

# Symbiotic Composition and Evolvability

Richard A. Watson, and Jordan B. Pollack

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

**Abstract.** Several of the Major Transitions in natural evolution, 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. In this paper we present a very abstract model of ‘symbiotic composition’ to explore its possible impact on evolvability. A particular adaptive landscape is used to exemplify a class where symbiotic composition has an adaptive advantage with respect to evolution under mutation and sexual recombination. Whilst innovation using conventional evolutionary algorithms becomes increasingly more difficult as evolution continues in this problem, innovation via symbiotic composition continues through successive hierarchical levels unimpeded.

## Introduction

The Major Transitions in evolution [1,2,3] involve the creation of new higher-level complexes of simpler entities. Summarised by Michod for example [3], 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 Modern Synthesis preoccupation with the individual as the unit of selection, they involve the adoption of new modes of transmitting information, they address fundamental questions about individuality, cooperation, fitness, and not least, the origins of life [1,2,3].

In several of the transitions “entities that were capable of independent replication before the transition can replicate only as part of a larger whole after it” [2]. Although Maynard Smith and Szathmary identify some transitions which do not fit what they describe as “symbiosis followed by compartmentation and synchronised replication”, several of the transitions including the origin of eukaryotes from prokaryotes [4], and the origin of chromosomes from independent genes, do involve the quite literal union of previously free-living entities into a new whole. This paper focuses on the evolutionary impact of this mechanism, which we shall refer to as ‘symbiotic composition’, or simply ‘composition’. We are concerned with an *algorithmic* understanding of this mechanism the Major Transitions: What kind of adaptation does the formation of higher-level complexes from simpler entities afford? And we seek to understand the class of adaptive problem, the kind of fitness landscape, for which this mechanism is well suited.

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 an entity. But 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. This immediately suggests some concepts impacting evolvability.

First, a composition mechanism may potentially allow ‘jumps’ in feature space that may cross ‘fitness saddles’ [5] in the original adaptive landscape (defined by the mutation neighbourhood). Moreover, since these higher-level aggregations of features are not arbitrary but rather are shaped by prior adaptation, these jumps are not equivalent to large random mutations, but rather are ‘informed’ by prior or parallel adaptation.

Crossing fitness saddles has been a central issue in evolvability. There are many possible scenarios for how adaptation may overcome a fitness saddle: for example, genetic drift and ‘Shifting Balance Theory’ [5,6], exaptation [7], neutral networks [8], extra-dimensional bypass [9], ontogenic processes [10], or landscape dynamics [11]. Each of these affords an increase in the width of fitness saddle that might be crossed (with respect to that which may be crossed by simple mutation). And conceivably, some of them may produce saddle-crossing ability that is not arbitrary, but informed by prior or parallel adaptation. But, what is the size of fitness saddle that we should expect to encounter? It seems likely that as one scale of ruggedness is overcome, a larger scale of ruggedness becomes the limiting characteristic of a landscape. Under composition, the entities resulting from one level of organisation provide a new ‘unit of variation’ for compositions at the next level, and thus the size of jumps is proportional to extant complexity. In this sense, composition suggests a scale-invariant adaptive mechanism.

Second, in composition, the sets of features that are composed are pre-adapted in independent entities. The components of the union arise from entities at a lower level ‘unit of selection’. This independence provides a ‘divide and conquer’ treatment of the feature set. Intuitively, the hope is that a generalist entity, utilising two different niches, resources, or habitats, for example, can be created by the composition of two specialist entities each independently adapted to one of these niches, resources or habitats. Thereby, the problem of being well adapted to the general habitat is divided into the independent problems of being well adapted to component habitats. This decomposition of a problem into smaller problems is known algorithmically as ‘divide and conquer’ (e.g. [12]); so named because of the significant algorithmic advantage it offers when applicable. Such divide and conquer advantage is not available to natural selection when features are adapted within a single reproductive entity.

The model we describe below develops these two concepts – a scalable mechanism of saddle-crossing, and divide and conquer advantage – both applied in a scale-invariant hierarchical fashion. We do not attempt to model biological mechanisms in any detailed way; our model is deliberately very abstract. For example, we assume a mechanism of symbiotic composition that simply produces the union of features from two organisms. And, the fitness landscape that we use for our experiments is deliberately chosen to exemplify the adaptive utility of composition as contrasted with conventional evolutionary algorithms. However, by using an abstract model we can focus on the combinatorial aspects of the processes, and an algorithmic model such as this provides an important facet to our understanding of the Major Evolutionary Transitions and the adaptational significance of symbiotic composition.

