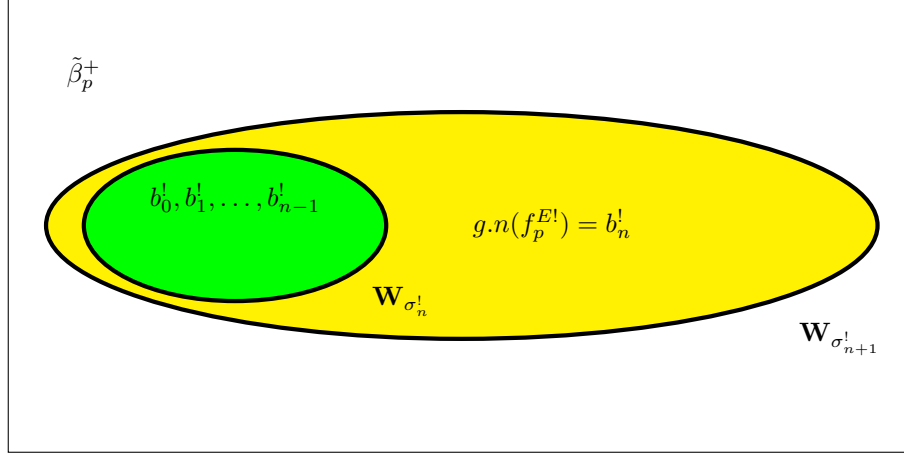
FIGURE 4. Arms Race in Surprises/Innovations: Productive Function Growth of the Surprise Strategy Set (see equation (25))(NB g.n: Gödel number)

point $\sigma(b_a^\neg, b_a^\neg)$ entailed in the Gödel sentence in (18) is the code of a self-referential halting TM. Also then, $\sigma_n^\neg \in \mathbf{W}_{\sigma_n}$ implying for some $b_a^\neg = b_p^{1\neg}$, $b_a^\neg \in \boldsymbol{\beta}_p^1$ which is a contradiction from (21) and (22).   □

The significance of **Theorem 6.2** 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 $\mathbf{B}_p$ and outside of recursively enumerable disjoint sets $\mathbf{W}_{\sigma_n} \cup \mathbf{W}_{\sigma_n^!}$ that, respectively, include all g.ns of actions/phenotypes on which both agents cooperated with and those that produced innovations in the past. 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 $\mathfrak{R}$ of best response functions in (8) should provide an inexhaustible source of surprise or innovative strategies. However, by the same token, by **Theorem 6.2**, 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 produce an output q though both players can mutually identify that $\sigma(b_p^{E!}, b_p^{E!})$ is the fixed point of the surprise Nash equilibrium best response function $f_p^{E!}$. Indeed, $\sigma(b_p^{E!}, b_p^{E!})$ says that this is so self-referentially. In a nutshell 'innovate or be destroyed' describes this Nash equilibrium in which neither party can unilaterally deviate without drastically impairing their prospects. The Gödel numbers $b_0^!, b_1^!, b_3^!, \ldots, b_n^!$ in **Figure 4** with a full digital encoding of it in (25) can be interpreted to be the g.ns of the 'antibodies' or junk DNA produced in the oppositional encounters. Indeed, the inverse image recursive construction of the reduction function in condition (ii) of **Lemma 3.8** as implemented in **Theorem 6.2** will identify the code that was under 'attack' which triggered the novelty production at points like $b_n^!$. Following from part (ii) of **Theorem 6.2** and as seen in **Figure 4**, once players are locked into an oppositional structure, the strategy set of each player will grow utilizing the formalism of an arms race in novelty.

7. **Concluding remarks.** The paper is original in proposing that the issue of novelty production by digital agents should be viewed as the capacity to exit from recursively enumerable sets using the Post [58] productive functions that generate the Gödel sentences or some non-trivial recursive reduction of it (see **Lemmas 3.8** and **6.1**) to yield constructive 'witness' of Gödel incompleteness. A key objective of this paper is to demonstrate that the Gödel sentence far from being a 'funky' and esoteric mathematical construction of little relevance beyond the foundations of mathematics, is an ubiquitous phenomena which can be seen to be the driving force behind the complex protean phenotypes associated with genomic evolution and in the form of novel artifacts or extended phenotypes in organisms and humans. The capacity for syntactic encoding of a mutual state of hostility, negation or deceit, viz. the embedding of the Gödel sentence is what is logically necessary for novelty producing behaviours in digital agents.

