Biology of Applied Digital Ecosystems

Gerard Briscoe
Intelligent Systems and Networks Group
Department of Electrical and Electronic Engineering
Imperial College London
London, United Kingdom
e-mail: gerard.briscoe@ic.ac.uk

Suzanne Sadedin, Greg Paperin
Clayton School of Information Technology
Monash University
Melbourne, Victoria 3800, Australia
e-mail: suzanne.sadedin@infotech.monash.edu.au,
gregory.paperin@infotech.monash.edu.au

Abstract— A primary motivation for our research in Digital Ecosystems is the desire to exploit the self-organising properties of biological ecosystems. Ecosystems are thought to be robust, scalable architectures that can automatically solve complex, dynamic problems. However, the biological processes that contribute to these properties have not been made explicit in Digital Ecosystems research. Here, we discuss how biological properties contribute to the self-organising features of biological ecosystems, including population dynamics, evolution, a complex dynamic environment, and spatial distributions for generating local interactions. The potential for exploiting these properties in artificial systems is then considered. We suggest that several key features of biological ecosystems have not been fully explored in existing digital ecosystems, and discuss how mimicking these features may assist in developing robust, scalable self-organising architectures. An example architecture, the Digital Ecosystem, is considered in detail. The Digital Ecosystem is then measured experimentally through simulations, with measures originating from theoretical ecology, to confirm its likeness to a biological ecosystem. Including the responsiveness to requests for applications from the user base, as a measure of the ecological succession (development).

Index Terms— evolution, ecosystem, complexity

I. INTRODUCTION

Is mimicking ecosystems the future of information systems? A key challenge in modern computing is to develop systems that address complex, dynamic problems in a scalable and efficient way, because the increasing complexity of software makes designing and maintaining efficient and flexible systems a growing challenge [1], [2], [3]. What with the ever expanding number of services being offered online from Application Programming Interfaces (APIs) being made public, there is an ever growing number of computational units available to be combined in the creation of applications. However, this is currently a task done manually by programmers, and it has been argued that current software development techniques have hit a complexity wall [4], which can only be overcome by automating the search for new algorithms. There are several existing efforts aimed at achieving this automated service composition [5], [6], [7], [8], the most prevalent of which is Service-Oriented Architectures and its associated standards and technologies [9], [10].

Alternatively, nature has been in the research business for 3.8 billion years and in that time has accumulated close to 30 million well-adjusted solutions to a plethora of design challenges that humankind struggles to address with mixed results [11]. Biomimicry is a discipline that seeks solutions by emulating nature’s designs and processes, and there is considerable opportunity to learn elegant solutions for human-made problems [11]. Biological ecosystems are thought to be robust, scalable architectures that can automatically solve complex, dynamic problems, possessing several properties that may be useful in automated systems. These properties include self-organisation, self-management, scalability, the ability to provide complex solutions, and automated composition of these complex solutions [12].

Therefore, an approach to the aforementioned challenge would be to develop Digital Ecosystems, artificial systems that aim to harness the dynamics that underlie the complex and diverse adaptations of living organisms in biological ecosystems. While evolution may be well understood in computer science under the auspices of evolutionary computing [13], ecological models are not. The possible connections between Digital Ecosystems and their biological counterparts are yet to be closely examined, so potential exists to create an Ecosystem-Oriented Architecture with the essential elements of biological ecosystems, where the word ecosystem is more than just a metaphor. We propose that an ecosystem inspired approach, would be more effective at greater scales than traditionally inspired approaches, because it would be built upon the scalable and self-organising properties of biological ecosystems [12].

Our focus is in creating the digital counterpart of biological ecosystems. However, the term digital ecosystem has been used to describe a variety of concepts, which it now makes sense to review. Some of these refer to the existing networking infrastructure of the internet [14], [15], [16], while several companies offer a digital ecosystem service or solution, which involves enabling customers to use existing e-business solutions [17], [18], [19]. The term is also being increasingly linked, yet undefined, to the future developments of Information and Communications Technology (ICT) adoption for e-business and e-commerce, to create so called business ecosystems [20], [21], [22]. However, perhaps the most frequent references to digital ecosystems arise in Artificial Life research, where they are created primarily to investigate aspects of biological and other complex systems [23], [24], [25].

The extent to which these disparate systems resemble bio-
logical ecosystems varies, and frequently the word *ecosystem* is merely used for branding purposes without any inherent ecological properties. We consider Digital Ecosystems to be software systems that exploit the properties of biological ecosystems, and suggest that several key features of biological ecosystems have not been fully explored in existing *digital ecosystems*. So, we will now discuss how mimicking these features can create Digital Ecosystems, which are robust, scalable, and self-organising.

Arguably the most fundamental differences between biological and digital ecosystems lie in the motivation and approach of their respective researchers. Biological ecosystems are ubiquitous natural phenomena whose maintenance is crucial to our survival, developing through the process of *ecological succession* [26]. In contrast, Digital Ecosystems will be defined here as a technology engineered to serve specific human purposes, developing to solve dynamic problems in parallel with high efficiency.