The remainder of the paper is structured as follows: Section 2 describes a scale-invariant fitness landscape; Section 3 describes our composition model, the Symbiogenic Evolutionary Adaptation Model (SEAM); Section 4 describes the results of some experiments with SEAM and this scale-invariant fitness landscape; Section 5 concludes.

## 2 A Scale-Invariant Fitness Landscape

In this section we examine a fitness landscape that we will use to exemplify the characteristics of the composition model we describe later. Of interest to us here is that this landscape has saddles at all scales, resulting from its hierarchical construction [13].

Sewell Wright [5] 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 concept of evolution as a combinatoric optimisation process on a rugged landscape [14], provides the now familiar model at the heart of the issues addressing evolvability. 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,  $F_1$  and  $F_2$ , each with two possible discrete values, 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 overlaid arrows in each case show possible paths of adaptation that improve in fitness by changing one feature at a time.

	Case 1		Case 2		Case 3		Case 4		Case 4b	
	$F_{2a}$	$F_{2b}$	$F_{2a}$	$F_{2b}$	$F_{2a}$	$F_{2b}$	$F_{2a}$	$F_{2b}$	$F_{2a}$	$F_{2b}$
$F_{1a}$	1	3	1	3	1	4	3	2	1	0
$F_{1b}$	2	4	2	5	2	3	1	4	0	1

**Table 1.** Example selective values 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  $F_1$ . Cases 2, 3 and 4 each show some epistasis but with different effects. In Cases 2 and 3, the landscape is not planar, and the possible routes of single-feature variation are different in Case 3, but the landscapes still only have 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  $F_1$ , does epistasis create two optima and a resultant fitness saddle. Changing from  $F_{1a}F_{2a}$  to  $F_{1b}F_{2b}$  without going through a lower fitness configuration requires changing two features at once. Lewontin ([11], 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 further simplification, we make the fitness values symmetric (Case 4b).

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.

		$F_1/F_2$	
		$F_2a$	$F_2b$
$F_1a$	1	0	
$F_1b$	0	1	

		$F_3/F_4$	
		$F_4a$	$F_4b$
$F_3a$	1	0	
$F_3b$	0	1	

		$MF_1/MF_2$	
		$MF_2e$	$MF_2f$
$MF_1c$	2	0	
$MF_1d$	0	2	

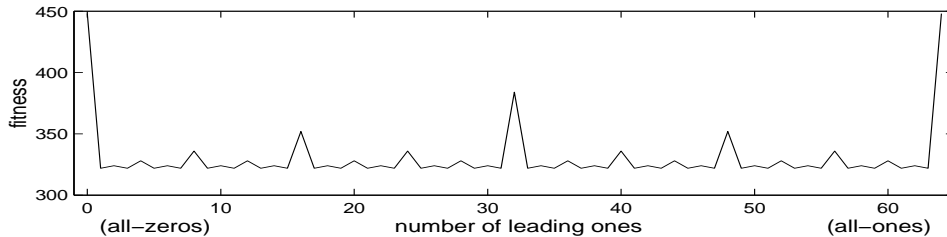
**Table 2.** Abstracting the interaction of two pairs of features,  $F_1/F_2$  and  $F_3/F_4$ , into the interaction of two ‘meta-features’  $MF_1/MF_2$ .

The fitness landscape resulting from this interaction at the bottom level, together with the interaction of pairs at the abstracted level, produces four local optima altogether. Using  $a=0$  and  $b=1$ , we can write these four optima as the strings 0000 and 1111, which are equally preferred, and, 0011 and 1100, which are equally preferred but less so. Naturally, we can take the two best-preferred configurations from the  $F_{1..4}$  system and describe a similar interaction with an  $F_{5..8}$  system, and so on. Equation 1 below, describes the fitness of a string of bits (corresponding to binary feature values, as above) using this construction. This function, which we call Hierarchical If-and-Only-If (HIFF), was first introduced in previous work as an alternative to functions such as ‘The Royal Roads’ and ‘N-K landscapes’, (see [15]).