The deeply contextual points of exit from recursively enumerable sets, with innovations following in lock step, have been demonstrated to occur in **Theorem 6.2**. It has been claimed that the novel digital objects generated in the bootstrapped growth of the recursively enumerable subset of the Post productive set, as featured in (25), at the genomic level can be considered to be digital records of phenotypical responses to counter hostile agents that is analogous to 'junk DNA' or antibodies.

The remarkable significance of the Gödel incompleteness theorem is that to date there is only one mathematical exit route from known listable sets in order for digital agents to construct novel objects that fall outside of these sets. Without the wherewithal to construct the syntactic object called the Gödel sentence, digital agents will be entrapped within given recursively enumerable repertoires with no autonomy in bootstrapping complex and novel outcomes. The paper gives a full elucidation of of the GTP logic behind the Wolfram-Chomsky Type 4 dynamics that underpin the respective conjectures of McClintock [47] on the 'dynamic genome' and the Ben-Jacob [11] on the 'creative genome'.

Despite the seminal insights of Binmore [13], **Type 4** dynamics with a Nash equilibrium of novelty and surprises lies outside the ambit of extant game theory. The game involving digital agents with GTP powers for encoding, self-reference, recursion and incorporation of the oppositional structure qua the Liar has been shown to lead to strategic innovations. The GTP meta-mathematics has thrown light on the significance of the discoveries of the Parma Group of the mirror neuron system and their hypothesis that the MNS provides embodied offline simulations of online action executions by self and of those based on the observations of others. Further, the discovery by Kelso and co-authors (Tognoli et. al. [69], Naeem et. al. [53]) that negation inherent to anti-coordination in social interaction is part of MNS, gives evidence for the second of the logical conditions needed by the GTP logic for complex protean behaviours. This calls for considerable future work covering GTP logic in digital agents, their powers of recursion and the neuro-physiology of adversarial agents to understand the nexus between social coordination, anti-coordination and innovation.

REFERENCES

[1] P. Albin, *The metalogic of economic predictions, calculations and propositions*, *Mathematical Social Sciences*, **3** (1982), 329–358.

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

[3] L. Anderlini, Some notes on Church's thesis and the theory of games, *Theory and Decision*, **29** (1990), 19–52.

[4] L. Anderlini and H. Sabourian, Cooperation and effective computability, *Econometrica*, **63** (1995), 1337–1369.

[5] M. Arbib and A. Fagg, Modeling parietal-premotor interactions in primate control of grasping, *Neural Networks*, **11** (1998), 1277–1303.

[6] W. B. Arthur, On The evolution of complexity, *Santa Fe Institute Working Paper*, **93-11-070** (1993).

[7] R. Axelrod, *Risk in Networked Information Systems*, Mimeo, Gerald R. Ford School of Public Policy, University of Michigan, 2003.

[8] W. Baumol, *The Free Market Innovation Machine*, Princeton University Press, 2002.

[9] W. Baumol, Red-Queen games: Arms races, rule of law and market economies, *Journal of Evolutionary Economics*, **14** (2004), 237–247.

[10] E. Beinhocker, Evolution as computation: Integrating self-organization with generalized Darwinism, *Journal of Institutional Economics*, **7** (2011), 393–423.

[11] E. Ben-Jacob, Bacterial wisdom, Gödel's theorem and creative genomic webs, *Physica A: Statistical Mechanics and Its Applications*, **248** (1998), 57–76.

[12] M. Bhatt and C. Camerer, Self-referential thinking and equilibrium as states of mind in games: FMRI evidence, *Games and Economic Behaviour*, **52** (2005), 424–459.

[13] K. Binmore, Modelling rational players: Part 1, *Economics and Philosophy*, **3** (1987), 179–214.

[14] E. Brynjolfsson and A. McAfee, *The Second Machine Age: Work, Progress and Prosperity in a Time of Brilliant Technologies*, W.W. Norton, New York, 2014.

[15] R. Byrne and A. Whiten, *Machiavellian Intelligence: Social Expertise and the Evolution of Intellect in Monkeys, Apes, and Humans*, Oxford University Press, Oxford, 1988.

[16] D. Canning, Rationality, computability and Nash equilibrium, *Econometrica*, **60** (1992), 877–888.

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

[18] N. Chomsky, On certain formal properties of grammars, *Information and Control*, **2** (1959), 137–167.

[19] N. Chomsky, Three models for the description of language, *IRE Transactions on Information Theory*, **2** (1956), 113–124.

[20] M. Corballis, *The Recursive Mind: The Origins of Human Language, Thought, and Civilization*, Princeton University Press, 2014.

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

[22] R. Dawkins, *The Extended Phenotype*, Oxford University Press, Oxford, 1989.

[23] P. M. Driver and D. A. Humphries, Protean behaviour: The biology of unpredictability, *Ethology and Sociobiology,* **10** (1989), 393–394.

