# A Computational Model of Symbiotic Composition in Evolutionary Transitions

Richard A. Watson Jordan B. Pollack

Dynamical and Evolutionary Machine Organization Group
Volen Center for Complex Systems – Brandeis University – Waltham, MA – USA
richardw@cs.brandeis.edu

**Abstract.** Several of the major transitions in evolutionary history, such as the symbiogenic origin of eukaryotes from prokaryotes, share the feature that existing entities became the components of composite entities at a higher level of organisation. This composition of pre-adapted extant entities into a new whole is a fundamentally different source of variation from the gradual accumulation of small random variations, and it has some interesting consequences for issues of evolvability. Intuitively, the pre-adaptation of sets of features in reproductively independent specialists suggests a form of 'divide and conquer' decomposition of the adaptive domain. Moreover, the compositions resulting from one level may become the components for compositions at the next level, thus scaling-up the variation mechanism. In this paper, we explore and develop these concepts using a simple abstract model of symbiotic composition to examine its impact on evolvability. To exemplify the adaptive capacity of the composition model, we employ a scale-invariant fitness landscape exhibiting significant ruggedness at all scales. Whilst innovation by mutation and by conventional evolutionary algorithms becomes increasingly more difficult as evolution continues in this landscape, innovation by composition is not impeded as it discovers and assembles component entities through successive hierarchical levels.

**Keywords:** symbiogenesis, major evolutionary transitions, evolutionary computation, evolutionary algorithms, Symbiogenic Evolutionary Adaptation Model, Hierarchical-if-and-only-if, (HIFF).

## 1 Introduction

### 1.1 The major evolutionary transitions and symbiotic composition

The major evolutionary transitions (Buss 1987, Maynard Smith & Szathmary 1995, Michod 1999) involve the creation of new higher-level complexes of simpler entities. Summarised by Michod for example, they include the transitions "from individual genes to networks of genes, from gene networks to bacteria-like cells, from bacteria-like cells to eukaryotic cells with organelles, from cells to multicellular organisms, and from solitary organisms to societies". There are many good reasons to be interested in the evolutionary transitions: they challenge the

Preprint: To appear in BioSystems Journal, 2002. ©

Modern Synthesis preoccupation with the individual as the unit of selection, they involve the adoption of new modes of transmitting information, and they address fundamental questions about individuality, cooperation, fitness, and not least, the origins of life (Buss 1987, Maynard Smith & Szathmary 1995, Michod 1999).

In several of the major evolutionary transitions "entities that were capable of independent replication before the transition can replicate only as part of a larger whole after it" (Maynard Smith & Szathmary 1995). Although Maynard Smith and Szathmary identify several transitions which do not fit what they describe as "symbiosis followed by compartmentation and synchronised replication", several of the transitions do involve the quite literal joining of previously free-living entities into a new whole. We shall refer to this mechanism as 'symbiotic composition', or simply 'composition'. Well known examples include the origin of eukaryotes from prokaryotes via symbiogenesis (the genesis of new species through the genetic integration of symbionts), (Margulis 1993a & 1993b), and the origin of chromosomes from independent genes (Maynard Smith & Szathmary 1993).

Composition presents some obvious contrasts with how we normally understand the mechanisms of neo-Darwinist evolution. The ordinary (non-transitional) view of evolutionary change involves the accumulation of random variations in genetic material within a single lineage, whereas innovation by composition involves the union of different entities, each containing relatively large amounts of genetic material, that are independently pre-adapted as entities in their own right, if not in their symbiotic role. We will use the term 'accretive adaptation' to refer to the normal view of evolutionary change occurring by accumulating variations within one lineage.

This paper, and our previous research (e.g. Watson et al. 1998, Watson & Pollack 1999b, 2000 & 2001b), is directed toward an adaptational understanding of composition: What kind of adaptation does the formation of higher-level complexes from simpler entities afford in an evolutionary system? Following the conception of evolution as a combinatorial optimisation process on a fitness landscape (Wright 1967), we seek to understand the kind of adaptive domain, the kind of fitness landscape, for which composition is well suited, and to elucidate the adaptive potential of composition as contrasted with accretive adaptation. The model we develop is an abstract theoretical model of biological composition, and the major evolutionary transitions, complementing other abstract models such as the algebraic model of Nehaniv & Rhodes, (2000). It is also an algorithmic model motivated by, and contributing to, the concepts of modularity and abstraction in computational artificial evolution methods, i.e. evolutionary computation, (e.g. Holland 1975, Mitchell 1996, Spears et al. 1993).

We acknowledge that the evolutionary transitions and compositional mechanisms do not necessarily have an adaptive role, and that the notion of an objective fitness landscape and the 'problem/solution' metaphor are not necessarily appropriate for evolutionary processes in general (e.g. see Lewontin 2000 for discussion). Nonetheless, an examination of the *potential* for adaptive change offered by a biological phenomenon, and a better understanding of the circumstances, if any, where an adaptive advantage is conceivable, are likely to be a useful component of our understanding. More specifically, when a phenomenon has been involved in several major innovations in evolutionary history, we suggest an adaptational stance is one of those that should be investigated. Since we find in the following models, under certain circumstances, compositional mechanisms can enable the evolution of complex adaptations that are not evolvable via accretive mechanisms, it behoves us to examine these conditions and the kind of adaptations for which this is possible.

### 1.2 Composition and scalable evolvability

Composition immediately suggests two complementary concepts that impact evolvability: a scaling-up in the 'unit of selection', and a scaling-up in the 'unit of variation'. The creation of higher-level organizational units in the major evolutionary transitions is often associated with new units of selection, hence hierarchical selection, (e.g. Michod 1999). But we want also to point out that the entities created by a union may create a new higher-level unit of variation—a kind of 'coarse graining' for the formation of groups at the next level of organisation. That is, since the entities involved in a composition each contribute a number of features, variation in the space of their combinations constitutes a higher-level variation mechanism (than mutational change), and each new entity creates a new unit of variation for subsequent composition. Put from the viewpoint of each entity involved, a new partner introduces a large set of features simultaneously. Moreover, this is not an entirely arbitrary set of features but a set that has been pre-adapted by parallel adaptation in (semi-)independent lineages. Thus the results of selection at one level of organisation provide the components for variation at a higher level of organisation. Our intuition is therefore that composition permits an adaptive scaling-up in the mechanism of adaptation. Algorithmically, the unit of variation impacts *modularity*—the identification of meaningful components that can be re-used to make subsequent variation more 'informed',—and the unit of selection impacts *division of labour*—the decomposition of a complex adaptation into simpler adaptations such that each can be evolved by semi-independent processes.

We can view the entities involved in composition as an abstraction of the feature space into higher-level units—thus enabling variation to search in the space of successful combinations of organisms rather than combinations of the original 'atomic' features. Naturally, most combinations of organisms do not make a successful composite—but then, neither do most mutations, for example. The point is whether combinations of entities, that are themselves successful combinations of features, are more likely to produce a viable variant than the 'raw' combinations of features provided by mutational events. The fact that the component entities are pre-adapted in separate lineages provides a 'divide and conquer' treatment of the feature set. Intuitively, an example might be provided by the notion of a generalist entity, utilising two different niches, resources, or habitats, that can be formed by the composition of two specialist entities each independently adapted to one of these niches, resources or environments. Thereby, the problem of being well-adapted to the general environment is divided into the sub-problems of being well-adapted to component environments. This decomposition of a problem into smaller problems is know as 'divide and conquer' (e.g. Cormen et al. 1991); so named because of the significant algorithmic advantage it offers when applicable. Such divide and conquer advantage is not available to a process that optimises systems monolithically and thus is not available to natural selection when features are adapted within a single unit of selection.

### 1.3 Models, and paper structure

To illustrate these concepts we devise an abstract algorithmic model which we call "The Symbiogenic Evolutionary Adaptation Model", or "SEAM", to invoke the notion of symbiotic union or joining. Our intent is to provide a model in which the combinatorics of composition can be clearly seen, showing a concrete illustration of a mechanism that scales-up the unit of

variation and enables a divide-and-conquer algorithmic advantage. To contrast with this model we use the regular Genetic Algorithm, GA, (Holland 1975) as a model of accretive adaptation.

An important complementary part of our model is a characterisation of an adaptive landscape to which composition is well suited. To this end we introduce an adaptive landscape that is 'composition-easy' but very difficult for accretive adaptation. The landscape results from a system of interdependent variables that have a hierarchically clustered structure. This interdependency structure produces a fractal fitness landscape exhibiting significant ruggedness at all scales. The purpose of using this landscape for our experiments is not to suggest that all adaptive problems encountered in nature have this structure. Our purpose is to exemplify the kind of adaptation that is enabled by composition and contrast this with that which is possible under accretive adaptation. That said, this kind of scale-invariant problem structure does have some interesting properties that are quite general and potentially related to scale-invariance often found in natural self-organised dynamical systems (Bak 1996).

In our experiments using our scale-invariant fitness landscape we investigate the ability of both a mutation only algorithm and the GA to cross fitness saddles of increasing size. More exactly, as adaptation continues and the distance to the next-best optimum increases, we would expect that adaptation by these methods would become increasingly difficult. In contrast, we will use SEAM to investigate whether composition is able to overcome the epistasis structure in the landscape by searching combinations of coevolved entities through many hierarchical levels. On this class of adaptive landscape, we expect that evolvability under mutation and sexual recombination within the accretive model of adaptation will be inherently limited, whereas innovation by composition offers the possibility of inherently scalable, open-ended evolvability.

The remainder of the paper is structured as follows. In the following section we outline some related evolvability issues in both evolutionary biology and evolutionary computation. These provide a larger context for the concepts introduced above and detailed in the remaining sections. In Section 3 we describe the Symbiogenic Evolutionary Adaptation Model, as a simple abstract model of composition that we use to explore the ideas we have introduced. Section 4 notes some comparisons between this model and well known evolutionary computation algorithms. In Section 5 we describe the scale-invariant adaptive landscape we will use for our experiments. The experimental results are described in Section 6. Section 7 concludes.

## 2 Related issues in biological and computational models

In this section we outline some related evolvability issues in both theoretical evolutionary biology and evolutionary computation. These provide a larger context for the theoretical concepts introduced above and detailed in the remaining sections of the paper.

### 2.1 Some related models impacting biological evolvability

Sewell Wright (1931) stated that "the central problem of evolution... is that of a trial and error mechanism by which the locus of a population may be carried across a saddle from one peak to another and perhaps higher one." This conception of evolutionary difficulty, and the now familiar concept of evolution as a combinatoric optimisation process on a rugged landscape (Wright 1967), are central in issues of evolvability. In keeping with this view, there are many models of how evolvability can be enhanced by increasing the ability of adaptation to escape or

otherwise avoid local optima—configurations of features where no small change in features will produce a fitter variant.