$$F(\mathbf{B}) = \begin{cases} 1, & \text{if } |\mathbf{B}|=1, \\ |\mathbf{B}| + F(\mathbf{B}_L) + F(\mathbf{B}_R), & \text{if } (|\mathbf{B}|>1) \text{ and } (\forall i\{b_i=0\} \text{ or } \forall i\{b_i=1\}) \\ F(\mathbf{B}_L) + F(\mathbf{B}_R), & \text{otherwise.} \end{cases} \quad \text{Eq.1}$$

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

A 64-feature landscape using HIFF, as used in our experiments, has  $2^{32}$  local optima (for adaptation changing one feature at a time) [16], only two of these give maximal fitness. To jump from either of the second-best optima (at half-ones and half-zeros) to either global optimum requires changing 32 features at once. Thus, an algorithm using only mutation cannot be guaranteed to succeed in less than time exponential in the number of features [16]. A particular section through the fitness landscape is shown in Figure 1 – the section runs from one global optimum to the other at the opposite corner of the hyperspace (see [13]). As is clearly seen in the fractal nature of the curve in Figure 1, the local optima create fitness saddles that are scale-invariant in structure: that is, the nature of the ruggedness is the same at all resolutions.



**Fig. 1.** A section through a ‘scale-invariant’ HIFF fitness landscape.

The hierarchical structure of the HIFF landscape makes the problem recursively decomposable. For example, a 64-feature problem is composed of two 32-feature problems, each of which has two optima. If both of these optima can be found for both of these subproblems then a global optima 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  $2^{32} + 2^{32} + 4$ . Compared with the original  $2^{64}$  configuration space, even this two-level decomposition is a considerable saving. In addition, each size 32-problem may be recursively decomposed giving a further reduction of the search space. In [16] 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. We call this the ‘Shuffled-HIFF’ landscape [15] 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 [17]).

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

### 3 The Composition model

In this section we examine a simple abstraction of composition, the Symbiogenic Evolutionary Adaptation Model, or SEAM, (to invoke the idea of symbiotic union). This model was first introduced in [17] as a variant of abstract genetic algorithms.

SEAM manipulates strings of binary features, corresponding to feature combinations as discussed in the previous section. These strings, each representing a type of entity, have variable ‘size’ in that the number of specified features on the string is variable: for example, in an 8-feature system,  $F_3bF_5a$  is represented by the string “--1-0---”, where “-” represents a “null” feature for which the entity has no value, or is neutral. The primary mechanism of the model is a mechanism of composing two partial strings together to form a new composite entity. This is illustrated in Figure 2. Such a mechanism assumes the availability of a means of symbiotic union such as the formation of a cell membrane, horizontal gene-transfer [19,20], endosymbiosis [4], allopolyploidy (having chromosome sets from different species) [21], or any other means of encapsulating genetic material from previously independent entities such that they will be co-dispersed in future. However, assuming the availability of such a mechanism does not preclude the interesting biological question of when such a mechanism would provide an adaptive advantage, and when such a union will be selected for.

$$\begin{array}{r}
 \text{A: } \text{---1---0-1--} \\
 \text{B: } \text{--1--0--0-----} \\
 \hline
 \text{A+B: } \text{--1-10--00--1--}
 \end{array}
 \qquad
 \begin{array}{r}
 \text{A: } \text{---1---00-1--} \\
 \text{B: } \text{--1-0--0-1----} \\
 \hline
 \text{A+B: } \text{--1-1---000-1--}
 \end{array}$$

**Fig. 2.** ‘Symbiotic composition’. Left) Union of two variable size entities, A and B, produces a composite, A+B, that is twice the size of the donor entities with the sum of their characteristics. The composite is created by taking specified features from either donor where available. Right) Where conflicts in specified features occur we resolve all conflicts in favour of one donor, e.g. A.