[24] H. Gaifman, Naming, Diagonalization, from Cantor, to Gödel to Kleene, *Logic Journal of the IGPL*, **14** (2006), 709–728.

[25] V. Gallese, Mirror neurons, embodied simulation, and the neural basis of social identification, *Psychoanalytic Dialogues*, **19** (2009), 519–536.

[26] V. Gallese, L. Fadiga, L. Fogassi and G. Rizzolatti, Action recognition in the premotor cortex, *Brain*, **119** (1996), 593–609.

[27] V. Gallese and C. Sinigaglia, What is so special about embodied simulation, *Trends in Cognitive Sciences*, **15** (2011), 512–519.

[28] K. Gödel, *On Formally Undecidable Propositions of Principia Mathematica and Related Systems,* (English) [Gödel's Theorem in Focus, (Ed. S.G Shanker)], Publishers, New York, 1963.

[29] D. E. Goldberg, The existential pleasures of genetic algorithms, in *Genetic Algorithms in Engineering and Computer Science* (eds. G. Winter, J. Perioux, M. Galan and P. Cuesta), John Wiley and Sons (1995), 23–31.

[30] A. Gunnthorsdottir, K. McCabe and V. Smith, Using the Machiavellianism instrument to predict trustworthiness in a bargaining game, *Journal of Economic Psychology*, **23** (2002), 49–66.

[31] A. Haldane, *Financial Arms Races*, Speech delivered at the Institute for New Economic Thinking, Berlin, 14 April 2012.

[32] M. Hauser, N. Chomsky and W. Fitch, The faculty of language: What is it, who has it, and how did it evolve?, *Science*, **298** (2002), 1569–1579.

[33] F. A. Hayek, *The Sensory Order*, The University of Chicago Press, Chicago, 1999.

[34] F. A. Hayek, *The Constitution of Liberty*, The University of Chicago Press, Chicago, 2011.

[35] F. A. Hayek, The Theory of Complex Phenomena, In *Studies In Philosophy, Politics, and Economics*, The University of Chicago Press, Chicago, 1967.

[36] J. Holland, *Adaptation in Natural and Artificial Systems*, MIT Press, New York, 1975.

[37] A. Kashiwagi and T. Yomo, Ongoing Phenotypic and Genomic Changes in Experimental Coevolution of RNA Bacteriophage Qβ and Escherichia coli, *PLoS Genet*, **7** (2011), e1002188.

[38] R. Koppl and B. Rosser, Everything I Might Say Will Already Have Passed Through Your Mind, *Metroeconomica*, **53** (2002), 339–360.

[39] C. Langton, Life at the edge of chaos, (Ed.) [C. Langton, C. Taylor, D. Farmer, and S, Rasmussen (Eds.) Artificial Life II], *Santa Fe Institute Studies in the Sciences of Complexity*, **10** (1992).

[40] E. Lieberman-Aiden, N. van Berkum, L. Williams, M. Imakaev, T. Ragoczy, A. Telling, I. Amit, B. Lajoie, P. Sabo, M. Dorschner, R. Sandstrom, B. Bernstein, M. Bender, M. Groudine, A. Gnirke, J. Stamatoyannopoulos, L. Mirny, E. Lander and J. Dekker, Comprehensive mapping of long-range interactions reveals folding principles of the human genome, *Science*, **326** (2009), 289–293.

[41] J. MacInnes, *Modelling the Enemy: Recursive Cognitive Models in Dynamic Environments*, Division of Life Sciences University of Toronto (Scarborough), Paper presented at 2006 Cognitive Science Conference, 2006.

[42] B. Mandelbrot, *The Fractal Geometry of Nature*, W.H. Freeman, San Francisco, 1982.

[43] S. M. Markose, *Mirroring, Offline Simulation and Complex Strategic Interactions: Coordination, Anti-Coordination and Innovation*, University of Essex Mimeo, 2015.

[44] S. M. Markose, Computability and evolutionary complexity: Markets as complex adaptive systems (CAS), *Economic Journal*, **115** (2005), F159–F192.

[45] S. M. Markose, Novelty in complex adaptive systems (CAS): A computational theory of actor innovation, *Physica A: Statistical Mechanics and Its Applications*, **344** (2004), 41–49.

[46] S. M. Markose, 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*, (Ed. Shu-Heng Chen), Kluwer Academic Publishers, (2002), 443–484.

[47] B. McClintock, The significance of responses of the genome to challenge, *Science* **226**(4676) (1984), 792–801.