Genetic algorithms are a form of evolutionary computing, and like all forms uses natural selection to evolve solutions [27]; started with a set of possible solutions chosen arbitrarily, then selection, replication, recombination, and mutation are applied iteratively. Selection is based on conforming to a fitness function which is determined by a specific problem of interest, and so over time better solutions to the problem can thus evolve [27]. As Digital Ecosystems will likely solve problems by evolving solutions, they will probably incorporate some form of evolutionary computing. However, we suggest that Digital Ecosystems should also incorporate additional features, providing it with a closer resemblance to biological ecosystems. Including features such as complex dynamic fitness functions, a distributed or network environment, and self-organisation arising from interactions among organisms and their environment, which we will discuss later.

### II. Fitness Landscapes and Agents

An ecosystem comprises both an environment and a set of interacting, reproducing entities (or agents) in that environment; with the environment acting as a set of physical and chemical constraints on reproduction and survival [26]. These constraints can be considered in abstract using the metaphor of the fitness landscape, in which individuals are represented as solutions to the problem of survival and reproduction [28]. All possible solutions are distributed in a space whose dimensions are the possible properties of individuals. An additional dimension, height, indicates the relative fitness (in terms of survival and reproduction) of each solution. The fitness landscape is envisaged as a ruged, multidimensional landscape of hills, mountains, and valleys, because individuals with certain sets of properties are *fitter* than others [28].

In biological ecosystems, fitness landscapes are virtually impossible to identify. This is both because there are large numbers of possible traits that can influence individual fitness, and because the environment changes over time and space [26]. In contrast, within a digital environment, it is normally possible to specify explicitly the constraints that act on individuals in order to evolve solutions that perform better within these constraints. Within genetic algorithms, exact specification of a fitness landscape or function is common practice [27]. However, within a Digital Ecosystem the ideal constraints are those that allow solution populations to evolve to meet user needs with maximum efficiency. User needs will change from place to place and time to time. In this sense the fitness landscape of a Digital Ecosystem is complex and dynamic, and more like that of a biological ecosystem than like that of a traditional genetic algorithm [29], [27]. The designer of a Digital Ecosystem therefore faces a double challenge: firstly, to specify rules that govern the shape of the fitness function/landscape in a way that meaningfully maps landscape dynamics to user requests, and secondly, to evolve within this space, solution populations that are diverse enough to solve disparate problems, complex enough to meet user needs, and efficient enough to be preferable to those generated by other means.

The agents within a Digital Ecosystem will need to be like biological individuals in the sense that they reproduce, vary, interact, move, and die [26]. Each of these properties contributes to the dynamics of the ecosystem. However, the way in which these individual properties are encoded may vary substantially depending on the intended purpose of the system [30].

### III. Networks and Spatial Dynamics

A key factor in the maintenance of diversity in biological ecosystems is spatial interactions, and several modelling systems have been used to represent these spatial interactions. Including metapopulations, diffusion models, cellular automata and agent-based models (termed individual-based models in ecology) [34]. The broad predictions of these diverse models are in good agreement. At local scales, spatial interactions favor relatively abundant species disproportionately. However, at a wider scale, this effect can preserve diversity, because different species will be locally abundant in different places. The result is that even in homogeneous environments, population distributions tend to form discrete; long-lasting patches that can resist an invasion by superior competitors [34]. Population distributions can also be influenced by environmental variations such as barriers, gradients, and patches. The possible behaviour of spatially distributed ecosystems is so diverse that scenario-specific modelling is necessary to understand any real system [35]. Nonetheless, certain robust patterns are observed. These include the relative abundance of species, which consistently follows a roughly log-normal relationship [36], and the relationship between geographic area and the number of species present, which follows a power law [37]. The reasons for these patterns are disputed, because they can be generated by both spatial extensions of simple Lotka-Volterra competition models [38], and more complex ecosystem models [39].

1 A metapopulation is a collection of relatively isolated, spatially distributed, local populations bound together by occasional dispersal between populations. [31], [32], [33]
Landscape connectivity plays an important part in ecosystems. When the density of habitats within an environment falls below a critical threshold, widespread species may fragment into isolated populations. Fragmentation can have several consequences. Within populations, these effects include loss of genetic diversity and detrimental inbreeding [40]. At a broader scale, isolated populations may diverge genetically, leading to speciation.

From an information theory perspective, this phase change in landscape connectivity can mediate global and local search strategies [41]. In a well-connected landscape, selection favors the globally superior, and pursuit of different evolutionary paths is discouraged, potentially leading to premature convergence. When the landscape is fragmented, populations may diverge, solving the same problems in different ways. Recently, it has been suggested that the evolution of complexity in nature involves repeated landscape phase changes, allowing selection to alternate between local and global search [42].

In a digital context, we can have spatial interactions by using a distributed system that consists of a set of interconnected locations, with agents that can migrate between these connected locations. In such systems the spatial dynamics are relatively simple compared with those seen in real ecosystems, which incorporate barriers, gradients, and patchy environments at multiple scales in continuous space [26]. Nevertheless, depending on how the connections between locations are organised, such Digital Ecosystems might have dynamics closely parallel to spatially explicit models, diffusion models, or metapopulations [35]. We will discuss later the use of a dynamic non-geometric spatial network, and the reasons for using this approach.