Algebraically, we define the composition of two entities A and B, as the superposition of A on B, below.  $A=\{A_1,A_2,\dots,A_n\}$ , 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:

$$S(A,B)=S(\{A_1,A_2,\dots,A_n\},\{B_1,B_2,\dots,B_n\})=\{s(A_1,B_1),s(A_2,B_2),\dots,s(A_n,B_n)\}, \quad \text{Eq.2.}$$

where,  $s(a,b)=a$ , if  $a \neq \text{null}$ ,  $s(a,b)=b$ , otherwise. Having defined a variation operator that defines a union of two entities we need to determine whether such a union would be selected for. 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. But, 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 symbiont may depreciate their 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 simply a complete set of features (Figure 3).

$$\begin{array}{ll}
 \text{---0-11---110--- } & \mathbf{x}, \quad \text{an entity specifies a partial set of feature values.} \\
 \underline{0110101100010011} & \boldsymbol{\theta}, \quad \text{an ‘environmental context’ is a complete set of features.} \\
 0110111100110011 & S(\mathbf{x},\boldsymbol{\theta}), \quad \text{the entity } \mathbf{x} \text{ superimposed on the context } \boldsymbol{\theta}.
 \end{array}$$

**Fig. 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. 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. But fortunately, the concept of *Pareto dominance* is specifically designed for application in cases where the relative importance of a number of factors is unknown [21]. 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 ‘ $x \gg y$ ’, and:

$$x \gg y \Leftrightarrow \{ \forall \theta : csf(x, \theta) \geq csf(y, \theta) \} \wedge \{ \exists \theta : csf(x, \theta) > csf(y, \theta) \}$$

where, for our ecological application,  $csf(p, q)$  is the ‘context sensitive fitness’ of entity  $p$  in context  $q$ . So crudely, if  $x$  is fitter (or at least as fit as)  $y$  in all possible environments then, regardless of the weightings of the environments for each entity we know that the overall fitness of  $x$  is greater than that of  $y$ . This pair-wise comparison of two entities over a number of contexts will be used to determine whether a symbiotic union produces a stable composite. If we write the union of entities  $a$  and  $b$  as  $a+b$ , then using the notion of Pareto dominance,  $a+b$  is stable if and only if  $a+b \gg a$ , and  $a+b \gg b$ . In other words,  $a+b$  is *unstable* if there is any context in which either  $a$  or  $b$  is fitter than  $a+b$ .

$$\text{i.e. } stable(a+b, a, b) \equiv (a+b \gg a) \wedge (a+b \gg b),$$

$$\text{i.e. } unstable(a+b, a, b) \Leftrightarrow \{ \exists \theta \in C : (csf(a, \theta) > csf(a+b, \theta)) \vee (csf(b, \theta) > csf(a+b, \theta)) \},$$

where  $C$  is a set of complete feature specifications. For our purposes,  $csf(x, \theta) > csf(y, \theta) \Leftrightarrow f(S(x, \theta)) > f(S(y, \theta))$ , where  $f(w)$  is the objective fitness of the complete feature set  $w$  as given by the fitness function. Thus our condition of instability becomes:

$$unstable(a+b, a, b) \Leftrightarrow \{ \exists \theta \in C : (f(S(a, \theta)) > f(S(a+b, \theta))) \vee (f(S(b, \theta)) > f(S(a+b, \theta))) \}$$

**Eq.3.**

We will build each context in the set of contexts by the temporary superposition of other members of the ecosystem. Algebraically, we define a context using the recursive function  $S^*$ , from a set of  $n \geq 2$  entities  $X_1, X_2, \dots, X_n$ , as follows:

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

where  $S(X_1, X_2)$  is the superposition of two entities as per Eq.1 above. Some contexts may require more or less 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.

An interesting analogy for this group evaluation is made with the Baldwin effect [22], and ‘Symbiotic Scaffolding’ [23,24]. 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). In other words, evaluation of entities in contextual groups ‘primes’ them for subsequent joins, or equivalently, solutions found first by groups are later *canalised* [9] by permanent composite entities.

Figure 4 uses Equations 2, 3 and 4 to define a simple version of the Symbiogenic Evolutionary Adaptation Model.

- Initialise ecosystem,  $E$ , with random, single-feature, entities.
- Repeat until *stopping condition*:
  - Remove two entities at random from the ecosystem  $\rightarrow a \ \& \ b$ .
  - Produce their union,  $a+b=S(a,b)$ , using composition (see Eq.2).
  - If *unstable*( $a+b, a, b$ ) return  $a$  and  $b$  to ecosystem, else add  $a+b$  to ecosystem.

where (as in Eq. 3)

$unstable(a+b, a, b) \Leftrightarrow \{\exists \theta \in C: (f(S(a, \theta)) > f(S(a+b, \theta))) \vee (f(S(b, \theta)) > f(S(a+b, \theta)))\}$   
 and  $C$  is a random set of contexts each built using  $S^*(E)$  (see Eq.4).