Some models suggest that local optima are not as prevalent as might be expected naively. For example, neutral networks (Huynen et al. 1996) are pathways through genotype space that enable neutral variation (Kimura 1983) to arrive at configurations that are genotypically close to a large number of different phenotypes. By this means, the number of fitter phenotypes that are reachable (without passing through phenotypes that are less fit) is increased with respect to a substrate without such pathways. Extra-dimensional bypass (Conrad 1990), recognises that the number of features an entity exhibits changes over evolutionary time, and put crudely, although an entity might be stuck on a local optimum in 4-feature space, it might be able to move around the impasse in the extra 'degree of freedom' provided by a 5-feature space.

Other issues impacting evolvability include the fact that although there may be local optima in phenotype space, small variations in genotype can provide large changes in phenotype. Sophisticated ontogenic processes (e.g. Waddington 1942) provide a complex mapping from genotype to phenotype and the structure of this mapping is critical to understanding how small random changes in genotype might enable large changes in phenotype. Exaptation (Gould & Vrba 1982) refers to cases where a collection of features adapted for some purpose is co-opted for some other purpose or function; with respect to the function of interest, a large set of phenotypic features may be introduced simultaneously.

Each of these models/issues has some impact on the ability of adaptation to 'tunnel across', 'by-pass', 'jump over', or otherwise traverse fitness saddles and escape local optima. Some of them even offer the potential for a mechanism of variation that improves adaptively. For example, the structure and effect of ontogenic processes are subject to selection, so the large changes in phenotype that they facilitate are not arbitrary and could conceivably scale-up adaptively as they are 'tuned' by further selection. But, none of these models involve a scaling-up in the reproductive unit. The ontogenic mechanisms, neutral networks, extra-dimensions, and exapted features occur within single lineages, and do not involve (in and of themselves) sets of features being adapted in parallel in different reproductive lineages and or being subsequently combined or assembled together into a new reproductive unit. Algorithmically, this means that they do not afford any opportunity for divide and conquer algorithmic advantage. Biologically, this means that these models apply to adaptation between transitions, not to transitional changes themselves.

Nonetheless, it is quite possible that all of the advantages and possibilities that these mechanisms offer can be multiplied by the opportunity to compose entities together during a transition. The model of composition we present here does not include these mechanisms—we use a direct one-to-one map between genotype and phenotype that does not allow the possibility of neutral networks, or ontogenic processes, for example. However, the model is quite compatible with these possibilities—these non-transitional issues are mostly orthogonal to the possibility of composition, changing the unit of selection, and divide and conquer algorithmic advantage.

### 2.2 Biological mechanisms relating to composition

Unlike the above mechanisms, there is a family of mechanisms that copy, or otherwise re-use, pre-adapted feature sets—e.g. gene duplication (Ohno 1970), horizontal gene-transfer (Mazodier & Davies 1991, Smith et al. 1992), and sexual recombination. Clearly, these mechanisms involve

variations in higher-level aggregations of genetic material (with respect to mutation). And arguably, to the extent that a gene, or a section of chromosome, can be duplicated, or be propagated through reproductive events, without the whole chromosome being reproduced, these mechanisms do involve a unit of selection smaller than the individual (Dawkins 1976). If we accept a two-level model of selection for these mechanisms—the sub-organismic level (e.g. gene or subsection of chromosome) and the organismic level—then these mechanisms constitute a limited form of composition. That is, the components at the organism level have been pre-adapted in parallel (semi)-independently and subsequently brought together.

Allopolyploidy (having chromosome sets from different species) (Werth et al. 1985) is also a form of composition; limited only in the sense that it usually occurs between closely related species.

However, these mechanisms do not provide a clear hierarchical model moving through many successive levels. Moreover, gene duplication, horizontal gene-transfer, and sexual recombination also depend on a specific effect to maintain the coherence of lower level components, or modules—namely, *genetic linkage*. That is, if the nucleotides of a gene were somehow distributed along the chromosome then they could not maintain their integrity through sexual recombination events, they would not be likely to be copied as a unit, nor be transferred horizontally as a unit. The usefulness of modules represented in sections of chromosome depends on the correspondence of genetic linkage with epistatic linkage (Watson & Pollack 1999c, Wagner and Altenberg 1996)—which must not be taken for granted.

In the work we present here we wanted to present an open-ended multi-level model where there is no *a priori* definition of different levels of units/modules, and we did not want assume favourable genetic linkage. Accordingly, we do not include gene duplication, horizontal gene transfer, or polyploidy explicitly in our model—but, as stated, we view composition as a general form of these mechanisms. Previous work has explored the operation of sexual recombination with and without the assumption of favourable genetic linkage (Watson and Pollack 2000, see also, Watson 2002). For the purposes of contrast, we include experiments using sexual recombination (without the assumption of favourable genetic linkage) in our experimental section.

### 2.3 Modularity and credit assignment in artificial evolution

Concepts of modularity are familiar in artificial evolution methods. The notion of 'building-blocks'—aggregations of features that form useful components for subsequent adaptation—has been present since the inception of Genetic Algorithms, GAs, (Holland 1975). We will not attempt to provide a comprehensive review here, but we mention that there are many methods and techniques used in artificial evolution models that address modularity and division of labour.

For example, modularity is addressed implicitly by the use of variable-length, moving-locus, non-linear, and generative encodings—for example, Messy GA (Goldberg et al. 1989), Linkage learning GA (Harik & Goldberg 1996), Genetic Programming (Koza 1992), and cellular encoding (Gruau 1994). And modularity is addressed explicitly in mechanisms that 'encapsulate' subsets of features for subsequent re-use during the search process—for example, 'Automatic module acquisition' (Angeline and Pollack, 1993), 'automatically defined functions' (Koza 1994), and 'adaptive representation' (Rosca 1997). The advantage of these explicit methods is that "the modularization of representational components and their protection from mutation

[/internal variation] can be viewed as removing unnecessary dimension[s] from the search space…" (Angeline and Pollack, 1993).

Methods explicitly addressing the division of labour include Learning Classifier Systems (Holland & Reitman 1978), Cooperative Coevolution (Potter 1997), and Evolutionary Divide and Conquer (Rosca 1997), as well as techniques embedded in the modularity methods listed above. In all these methods, the same question re-occurs: How do we evaluate the value of a module? We want to promote modules because they are useful 'building-blocks' even though they may not necessarily be valuable in isolation. Since the module is not a complete solution but a partial solution it must be evaluated in some context—for example, an assembly of modules. If it is evaluated in an assembly of modules then how do we apportion fitness to the modules involved? This is known as the 'credit assignment problem'. In genetic algorithms using crossover, (where the implicit modules are sections of chromosome and the context is the individual), the credit assignment problem is manifest in 'parasites' (Goldberg et al. 1989) and 'hitch-hiking' (Forrest & Mitchell 1993).

Different methods take different approaches to the credit assignment problem; for example, in previous work we have used fitness sharing methods to successfully evolve modular solutions to Genetic Programming problems (Juille & Pollack 1996). In the model presented here, we attempt to find a principled way to make selection over different sized entities provide the characteristics of modularity and division of labour hierarchically and in a principled fashion following from the inspiration of the evolutionary transitions. Some comparisons between the Messy GA and composition are given in (Watson & Pollack 2000, Watson & Pollack 1999c).

Related computational models specifically addressing symbiosis and compositional mechanisms include (Bull 1999a & 1999b, Bull & Fogarty 1995). These models offer important insights into compositional mechanisms and their role in the evolutionary transitions, and the subject of our models is closely aligned with these. Our work also contrasts with these previous works in several respects: we address the serial application of the relevant mechanisms through several hierarchical levels in a unified model; we place an emphasis on the composition of lineages that are not genetically related (or similar); we model the interaction of more than two lineages; and we do not pre-define the identity of possible groups. Dumeur (1995) offers an algorithmic model that is conceptually allied with ours, but of quite a different procedural style. An additional distinction from previous work comes from our work on characterising the class of adaptive landscape that exemplifies the adaptive capacity of compositional mechanisms. This landscape is an important part of our computational model and helps us to understand the potential impact of compositional mechanisms on the evolvability of adaptive processes.

Some additional, brief comparisons with existing evolutionary computation methods are given in Section 4, after our composition model is introduced.

# 3 The Composition model

## 3.1 Overview of the composition model

In this section we examine a simple abstraction of composition which we call the *Symbiogenic Evolutionary Adaptation Model*, or "SEAM", to invoke the idea of symbiotic unions or joins (see Watson & Pollack 2000). This model is a population based computer simulation sharing some characteristics with Evolutionary Algorithms, (EAs), and in particular, Genetic Algorithms,

(GAs). For example, the model uses a population of entities, a variation operator to create new entities, and a 'fitness function' to select between variants.

However, there are important differences. The entities in the population of an EA are usually interpreted as variants of the same species, but in SEAM the set of entities represent an ecosystem of different species. This has implications for how we perform selection, outlined below. The variation operator is the central component of the model and is based on an abstraction of composition. Instead of the usual genetic operators of mutation or sexual recombination, the variation operator in SEAM can be thought of as a means for joining two (randomly picked) entities in a symbiotic union. The use of this operator asserts the possibility of mechanisms that support the creation of a union between two entities, e.g. the formation of cell membranes, or the instantiation of an endosymbiotic relationship; the critical remaining factor is to determine whether such a union, if it should occur, would be selected for.[1]

The fitness function is a mapping from a set of feature values to some scalar value that represents the 'adaptive utility' of a variant. As in most EAs, this fitness value goes through some manipulation, at least scaling, to determine the number of offspring a variant will have. But in SEAM, the treatment of fitness is quite different from normal EAs. The central question that our fitness assessments need to answer is whether an entity is fitter when composed with some symbiont or when it is alone. We will assume that a join between two entities is 'unstable' if either of the component entities is fitter alone than when joined with the proposed symbiotic partner. In this case, the composition will be dismantled, otherwise the pair will always co-occur in future and thereby become a new higher-level entity that may participate in subsequent joins.

Of course, the benefit of a symbiotic composition is dependent on what environment it is in, or what alternative environments each component entity might occupy. Accordingly, a fundamental aspect of the model must be that the fitness of an entity changes in different environments, and we need to assess the fitness of an entity, not in isolation, but in some 'environmental context'. For the purposes of our model, we will not consider adaptation to static environmental factors but rather focus on the interaction with the biotic environment provided by other coevolving entities, since it is changes in their biotic associations, if any, that are the subject of interest. In short, the environmental contexts will be formed from transient groups of other individuals in the ecosystem. The 'overall fitness' of an entity will then be some function of many 'context sensitive' fitness measures.