IV. SELECTION AND SELF-ORGANIZATION

The major hypothetical advantage of Digital Ecosystems over other complex organisational models is their potential for dynamic adaptive self-organisation. However, for the solutions evolving in Digital Ecosystems to be useful, they must not only be efficient in a computational sense, but they must also solve purposeful problems. That is, the fitness of agents must translate in some sense to real-world usefulness as demanded by the users [43].

Constructing a useful Digital Ecosystem therefore requires a balance between freedom of the system to self-organise, and constraint of the system to generate useful solutions. These factors must be balanced because the more the system’s behaviour is dictated by its internal dynamics, the less it may respond to fitness criteria imposed by the users. At one extreme, when system dynamics are mainly internal, agents may evolve that are good at survival and reproduction within the digital environment, but useless in the real world [43]. At the other extreme, where the users’ fitness criteria overwhelmingly dictates function, we suggest that dynamic exploration, of the solution space and complexity, is likely to be limited. The reasoning behind this argument is as follows. Consider a multidimensional solution space which maps to a rugged fitness landscape [28]. In this landscape, competing solution lineages will gradually become extinct through chance processes. So, the solution space explored becomes smaller over time as the population adapts and the diversity of solutions decreases. Ultimately, all solutions may be confined to a small region of the solution space. In a static fitness landscape, this situation is not undesirable because the surviving solution lineages will usually be clustered around an optimum [27]. However, if the fitness landscape is dynamic, the location of optima varies over time, and should lineages become confined to a small area of the solution space, then subsequent selection will locate only optima that are near this area [29]. This is undesirable if new, higher optima arise that are far from pre-existing ones. A related issue is that complex solutions are less likely to be found by chance than simple ones. Complex solutions can be visualised as sharp, isolated peaks on the fitness landscape. Especially for dynamic landscapes, these peaks are most likely to be found when the system explores the solution space widely [29]. Therefore, a self-organising mechanism other than the fitness criteria of users is required to maintain diversity among competing solutions in a Digital Ecosystem.

V. COMPLEX ADAPTIVE SYSTEMS: STABILITY/DIVERSITY

Ecosystems are often described as Complex Adaptive Systems (CAS), because like them, they are systems made from diverse, locally interacting components that are subject to selection. Other CAS include brains, individuals, economies, and the biosphere. All are characterised by hierarchical organisation, continual adaptation and novelty, and non-equilibrium dynamics. These properties lead to behaviour that is non-linear, historically contingent, subject to thresholds, and contains multiple basins of attraction [12].

In the previous subsections, we have advocated Digital Ecosystems that include agent populations evolving by natural selection in distributed environments. Like real ecosystems, digital systems designed in this way fit the definition of CAS. The features of these systems, especially non-linearity and non-equilibrium dynamics, offer both advantages and hazards for adaptive problem-solving. The major hazard is that the dynamics of CAS are intrinsically hard to predict because of the non-linear emergent self-organisation [44]. This observation implies that designing a useful Digital Ecosystem will be partly a matter of trial and error. The occurrence of multiple basins of attraction in CASs suggests that even a system that functions well for a long period may suddenly at some point transition to a less desirable state [45]. For example, in some types of system self-organising mass extinctions might result from interactions among populations, leading to temporary unavailability of diverse solutions [46]. This concern may be addressed by incorporating negative feedback or other mechanisms at the global scale. The challenges in designing an effective Digital Ecosystem are mirrored by the system’s potential strengths. Non-linear behaviour provides the opportunity for scalable organisation and the evolution of complex hierarchical solutions, while rapid state transitions
potentially allow the system to adapt to sudden environmental changes with minimal loss of functionality [12].

A key question for designers of Digital Ecosystems is how the stability and diversity properties of biological ecosystems map to performance measures in digital systems. For a Digital Ecosystem the ultimate performance measure is user satisfaction, a system-specific property. However, assuming the motivation for engineering a Digital Ecosystem is the development of scalable, adaptive solutions to complex dynamic problems, certain generalisations can be made. Sustained diversity [45], is a key requirement for dynamic adaptation. In Digital Ecosystems, diversity must be balanced against adaptive efficiency because maintaining large numbers of poorly-adapted solutions is costly. The exact form of this tradeoff will be guided by the specific requirements of the system in question. Stability [12], is likewise, a trade-off: we want the system to respond to environmental change with rapid adaptation, but not to be so responsive that mass extinctions deplete diversity or sudden state changes prevent control.

VI. THE DIGITAL ECOSYSTEM

We are concerned with the digital counterpart of biological ecosystems. However, the term digital ecosystem has been used to describe a variety of concepts, which it now makes sense to review. Some of these refer to the existing networking infrastructure of the internet [14], [15], [16], while several companies offer a digital ecosystem service or solution, which involves enabling customers to use existing e-business solutions [17], [18], [19]. The term is also being increasingly linked, yet undefined, to the future developments of ICT adoption for e-business and e-commerce, to create so called business ecosystems [20], [21], [22]. However, perhaps the most frequent references to digital ecosystems arise in Artificial Life research, where they are created primarily to investigate aspects of biological and other complex systems [23], [24], [25]. The extent to which these disparate systems resemble biological ecosystems varies, and frequently the word ecosystem is merely used for branding purposes without any inherent ecological properties.