**Fig. 4.** Pseudocode for a simple version of SEAM.

In the implementations used in the following experiments the maximum number of contexts used in the stability test is 50 (although unstable joins are usually detected in about 6 trials on average). Since the Pareto test of stability abstracts away all population dynamics, only one instance of each type of entity need be maintained in the ecosystem. Initialisation needs to completely cover the set of single-feature ‘atoms’ so that all values for all features are available in the ecosystem. This can be done systematically, or as in our experiments, by over-generating randomly and then removing duplicates.

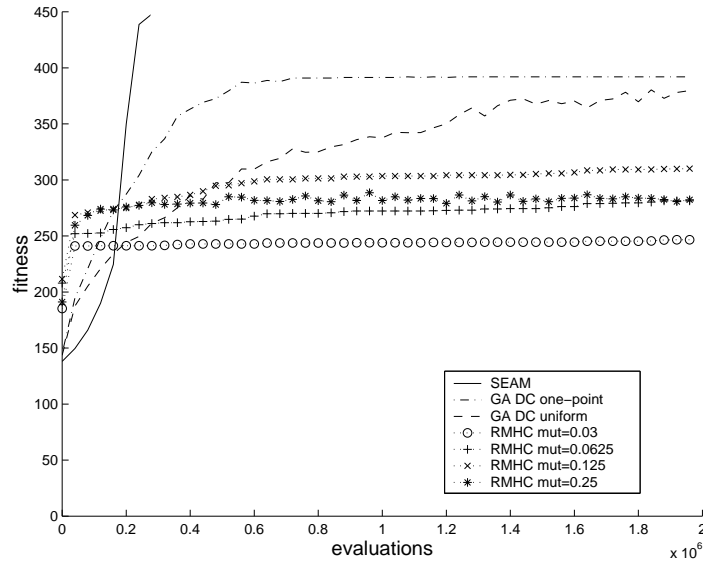
## 4 Experimental Results

We show empirical results of SEAM applied to a 64-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 accumulation of mutation and conventional abstractions of population-based evolutionary search. 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. RMHC repeatedly applies mutation to the features of a single string and accepts a variant if it is fitter [25]. We show results for various mutation rates (probability of assigning a new random value  $\{0,1\}$  to each feature). The GA is steady state using Deterministic Crowding to maintain diversity in the population (see [17], for details). The GA is tested using uniform and one-point crossover. A population size of 1000 is used; crossover is applied with probability 0.7; and mutation is applied with 0.03 probability of assigning a new random allele to each feature (0 or 1 with equal probability). SEAM uses the algorithm described in Figure 4. An ecosystem is initialized to the 128 unique entities (for a 64-bit problem). Symbiotic composition (Figure 2) is applied in all cases, no mutation is required. Performance is measured by the fitness of the best string evaluated (in the preceding 500 evaluations) averaged over 30 runs for each algorithm. The



problem size of 64 bits gives a maximum fitness of 448. The data for SEAM are terminated when all runs have found both global optima.

As Figure 5 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 progresses unimpeded through successive hierarchical levels, and actually shows an inverted performance curve compared to the other methods.



**Fig. 5.** Performance of SEAM, regular GA with Deterministic Crowding (using one-point and uniform crossover), and Random Mutation Hill-Climbing, on Shuffled HIFF. Error bars are omitted for clarity, but the differences in means are, as they appear, significant.

## 5 Discussion and Conclusions

Scale-invariance is a property observed in many natural systems, but whether the natural adaptive environment has characteristics like those of the particular landscape that we use is an empirical matter. Also, SEAM abstracts away all population dynamics and uses a simple multi-context test to determine whether a composite will be stable. Resulting joins are compatible with a selfish model of the entities, but whether this is an appropriate model for multi-species competition in an ecosystem needs to be justified. The experiments reported here involve only one adaptive landscape, selected to exemplify the utility of composition, and composition is modeled using one particular set of assumptions. The behaviour of the models must be examined in more general cases. In the meantime, the model we have outlined provides a concrete illustration of some important concepts raised in the introduction. Specifically: informed jumps in feature space – the fitness saddles in the landscape require large jumps in feature space and no amount of random mutation can provide this appropriately; a variation mechanism (and

contextual evaluation) that scales-up with the size of extant entities; and divide and conquer problem decomposition by combining together solutions to small sets of features to find solutions to larger sets of features.