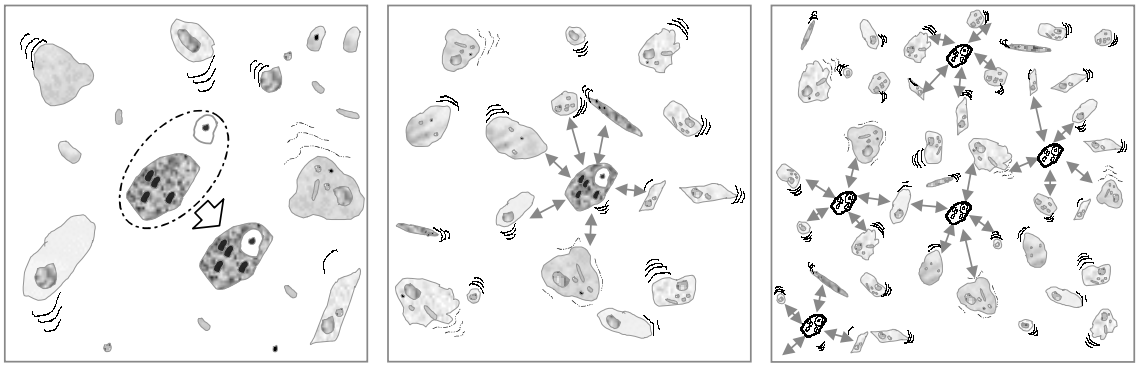
However, determining such an overall fitness will require knowledge of the 'weighting' of each context for each entity, (for example, the probability with which each entity may occupy each environment). Such weightings are the effect of many complex ecosystem factors that are beyond the scope of what we want to model here. Nonetheless, we make some conservative simplifications that, in some cases, allow the determination of fitness superiority even without the knowledge of environmental weightings. Our method borrows ideas from multi-objective optimisation (e.g. Fonseca and Fleming 1995), and particularly the concept of 'Pareto dominance', which is specifically developed for determining superiority where the weighting of a number of objectives is unknown. This enables us to determine the preference for a symbiotic join in a way that is fundamentally sensitive to environmental context, yet we can do this using a

[1] There are two complementary views concerning the mechanisms of the major evolutionary transitions: a) organisms cooperate because they have been encapsulated by some mechanism into a single reproductive unit (e.g. co-dispersal (Frank 1997), shared genetic transmission mechanisms (Dawkins 1976)), b) organisms become encapsulated as a single reproductive unit because this is a canalisation (Waddington 1942) of an existing cooperative relationship. In a sense, our model follows the latter view since entities do not have an opportunity to change their behaviour after encapsulation. But the hypothetical pay-off matrices we will introduce could be conditioned on the effects of co-dispersal or co-transmission, for example. Accordingly, the distinction is not essential in an abstract model such as ours.

simple and abstract model that does not explicitly contain complex factors of environmental structure.

In overview, the composition model, SEAM, that we will use in our experiments develops as follows. The ecosystem is initialised with many different small entities. Pairs of entities are then picked at random to see if they might form a stable symbiotic join. If the overall fitness of either entity alone could be, dependent on environmental contexts, greater than the fitness of the entity with the proposed symbiotic partner then the composition is deemed *unstable* and the original entities are returned to the ecosystem. Otherwise the composition is deemed *stable* and the two entities always co-occur together in future. The process of building and selecting compositions of entities is repeated, eventually building larger and larger composite entities.

Three main features of SEAM are depicted in Figure 1. Frames (a) through (c) in the figure loosely correspond to variation, evaluation, and selection, respectively. These processes, outlined in the figure, are detailed in the subsequent subsections.



a) New entities are created by joining two existing entities together.

b) The fitness of an entity is dependent on its environmental context.

c) An entity is placed in many contexts to test the stability of a new join.

**Figure 1**: A caricature of processes in SEAM. a) New entities are created from the composition or joining of two randomly selected extant entities (Section 3.2, Figure 2). b) The fitness of any entity (possibly the result of a previous join) has dependencies with its environmental context, i.e. a random selection of other entities from the ecosystem (Section 3.4, Figures 3 & 4). c) The new pairing is subject to many such contexts. If there is some environment of other entities in which either component of the join is fitter individually than when it is with its proposed partner, then the join is deemed unstable and is dismantled. This follows the assumption that the partnership must be in the 'selfish' interest of the partners involved. In our implementation, the stability of a proposed join is tested in many contexts and is immediately undone if found to be unstable. This models the assumption that competition between joined and non-joined variants of an entity occurs rapidly such that only reliably successful joins persist long enough to be involved in a subsequent join (Section 3.3, Eq. 3). A join that persists through (c) is treated as a new entity that may participate in further joins as the cycle of the model repeats.

### 3.2 Entities and their composition

SEAM is an abstract representation of an ecosystem incorporating many different entities. These entities may be interpreted as genes, bacterial cells, more complex cells, or any other higher level of organization—the intent is to model transitions between these levels in an integrated model of 'entities'. We will use the word 'species' to refer to types of entity at any level. Entities are represented only by their *features values*, and for now, species are simply the set of entities with identical feature values. These features may be interpreted as genes, as phenotypic features corresponding to genes, or as higher level features of an organism such as resource usage or a behavioural strategy. In general, they are the set of characteristics that affect the fitness of an entity and the fitness-sensitive interactions of the entity with its environment and other entities. Our model abstracts away all population dynamics within a species and therefore the ecosystem will only incorporate one representative entity of each species.

The basis of our composition model will be that a composite is created from the joining of features from two different species of entity. Accordingly, it is necessary that different entities will specify different subsets of features (not just different values of the same set of features). To provide a simple example: let each feature take one of two values, "0" or "1", and let the features be identified by an index, $F_n$. Then one species might specify features $F_3$=0, $F_7$=0, $F_{12}$=1, and a second species might specify $F_1$=1, $F_9$=0, $F_{10}$=1, $F_{15}$=0. Then their join may create the new species with features $F_1$=1, $F_3$=0, $F_7$=0, $F_9$=0, $F_{10}$=1, $F_{12}$=1, $F_{15}$=0.

We will use a large finite set of possible features for simplicity in the implementation of the model,[2] but the number of features could be flexible in alternate implementations. The number of features specified by any one entity may be anything from one to the full set. In this way it is simple to write the specification of a species using a fixed length string. For example, working in a 16-feature space, we may write the two entities given in the example above as ***A*** and ***B*** in the left of Figure 2 below, and their composition may be written as ***A+B***.

The 'null features', "**–**", in this representation are 'placeholders' for features that are not (currently) specified by an entity. We will refer to the 'size' of an entity to mean the number of non-null features—for example, the entities used above have sizes 3, 4 and 7, respectively. Figure 2 also illustrates how we will deal with conflicting specifications when they arise.

[2] Note that in the GA individuals also generally use a finite set of binary features, 'genes', but unlike the entities in SEAM, individuals in the GA must generally specify a value for every possible gene. This is natural for a model of evolution within a single lineage where every individual has basically the same features but varies in the values of these features. The 'null' value used in the implementation of SEAM, detailed shortly, cannot reasonably be characterised as a third allele since it is not heritable in the same way as non-null values (see Figure 2).

**Figure 2**: 'Symbiotic composition' (an abstraction of Figure 1, (a)). Left) Composition of two variable size entities, A and B, produces a composite, C, that is twice the size of the donor entities with the union of their features. Here we represent unspecified features by "**–**". The composite is created by taking specified (i.e. non-null) features from either donor where available. Right) Where conflicts in specified features occur we resolve all conflicts in favour of one donor, e.g. the first.

Algebraically, we define the composition of two entities A and B, as the superposition of A on B, below. $A=(A_1,A_2,\ldots A_n)$, is the entity where feature $F_i$ takes value $A_i$. S(A,B) is the superposition of entity A on entity B, and s(a,b) is the superposition of two individual feature values, as follows:

$$S(A,B)= S((A_1,A_2,\ldots A_n),(B_1,B_2,\ldots,B_n)) = (s(A_1,B_1),s(A_2,B_2),\ldots,s(A_n,B_n)), \quad \text{Eq.1.}$$

$$\text{where, } s(a, b) = \begin{cases} a, \text{ if } a \neq \text{null}, \\ b, \text{ otherwise.} \end{cases}$$

This composition will be the only mechanism of variation in our model. The intent is that the model will start from 'atomic', i.e. single-feature, entities and compose them together into larger composites, and compose these together, and so on. When small entities are composed with relatively large entities, their effect is like single-feature mutations, but as entities become larger, their composition enables variations that scale-up with their size.

Note that the way we use species in this model has no implication of restricting possible unions based on type—in principle, new entities may be created by the composition of any two existing entities regardless of their species, i.e. regardless of the features they represent.

### 3.3 'Pareto dominance' to determine whether a symbiotic composition is preferred

Having defined a variation operator that defines a join of two entities, we need to determine whether such a join would be adaptive. Our basic assumption is that the symbiotic relationship must be in the 'selfish' interest of both the component entities involved. That is, if the fitness of either component entity is greater without the proposed partner than it is with the proposed partner then the composite will not be selected for. If, on the other hand, the fitness of both component entities is greater when they co-occur then the relationship is deemed stable and will persist. However, the fitness of any entity is dependent on its environmental context; possibly, in one environment an entity may be fitter when co-occurring with the proposed symbiont, and in another context the symbiosis may depreciate its fitness. Thus whether a symbiotic relationship is preferred or not depends on what environmental contexts are available.

For our purposes, the set of possible environmental contexts is well defined: an environmental context is a complete set of features (in which some partially specified entity, which may be the result of many joins, can be assessed). See Figure 3.

```
---0-11---110---   x,   an entity specifies a partial set of feature values.
0110101100010011   θ,   an 'environmental context' is a complete set of feature values.
011011110011001011 S(x,θ), the entity x superimposed on the context θ.
```

**Figure 3:** A partially specified entity must be assessed in a context.

We assume that the overall fitness of the entity will be a sum of its fitness over different environmental contexts weighted by the frequency with which each environment is encountered. But, we would not generally suppose that the frequencies with which different environments are encountered by one type of entity would be the same as the frequencies relevant to a different type of entity. That is, we imagine that different species have different distributions over possible environments. Let us assume that we have a measure of the 'context sensitive fitness', $csf(x,\theta)$, of an entity, $x$, in any given environmental context, $\theta$, and that the overall fitness of the entity $x$, will be $F(x)$ which is the sum of its fitness over all environments weighted by the frequency of that environment for that species, as below.

$$F(p) = \sum_{\theta \in Contexts} \left( \lambda_{(\theta,p)} csf(p,\theta) \right) \qquad \text{Eq.2.}$$

where $\lambda_{(\theta, p)} \geq 0$ is the weighting of the environmental context $\theta$, for entity $p$.