We consider Digital Ecosystems [47], [48], [49], [50] to be software systems that exploit the properties of biological ecosystems, which are robust, scalable, and self-organising [12]. So, Digital Ecosystems provide a two-level optimisation scheme inspired by natural ecosystems, in which a decentralised peer-to-peer network forms an underlying tier of distributed agents. These agents then feed a second optimisation level based on an evolutionary algorithm that operates locally on single habitats (peers), aiming to find solutions that satisfy locally relevant constraints. The local search is sped up through this twofold process, providing better local optima as the distributed optimisation provides prior sampling of the search space by making use of computations already performed in other peers with similar constraints [47], [48]. The agents consist of an executable component and an ontological description [51]. So, the Digital Ecosystem can be considered a Multi-Agent System (MAS) [51] which uses distributed evolutionary computing [52], [53] to combine suitable agents in order to meet user requests for applications.

The motivation for using parallel or distributed evolutionary algorithms is twofold. First, improving the speed of evolutionary processes by conducting concurrent evaluations of individuals in a population. Second, improving the problem-solving process by overcoming difficulties that face traditional evolutionary algorithms, such as maintaining diversity to avoid premature convergence [54], [53]. The fact that evolutionary computing manipulates a population of independent solutions actually makes it well suited for parallel computation architectures [52]. There are several variants of distributed evolutionary computing, leading some to propose a taxonomy for their classification [55], with there being two main forms [52], [53]: multiple-population/coarse-grained migration/island models [56], [52], and single-population/fine-grained diffusion/neighbourhood models [57], [53]. Fine-grained diffusion models [57], [53] assign one individual per processor. A local neighbourhood topology is assumed, and individuals are allowed to mate only within their neighbour, called a deme. The demes overlap by an amount that depends on their shape and size, and in this way create an implicit migration mechanism. Each processor runs an identical evolutionary algorithm which selects parents from the local neighbourhood, produces an offspring, and decides whether to replace the current individual with an offspring. In the coarse-grained island models [56], [52], evolution occurs in multiple parallel sub-populations (islands), each running a local evolutionary algorithm, evolving independently with occasional migrations of highly fit individuals among sub-populations. This model has also been used successfully in the determination of investment strategies in the commercial sector, in a product known as the Galapagos toolkit [58], [59]. However, all the islands in this approach work on exactly the same problem, which makes it less analogous to biological ecosystems in which different locations can be environmentally different [26].

The landscape, in energy-centric biological ecosystems, defines the connectivity between habitats [26]. Connectivity of nodes in the digital world is generally not defined by geography or spatial proximity, but by information or semantic
proximity. For example, connectivity in a peer-to-peer network is based primarily on bandwidth and information content, and not geography. The island-models of distributed evolutionary computing use an information-centric model for the connectivity of nodes (islands) [56]. However, because it is generally defined for one-time use (to evolve a solution to one problem and then stop) it usually has a fixed connectivity between the nodes, and therefore a fixed topology [52]. So, supporting evolution in the Digital Ecosystem, with a multi-objective selection pressure (fitness landscape [28] with many peaks), requires a re-configurable network topology, such that habitat connectivity can be dynamically adapted based on the observed migration paths of the agents between the users within the habitat network. Based on the island-models of distributed evolutionary computing [56], each connection between the habitats is bi-directional and there is a probability associated with moving in either direction across the connection, with the connection probabilities affecting the rate of migration of the agents. However, additionally, the connection probabilities will be updated by the success or failure of agent migration using the concept of Hebbian learning [60]: the habitats which do not successfully exchange agents will become less strongly connected, and the habitats which do successfully exchange agents will achieve stronger connections. This leads to a topology that adapts over time, resulting in a network that supports and resembles the connectivity of the user base. If we consider a business ecosystem, network of Small and Medium sized Enterprises, as an example user base; such business networks are typically small-world networks [61], [62]. They many strongly connected clusters (communities), called sub-networks (quasi-complete graphs), with a few connections between these clusters (communities) [63]. Graphs with this topology have a very high clustering coefficient and small characteristic path lengths [63]. So, the Digital Ecosystem will take on a topology similar to that of the user base.

The novelty of our approach comes from the evolving populations being created in response to similar requests. So whereas in the island-models of distributed evolutionary computing there are multiple evolving populations in response to one request [56], here there are multiple evolving populations in response to similar requests. In our Digital Ecosystems different requests are evaluated on separate islands (populations), and so adaptation is accelerated by the sharing of solutions between evolving populations (islands), because they are working to solve similar requests (problems).

The users will formulate queries to the Digital Ecosystem by creating a request as a semantic description, like those being used and developed in Service-Oriented Architectures (SOAs) [64], specifying an application they desire and submitting it to their local peer (habitat). This description defines a metric for evaluating the fitness of a composition of agents, as a distance function between the semantic description of the request and the agents' ontological descriptions. A population is then instantiated in the user’s habitat in response to the user’s request, seeded from the agents available at their habitat. This allows the evolutionary optimisation to be accelerated in the following three ways: first, the habitat network provides a subset of the agents available globally, which is localised to the specific user it represents; second, making use of agent-sequences previously evolved in response to the user’s earlier requests; and third, taking advantage of relevant agent-sequences evolved elsewhere in response to similar requests by other users. The population then proceeds to evolve the optimal agent-sequence(s) that fulfils the user request, and as the agents are the base unit for evolution, it searches the available agent-sequence combination space. For an evolved agent-sequence that is executed (instantiated) by the user, it then migrates to other peers (habitats) becoming hosted where it is useful, to combine with other agents in other populations to assist in responding to other user requests for applications.