Some or all of these characteristics could conceivably be provided by other mechanisms. Investigation of these concepts provides valuable insight for issues of evolvability and our model shows that there is a way in which these interesting characteristics could be provided by composition. 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 evolutionary impact of the Major Evolutionary Transitions.

## References

1. Buss, LW, 1987, *The Evolution of Individuality*, Princeton University Press, New Jersey.
2. Maynard Smith, JM & Szathmari, E, 1995 *The Major Transition in Evolution*, WH Freeman.
3. Michod, RE 1999. *Darwinian Dynamics, Evolutionary Transitions in Fitness and Individuality*. Princeton Univ. Press.
4. Margulis, L, 1993 *Symbiosis in cell evolution*, 2nd ed, WH Freeman and Co.
5. Wright, S. 1931, Evolution in Mendelian populations. *Genetics* **16**: 97-159.
6. Wright, S, 1977, *Evolution and the Genetics of Populations, Volume 3*, U. of Chicago Press.
7. Gould, SJ & Vrba, E, 1982, "Exaptation - a missing term in the science of form", *Paleobiology*, **8**: 4-15. Ithaca
8. Kimura, M. 1983, *The Neutral Theory of Molecular Evolution*, Cambridge University Press.
9. Conrad, M, 1990, "The Geometry of Evolution." *BioSystems* **24** (1990): 61-81.
10. Waddington, CH, 1942, "Canalization of development and the inheritance of acquired characters", *Nature* **150**: 563-565.
11. Lewontin, RC, 2000 *The Triple Helix: Gene, organism and Environment*, Harvard U. Press.
12. Cormen, TH, Leiserson, CE, Rivest, RL, 1991 *Introduction to Algorithms*, MIT Press.
13. Watson, RA, & Pollack, JB, 1999, "Hierarchically-Consistent Test Problems for Genetic Algorithms", *procs. of 1999 CEC*, Angeline, et al. eds. IEEE Press, pp.1406-1413.
14. Wright, S, 1967, "Surfaces of selective value", *Proc. Nat. Acad. Sci.* **58**, pp165-179 (1967).
15. Watson, RA, Hornby, GS & Pollack, JB, 1998, "Modeling Building-Block Interdependency", *procs. of PPSN V*, Springer 1998, pp.97-106 .
16. Watson RA, 2001, "Analysis of Recombinative Algorithms on a Non-Separable Building-Block Problem", *Foundations of Genetic Algorithms VI*, 2000, Martin WN, & Spears WM, Morgan Kaufmann.
17. Watson, RA, & Pollack, JB, 2000, "Symbiotic Combination as an Alternative to Sexual Recombination in Genetic Algorithms", *procs. PPSN VI*, Schoenauer et al. Springer, 425-434.
18. Fonseca, CM, Fleming PJ, 1995, "An Overview of Evolutionary Algorithms in Multiobjective Optimization", *Evolutionary Computation*, Vol. 3, No.1, pp.1-16.
19. Mazodier, P, and Davies, J, 1991, "Gene transfer between distantly related bacteria", *Annual Review of Genetics* **25**:147-171.
20. Smith, MW, Feng, DF, & Doolittle, RF, 1992, "Evolution by acquisition: the case for horizontal gene transfers", *Trends in Biochemical Sciences* **17**(12):489-493.
21. Werth, C.R., Guttman, S.I. & Eshbaugh, W.H. (1985) "Recurring origins of allopolyploid species in *Asplenium*" *Science* **228**, 731-733.
22. Baldwin, JM, 1896, "A New Factor in Evolution," *American Naturalist*, **30**, 441-451.
23. Watson, RA, & Pollack, JB, 1999, "How Symbiosis Can Guide Evolution", *procs. of ECAL V*, Floreano, D, Nicoud, J-D, Mondada, F, eds. Springer, 1999.
24. Watson, RA, Reil, T, & Pollack JB 2000, "Mutualism, Parasitism, and Evolutionary Adaptation", *procs. of Artificial Life VII*, Bedau, M, et al. eds.
25. 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.