Now, whether a symbiotic relationship is preferred or not depends on the relative weighting of each context to each entity involved, and many factors could influence this. For example, a biased distribution over environmental contexts may be 'inherited' by virtue of the collocation of parents and offspring, or affected by the behavioural migration of organisms during their lifetime, or the selective displacement of one species by another in short term population dynamics. We did not wish to introduce such factors and accompanying assumptions into our model. Fortunately, the concept of *Pareto dominance* is specifically designed for application in cases where the relative importance of a number of factors is unknown (e.g. see Fonseca & Flemming 1995). Put simply, this concept states intuitively that, even when the relative weighting of dimensions is not known, the overall superiority of one candidate with respect to another can be confirmed in the case that it is non-worse in all dimensions and better in at least one. More exactly, '$x$ Pareto dominates $y$' is written '$x >> y$', and:

$x >> y \Leftrightarrow (\forall\theta : csf(x,\theta) \geq csf(y,\theta)$ AND $\exists\theta : csf(x,\theta) > csf(y,\theta)$.

or equivalently, given that $x$ and $y$ are different in at least one dimension:

$x >> y \Leftrightarrow \nexists\, \theta : csf(y,\theta) > csf(x,\theta)$.

In cases where there is some $\theta$ such that $csf(x,\theta) > csf(y,\theta)$ and some other $\theta$ such that $csf(x,\theta) < csf(y,\theta)$, we say that $x$ and $y$ are *non-sorted*. And in cases where $\exists x: x >> y$ we say that $y$ is *dominated*, else $y$ is *non-dominated*. For our ecological domain, these simple rules are easily interpreted. In the case where $x$ is better in some environments than $y$, and $y$ is better in some environments than $x$, then we do not know which is fitter overall unless we know the relative weighting of the environments for each entity. But, if $x$ is always fitter (or at least as fit as) $y$, then regardless of the weightings of the environments for each entity, we know that the overall fitness of $x$ is greater than that of $y$ (assuming $x$ and $y$ are different in at least one dimension).

This pair-wise comparison of two entities over a number of contexts will be used to determine whether a symbiotic join produces a stable composite. If we write the composition of entities ***a*** and ***b*** as ***a+b***, then, using the notion of Pareto dominance, ***a+b*** is stable iff ***a+b*** **>>** ***a***, and ***a+b*** **>>** ***b***. In other words, ***a+b*** is unstable if there is any context in which either ***a*** or ***b*** is fitter than ***a+b.***

i.e. ***stable***(***a+b***, ***a***, ***b***) ≡ ***a+b*** **>>** ***a*** AND ***a+b*** **>>** ***b***,

i.e. ***unstable***(***a+b***, ***a***, ***b***) ⇔ ∃θ∈***Contexts***: (***csf***(***a***,θ) > ***csf***(***a+b***,θ) OR ***csf***(***b***,θ))> ***csf***(***a+b***,θ))

where ***Contexts*** is a set of complete feature specifications.

We should note that there is a subtle distinction between 'the fitness of an entity *in* an environment' and 'the fitness of the entity *and* environment together' i.e. ***csf***(***x,***θ) ≠ ***f***(***S***(***x,***θ)). However, our method precludes the need to separate the former from the latter because the pair-wise comparison of two entities in the same environmental context implicitly 'differences away' the contribution of the environment. That is, ***csf***(***x,***θ) > ***csf***(***y,***θ) ⇔ ***f***(***S***(***x,***θ)) > ***f***(***S***(***y,***θ)), where ***f***(***w***) is the objective fitness of the complete feature set ***w*** as given by the fitness function. This assumes that although we can only measure the fitness of a complete feature specification (organism and environment together) we can determine the information we need by differencing away the fitness contributions coming from the environment by including them in both sides of the inequality.

Thus our condition of instability becomes:

***unstable***(***a+b***, ***a***, ***b***) ⇔ ∃θ∈***Contexts***: (***f***(***S***(***a***,θ)) > ***f***(***S***(***a+b***,θ)) OR ***f***(***S***(***b***,θ))> ***f***(***S***(***a+b***,θ)))

Eq.3.

Equation 3 becomes our abstraction for Figure 1 (c).

### 3.4 Building environmental contexts

In our model, the environmental contexts, used in determining Pareto dominance and the stability of a proposed composition, will be formed entirely from other members of the ecosystem. The intent here is that the assessment of a new composition involves selecting between being in permanent association with some particular member of the ecosystem or being in transient association with members of the ecosystem. If we were to employ the naive alternative, selecting between being in permanent association with some particular member of the ecosystem or remaining in entirely random environmental contexts, then it would be likely that many more proposed associations would be preferred. This would result in many sub-optimal associations. Additionally, if entities are evaluated in transient groups of other entities then there is the potential that they may become co-adapted to one-another, and thereby 'primed' to make successful permanent joins by composition. Figure 4 illustrates how to build a context from a randomly selected set of entities.

| | |
|---|---|
| a: | --0---1- |
| b: | 01------ |
| c: | -0---0-- |
| d: | ----1-0- |
| e: | ------10 |
| f: | ---0-00- |
| Resultant context | 01001010 |

**Figure 4:** Building a context from other entities, (an abstraction for Figure 1 (b)). In this example, six entities **a** through **f**, are needed to complete a fully-specified feature set of eight features. Where specified features conflict, the specifications of the topmost entity take precedence, as in Figure 2.

Algebraically, we define a context, using the recursive function $\boldsymbol{S}^*$, from an ordered set of n≥2 entities $X_1, X_2, \ldots X_n$, as follows:

$$\boldsymbol{S}^*(X_1, X_2, \ldots X_n) = \begin{cases} \boldsymbol{S}(X_1, \boldsymbol{S}^*(X_2, \ldots X_n)), & \text{if } n>2, \\ \boldsymbol{S}(X_1, X_2), & \text{otherwise.} \end{cases} \quad \text{Eq.4.}$$

where $\boldsymbol{S}(X_1, X_2)$ is the superposition of two entities as per Eq.1 above.

Some contexts may require more or fewer entities to provide a fully-specified feature set. In principle, we may use all entities of the ecosystem, in random order, to build a context—but, after the context is fully-specified, additional entities will have no effect. This allows us to write a context as S*(E), where E is all members of the ecosystem in random order. Implementationally, we may simply add entities until a fully-specified set is obtained.

### 3.5 The Symbiogenic Evolutionary Adaptation Model (SEAM)

We may now put together the components we have introduced above to provide a complete model. To summarise, the model includes the following features:

- Variable size entities and a variation operator based on composition.
- Building environmental contexts from other co-adapting entities in the ecosystem.
- Testing (in)stability of compositions by testing for Pareto dominance of the composition over the component entities.

Although each of these features is conceptually somewhat involved, the overall simulation model is not that complicated. Figure 5 overviews the operation of SEAM.

- Initialise ecosystem, E, to random, single-feature, entities.[(1)]
- Repeat until *stopping condition*:
  - Remove two entities at random from the ecosystem $\rightarrow \boldsymbol{a}$ & $\boldsymbol{b}$.
  - Produce $\boldsymbol{a+b}=\boldsymbol{S}(\boldsymbol{a},\boldsymbol{b})$, using composition (see Eq.1).
  - If ***unstable***$(\boldsymbol{a+b}, \boldsymbol{a}, \boldsymbol{b})$ return $\boldsymbol{a}$ and $\boldsymbol{b}$ to ecosystem, else add $\boldsymbol{a+b}$ to ecosystem.

where ***unstable***$(\boldsymbol{a+b}, \boldsymbol{a}, \boldsymbol{b}) \Leftrightarrow$

$$\exists\theta \in \boldsymbol{Contexts}: (\boldsymbol{f}(\boldsymbol{S}(\boldsymbol{a},\theta)) > \boldsymbol{f}(\boldsymbol{S}(\boldsymbol{a+b},\theta)) \text{ OR } \boldsymbol{f}(\boldsymbol{S}(\boldsymbol{b},\theta)) > \boldsymbol{f}(\boldsymbol{S}(\boldsymbol{a+b},\theta)))$$

where ***Contexts*** is a random set of contexts each built by composing together other members of the current ecosystem, E, using $\boldsymbol{S}^*(E)$ (see Eqs. 3 & 4).

(1) Initialisation needs to completely cover the set of single-feature 'atoms' so that all values for all features are available in the ecosystem.

**Figure 5:** Pseudocode for a simple implementation of SEAM.

# 4 Comparisons of SEAM with Genetic Algorithms

## 4.1 Comparison of Pareto dominance with selection in Genetic Algorithms

The use of Pareto dominance in SEAM explicitly respects the *multi-dimensional* nature of fitness: That is, the fitness of an entity is different in different environments, and each environmental context provides a *dimension* of its fitness. In a multi-dimensional view of fitness we immediately lose the notion of an absolute 'best' individual—'best' is undefined without specifying a context, i.e. one individual might be the best in some context, and some other may be best in some other context. Interestingly, we may instead use the notion of a *Pareto optimal set* (e.g. see Fonseca & Fleming 1995) of individuals which are optimal in the sense that no individual in the set can be improved in any context without necessarily being degraded in some other context.

In contrast, selection in the simple GA assumes a one-dimensional fitness metric against which all entities will be compared. The usual method of dealing with a context sensitive fitness measure is to average its performance over a number of sample contexts.[3] But, this immediately

[3] The notion of 'schema fitness' in genetic algorithm theory is the average fitness of a partial specification over all possible contexts, and when the fitness of an individual is dependent on coevolving individuals, as in coevolved players of a game, an overall fitness is usually acquired from an average score against many opponents (see Watson & Pollack 2001).

collapses the multi-dimensional information back to a single dimension i.e. 'how fit is the entity on average'. Repeated selection in a single dimension of fitness has the consequence that the population will tend to converge to the variants of the 'best on average' individual found. Thus it is not surprising that the problem of maintaining diversity in EAs and premature convergence of the population is ubiquitous (e.g. Mahfoud 1995).

Since SEAM utilises a multi-dimensional treatment of fitness, and accepts a partial ordering of the entities, it can apply useful selection to converge toward the Pareto set without converging immediately to a single type. By using Pareto domination over a number of contexts as a selection criteria, rather than an equally weighted sum of fitnesses over a number of contexts, SEAM provides exactly the balance of competition and coexistence that we were looking for to maintain an ecosystem of complementary specialists, while still permitting selection for good joins in a principled manner.

### 4.2 Comparison of ecosystem contexts with evaluation in Genetic Algorithms

In the Messy GA, (Goldberg et al. 1989), partially specified individuals are evaluated with the use of a 'template' having the role of the context in SEAM. Goldberg et al. correctly suggest that 'locally optimised templates' are useful in revealing the epistatic interactions of a partial feature set, and that a sample of random templates would be problematic because: a) they would not be likely to encounter all the important epistatic interactions with features of the environment (the required number of templates/contexts is exponential in the number of unspecified features); and b) the 'signal to noise ratio' is low in templated context measures (most fitness differentiation would come from the 'background noise' of the template).