[48] G. F. Miller, Protean primates: The evolution of adaptive unpredictability in competition and courtship, in *Machiavellian Intelligence: II. Extensions and Evaluations*, Cambridge University Press, (1997), 312–340.

[49] P. Mirowski, *Machine Dreams: How Economics Became a Cyborg Science*, Cambridge University Press, New York, 2012.

[50] Y. N. Moschovakis, Kleene's amazing second recursion theorem, *Bulletin of Symbolic Logic*, **16** (2010), 189–239.

[51] G. B. Müller and S. A. Newman, The innovation triad: An EvoDevo agenda, *Journal of Experimental Zoology Part B Molecular and Developmental Evolution*, **304** (2005), 487–503.

[52] J. H. Nachbar and W. R. Zame, Non-computable strategies and discounted repeated games, *Economic Theory*, **8** (1996), 103–122.

[53] M. Naeem, G. Prasad, D. R. Watson and J. A. S. Kelso, Electrophysiological signatures of intentional social coordination in the 10-12 Hz range, *Neuroimage*, **59** (2), 1795–1803.

[54] *New Scientist, Feature, You Are Junk: Why It's Not Your Genes that Make You Human*, 2016. Available from:https://www.newscientist.com/article/mg23130840-400-the-junk-that-makes-you-human/?cmpid=NLC.

[55] S. Ohno, So much 'junk' DNA in our genome, *Evolution of Genetic Systems, Brookhaven Symposia in Biology*, **23** (1972), 366–370.

[56] M. Osborne and A. Rubinstein, *A Course in Game Theory*, MIT Press, 1994.

[57] A. J. Pellionisz, The principle of recursive genome function, *Cerebellum,* **7** (2008), 348–359. PMID:18566877,

[58] E. Post, Recursively enumerable sets of positive integers and their decision problems, *Bulletin of American Mathematical Society*, **50** (1944), 284–316.

[59] V. S. Ramachandran, Mirror Neurons and imitation learning as the driving force behind the great leap forward in human evolution, EDGE, 2000. Available from: https://www.edge.org/conversation/mirror-neurons-and-imitation-learning-as-the-driving-force-behind-the-great-leap-forward-in-human-evolution.

[60] G. Rizzolatti, L. Fadiga, V. Gallese and L. Fogassi, Premotor cortex and the recognition of motor actions, *Cognitive Brain Research*, **3** (1996), 131–141.

[61] H. Rogers, *Theory of Recursive Functions and Effective Computability*, Mc Graw-Hill, 1967.

[62] J. E. Samson, A. H. Magadan, M. Sabri and S. Moineau, Revenge of the phages: Defeating bacterial defences, *Nat Rev Microbiol*, **11** (2013), 675–687.

[63] J. A. Schumpeter, *Capitalism, Socialism and Democracy*, Routledge, London, 1994 [1942].

[64] K. Simmons, *Universality of the Liar*, Cambridge University Press, 1993.

[65] R. Smullyan, *Theory of Formal Systems*, Princeton University Press, 1961.

[66] R. I. Soare, Computability and recursion, *Bulletin of Symbolic Logic,* **2** (1996), 284–321.

[67] S. Spear, Learning rational expectations under computability constraints, *Econometrica*, **57** (1989), 889–910.

[68] P. Thagard, Adversarial problem solving: Modeling and opponent using explanatory coherence, *Cognitive Science*, **16** (1992), 123–149.

[69] E. Tognoli, J. Lagarde, G. C. de Guzman and J. A. S. Kelso, The phi complex as a neuromarker of human social coordination, *Proceedings of the National Academy of Sciences of the United States of America*, **104** (2007), 8190–8195.

[70] I. Tsuda, Logic dynamics for deductive inference – Its stability and neural basis, In *Chaos, Information Processing and Paradoxical Games: The Legacy of John S Nicolis*, (Eds, Nicolis Gregoire and Basios Vasileios), *World Scientific Publishing Co.*, (2014), 355–373.

[71] A. M. Turing, On computable numbers, with an application to the entscheidungsproblem, *Proceedings of the London Mathematical Society*, **2-42** (1936), 230–265.

[72] L. van Valen, A new evolutionary law, *Evolutionary Theory*, **1** (1973), 1–30.

[73] J. von Neumann, *Theory of Self-Reproducing Automata,* (Edited and comlpeted by A.W. Burks), University of Illinois Press, 1966.

[74] H. Youn, L. Bettencourt, D. Strumsky and J. Lobo, Invention as a combinatorial process: Evidence from U.S. patents, *Journal of the Royal Society Interface*, **12** (2015), 20150–20272.

[75] S. Wolfram, Cellular automata as models of complexity, *Nature*, **311** (1984), 419–424.