VII. Simulation and Results

We simulated the Digital Ecosystem, based upon our Ecosystem-Oriented Architecture, and recorded key variables to determine whether it displayed behaviour typical of biological ecosystems. We created a simulation, following the Ecosystem-Oriented Architecture from the previous section, using the Business Ecosystem of Small and Medium sized Enterprises from Digital Business Ecosystems [65] as an example user base.

Throughout the simulations we assumed a hundred users, which meant that at any time the number of users joining the network equalled those leaving. The habitats of the users were randomly connected at the start, to simulate the users going online for the first time. The users then produced agents (services) and requests for business applications. Initially, the users each deployed five agents to their habitats, for migration (distribution) to any habitats connected to theirs (i.e. their community within the Business Ecosystem). Users were simulated to deploy a new agent after the submission of three requests for business applications, and were chosen at random to submit their requests. A simulated user request consisted of an abstract semantic description, as a list of sets of numeric tuples to represent the properties of a desired business application. The use of the numeric tuples made it comparable to the semantic descriptions of the services represented by the agents; while the list of sets (two level hierarchy) and a much longer length provided sufficient complexity to support the sophistication of business applications.

The user requests were handled by the habitats instantiating evolving populations, which used evolutionary computing to find the optimal solution(s), agent-sequence(s). It was assumed that the users made their requests for business applications accurately, and always used the response (agent-sequence) provided.

Populations of agents, \([A_1, A_1, A_2, \ldots]\), were evolved to solve user requests, seeded with agents and agent-sequences from the agent-pool of the habitats in which they were instantiated. A dynamic population size was used to ensure exploration of the available combinatorial search space, which increased with the average length of the population’s
agent-sequences. The optimal combination of agents (agent-sequence) was evolved to the user request, $R$, by an artificial selection pressure created by a fitness function generated from the user request, $R$. An individual (agent-sequence) of the population consisted of a set of attributes, $a_1, a_2, \ldots$, and a user request essentially consisted of a set of required attributes, $r_1, r_2, \ldots$. So, the fitness function for evaluating an individual agent-sequence, $A$, relative to a user request, $R$, was,

$$\text{fitness}(A, R) = \frac{1}{1 + \sum_{r \in R} |r - a|}, \quad (1)$$

where $a$ is the member of $A$ such that the difference to the required attribute $r$ was minimised. Equation 1 was used to assign fitness values between 0.0 and 1.0 to each individual of the current generation of the population, directly affecting their ability to replicate into the next generation. The evolutionary computing process was encoded with a low mutation rate, a fixed selection pressure and a non-trapping fitness function (i.e. did not get trapped at local optima). The type of selection used fitness-proportional and non-elitist. Fitness-proportional meaning that the fitter the individual the higher its probability of surviving to the next generation [66]. Non-elitist meaning that the best individual from one generation was not guaranteed to survive to the next generation; it had a high probability of surviving into the next generation, but it was not guaranteed as it might have been mutated, [13]. Crossover (recombination) was then applied to a randomly chosen 10% of the surviving population, a one-point crossover, by aligning two parent individuals and picking a random point along their length, and at that point exchanging their tails to create two offspring [13]. Mutations were then applied to a randomly chosen 10% of the surviving population; point mutations were randomly located, consisting of insertions (an agent was inserted into an agent-sequence), replacements (an agent was replaced in an agent-sequence), and deletions (an agent was deleted from an agent-sequence) [67]. The issue of bloat was controlled by augmenting the fitness function with a parsimony pressure [68] which biased the search to shorter agent-sequences, evaluating longer than average length agent-sequences with a reduced fitness, and thereby providing a dynamic control limit which adapted to the average length of the ever-changing evolving agent populations.

A. Ecological Succession

We then compared some of the Digital Ecosystem’s dynamics with those of biological ecosystems, to determine if it had been imbibed with the properties of biological ecosystems. A biological ecosystem develops from a simpler to a more mature state, by a process of succession, where the genetic variation of the populations changes with time [26]. So, it becomes increasingly more complex through this process of succession, driven by the evolution of the populations within the ecosystem [69]. Equivalently, the Digital Ecosystem’s increasing complexity comes from the agent populations being evolved to meet the dynamic selection pressures created by the user requests.

The formation of a mature ecosystem, ecological succession, is the slow, predictable, and orderly changes in the composition and structure of an ecological community, for which there are defined stages in the increasing complexity [26], as shown. So, it becomes increasingly more complex through this process of succession, driven by the evolution of the populations within the ecosystem [69].

Fig. 2. Ecological Succession: The formation of a mature ecosystem is the slow, predictable, and orderly changes in the composition and structure of an ecological community, for which there are defined stages in the increasing complexity [26], as shown. So, it becomes increasingly more complex through this process of succession, driven by the evolution of the populations within the ecosystem [69].