The use of other entities to provide contexts in SEAM is an important heuristic for reducing the number of contexts we need to sample. The entities that are used for building contexts include some that are about the same size as the entities being tested and thereby they have the potential to provide templates that are optimised to an appropriate level at all stages of the process. Thus using other members of the ecosystem to build contexts is an important part of the scalable mechanism of SEAM. The issues of background noise are avoided in our method since assessment is carried out in pairwise comparisons of entities over the same set of contexts.

The use of group evaluation in SEAM is similar to the use of a 'shared domain model' in 'cooperative coevolution' (Potter 1997). However, cooperative coevolution is essentially like a single level of the problem decomposition used in SEAM.

## 5 A scale-invariant fitness landscape

A proper account of the evolutionary adaptation of an entity must fundamentally involve a description of the structure and nature of the interdependency of the variables that affect its survival. It must be stated which variables, which genes, environmental properties, and characteristics of other organisms, etc., are dependent on which other variables. Although it is difficult to see how we might attempt to investigate this structure in a specific case, perhaps it is possible to give some general qualitative description of the interdependency structure. For example, we might suppose that the dependency matrix is essentially random and hopefully sparse, as in N-K landscapes (Kauffman 1993). In this section, we describe a dependency structure that is more specific and which makes a significant difference to how adaptation may

take place. Specifically, the interdependency structure is hierarchically clustered in groups, and sub-groups of variables, through many levels. This structure is closely related to the ideas of Simon (1969) on 'nearly decomposable systems'. The resulting fitness landscape exemplifies the adaptive potential of the composition model. Of interest to issues of evolvability is the fact that this landscape is scale-invariant, in the sense that it has fitness saddles at all scales, or resolutions, resulting from its hierarchical construction.

### 5.1 Two-feature epistasis

Ruggedness in a fitness landscape is introduced by the *frustration* of adaptive features, or *epistasis* when referring to the interdependency of genes – that is, it occurs when the 'selective value' of one feature is dependent on the configuration of other features. Fitness saddles are created between local optima. The simplest illustration is provided by a model of two features, each with two possible discrete states, a and b, creating four possible configurations: $F_1a/F_2a$, $F_1a/F_2b$, $F_1b/F_2a$, $F_1b/F_2b$. Table 1, below, gives four exemplary cases for selective values, or fitnesses, for these four combinations. The overlayed arrows in each case show possible paths of adaptation that improve in fitness by changing one feature at a time.

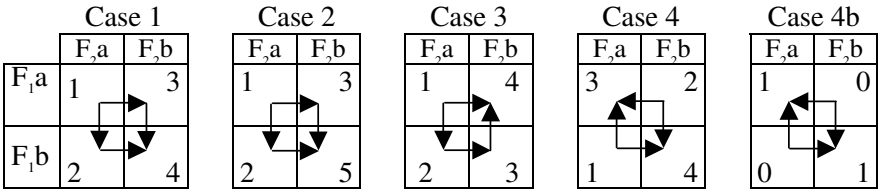



**Table 1:** Example fitness contributions for combinations of two features.

Case 1 shows no epistasis: the difference in selective value between $F_1a$ and $F_1b$ is the same regardless of the value of $F_2$; and the difference in selective value between $F_2a$ and $F_2b$ is the same regardless of the value of $F_1$. Cases 2, 3 and 4 each show some epistasis but with different effects. In Case 2, although the landscape is not planar, the possible routes of single-feature variation are the same as in Case 1, and the landscape still only has one optimum. In Case 3, the preference in selective value between $F_1a$ and $F_1b$ is reversed depending on the value of $F_2$. This forces adaptation into different routes through the landscape, but there is still only one optimum. Only in Case 4, where preference in selective value between $F_1a$ and $F_1b$ is reversed depending on the value of $F_2$, and the preference in selective value between $F_2a$ and $F_2b$ is reversed depending on the value of $F_1$, does epistasis create two optima and a resultant fitness saddle. Changing from $F_1aF_2a$ to $F_1bF_2b$ without going through a lower fitness configuration requires changing two features at once. Lewontin (2000, p84) identifies this same problematic case (for two diploid loci) in a concrete biological example. Accordingly, this form of epistasis provides the base case for the landscape we will use, but for the sake of further simplification, we make the fitness values symmetric (Case 4b), so the change in the sign of preference is retained without a change in magnitude, and the resultant local optima have equal value.

### 5.2 Scaling-up recursively

Having established an appropriate two-feature epistasis model, we need an appropriate way to extend it to describe epistasis between a larger number of features. In particular, we want to re-use the same structure at a higher level so as to create the same kind of epistasis between *sets* of features as we have here between *single* features; in this way, we can create a principled method for producing fitness saddles of larger scales. Our approach is to describe the interaction of four features $F_1$, $F_2$, $F_3$, $F_4$, using the interaction of $F_1$ with $F_2$ in one pair, as above, the interaction of $F_3$ and $F_4$ as a second pair similarly, and then, at a higher level of abstraction, describe the interaction of these two pairs in the same fashion. To do this abstraction we treat the two possible end states of the $F_1/F_2$ subsystem, i.e. its two local optima (labelled c and d, in Table 2), as two discrete states of an 'emergent variable', or 'meta-feature', $MF_1$. Similarly, the two possible end states of the $F_3/F_4$ subsystem (e and f) form two states for $MF_2$. If the original, 'atomic' features are interpreted as low-level features of an entity, then a meta-feature may be interpreted as a higher-level phenotypic feature of an entity, or some higher-level property of an entity that determines its interaction with other entities and/or its environment.

In this manner we may describe the interaction of the two subsystems as the additional fitness contributions resulting from the epistasis of $MF_1$ and $MF_2$. Since each meta-feature includes two 'atomic' features, we double the fitness contributions in the inter-group interaction. Table 2 illustrates.

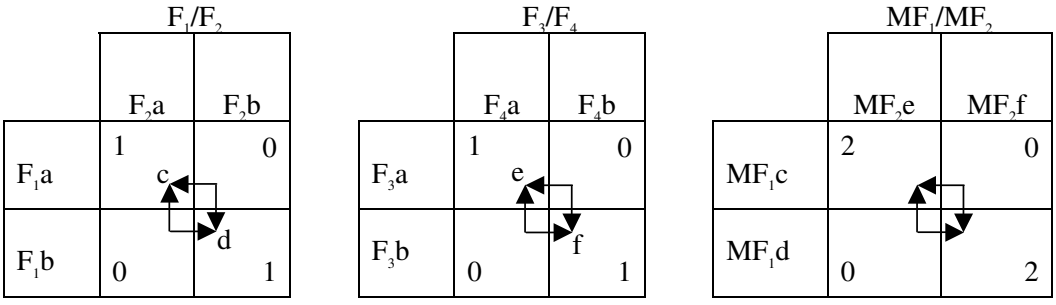



**Table 2:** Abstracting the interaction of two pairs of features, $F_1$/ $F_2$ and $F_3/F_4$, into the interaction of two 'meta-features' $MF_1/MF_2$.

The fitness landscape resulting from this interaction at the bottom level, together with the interaction of pairs at the abstracted level, produces four optima altogether. Using a=0 and b=1, these are 0000 and 1111, which are equally preferred optima, and the local optima 0011 and 1100, which are equally preferred but less so. All other configurations are not local optima.

Naturally, we can take the two best-preferred configurations from the $F_1F_2F_3F_4$ system and describe a similar interaction with an $F_5F_6F_7F_8$ system, and so on. Equation 1 below, describes the fitness of a string of bits (corresponding to binary feature states, as above) using this construction. This function, which we call Hierarchical If-and-Only-If (HIFF), was first introduced in (see Watson et al. 1998) as a building-block test function for genetic algorithms; specifically, providing an alternative to functions such as 'The Royal Roads' (Forrest & Mitchell 1993) and 'N-K landscapes' (Kauffman 1993). In contrast to Royal Roads, HIFF has difficult epistasis between blocks at all levels in the hierarchy, and in contrast to the N-K landscapes the epistatic structure of HIFF is modular.

$$F(B) = \begin{cases} 1, & \text{if } |B|=1, \\ |B| + F(B_L) + F(B_R), & \text{if } |B|>1 \text{ and } (\forall i{:}b_i=0 \text{ OR } \forall i{:}b_i=1) \\ F(B_L) + F(B_R), & \text{otherwise.} \end{cases} \quad \textbf{Eq.5.}$$

where B is a set of features, $(b_1,b_2,...b_k)$, $|B|$ is the size of the set=k, $b_i$ is the $i^{th}$ element of B, $B_L$ and $B_R$ are the left and right subsets of B, i.e. $B_L=(b_1,...b_{k/2})$, $B_R=(b_{k/2+1},...b_k)$. The length of the string evaluated must equal $2^p$ where p is an integer (the number of hierarchical levels).

### 5.3 The resultant landscape

A 128-feature landscape using HIFF (as used in our experiments) has $2^{64} \approx 10^{19}$ local optima (for adaptation that can change one feature at a time) (Watson 2001), only two of which are global optima. If an adaptive mechanism can jump fitness saddles by changing two features at once it still has $2^{32}$ local optima, and so on. To guarantee that an algorithm can escape from any local optimum to a position of higher fitness requires a variation mechanism that can change N/2=64 features at once. Thus, an algorithm using only mutation cannot be guaranteed to succeed in less than time exponential in the number of features (Watson 2001). A particular section through the fitness landscape is shown in Figure 6—the section runs from one global optimum to the other at the opposite corner of the hyperspace (see Watson and Pollack 1999a). As is clearly seen in the fractal nature of the curve in Figure 6, the local optima create fitness saddles that are scale-invariant in structure: that is, the nature of the ruggedness is the same at many successive scales.

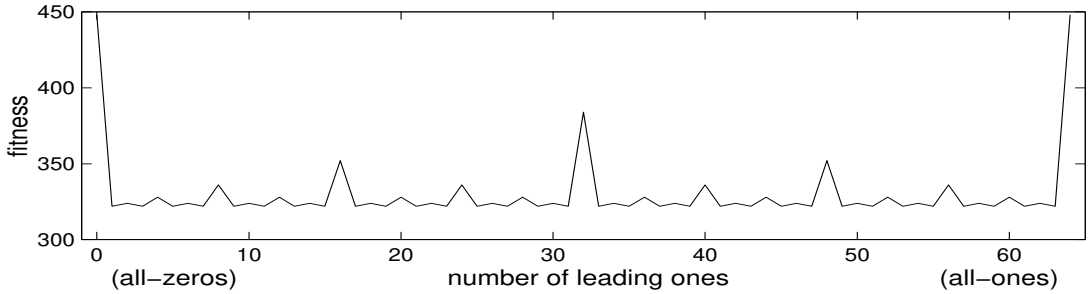



**Figure 6:** A section through a 64-feature HIFF landscape. The two global optima (fitness 448 for 64-feature landscape) are seen at opposite extremes of the space.

The modular structure of the HIFF landscape makes the problem recursively decomposable. For example, a 128-feature problem is composed of two interacting 64-feature problems, each of which has two optima. If both of these optima can be found for both of these subproblems then a global optimum will be found in 2 of their 4 possible combinations. Thus, if this decomposition is known, then the search space that must be covered is at most $2^{64} + 2^{64} + 4 \approx 2^{65}$ configurations. Compared with the original $2^{128}$ configuration space, even this two-level decomposition is a considerable saving. In addition, each size-64 problem may be recursively decomposed giving a further reduction of the search space. In (Watson 2001) we describe how an algorithm having some bias to exploit the decomposition structure (using the adjacency of features on the string) can solve HIFF in polynomial time. Here however, we are interested in the case where the

decomposition structure is not known to the adaptive mechanism. We call this the 'Shuffled-HIFF' landscape (Watson et al. 1998) because this preferential bias is prevented by randomly re-ordering the position of features on the string such that their genetic linkage does not correspond to their epistatic structure (see Watson & Pollack 1999c).

In summary, this landscape exhibits local optima at all scales, which makes it very challenging to accretive adaptation, and fundamental to the issues of saddle-crossing and scalable evolvability. Yet, it is amenable to a 'divide and conquer' approach *if* the decomposition of the problem can be discovered and sub-solutions can be manipulated and recombined appropriately.

### 5.4 Dissolving the distinction between epistasis and multi-player evolutionary games

An alternative interpretation of the two-feature epistasis model above is obtained by viewing the two different features as two different players in a symmetric two-player game, and the feature values as their possible strategies. In this view, the fitness contributions become the values of a pay-off matrix and the salient characteristic of Case 4 is that the optimal strategy for player one is dependent on the behaviour of player two, and vice versa. The particular matrix we arrive at in Case 4b is analogous to the 'mutual benefit' matrix from (Maynard Smith and Szathmary 1995, p.262), but here there is not yet any distinction between the two attractors of the system i.e. which is the 'defect' and which is the 'cooperate' strategy, because we assign them equal value.

As we recursively re-apply the two-feature model we apply the two-player matrix in a recursive fashion to define a four-player game. Note that now, in the context of $F_1aF_2a$, $F_3bF_4b$ is a 'defect/defect' result for players 3 and 4, because it is in their selfish interest for each player not to change from this strategy, but if they both changed to $F_3aF_4a$, this would provide a higher payoff. Conversely, in the context of $F_1bF_2b$, $F_3aF_4a$ is 'defect/defect' and $F_3bF_4b$ is 'cooperate/cooperate'. In other words, whether a strategy provides mutual benefit or not depends on the context in which the game is played.

Thus HIFF describes a hierarchical cooperate/defect game. The nature of the pay-off values is such that maximising the payoff for all (e.g. 128) players is achieved when two-subgroups (of 64) players are compatible. Other attractors in the evolutionary game occur when particular subsets of players are compatible intra-group but not inter-group. Accordingly, optimising HIFF requires the induction of hierarchical cooperation. The pay-off values at every level of resolution help to identify good combinations of strategies—but, which of the two optima at every level is best does not become clear until the context of other players is stabilised. HIFF deliberately dissolves the distinction between epistasis (the interdependency of genes within an individual) and multi-player evolutionary games (the interdependency of features of one entity with those of another) as is required from a model incorporating changes in the unit of selection.

HIFF may be contrasted with previous models of landscapes designed to model the interaction of coevolving species, for example, the NKC models of Kauffman (1993), as used in (Bull 1999). First, whereas the NKC model represents a modular landscape with a single level of modular organisation (composed of two coupled semi-independent landscapes, or several in 'NKCS' models), HIFF depicts a hierarchically modular landscape defined recursively. Second, whereas the NKC model (like other N-K models) uses random epistatic interactions between variables, HIFF uses a specific, and difficult, kind of epistatic dependency (Case 4b) that enables us to control what the consequences of these dependencies are in terms of local optima, and the width of fitness-saddles.

### 5.5 The HIFF landscape and natural hierarchy

HIFF is used in our experiments to exemplify the class of adaptive landscape in which the evolvability of composition can be contrasted with the evolvability of accretive evolution. We do not claim that HIFF is representative of the structure of adaptive landscapes in general. However, the problem of defining appropriate models for adaptive landscapes is an open one and, in passing, we note that HIFF exhibits some interesting landscape characteristics with respect to hierarchy in natural systems (Simon 1969). In particular, dynamical systems exhibiting an interdependency structure that is similar at many scales might be a natural product of self-organized dynamical systems—as evidenced by 'power law' signatures in their dynamics (e.g. Bak 1996). Then, to the extent that natural adaptive landscapes are the result of such systems—scale-invariant fitness landscapes, such as that which HIFF defines, might not be entirely hypothetical.

## 6 Experimental Results

In this section we show empirical results of SEAM applied to a 128-bit Shuffled HIFF. Our intent is to illustrate the qualitative difference in the way that composition operates in this scale-invariant problem as compared to the operation of ordinary (non-transitional) evolutionary change. Accordingly, we contrast the operation of SEAM with the results of a mutation only algorithm, Random Mutation Hill-Climbing, (RMHC), and a genetic algorithm, GA, using sexual recombination.

### 6.1 Experiments

RMHC repeatedly applies mutation to the features of a single binary string (a fully specified feature set) and accepts a variant if it is fitter (Forrest & Mitchell 1993). We conducted experiments with various mutation rates (probability of assigning a new random state {0,1} to each feature)—specifically, ***mut*** = 1/128, 2/128, 4/128, 6/128, 8/128, 12/128, 16/128, 24/128, 32/128 and 40/128. In the following results we show the performance of RMHC with ***mut***=16/128=0.125 which gives the best maximum average maximum fitness over all these values. (See Oates & Corne, 2001, for an investigation of the mutation landscape for HIFF).

The genetic algorithm is a steady state algorithm, using deterministic crowding, (DC), (Mahfoud 1995) to maintain diversity in the population—Figure 7. Previous work (Watson & Pollack 2000) indicates that DC is very effective at maintaining diversity in this problem, and this method provides the best performance of the GA we have found. (A GA using fitness proportionate selection or rank selection with no diversity maintenance method gives significantly inferior performance.) Notice that deterministic crowding has some algorithmic characteristics in common with SEAM.

- Initialize population.
- Repeat until stopping condition:
  - Pick two parents, p1 & p2, at random from the population.
  - Produce a pair of offspring, c1 & c2, using recombination, and mutation.
  - Pair-up each offspring with one parent according to the pairing rule below.
  - For each parent/offspring pair, if the offspring is fitter than the parent then replace the parent with the offspring.

**Pairing rule**: if H(p1,c1)+H(p2,c2) <H(p1,c2)+H(p2,c1) then pair p1 with c1, and p2 with c2, else pair p1 with c2, and p2 with c1, where H gives the genotypic Hamming distance between two individuals.[4]

**Figure 7:** Pseudocode for a simple form of a GA using deterministic crowding.

The GA is tested using uniform and one-point crossover.[5] A population size of 2000 is used; crossover is applied with probability 0.7. We tested mutation rates of ***mut***=0/128, 1/128, 2/128, 4/128, 8/128, and 16/128. The best performance for uniform crossover was with ***mut***=0 because (since DC maintains diversity appropriately) the mixing of bits from strings that disagree on building-blocks provides appropriate variation. The best performance of one-point crossover was with ***mut***=4/128=0.031.

The pseudocode for SEAM was given in Figure 5. The parameters we use are: number of features, N=128, alphabet of features, S={0,1}, initial population size 256 one-feature entities covering all alleles at all loci[6], number of contexts used for dominance test, t=200, (empirically, on average less than 10 of these are required to reveal the instability of a proposed join). The stopping condition is that $3 \cdot 10^6$ calls to the fitness function have been used. ***f***, the fitness function, is provided by Shuffled HIFF with 128 binary features.

### 6.2 Control experiments

Recall that the three main conceptual features of SEAM are: the use of variable size entities and a variation operator based-on composition; testing (in)stability of compositions by testing for Pareto dominance of composition over the component entities; and, building environmental contexts from other co-adapting entities in the ecosystem.

[4] Deterministic crowding explicitly uses genotypic similarity as a metric for similarity. This is fortuitously appropriate for maintaining diversity in HIFF. In contrast, SEAM makes no such assumption and uses no such measure on genotypic similarity.

[5] One-point crossover takes genes from parent 1 on the left of a single randomly positioned crossover point, and from parent 2 on the right of this crossover point, or vice versa. Uniform crossover takes each gene from either parent with equal probability independent of position. One-point crossover is a model of strong genetic linkage, and uniform crossover models no genetic linkage (see Watson 2002 for discussion).

[6] This may be done systematically for practical purposes, but may in principle be done without knowledge of the encoding dimensions by 'over-generating' the initial population and then removing duplicates—more specifically, by removing entities that behave the same (produce identical fitness changes) over a sample of random contexts (see Watson & Pollack 2001b).

RMHC and the GA provide some controls for the first of these. That is, they use a fully-specified feature set for each entity/individual, and use mutation and sexual recombination instead of composition. In preliminary work we also tested the operation of an algorithm that is the same as SEAM except that instead of using the Pareto dominance test, the second feature of SEAM, it simply determines that the join is unstable if the average fitness of either component is greater than the average fitness of the composite over the set of equally weighted environmental contexts. We also tested a control for the third feature of SEAM, by using an algorithm that is the same as SEAM except that it uses random feature sets for the contexts instead of contexts built from other members of the ecosystem. In both these latter two controls, sub-optimal associations are made and the entities 'fill-up', or 'bloat', with sub-optimal feature values—thus defeating the composition operator (Watson & Pollack 1999c). Overall, their performance is much like that of RMHC. A number of variations on SEAM, related experiments, and discussion are provided in (Watson 2002).

### 6.3 Results

Performance is measured by the fitness of the best string evaluated (in the preceding 1000 evaluations) averaged over 30 runs for each algorithm. For SEAM the strings evaluated are the groups of entities (i.e. an entity with its contextual environment), forming a complete feature specification. The problem size of 128 bits gives a maximum fitness of 1024. The performance curve for SEAM is truncated when 95% of runs (29/30) have found either of the two global optima.

As Figure 8 shows, the results for SEAM are clearly qualitatively different from the other algorithms: Whereas innovation by mutation and by conventional evolutionary algorithms becomes increasingly more difficult as evolution continues in this problem, innovation by composition is not impeded, and actually shows an inverted performance curve compared to all other methods tested. SEAM finds *both* global optima in *all* 30 runs. None of the other methods find *either* global optimum in *any* of the 30 runs.