The trajectory of ecological change can be influenced by site conditions, by the interactions of the species present, and by more stochastic factors such as availability of colonists or seeds, or weather conditions at the time of disturbance. Some of these factors contribute to predictability of successional dynamics; others add more probabilistic elements [70]. Trends in ecosystem and community properties of succession have been suggested, but few appear to be general. For example, species diversity almost necessarily increases during early succession upon the arrival of new species, but may decline in later successions as competition eliminates opportunistic species and leads to dominance by locally superior competitors [69]. Net Primary Productivity\(^2\), biomass, and trophic level properties all show variable patterns over succession, depending on the particular system and site [70]. Generally, communities in early succession will be dominated by fast-growing, well-dispersed species, but as the succession proceeds these species will tend to be replaced by more competitive species [26].

We then considered existing theories of complexity for ecological succession and how it would apply to Digital Ecosystems, seeking a high-level understanding that would apply equally to both biological and digital ecosystems. As

\(^2\)Net Primary Productivity (NPP) is defined as the net flux of carbon from the atmosphere into green plants per unit time [67].
succession leads communities, of an ecosystem, to states of dynamic equilibrium within the environment [26], the complexity has to increase initially or there would not be an ecosystem, and presumably this increase eventually stops, because there must be a limit to how many species can be supported. The period in between is more complicated. If we consider the neutral biodiversity theory [38], which basically states network aspects of ecosystems are negligible, we would probably get a relatively smooth progression, because although you would get occasional extinctions, they would be randomly isolated events whose frequency would eventually balance arrivals, not self-organised crashes like in systems theory. In systems theory [71], when a new species arrives in an ecological network, it can create a positive feedback loop that destabilises part of the network and drives some species to extinction. Ecosystems are constantly being perturbed, so it is reasonable to assume that a species that persists will probably be involved in a stabilising interaction with other species. So, the whole ecological network evolves to resist invasion. That would lead to a spiky succession process, perhaps getting less spiky over time.

So, which theory is more applicable to the Digital Ecosystem depends on the extent that a species in the ecosystem acts independently, competing entities (smooth succession) [38] versus tightly co-adapted ecological partners (spiky succession) [71]. Our Digital Ecosystem despite its relative complexity, is quite simplistic compared to biological ecosystems. It has the essential and fundamental processes, but no sophisticated social mechanisms. Therefore, the smooth succession of the neutral biodiversity theory [38] is more probable.

As the Digital Ecosystem’s increasing complexity comes from the Agent populations being evolved to meet the user requests, the effectiveness of these responses (agent-sequences) to user requests is the best available estimate of the Digital Ecosystem’s complexity. So we measured the responses in the simulation of the Digital Ecosystem over a thousand user requests, i.e. until it had reached a mature state like a biological ecosystem [26], and graphed a typical run in Figure 3. The range and diversity of agents at initial deployment were such that 70% fulfilment of user requests was possible, increasing to 100% fulfilment as more agents were deployed.

The Digital Ecosystem performed as expected, adapting and improving over time, reaching a mature state as seen in the graph of Figure 3. The succession of the Digital Ecosystem followed the smooth succession of the neutral biodiversity theory [38], shown by the right distribution, and equal density, of the points around the best fit curve of the graph in Figure 3. At the end of the simulation run, the agent-sequences had evolved and migrated over only ten generations, on average, and collectively had already reached near 70% effectiveness for the user requests. The formation of a mature biological ecosystem, ecological succession, is a relatively slow process [26], and the simulated Digital Ecosystem acted similarly in reaching a mature state.

B. Species Abundance

We then considered relative abundance, which is the proportion of all organisms, in a community, belonging to a particular species [36]. Relative abundance distributions provide a measure of inequalities in population size within an ecosystem, and in most biological ecosystems this distribution takes a log-normal form [36].

A snapshot of the agents (organisms) within the Digital Ecosystem, for a typical simulation run, was taken after a thousand user requests, i.e. once it had reached a mature state. In biology a species is a series of populations within which significant gene flow can and does occur, so groups of organisms showing a very similar genetic makeup [67]. We therefore chose to define species within Digital Ecosystems similarly, as a grouping of genetically similar digital organisms (based on their semantic descriptions), with no more than 10% variation within the species group. Relative abundance was
calculated for each species and grouped by frequency in Figure 4. In contrast to expectations from biological ecosystems, relative abundance in the Digital Ecosystem did not conform to the expected log-normal [36]. We speculate that the high frequency for the lowest relative abundance was caused by the dynamically re-configurable topology of the habitat network, which allowed species of small abundance to survive as their respective habitats were clustered by the Digital Ecosystem. It is also quite possible that this effect also skewed other frequencies for the relative abundance measure.

C. Species-Area Relationship
The species-area relationship measures diversity relative to spatial scale [37]. In the Digital Ecosystem, this relationship represents how similar solutions are to one another at different habitat scales. The species-area relationship is commonly found to follow a power law in biological ecosystems [37].

![Graph of Species-Area in the Digital Ecosystem](image)

Fig. 5. Graph of Species-Area in the Digital Ecosystem: The species-area relationship measures diversity relative to spatial scale [37]. In the Digital Ecosystem, this relationship represents how similar solutions are to one another at different habitat scales. The species-area relationship is commonly found to follow a power law in biological ecosystems, which the Digital Ecosystem also demonstrates.