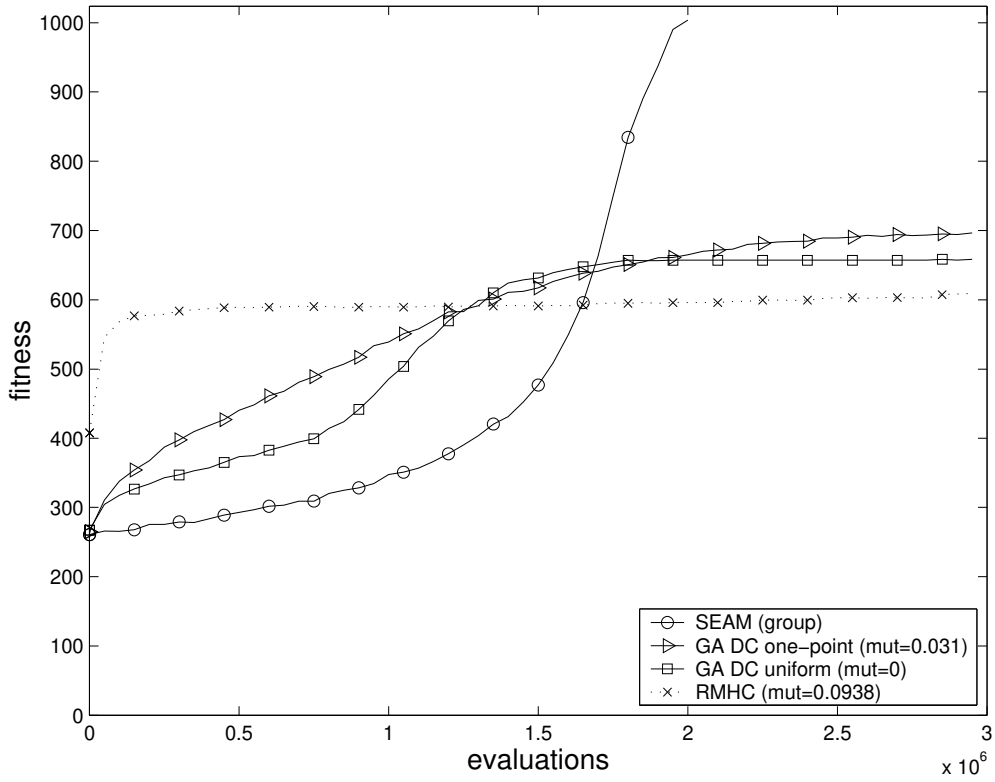



**Figure 8:** Performance of SEAM, GA with Deterministic Crowding (using one-point and uniform crossover), and Random Mutation Hill-Climbing, on Shuffled HIFF.

In Figure 9 we show the size of the largest correct sub-block discovered by each method. The 'group' curve for SEAM is the size of the largest correct building-block in any group of entities when they are evaluated together as a contextual environment, (this corresponds to the fitness curve in Figure 8). The 'indiv.' curve for SEAM is the size of the largest correct building-block in any stable individual entity. We use a log scale on the size axis—thus, if the increase in size is proportional to extant size the curve would appear as a straight line. We can see clearly in this figure that unlike the conventional evolutionary algorithms, innovation by composition continues steadily in this problem, approaching a scale-invariant increase in size of correct building-blocks in individual entities.

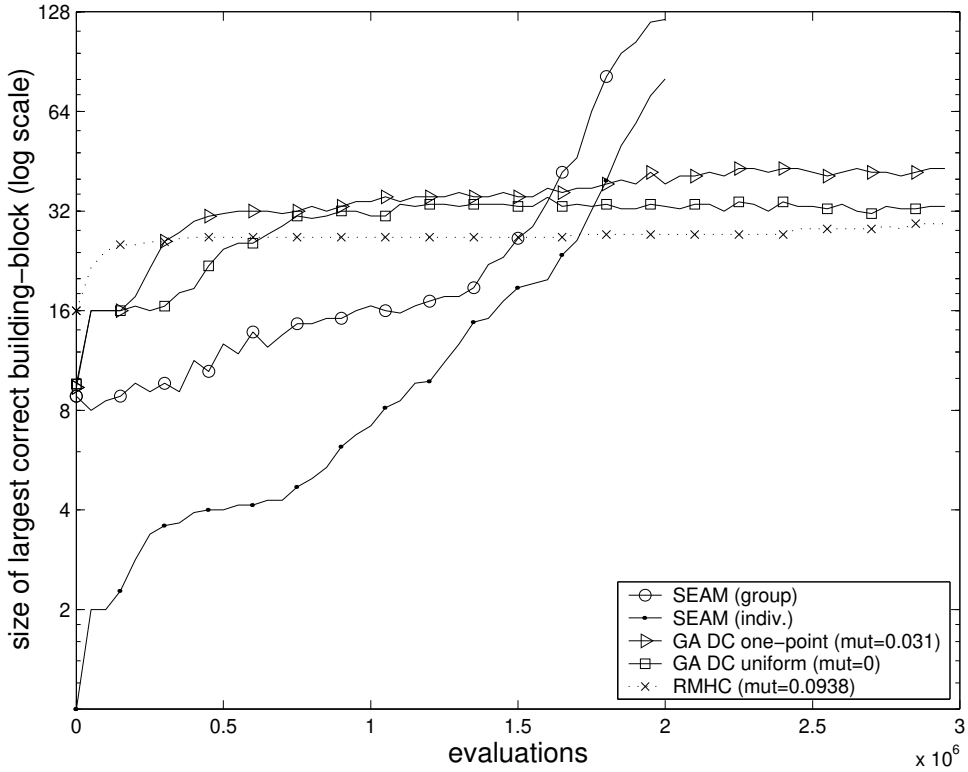



**Figure 9:** Size of largest correct building-block of features evolved (log scale) using SEAM, GA with Deterministic Crowding (using one-point and uniform crossover), and Random Mutation Hill-Climbing, on Shuffled HIFF.

### 6.4 Discussion

The SEAM model provides a concrete illustration of changing the unit of variation, and changing the unit of selection—as new entities are created they are selected for their abilities at that level of organisation and provide the components for entities at the next level of organisation. Clearly, the SEAM model operates by using a variation mechanism that scales-up with the size of extant entities, as illustrated in Figure 2. The model also illustrates how composition provides a divide and conquer problem decomposition of this class of problem by combining together solutions to small sets of features to find solutions to larger sets of features. Our RMHC results, and proofs in previous work (Watson 2001), show that no degree of random variation can provide continued innovation in this problem class. This indicates that the units of variation discovered by SEAM are not merely larger but are usefully informed by prior adaptation. Additionally, in the genetic algorithm experiments, the features of the individuals were subject to selection but not as independent entities—only as parts of a larger fully-specified feature set. This means that the subsets of features exchanged in sexual recombination are arbitrary, and accordingly do not provide meaningful modules. In contrast, because the entities in SEAM permit the unit of selection to scale lock-step with the unit of variation, the sets of

features provided by composition are not arbitrary, they are subject to selection as an integral whole, and provide meaningful units of variation.

In summary, the results show that mutation and sexual recombination are unable to exploit the decomposable structure of Shuffled HIFF or otherwise overcome the large-scale fitness saddles in the landscape. In contrast, the variables-sized entities in SEAM are able to each identify and represent a correct assembly of compatible features forming a useful building-block for the next hierarchical level. In evolutionary computation terms, SEAM describes an evolutionary algorithm where schemata of all sizes coevolve with one another, as if in a multi-player game, and cooperative groups are found incrementally from individual features through larger and larger schemata. With respect to the biological analogues, SEAM describes an ecosystem of entities that coevolve with one another, finding stable symbiotic relationships that satisfy their fitness dependencies with one another, and progress through successive evolutionary transitions, each occurring via the composition of simpler extant entities into more complex organisations.

### 6.5 Canalisation of successful groups

There is an interesting analogy between SEAM, the Baldwin effect (Baldwin 1896), and 'Symbiotic Scaffolding' (Watson & Pollack 1999b, Watson et al. 2000). That is, these scenarios have in common the feature that rapid non-heritable variation (lifetime learning or the temporary groups formed for contexts) guides a mechanism of relatively slow heritable variation (genetic mutation or composition, respectively). In other words, evaluation of entities in contextual groups 'primes' them for subsequent joins, or equivalently, solutions found first by groups are later *canalised* (Waddington 1942) by composite entities (see also Bull 1995). In Figure 9, the 'indiv.' curve shows how the discovery of correct building-blocks by individuals follows behind the discovery of correct building-blocks by groups.

## 7 Conclusions

Heritable variation is one of the fundamental axioms of evolutionary theory. However, it is a familiar irony that random variation is the source of new innovation but also inherently opposed to the heritability of extant complexity. Evolution has created mechanisms, such as enzymatic repair, that reduce the *error rate* (Nowak & Schuster 1989) and increase reproductive fidelity, but still, the question remains: How can it be the case that variation may be suppressed (by whatever mechanism) without also suppressing the opportunity for innovation? Differential reproduction is also not such a simple concept as it might first appear. Specifically, it requires us to delineate the entities involved—to identify the entities whose reproduction could be differentiated. There are many biological cases where the relevant reproductive units are not so obvious—more importantly, it may be in principle inaccurate to draw such boundaries.

Symbiotic composition offers an intriguing perspective on these issues. It is perfectly reasonable that a number of entities may each be individually stable and yet, via the discovery of successful compositions of these entities, there is still opportunity for innovation at a higher-level of organisation. Thus, composition presents no opposition between the stability or heritability of the component entities, and the opportunity for innovation in entities at the next level of organisation. And significantly, this view is enabled by a willingness to repeatedly re-define the boundary of the entities involved.

More concretely, the separation of a local optimum from the next best configuration of features is a fundamental limiting characteristic of adaptive landscapes, and saddle-crossing is a useful way to conceptualise the ability of an adaptive mechanism. But, what scale of fitness-saddle should we expect in a natural adaptive landscape? Intuitively, we might suspect that as one scale of ruggedness is overcome, a larger scale of ruggedness becomes the limiting characteristic of the adaptive landscape. If this is so then there is no fixed scale of saddle-crossing ability that is sufficient, and open-ended evolvability requires an adaptive mechanism that scales-up as evolution continues, enabling larger and larger 'jumps' in feature space.

In our experiments using a scale-invariant fitness landscape, we find that, as expected, both a mutation only algorithm and the GA have a limit to the size of fitness saddle that they can cross. More exactly, as adaptation continues and the distance to the next-best optimum increases, adaptation by these methods becomes increasingly difficult. In contrast, SEAM is able to discover the epistasis structure in the problem, use collections of features in different entities to represent it explicitly, and by searching combinations of these entities is able to continue to find successful combinations of features through many hierarchical levels. Accordingly, these experiments show that on this class of adaptive landscape, evolvability under mutation and sexual recombination within the accretive model of adaptation is inherently limited, whereas innovation by composition offers the possibility of inherently scalable, open-ended evolvability.