Again, a snapshot of the agents (organisms) within the Digital Ecosystem, for a typical simulation run, was taken once it had reached a mature state, after a thousand user requests. For this experiment, we assumed each habitat to have an area of one unit. Then, the number of species, at $n$ randomly chosen habitats, was measured, where $n$ ranged between one and a hundred. For each $n$, ten sets of measurements were taken at different random sets of habitats to calculate averaged results, and the log$_{10}$ values of these results are depicted in the graph of Figure 5. The distribution of species diversity over a spatial scale in the Digital Ecosystem demonstrates behaviour similar to biological ecosystems, also following a power law [37]. However, diversity at fine spatial scales appears to be lower than predicted by the line of best fit. This may be explained by higher specialisation at some habitats, making them more like micro-habitats in terms of a reduced species diversity [67].

VIII. CONCLUSION

In simulation, we compared the Digital Ecosystem’s dynamics to those of biological ecosystems. The ecological succession, measured by the responsiveness to user requests, conformed to expectations from biological ecosystems [38]: improving over time, before approaching a plateau. As the evolutionary self-organisation of an ecosystem is a slow process, even the accelerated form present in Digital Ecosystems, it reached only 70% responsiveness, showing potential for improvement. In the species abundance experiment the Digital Ecosystem did not conform to the log-normal distribution usually found in biological ecosystems [36]. The high frequency for the lowest relative abundance was probably caused by the dynamically re-configurable topology of the habitat network, which allowed species of small abundance to survive as their habitats were clustered by the Digital Ecosystem. In the species-area experiment, which measures diversity relative to spatial scale, the Digital Ecosystem did follow the power law commonly found in biological ecosystems [37]. The species diversity at fine spatial scales was lower than predicted by the line of best fit, and may be explained by the high specialisation at some habitats, making them more like micro-habitats, including a reduced species diversity [67]. The majority of the experimental results indicate that Digital Ecosystems behave like their biological counterparts, and suggest that incorporating ideas from theoretical ecology can contribute to useful self-organising properties in Digital Ecosystems, which can assist in generating scalable solutions to complex dynamic problems.

By comparing and contrasting the relevant theoretical ecology, with the anticipated requirements of Digital Ecosystems, we examined how ecological features may emerge in some systems designed for adaptive problem solving. Specifically, we suggested that Digital Ecosystems, like a biological ecosystem, will usually consist of self-replicating agents that interact both with one another and with an external environment [26]. Population dynamics and evolution, spatial and network interactions, and complex dynamic fitness landscapes, will all influence the behaviour of these systems. Many of these properties can be understood via well-known ecological models [72], [38], with a further body of theory that treats ecosystems as Complex Adaptive Systems [12]. These models provide a theoretical basis for the occurrence of self-organisation, in digital and biological ecosystems, resulting from the interactions among the agents and their environment, leading to complex non-linear behaviour [72], [38], [12]; and it is this property that provides the underlying potential for scalable problem-solving in digital environments.

Creating the digital counterpart of biological ecosystems was not without apparent compromises; the information-centric dynamically re-configurable network topology, and the species abundance result inconsistent with biological ecosystems. The Digital Ecosystem requires a re-configurable network topology, to support the constantly changing multi-objective information-centric selection pressures of the user base. Hence, using the concept of Hebbian learning [60], habitat connectivity is dynamically adapted based on the observed migration paths of the agents within the habitat network. The dynamically re-configurable network topology
probably caused the Digital Ecosystem not to conform, in the species abundance experiment, to the log-normal distribution expected from biological ecosystems [36]. We would argue that these differences are not compromises, but features unique to Digital Ecosystems. As we discussed earlier, biomimicry, when done well, is not slavish imitation; it is inspiration using the principles which nature has demonstrated to be successful design strategies [11]. Hypothetically, if there were an abstract definition of an ecosystem, defined as an abstract ecosystem class, then the Digital Ecosystem and biological ecological classes would both inherit from the abstract ecosystem class, but implement its attributes differently. So, we would argue that the apparent compromises in mimicking biological ecosystems are actually features unique to Digital Ecosystems.

Service-oriented architectures promise to provide potentially huge numbers of services that programmers can combine via standardised interfaces, to create increasingly sophisticated and distributed applications [73]. The Digital Ecosystem extends this concept with the automatic combining of available and applicable services in a scalable architecture to meet user requests for applications. This is made possible by a fundamental paradigm shift, from a pull-oriented approach to a push-oriented approach. So, instead of the pull-oriented approach of generating applications only upon request in Service-Oriented Architectures [74], the Digital Ecosystem follows a push-oriented approach of distributing and composing applications pre-emptively, as well as upon request. Although the use of Service-Oriented Architectures in the definition of Digital Ecosystems provides a predisposition to business [75], it does not preclude other more general uses. The Ecosystem-Oriented Architecture definition of Digital Ecosystems is intended to be inclusive and interoperable with other technologies, in the same way that the definition of Service-Oriented Architectures is with grid computing and other technologies [74]. For example, habitats could be executed using a distributed processing arrangement, such as cloud computing [76], which would be possible because the habitat network topology is information-centric (instead of location-centric).