The Symbiogenic Evolutionary Adaptation Model provides a concrete illustration of *one* way to realise a scaling-up of the units of variation and selection characteristic of the major evolutionary transitions. SEAM abstracts away all population dynamics and uses a simple multi-context test to determine whether a composite will be stable. Resulting compositions are compatible with a selfish model of the entities, and the mechanism has appealing analogies with natural ecosystems, but the appropriateness of this model for multi-species competition in an ecosystem needs to be qualified. Also, scale-invariance is a property observed in many natural systems, but whether the natural adaptive environment has characteristics like those of the particular model that we developed in HIFF is an empirical matter.

In the meantime, we suggest that this algorithmic perspective on the formation of higher-level entities from the composition of simpler entities provides a useful facet in our understanding of the impact of the major evolutionary transitions.

## Acknowledgments

The authors are indebted to many members of the DEMO at Brandies, especially Sevan Ficici, Shivakumar Viswanathan, Hod Lipson, Edwin de Jong and Anthony Bucci, and also to John Wakeley for much detailed discussion on related issues. We would also like to thank the reviewers for detailing significant improvements.

## References

Angeline, PJ, & Pollack, JB, 1993, "Evolutionary Module Acquisition", *Procs. of The Second Annual Conference on Evolutionary Programming*, eds. Fogel, D. and Atmar, W, pp. 154-163.

Baldwin, JM, 1896, "A New Factor in Evolution," *American Naturalist*, **30**, 441-451.

Bak, P, 1996, *How nature works: the science of self-organized criticality*, NY, USA: Copernicus.

Bull, L, & Fogarty T, 1995, "Artificial Symbiogenesis", *Artificial Life 2*(3): 269-292.

Buss, LW, 1987, *The Evolution of Individuality*, Princeton University Press, New Jersey.

Conrad, M, 1990, "The Geometry of Evolution." *BioSystems* **24**: 61-81.

Cormen, TH, Leiserson, CE, Rivest, RL, 1991 *Introduction to Algorithms,* MIT Press, Camb. MA, McGraw-Hill Book Company, NY.

Dawkins, R, 1976, *The Selfish Gene*, Oxford University Press, NY.

Demur, R, 1995, "Evolution Through Cooperation: The Symbiotic Algorithm", *Procs. of Artificial Evolution Euro. Conf. 1995 (AE95).*

Frank, SA, "Models of symbiosis", *American Naturalist* 150:S80--S99.

Fonseca, CM, Fleming PJ, 1995, "An Overview of Evolutionary Algorithms in Multiobjective Optimization", *Evolutionary Computation*, Vol. 3, No.1, pp.1-16.

Forrest, S & Mitchell, M, 1993, "What makes a problem hard for a Genetic Algorithm? Some anomalous results and their explanation" *Machine Learning 13*, pp.285-319.

Gould, SJ & Vrba, E, 1982, "Exaptation—a missing term in the science of form", *Paleobiology*, **8**: 4-15. Ithaca

Goldberg, DE, Korb, B, & Deb K, 1989 "Messy Genetic Algorithms: Motivation, Analysis and First Results" *Complex Systems* 3, pp.493-530.

Gruau F, 1994, *Neural Network Synthesis Using Cellular Encoding and the Genetic Algorithm*. PhD thesis, Ecole Normale Superieure de Lyon.

Harik, GR, & Goldberg, DE, 1996, "Learning Linkage", *Foundations of Genetic Algorithms 4*, Morgan Kaufmann, San Mateo, CA.

Holland, JH, 1975, *Adaptation in Natural and Artificial Systems*, Ann Arbor, MI: The University of Michigan Press.

Holland, JH, & Reitman, JS, 1978, "Cognitive systems based on adaptive algorithms", in *Pattern Directed Inference Systems*, Waterman, DA, & Hayes-Roth, F, eds., New York: Academic Press, pp. 313-329.

Huynen, MA, Stadler, PF & Fontana, W, 1996, "Smoothness Within Ruggedness: The Role of Neutrality in Adaptation." *Proc. Natl. Acad. Sci. (USA),* 93: pp.397-401.

Juille, H, & Pollack, JB, 1996, "Coevolving Intertwined Spirals", *Procs. of fifth annual conference on Evolutionary Programming*, pp.461-468. MIT press.

Kauffman, S, 1993, *The Origins of Order*, Oxford University Press.

Kimura, M, 1983, *The Neutral Theory of Molecular Evolution*, Cambridge University Press.

Koza, JR, 1992, *Genetic Programming: On the Programming of Computers by Means of Natural Selection*, Cambridge, MA: MIT Press.

Koza, JR, 1994, *Genetic Programming II: Automatic Discovery of Reusable Programs,* Cambridge, MA: MIT Press.

Lewontin, RC, 2000 *The Triple Helix: Gene, Organism and Environment,* Harvard U., Camb., MA.

Mahfoud, S, 1995, *"Niching Methods for Genetic Algorithms"*, PhD thesis, University of Illinois, also IlliGAl Report No. 95001.

Margulis, L, 1993a, *Symbiosis in Cell Evolution*, 2nd ed, WH Freeman and Co.

Margulis, L, 1993b, "Origins of Species: Acquired Genomes and Individuality", *BioSystems*, Vol. 31 (2-3) pp. 121-125.

Maynard Smith, J, & Szathmary, E, 1993, "The origin of chromosomes I. Selection of linkage", *Journal of Theoretical Biology* **164**: 437-66.

Maynard Smith, JM & Szathmary, E, 1995 *The Major Transitions in Evolution*, WH Freeman.

Mazodier, P, & Davies, J, 1991, "Gene transfer between distantly related bacteria", *Annual Review of Genetics* 25:147-171.

Michod, RE, 1999. *Darwinian Dynamics, Evolutionary Transitions in Fitness and Individuality*. Princeton Univ. Press.

Mitchell, M, 1996, *An Introduction to Genetic Algorithms*, MIT Press, Cambridge, MA.

Nehaniv, C, & Rhodes, JL, 2000, "The evolution and understanding of hierarchical complexity in biology from an algebraic perspective", *Artificial Life 6,* 1: 45-67, MIT press.

Nowak M, & Schuster, P, 1989, "Error Thresholds of Replication in Finite Populations, Mutation Frequencies and the Onset of Muller's Ratchet", *J. Theor. Biol.*, 137:375-395.

Oates M, & Corne, D, 2001, "Overcoming Fitness Barriers in Multi-modal Search Spaces", *Foundations of Genetic Algorithms VI,* (2000) Martin WN and Spears WM, eds., Morgan Kaufmann, published 2001, pp 37--50.

Ohno & Susumu, 1970. *Evolution by Gene Duplication,* New York: Springer-Verlag.

Potter, M, 1997, *The Design and Analysis of a Computational Model of Cooperative Coevolution*, PhD thesis, George Mason University, Fairfax, Virginia.

Rosca, JP, *Hierarchical Learning with Procedural Abstraction Mechanisms*, Ph.D. dissertation, University of Rochester, February 1997.

Simon, HA, 1969, *The Sciences of the Artificial,* Cambridge, MA. MIT Press.

Smith, MW, Feng, DF, & Doolittle, RF, 1992, "Evolution by Acquisition: the Case for Horizontal Gene Transfers", *Trends in Biochemical Sciences* 17(12):489-493.

Spears, WM, DeJong, KA, Baeck, T, Fogel, D, & de Garis, H, 1993, "An Overview of Evolutionary Computation", in *European Conference on Machine Learning (ECML93)*, pp. 442-459.

Waddington, CH, 1942, "Canalization of development and the inheritance of acquired characters", *Nature* **150**: 563-565.

Wagner, GP, & Altenberg, L, 1996, "Complex adaptations and the evolution of evolvability", *Evolution*, Vol. 50, No. 3, pp. 967-976.

Watson RA, 2001, "Analysis of Recombinative Algorithms on a Non-Separable Building-Block Problem", *Foundations of Genetic Algorithms VI,* (2000), eds. Martin WN and Spears WM, Morgan Kaufmann, published 2001, pp. 69-89.

Watson, RA, 2002, "Compositional Evolution: Interdisciplinary Investigations in Evolvability, Modularity, and Symbiosis, in Natural and Artificial Evolution", PhD dissertation, Brandeis University, in preparation.

Watson, RA, Hornby, GS & Pollack, JB, 1998, "Modeling Building-Block Interdependency", *Procs. of Parallel Problem Solving from Nature (PPSN) V*, Springer, pp.97-106.

Watson, RA, & Pollack, JB, 1999a, "Hierarchically-Consistent Test Problems for Genetic Algorithms", *Procs. of 1999 Congress on Evolutionary Computation (CEC)*, Angeline, et al. eds. IEEE Press, pp.1406-1413.

Watson, RA, & Pollack, JB, 1999b, "How Symbiosis Can Guide Evolution", *Procs. of European Conf. on Artificial Life (ECAL) V*, Floreano, D, Nicoud, J-D, Mondada, F, eds. Springer. pp. 29-38.

Watson, RA, & Pollack, JB, 1999c, "Incremental Commitment in Genetic Algorithms", *Procs. of Genetic and Evolutionary Computation Conference (GECCO) 1999*. Banzhaf, et al. eds., Morgan Kaufmann, pp.710-717.

Watson, RA, & Pollack, JB, 2000, "Symbiotic Combination as an Alternative to Sexual Recombination in Genetic Algorithms", *Procs. of Parallel Problem Solving from Nature (PPSN) VI*, Schoenauer et al. Springer pp.425-434.

Watson RA, & Pollack, JB, 2001a, "Coevolutionary Dynamics in a Minimal Substrate", *Procs. of Genetic and Evolutionary Computation Conference (GECCO) 2001*, eds. Spector, L., et al., Morgan Kaufmann. pp. 702-709.

Watson RA, & Pollack, JB, 2001b, "Symbiotic Composition and Evolvability", *Procs. of* , *European Conf. on Artificial Life (ECAL) VI,* Kelemen,, J, & Sosík, P, eds., Springer (LNAI 2159), pp. 480 -490.

Watson, RA, Reil, T, & Pollack JB 2000, "Mutualism, Parasitism, and Evolutionary Adaptation", *Procs. of Artificial Life VII*, Bedau, M, McCaskill, J, Packard, N, Rasmussen, S (eds.), 2000. MIT Press, pp. 170-178.

Werth, CR, Guttman, S.I. & Eshbaugh, W.H. (1985) "Recurring origins of allopolyploid species in Asplenium" *Science* 228, pp. 731-733.

Wright, S, 1967, "Surfaces of selective value", *Proc. Nat. Acad. Sci*. **58**: pp. 165-179.

Wright, S. 1931, "Evolution in Mendelian populations", *Genetics* **16**: pp. 97-159.