We have confirmed the fundamentals for a new class of system, Digital Ecosystems, created through combining understanding from theoretical ecology, evolutionary theory, Multi-Agent Systems, distributed evolutionary computing, and Service-Oriented Architectures. Digital Ecosystems, where the word ecosystem is more than just a metaphor, being the digital counterpart of biological ecosystems, and therefore having their desirable properties, such as scalability and self-organisation. It is a complex system that shows emergent behaviour, being more than the sum of its constituent parts.

Acknowledgments

We thank P. Dini and P. De Wilde for constructive comments. This work was supported by the European Commission under the EU project Digital Business Ecosystems (contract number 507953 [77]).
S. Wright, “The roles of mutation, inbreeding, crossbreeding and selection in evolution,” in International Congress on Genetics, D. Jones, Ed. Brooklyn botanic garden, 1932, pp. 356–366.

R. Morrison, Designing Evolutionary Algorithms For Dynamic Environments. Springer, 2004.

L. Chambers, The practical handbook of genetic algorithms: applications. CRC Press, 2001.

R. Levins, “Some demographic and genetic consequences of environmental heterogeneity for biological control,” Bulletin of the Entomological Society of America, vol. 15, pp. 237–240, 1969.

I. Hanski, Metapopulation Ecology. Oxford University Press, 1999.

I. Hanski and O. Ovaskainen, “Metapopulation theory for fragmented landscapes,” Theoretical Population Biology, vol. 64, pp. 119–127, 2003.

D. Green, N. Klomp, G. Rimmington, and S. Sadedin, Complexity in Landscape Ecology. Springer, 2006.

D. Green and S. Sadedin, “Interactions matter: complexity in landscapes and ecosystems,” Ecological Complexity, vol. 2, pp. 117–130, 2005, cr GreenSadedin.

S. Bell, “The distribution of abundance in neutral communities,” American Naturalist, vol. 396, pp. 606–617, 2000.

A. Szollosi and D. Storch, “Power-law species-area relationships and self-similar species distributions within finite areas,” Ecology Letters, vol. 7, pp. 60–68, 2004.

S. Hubbell, The Unified Neutral Theory of Biodiversity and Biogeography. Princeton University Press, 2001.

R. Solé, D. Alonso, and A. McKane, “Self-organized instability in complex ecosystems,” Philosophical Transactions: Biological Sciences, vol. 357, pp. 667–681, 2002.

D. Green and M. Kirley, “Adaptation, diversity and spatial patterns,” International Journal of Knowledge-Based Intelligent Engineering Systems, vol. 4, pp. 184–190, 2000.

D. Green, D. Newth, and M. Kirley, “Connectivity and catastrophe - towards a general theory of evolution,” in International Conference on Artificial Life, M. Bedau, Ed. MIT Press, 2000, pp. 153–161.

D. Green, T. Leishman, and S. Sadedin, “Dual phase evolution: a mechanism for self-organization in complex systems,” in International Conference on Complex Systems, A. Minai, D. Braha, and Y. Bar-Yam, Eds. Springer, 2006.

E. Ducheyne, B. De Baets, and R. De Wulf, “Is fitness inheritance useful for real-world applications,” in Evolutionary Multi-criterion Optimization, C. Fonseca, Ed. Springer, 2003, pp. 31–42.

S. Levin, Fragile dominance: complexity and the commons. Perseus Books Group, 1999.

C. Folke, S. Carpenter, B. Walker, M. Scheffer, T. Elmqvist, L. Gunderson, and C. Holling, “Regime shifts, resilience, and biodiversity in ecosystem management,” Annual Review of Ecology, Evolution, and Systematics, vol. 35, pp. 557–581, 2004.

M. Newman, “A model of mass extinction,” Journal of Theoretical Biology, vol. 189, pp. 235–252, 1997.

G. Briscoe, M. Chi, and M. Vidal, “Creating a Digital Ecosystem: Service-oriented architectures with distributed evolutionary computing (BOE-0759),” in JavaOne Conference. Sun Microsystems, 2006. [Online]. Available: http://arxiv.org/0712.4159

G. Briscoe, “Digital Ecosystems: Evolving service-oriented architectures,” in Conference on Bio Inspired mOdelS of NETwork, Information and Computing Systems. IEEE Press, 2006.

C. Masuch. (2008) Evolutionary environment network. [Online]. Available: http://eventnet.sourceforge.net/

T. Kurz. (2008) EVe simulator. [Online]. Available: http://evesim.sourceforge.net/

M. Wooldridge, Introduction to MultiAgent Systems. Wiley, 2002.

E. Cantu-Paz, “A survey of parallel genetic algorithms,” Réseaux et systèmes répartis, Calculateurs Parallèles, vol. 10, pp. 141–171, 1998.

J. Stender, Parallel Genetic Algorithms: Theory and Applications. IOS Press, 1993.

H. Mühlenbein, “Evolution in time and space - the parallel genetic algorithm,” Foundations of Genetic Algorithms, vol. 1, pp. 316–337, 1991.

M. Nowostawski and R. Poli, “Parallel genetic algorithm taxonomy,” in International Conference on Knowledge-Based Intelligent Information Engineering Systems, L. Jain, Ed. IEEE Press, 1999, pp. 88–92.

S. Lin, W. Punch III, and E. Goodman, “Coarse-grain parallel genetic algorithms: categorization and new approach,” in Symposium on Parallel and Distributed Processing. IEEE Press, 1994, pp. 28–37.